Wavelet Techniques for Saddle Point Problems in Optimal Control

Angela Kunoth*
Institut für Angewandte Mathematik, Universität Bonn,
Wegelerstr. 6, 53115 Bonn, Germany
kunoth@iam.uni-bonn.de
www.iam.uni-bonn.de/∼kunoth
Fax: +49–228–73—7527

Andrew J. Kurdila
Department of Aerospace Engineering, Mechanics & Engineering Science
University of Florida, Gainesville, FL 32611-6250, USA
ajk@aero.ufl.edu
www.aero.ufl.edu/faculty/page/kurdila.html

June 4, 2001

Abstract

We consider optimal control problems involving the minimization of quadratic functionals subject to an elliptic partial differential equation as a constraint. As the control is exerted through the boundary conditions, the elliptic boundary value problem is formulated in weak form by appending the boundary conditions by means of Lagrange multipliers, leading to a saddle point problem. The cost functionals may contain different norms of the state and the natural norm for the boundary control. We use the concept of wavelets to derive an *equivalent* problem involving only ℓ_2 norms and operators. Consequently, certain finite submatrices associated with the minimization functionals are already well–conditioned in that their condition numbers do *not* depend on the discretization. This in turn entails that appropriate iterative solvers can be employed whose convergence speed does not deteriorate as the discretization step size becomes finer.

From the necessary conditions characterizing the solution of the optimal control problem, we derive a second saddle point problem as the adjoint problem. We show that the combined weakly coupled system consisting of the two saddle point problem defines an isomorphism. Furthermore, we derive circumstances under which this carries over to certain finite sections of the operator equations.

In order to solve the resulting coupled saddle point systems, we employ an iterative method consisting of an outer iteration based on a gradient method which alternatingly

*This author was funded by the Texas Higher Education Coordination Board, Advanced Research Program, and by the Partnership In Computational Science during her stays at the Institute for Scientific Computation, Texas A&M University, College Station, TX 77843–3404, U.S.A. This author also acknowledges support by the SFB256, Universität Bonn.

solves the two dual systems, and an inner iteration to solve each of the saddle point problems. We will see that the iterative strategy converges, provided that the inner systems are solved sufficiently well up to a certain accuracy. For the latter task, we use an incomplete Uzawa type algorithm. Finally we provide numerical results.

Key words: Optimal control, Lagrange multiplier, saddle point problem, wavelets, ℓ_2–problem, isomorphism, iterative method, outer iteration, inner iteration.

AMS subject classification: 65K10, 65F35, 49M07, 65N99.

1 Introduction

In order to solve numerically a stationary elliptic partial differential equation (with or without control), the discretization of its weak formulation typically leads, particularly in 3D, to a very large linear system of equations. The involved basis functions in a Galerkin formulation are often chosen to be finite elements of small support (relative to a grid of step size $h \sim 2^{-j}$) since the system matrix is highly structured and sparse with a total number of matrix entries proportional to the overall number of unknowns N. When the size of the matrix does not allow for the use of direct solvers due to storage restrictions, one has to resort to iterative techniques. Typically, these methods exploit the matrix structure and only need the same amount $\mathcal{O}(N)$ entries for storage as the matrix. In particular, one wants to employ methods that still provide the solution up to a desired degree of accuracy, like the accuracy provided by the discretization error, in a reasonable amount of iteration steps. Recall that the convergence speed of an iteration for symmetric positive definite systems like the Jacobi or the Gauss–Seidel method, or an acceleration based on the conjugate gradient method, is governed by the (spectral) condition number $\kappa(\mathbf{A})$ of the system matrix $\mathbf{A}$. That is, the methods become slower the greater the condition number $\kappa(\mathbf{A})$ is. For elliptic partial differential equations of order ρ on a domain in $\mathbb{R}^n$, a standard finite element discretization on a uniform grid yields a system matrix $\mathbf{A}$ where $\kappa(\mathbf{A})$ is proportional to $N^{2\rho/n}$. As the number of degrees of freedom grow with finer discretization step size $h \sim 2^{-j}$, the condition number grows like $2^{2j\rho}$ and the convergence speed starts to deteriorate.

For this reason, one often employs a *preconditioner* $\mathbf{C}$ for $\mathbf{A}$ which approximates the action of $\mathbf{A}^{-1}$ and is easy to implement and apply. In fact, a good preconditioner should be such that $\kappa(\mathbf{CA})$ is significantly smaller than $\kappa(\mathbf{A})$, that it requires only $\mathcal{O}(N)$ matrix entries for storage and that it can be applied with an amount of $\mathcal{O}(N)$ arithmetic operations. For elliptic boundary value problems there are by now two classes of techniques based on multilevel concepts that provide *asymptotically optimal preconditioners* in the sense that they satisfy the above requirements and are such that the resulting condition number does *not* depend on the discretization level j. Their common idea is that a finite–dimensional approximation of a function is made more efficient by simultaneously taking into account different discretization levels that capture coarse structures as well as finer details, without having to resolve everywhere at a very fine discretization level.

One class consists of multigrid methods, see e.g. [BHM, Ha2], which are classically formulated in terms of discretizations relative to grids having different grid spacing. The second class is comprised of basis–oriented approaches, approximating the solution of the underlying problem in terms of bases with respect to different resolutions. Of this latter class, one example is the finite–element based BPX-type preconditioner [BPX] which has been proven to be asymptotically optimal in [DK1, O1] and which has been developed primarily for dif-

ferential operators of positive order. In recent years also preconditioners based on *wavelets* [DK1, DPS, K1, Ja] have been increasingly employed.

The first results proving the optimality of the wavelet preconditioners were obtained in [DK1, O1, Ja] for elliptic boundary value problems discretized under the assumption that boundary conditions were incorporated into the solution and approximation spaces, thereby leading to symmetric positive definite systems of equations. In contrast, appending boundary conditions by Lagrange multipliers, a technique that was initiated in [Ba], leads to a saddle point problem. In [K1] an asymptotically optimal preconditioner based on wavelets has been constructed for the corresponding indefinite system. By now preconditioning based on wavelets has been used in many situations involving different kinds of differential and integral operators of both positive and negative order, see e.g. [D2].

In fact, one can isolate the question of how to construct a preconditioner for many kinds of linear operators using wavelets from the following observations. They are detailed here because the derivation is somewhat different from usual finite element techniques. Typical discretizations start out with a continuous operator equation (involving e.g. partial differential operators) of the form

$$\mathcal{N}\,U = F. \tag{1.1}$$

That is, given F, one wants to determine U belonging to some product of Hilbert spaces $\mathcal{H}$. Here $\mathcal{N} : \mathcal{H} \to \mathcal{H}'$ is assumed to be an *isomorphism* from $\mathcal{H}$ into its topological dual $\mathcal{H}'$ (with the corresponding dual form denoted by $\langle \cdot, \cdot \rangle_{\mathcal{H}' \times \mathcal{H}}$ or simply by $\langle \cdot, \cdot \rangle$). This means that $\mathcal{N}$ is invertible and satisfies the mapping property

$$c_{\mathcal{N}}\,\|V\|_{\mathcal{H}} \;\le\; \|\mathcal{N}V\|_{\mathcal{H}'} \;\le\; C_{\mathcal{N}}\,\|V\|_{\mathcal{H}}, \quad V \in \mathcal{H}, \tag{1.2}$$

where $c_{\mathcal{N}} \le C_{\mathcal{N}} < \infty$ are positive constants. We will often abbreviate equivalences of the form (1.2) as

$$\|\mathcal{N}V\|_{\mathcal{H}'} \;\sim\; \|V\|_{\mathcal{H}} \tag{1.3}$$

when the explicit form of the constants can be neglected. An equivalence of this sort, summarizing properties of the *operator*, is the first cornerstone in deriving asymptotically optimal preconditioners. The second ingredient exploits an important feature of wavelets, namely, the Riesz basis property. By *wavelets* we mean a collection of functions indexed from some infinite set $I\!I$

$$\Psi := \{\psi_\lambda : \lambda \in I\!I\} \subset \mathcal{H}. \tag{1.4}$$

The *Riesz basis property* now means that every function $v \in \mathcal{H}$ can be uniquely expanded in terms of Ψ,

$$v = \mathbf{v}^T\,\Psi := \sum_{\lambda \in I\!I} v_\lambda\,\psi_\lambda, \tag{1.5}$$

and its expansion coefficients satisfy the *norm equivalence*

$$\|v\|_{\mathcal{H}} \;\sim\; \|\mathbf{D}\mathbf{v}\|_{\ell_2} \tag{1.6}$$

where $\mathbf{D}$ is some diagonal matrix. Combining (1.6) with (1.3) yields a scaled discretization of the continuous operator $\mathcal{N}$ in terms of the (infinite) collection Ψ. The result is a discrete operator equation

$$\mathbf{N}\,\mathbf{U} = \mathbf{F} \tag{1.7}$$

satisfying

$$\|\mathbf{N}\,\mathbf{V}\|_{\ell_2} \;\sim\; \|\mathbf{V}\|_{\ell_2}, \quad \mathbf{V} \in \ell_2(I\!I), \tag{1.8}$$

see Section 3.2 below. Thus, $\mathbf{N}$ is well–posed in *Euclidean metric*. At this point, $\mathbf{N}$ is still infinite–dimensional.

In order to obtain a computable version of this operator equation, one has to extract certain finite sections of $\mathbf{N}$. This corresponds to choosing a finite–dimensional subspace of $\mathcal{H}$ on which $\mathcal{N}$ would be discretized by standard methods. Here one needs to assure that the finite sections of $\mathbf{N}$ are picked corresponding to *stable* discretizations. This means that the resulting finite–dimensional system

$$\mathbf{N}_\Lambda \, \mathbf{U}_\Lambda = \mathbf{F}_\Lambda \tag{1.9}$$

represented in terms of wavelets

$$\Psi_\Lambda = \{\psi_\lambda : \lambda \in \Lambda\} \tag{1.10}$$

corresponding so some *finite* subset $\Lambda \subset I\!I$ still satisfies the equivalence (1.8) with constants *independent* of Λ,

$$\|\mathbf{N}_\Lambda \, \mathbf{V}_\Lambda\|_{\ell_2} \;\sim\; \|\mathbf{V}_\Lambda\|_{\ell_2}, \quad \mathbf{V}_\Lambda \in \ell_2(\Lambda). \tag{1.11}$$

This entails that the spectral condition number $\kappa(\mathbf{N}_\Lambda)$ is bounded *uniformly* in Λ,

$$\kappa(\mathbf{N}_\Lambda) \;\lesssim\; 1. \tag{1.12}$$

Depending on the operator, there are different ways to ensure (1.11). If $\mathcal{N}$ is a symmetric elliptic system (involving differential or integral operators), the corresponding Galerkin scheme is already stable, implying (1.11). The problem that is considered in this paper is when $\mathcal{N}$ contains blocks of saddle point operators. In this case, additional conditions on the discretization involving an LBB condition have to be taken into account.

This general approach consists of the following four steps:

1.) Establishing the well–posedness of the continuous operator equation (1.3);

2.) Defining a discretization in terms of wavelets, leading to an infinite ℓ_2 system satisfying (1.8);

3.) Establishing the stability of the discretizations to ensure (1.11) for the finite system (1.9);

4.) Introducing an iterative solution of the system (1.9).

The above strategy has been used on several occasions [D2] and in particular in [K2, K4].

In the sequel, we will also follow this program. We will start out with formulating a general class of optimal control problems such that the resulting system provides on one hand an isomorphism. On the other hand, we try to be as flexible as possible with respect to the requirements of the control problem. Namely, we use the Lagrange multiplier approach to append the boundary control.

The strategy pursued here consists of treating the issues of discretization, preconditioning and discrete stability separately. We apply this method to treat single– and vector–valued elliptic partial differential equations in several space dimensions which are affected on (part of) the boundary by some control that, in turn, is determined by minimizing a certain quadratic functional.

We first recall some examples of two- and three-dimensional problems. Let $\Omega \subset \mathbb{R}^n$, $n \geq 2$, be a domain with Lipschitz boundary $\partial\Omega$. We will always assume that on some smooth subset $\Gamma \subseteq \partial\Omega$ of positive measure some boundary control of Dirichlet type u is imposed whereas on the remaining part $\Gamma_N = \partial\Omega \setminus \Gamma$ one has Neumann boundary conditions. Note that also Dirichlet boundary conditions can be included here by fixing the control u on that part of the boundary.

Example 1.1 On $\Omega \subset \mathbb{R}^3$ we consider a simple single–valued elliptic partial differential equation with governing equations given in strong form by

$$\begin{aligned} -\nabla^2 y + ky &= f && \text{in } \Omega, \\ y &= u && \text{on } \Gamma, \\ \partial_{\mathbf{n}} y &= 0 && \text{on } \Gamma_N. \end{aligned} \tag{1.13}$$

Here ∇ denotes the gradient, $k > 0$ is some fixed parameter, and $\partial_{\mathbf{n}}$ is the outward normal derivative at Γ_N. Given the right hand side f, the problem is to determine the control u such that (1.13) has a unique solution y and, in addition, a functional specified below is minimized.

Employing a *Lagrange multiplier* p [Ba], one can formulate the above equations in the following weak form: given $f \in (H^1(\Omega))'$, find $(y,p) \in H^1(\Omega) \times (H^{1/2}(\Gamma))'$ such that

$$\begin{aligned} \int_\Omega (\nabla y \cdot \nabla v \; + k\, yv)\, dx + \int_\Gamma p\, v\, d\Gamma &= \int_\Omega f\, v\, dx && \text{for all } v \in H^1(\Omega), \\ \int_\Gamma y\, q\, d\Gamma &= \int_\Gamma u\, q\, d\Gamma && \text{for all } q \in H^{-1/2}(\Gamma), \end{aligned} \tag{1.14}$$

is satisfied. Precise definitions of the Sobolev space $H^1(\Omega)$ and the dual of the trace space $H^{1/2}(\Gamma)$ of $H^1(\Omega)$, $(H^{1/2}(\Gamma))'$, and other Sobolev spaces used below can be found e.g. in [Ad, OC]. If $\Gamma = \partial\Omega$, one has $(H^{1/2}(\Gamma))' = H^{-1/2}(\Gamma)$. Note that (1.14) are the optimality conditions of the *saddle point problem*

$$\inf_{v \in H^1(\Omega)} \sup_{q \in (H^{1/2}(\Gamma))'} \tfrac{1}{2}\, a(v,v) - \langle f, v\rangle_\Omega + b(v,q) - \langle u, q\rangle_\Gamma \tag{1.15}$$

where

$$\begin{aligned} a(v,w) &:= \int_\Omega (\nabla v \cdot \nabla w \; + k\, vw)\, dx \\ b(v,q) &:= \int_\Gamma v\, q\, d\Gamma \end{aligned}$$

and $\langle \cdot, \cdot \rangle_\Omega$, $\langle \cdot, \cdot \rangle_\Gamma$, are the dual forms with respect to Ω and Γ. Now the governing equations can be rewritten as

$$\begin{aligned} a(y,v) + b(v,p) &= \langle f, v\rangle_\Omega && \text{for all } v \in H^1(\Omega), \\ b(y,q) &= \langle u, q\rangle_\Gamma && \text{for all } q \in (H^{1/2}(\Gamma))'. \end{aligned} \tag{1.16}$$

Assuming that $f \in (H^1(\Omega))'$ is given, the *optimal control problem* is then to find a boundary control $u \in U_{\text{ad}}$, $U_{\text{ad}} = H^{1/2}(\Gamma)$ denoting the set of admissible controls, such that a functional of the form

$$\mathcal{J}(y,u) := \tfrac{\omega}{2}\, \|y - y_\Omega\|^2_{H^s(\Omega)} + \tfrac{1}{2}\, \|u\|^2_{H^{1/2}(\Gamma)} \tag{1.17}$$

for some $s \in \mathbb{R}$ is minimized. Here y_Ω denotes a prescribed value on Ω, $\omega > 0$ is some fixed constant balancing the two norms and y is the solution of (1.14) generated by the control input u.

Example 1.2 As a second example, we consider the same problem as in Example 1.1 except for the functional $\mathcal{J}$ in (1.17) that is exchanged by

$$\mathcal{J}(y,u) := \tfrac{\omega}{2}\,\|y - y_\Gamma\|^2_{H^s(\Gamma_y)} + \tfrac{1}{2}\,\|u\|^2_{H^t(\Gamma)} \tag{1.18}$$

for $s, t \in \mathbb{R}$, and y_Γ is a fixed value on Γ. Such a functional is more appropriate to use when measurements of y can only be taken at the boundary $\Gamma_y \subset \partial\Omega$.

Example 1.3 Next we want to mention the problem of shape control of linearly elastic solids, see e.g. [OC], which is essentially a vector–valued generalization of the previous examples. Now $\Omega \subset \mathbb{R}^3$ denotes an elastic body (its bivariate copy is depicted in Figure 1 below). Employing the summation convention in this example, the equations governing the equilibrium can be written as

$$\frac{\partial \sigma_{ij}}{\partial x_i} + \varrho f_j = 0 \qquad \text{in } \Omega, \quad j = 1,2,3,$$

since Ω is an elastic continuum. As usual σ_{ij} are the stress components in the continuum, ϱ is the density and f_j are the components of the body force per unit volume. Applying the generalized Hooke's law, we write the stress tensor in terms of the strain tensor as

$$\sigma_{ij} = E_{ijkl}\varepsilon_{kl}$$

with ε_{kl} given by the linearized strain displacement relations

$$\varepsilon_{kl} = \tfrac{1}{2}\left(\tfrac{\partial y_k}{\partial x_l} + \tfrac{\partial y_l}{\partial x_k}\right),$$

and E_{ijkl} is Hooke's tensor, a tensor of elastic constants satisfying $E_{ijkl} = E_{jikl} = E_{ijlk} = E_{klij}$. With these definitions, the strong form of the equations governing the displacement $y = (y_1, y_2, y_3)^T$ in the elastic body can be derived as

$$-\tfrac{\partial}{\partial x_i}\left(E_{ijkl}\tfrac{\partial y_k}{\partial x_l}\right) = \varrho f_j, \qquad j = 1,2,3. \tag{1.19}$$

On the controlled surface boundary part Γ, displacement boundary conditions are imposed whereas on the remaining part Γ_N the displacements are uncontrolled and stress-free. Thus, with $\mathbf{n}$ denoting the outward normal unit vector along $\partial\Omega$, we can formulate the boundary conditions as

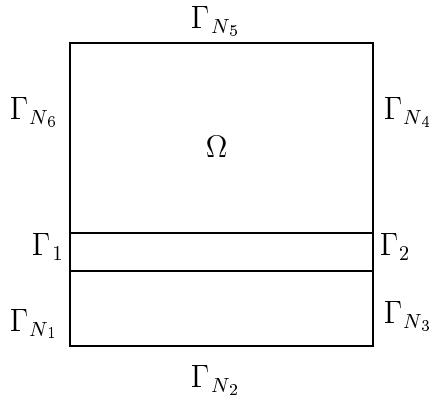
$$\begin{aligned} y &= u \qquad && \text{on } \Gamma, \\ n_i\, E_{ijkl}\,\tfrac{\partial y_k}{\partial x_l} &= 0 \qquad && \text{on } \Gamma_N,\ j = 1,2,3. \end{aligned} \tag{1.20}$$

The corresponding weak formulation of (1.19) with boundary conditions (1.20) appended by a Lagrange multiplier $p = (p_1, p_2, p_3)^T$ reads

$$\begin{aligned} \int_\Omega E_{ijkl}\tfrac{\partial y_k}{\partial x_l}\tfrac{\partial v_j}{\partial x_i}\,dx + \int_\Gamma p\cdot v\,d\Gamma &= \varrho\int_\Omega f\cdot v\,dx \qquad && \text{for all } v \in H^1(\Omega)^3, \\ \int_\Gamma y\cdot q\,d\Gamma &= \int_\Gamma u\cdot q\,d\Gamma \qquad && \text{for all } q \in (H^{1/2}(\Gamma)^3)', \end{aligned}$$

which can be written in compact form as

$$\begin{aligned} a(y,v) + b(v,p) &= \langle f, v\rangle_\Omega \qquad && \text{for all } v \in H^1(\Omega)^3, \\ b(y,q) &= \langle u, q\rangle_\Gamma \qquad && \text{for all } q \in (H^{1/2}(\Gamma)^3)' \end{aligned}$$



$$\Gamma = \Gamma_1 \cup \Gamma_2$$
$$\Gamma_N = \sum_{i=1}^{6} \Gamma_{N_i}$$

Figure 1: Sketch of Ω in two dimensions.

where dual forms are taken with respect to vector–valued quantities and

$$a(y,v) := \int_\Omega E_{ijkl} \frac{\partial y_k}{\partial x_l} \frac{\partial v_j}{\partial x_i}\, dx, \qquad b(v,q) := \int_\Gamma v \cdot q \, d\Gamma.$$

Given $f \in (H^1(\Omega)^3)'$ and a specified shape y_Γ on Γ_{N_5}, one shall determine the control $u \in H^{1/2}(\Gamma)^3 =: U_{\rm ad}$ such that the cost functional

$$\mathcal{J}(y,u) = \tfrac{\omega}{2} \, \|y - y_\Gamma\|^2_{H^s(\Gamma_{N_5})^3} + \tfrac{1}{2} \, \|u\|^2_{H^{1/2}(\Gamma)^3}$$

for some $s \in \mathbb{R}$ is minimized.

Example 1.4 Another example of the category we are interested in is the following specific transmission problem, a coupled *solid/fluid temperature control problem* similar to the one treated in [GL1].

Let the domain Ω with boundary $\partial\Omega$ consist of two parts, the *solid body domain* Ω_1 and the *fluid flow domain* Ω_2 which are connected by an interface wall Γ_y, see Figure 2 for a two-dimensional sketch of the domain and its boundaries. We want to find a control u along Γ ($\subset \partial\Omega_2$) in order to approximately match a desired temperature of a fluid along (part of) the interface Γ_y. That is, the temperature y is supposed to satisfy the energy equations

$$\begin{aligned} -\hat\kappa_1 \, \Delta y &= g_1 && \text{in } \Omega_1, \\ -\hat\kappa_2 \, \Delta y + (\vec w \cdot \nabla) y &= g_2(\vec w) && \text{in } \Omega_2, \end{aligned} \tag{1.21}$$

with boundary conditions

$$\begin{aligned} y &= u && \text{on } \Gamma, \\ \partial_{\mathbf{n}} y &= 0 && \text{on } \Gamma_r, \end{aligned} \tag{1.22}$$

where $\Gamma_r = \partial\Omega \setminus \Gamma$. The constants $\hat\kappa_1, \hat\kappa_2$ denote the thermal conductivity coefficients and g_1, g_2 are given functions where $g_2(\vec w)$ depends nonlinearly on $\vec w$. Here $\vec w : \Omega_2 \to \mathbb{R}^n$ is the fluid velocity field on Ω_2 which is determined beforehand (assuming that the viscous fluid is incompressible) by the Stokes equations on Ω_2. That is, $\vec w : \Omega_2 \to \mathbb{R}^n$ and the pressure $p : \Omega_2 \to \mathbb{R}$ are related by

$$\begin{aligned} -\nu \, \Delta \vec w + \nabla p &= \vec f && \text{in } \Omega_2, \\ \text{div } \vec w &= 0 && \text{in } \Omega_2, \end{aligned} \tag{1.23}$$

with boundary conditions

$$\begin{aligned} \vec{w} &= \vec{w}_0 && \text{on } \Gamma, \\ \vec{w} = \vec{0} \quad \text{on } \Gamma_y \cup \Gamma_b, \qquad \mathbf{n} \cdot \nabla \vec{w} &= \vec{0} && \text{on } \Gamma_o, \end{aligned} \tag{1.24}$$

where Γ and Γ_o are the control inflow and the outflow boundary, respectively. Furthermore, $\vec{w}_0$ is given, the constant ν is the kinematic viscosity coefficient of the fluid, and Γ_b the bottom of Ω_2. For the numerical solution of the Stokes equations (1.23) with (1.24) there are a variety of standard methods based on finite difference or finite element methods, or the Galerkin method using biorthogonal wavelets proposed in [DKU1] can be employed.

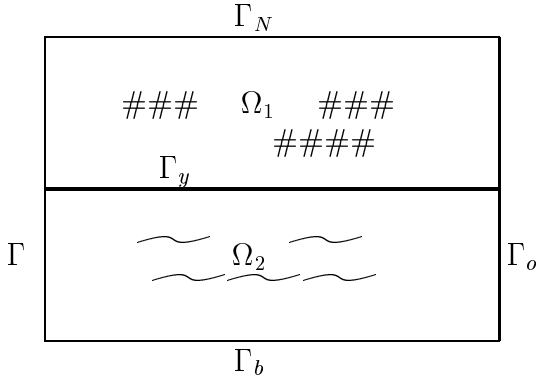


Figure 2: Sketch of Ω in two dimensions.

Then one is left with solving the equations for the temperature (1.21) in Ω, where on Ω_2 the velocity field $\vec{w}$ enters the differential equations as variable but known coefficient and in the source term on the right hand side.

In summary, the governing equations on $\Omega = \Omega_1 \cup \Gamma_y \cup \Omega_2$ read in weak form

$$\begin{aligned} a(y,v) + b(v,p) &= \langle g(\vec{w}), v \rangle_\Omega && \text{for all } v \in H^1(\Omega), \\ b(y,q) &= \langle u, q \rangle_\Gamma && \text{for all } q \in (H^{1/2}(\Gamma))' \end{aligned} \tag{1.25}$$

where

$$a(y,v) = \begin{cases} \hat{\kappa}_1 \int_{\Omega_1} \nabla y \cdot \nabla v \, d\vec{x} & \text{on} \quad \Omega_1, \\ \int_{\Omega_2} (\hat{\kappa}_2 \nabla y \cdot \nabla v \, + (\vec{w} \cdot \nabla) y \, v) d\vec{x} & \text{on} \quad \Omega_2, \end{cases} \tag{1.26}$$

the bilinear form $b(\cdot,\cdot)$ is just

$$b(v,q) = \langle v, q \rangle_\Gamma \tag{1.27}$$

and g comprises the right hand sides g_1 on Ω_1 and $g_2(\vec{w})$ on Ω_2.

The *optimal control problem* is now the following: find a control $u \in H^{1/2}(\Gamma) =: U_{\text{ad}}$ such that the cost functional

$$\mathcal{J}(y,u) = \tfrac{\omega}{2} \, \|y - y_{\Gamma_y}\|^2_{H^{1/2}(\Gamma_y)} + \tfrac{1}{2} \, \|u\|^2_{H^{1/2}(\Gamma)} \tag{1.28}$$

is minimized, where y_{Γ_y} is a prescribed value on Γ_y.

For such minimization problems like those discussed above, the specific choices for the form (1.16) or (1.25) are formulated in general terms as follows. The domain $\Omega \subset I\!R^n$ is an open bounded set with (Lipschitz) boundary $\partial\Omega = \Gamma \cup \Gamma_N$, where Γ is a smooth subset of

$\partial\Omega$ of positive Lebesgue measure on which the control is exerted. We explicitly express the restriction of a function v to Γ, usually denoted by $(\cdot)|_\Gamma$, in terms of the *trace operator* B,

$$Bv := v|_\Gamma. \tag{1.29}$$

Recall that for any $v \in H^s(\Omega)$, $1/2 < s < 3/2$, its trace Bv is known to be in $H^{s-1/2}(\Gamma)$, see e.g. [Gr] and Remark 2.2 below. For the special case $s = 1$, the bilinear form

$$b(v,q) := \langle Bv, q\rangle_{H^{1/2}(\Gamma)\times(H^{1/2}(\Gamma))'} = \langle Bv, q\rangle_\Gamma \tag{1.30}$$

is well–defined on $H^1(\Omega) \times (H^{1/2}(\Gamma))'$. We will always consider elliptic boundary value problems in the following weak form. Given $f \in (H^1(\Omega))'$ (and $u \in H^{1/2}(\Gamma)$ determined by the control problem), find $(y,p) \in H^1(\Omega) \times (H^{1/2}(\Gamma))'$ such that

$$\begin{aligned} a(y,v) + b(v,p) &= \langle f, v\rangle_\Omega \qquad \text{for all } v \in H^1(\Omega), \\ b(y,q) &= \langle u, q\rangle_\Gamma \qquad \text{for all } q \in (H^{1/2}(\Gamma))' \end{aligned} \tag{1.31}$$

holds, where $a(\cdot,\cdot)$ is a continuous bilinear form defined on $H^1(\Omega) \times H^1(\Omega)$. It will be convenient for later purposes to write (1.31) in operator form. Let a linear operator A be defined by

$$\langle Av, w\rangle_\Omega := a(v,w), \tag{1.32}$$

and let B' be the dual of B given by

$$\langle Bv, q\rangle_\Gamma = \langle v, B'q\rangle_\Omega = b(v,q). \tag{1.33}$$

Then (1.31) is rewritten as follows. Given $f \in (H^1(\Omega))'$ (and $u \in H^{1/2}(\Gamma)$ determined by the control problem), find $(y,p) \in H^1(\Omega) \times (H^{1/2}(\Gamma))'$ such that

$$\mathcal{L}\begin{pmatrix} y \\ p \end{pmatrix} := \begin{pmatrix} A & B' \\ B & 0 \end{pmatrix}\begin{pmatrix} y \\ p \end{pmatrix} = \begin{pmatrix} f \\ u \end{pmatrix}. \tag{1.34}$$

is satisfied. We call (1.34) in the sequel the *primal system.*

Remark 1.5 *In all cost functionals (1.17), (1.18), (1.28) we allow different norms for the observation of the state y on the domain Ω or on the boundary Γ_y, and also different norms for the control, including the natural ones. We want to stress that in principle the functionals may involve arbitrary Sobolev norms with smoothness s or t as long as they can be evaluated in terms of (biorthogonal) wavelets, i.e., a norm equivalence (1.6) holds. However, (unnecessarily) requiring higher smoothness than $s, t \le \frac{1}{2}$ for measuring on Γ_y, Γ or $s \le 1$ on Ω entails to assume higher regularity of the state y.*

Here we always require only the minimal amount of smoothness needed for having a well-posed problem.

Having derived this format of the constraints, the remainder of this paper is organized as follows. Section 2 deals with the continuous problem, treating a class of optimal control problems covering the previous examples. The constraints are formulated as the saddle problem (1.34). By standard arguments from control theory one can derive optimality conditions in terms of a second saddle point problem

$$\hat{\mathcal{L}}\begin{pmatrix} z \\ u \end{pmatrix} := \begin{pmatrix} A' & (\mathcal{R}_2 B)' \\ \mathcal{R}_2 B & 0 \end{pmatrix}\begin{pmatrix} z \\ u \end{pmatrix} = \begin{pmatrix} -T'\mathcal{R}_1(Ty - y_\Omega) \\ 0 \end{pmatrix} \tag{1.35}$$

called the *adjoint system*, and then show that the resulting weakly coupled system (1.34) and (1.35) admits a unique solution [K2].

However, we follow here a different path and first recall in Section 3 the necessary tools from wavelet theory which are employed in the sequel to discretize operator equations in terms of wavelets, leading to infinite ℓ_2 systems.

In Section 3.2, we will formulate the constraints (1.34) as an equivalent ℓ_2 problem and then derive a *representer* of the quadratic cost functionals, covering all examples mentioned above. We recall standard results from optimization to derive the optimality conditions for the discretized but yet infinite problem in Section 3.3. Conditions to ensure stability of discretizations for finite–dimensional sections of the involved operators are derived in Section 3.4. For the resulting weakly coupled system of saddle point problems, we describe in Section 4 an *inexact gradient method*, i.e., a fully iterative method to solve alternately the two saddle point problems by an Uzawa type method. We conclude in Section 5 with a numerical example.

2 Continuous Problems

This section introduces control problems in an abstract form with functionals covering the above examples. To this end, we firstly formulate the constraints (1.34) in weak form as a saddle point problem, recalling a few facts about abstract saddle point problems from e.g. [BF, GR].

2.1 An Abstract Saddle Point Problem

Let Y and Q be Hilbert spaces with their topological duals Y', Q' and dual forms $\langle\cdot,\cdot\rangle_{Y\times Y'}$, $\langle\cdot,\cdot\rangle_{Q\times Q'}$, respectively, which we often abbreviate by $\langle\cdot,\cdot\rangle$. We denote the norms on Y and Q and the induced inner products by $(\cdot,\cdot)_Y = \|\cdot\|_Y^2$ and $(\cdot,\cdot)_Q = \|\cdot\|_Q^2$. . Furthermore, the norm for the product space $\|\cdot\|_{Y\times Q}$ is always defined by

$$\left\| \begin{pmatrix} v \\ q \end{pmatrix} \right\|_{Y\times Q}^2 = \|v\|_Y^2 + \|q\|_Q^2. \tag{2.1}$$

Let $A : Y \to Y'$ and $B : Y \to Q'$ (with adjoint $B' : Q \to Y'$) be linear continuous operators. In view of the cases considered above in Section 1 where B is always a trace operator defined on some subset of $\partial\Omega$, it suffices here to restrict ourselves to the case of B being *surjective*, i.e., $\operatorname{range} B = Q'$ and $\ker B' = \{0\}$.

Assuming first for an instant that $u \in Q'$ is *given*, consider the following single saddle point problem: Given $(f,u) \in Y' \times Q'$, find the unique solution $(y,p) \in Y \times Q$ of

$$\begin{pmatrix} A & B' \\ B & 0 \end{pmatrix} \begin{pmatrix} y \\ p \end{pmatrix} = \begin{pmatrix} f \\ u \end{pmatrix}. \tag{2.2}$$

Concerning the existence and uniqueness of solutions of general saddle point problems, one has the following result, see e.g. [BF, GR].

Theorem 2.1 *Let the linear operator A be invertible on $\ker B \subseteq Y$, i.e., for some constant $\alpha_1 > 0$*

$$\inf_{v\in\ker B} \sup_{w\in\ker B} \frac{\langle Av, w\rangle_{Y'\times Y}}{\|v\|_Y\|w\|_Y} \;\geq\; \alpha_1, \qquad \inf_{v\in\ker B} \sup_{w\in\ker B} \frac{\langle A'v, w\rangle_{Y'\times Y}}{\|v\|_Y\|w\|_Y} \;\geq\; \alpha_1, \tag{2.3}$$

and let for some constant $\beta_1 > 0$ the inf–sup condition

$$\sup_{v \in Y} \frac{\langle Bv, q \rangle_{Q' \times Q}}{\|v\|_Y} \;\geq\; \beta_1 \|q\|_Q, \quad q \in Q, \tag{2.4}$$

hold. Then there exists a unique solution $(y, p) \in Y \times Q$ to problem (2.2) for given $f \in Y'$ and $u \in Q'$, i.e.,

$$\mathcal{L} := \begin{pmatrix} A & B' \\ B & 0 \end{pmatrix} \text{ is an isomorphism } Y \times Q \to Y' \times Q', \tag{2.5}$$

and one has the equivalence

$$\left\| \begin{pmatrix} v \\ q \end{pmatrix} \right\|_{Y \times Q} \;\sim\; \left\| \mathcal{L} \begin{pmatrix} v \\ q \end{pmatrix} \right\|_{Y' \times Q'} \tag{2.6}$$

for any $(v, q) \in Y \times Q$, where the constants are composed of α_1, β_1 and the continuity constants for A and B.

Recall that (2.4) is satisfied here because B is surjective.

So far the function u on the right hand side of (2.2) has been assumed to be known. We now treat the system (2.2) as constraints in an optimal control problem.

2.2 A Continuous Optimal Control Problem

In addition to Y and Q, two more Hilbert space Z, W are introduced with inner products $(\cdot,\cdot)_Z$, $(\cdot,\cdot)_W$ and induced norms $\|\cdot\|_Z$, $\|\cdot\|_W$, respectively. In order to treat the different norms for y and u in a unified way, $\|\cdot\|_Z$ and $\|\cdot\|_W$ will stand for the norms for the obversation of the state y and for u, that is, the cost functional (1.13) is firstly formulated in terms of the norms $\|\cdot\|_Z$ and $\|\cdot\|_W$. As for the role of Z, it will be assumed that there is a linear continuous operator

$$T : Y \to Z.$$

In view of the above examples where $Y = H^1(\Omega)$ and $Q = (H^{1/2}(\Gamma))'$ (or their vector–valued counterparts), we distinguish two choices of T, depending on whether the observation of the state is measured in a norm on Ω or on Γ_y.

Remark 2.2 *(i) Measuring state in a norm on Ω, then T is a Riesz map. A particular case is $Z = Y$, resulting in T just being the identity.*

(ii) When observations of the state can only be taken on (part of) the boundary $\partial\Omega$ denoted by Γ_y, Z is a trace space: $Z = H^{s-1/2}(\Gamma_y)$, and $T : H^s(\Omega) \to H^{s-1/2}(\Gamma_y)$ is the standard trace operator with respect to Γ_y, combined with a Riesz map. Recall that in this situation the Trace Theorem from e.g. [Gr] applies: for any $f \in H^s(\Omega)$, $1/2 < s < 3/2$, one can estimate

$$\|Tf\|_{H^{s-1/2}(\Gamma_y)} \;\lesssim\; \|f\|_{H^s(\Omega)}. \tag{2.7}$$

Conversely, for every $h \in H^{s-1/2}(\Gamma_y)$, there exists some $f \in H^s(\Omega)$ such that $Tf = h$ and

$$\|f\|_{H^s(\Omega)} \;\lesssim\; \|h\|_{H^{s-1/2}(\Gamma_y)}. \tag{2.8}$$

Extending previous results in [K4], one could now define the *minimization functional* as

$$\mathcal{J}(y,u) \;:=\; \tfrac{\omega}{2}\|y-y_\Omega\|_Z^2+\tfrac{1}{2}\|u\|_W^2 \tag{2.9}$$

involving some constant weight $0<\omega<\infty$ and some prescribed value y_Ω. The standard formulations of control problems would be to minimize (2.9) subject to the constraints (2.2).

However, at this point we deviate from this approach for the following reason.In many situations one is interested in the *qualitative* behavior of a minimization problem rather than in the specific values of particular norms used to express the cost functional. Consequently, then it suffices to replace the norms in the cost functional (2.9) by equivalent norms which are numerically easier to evaluate. In fact, we will formulate below a representer of (2.9) in terms of wavelets. Before we do so, we recall the necessary facts about wavelets.

3 Equivalent ℓ_2 Control Problems

3.1 Wavelets

We briefly collect from [D2] the mayor properties of wavelets that will be needed for the remainder of this paper.

To this end, let $I\!I := I\!I_1 \times \cdots \times I\!I_M$ be a product of infinite index sets. For each i, its elements λ of $I\!I_i$ consist of different types of indices such as the *level of resolution*, *refinement* or *discretization level* denoted by $|\lambda|$ and the spatial location. What we call *wavelets* for $\mathcal{H}=H_1\times\cdots\times H_M$ is a catenated collection of functions

$$\mathbf{\Psi} := \{{}^1\Psi,\ldots,{}^M\Psi\} \tag{3.1}$$

where for each $i=1,\ldots,M$ the collection ${}^i\Psi$ defined by

$${}^i\Psi := \{{}^i\psi_\lambda : \lambda\in I\!I_i\}\subset H_i \tag{3.2}$$

has the following properties:

(I) *Riesz basis property:* Every function v in H_i can be uniquely expanded in terms of ${}^i\Psi$,

$$v=\mathbf{v}^T\,{}^i\Psi := \sum_{\lambda\in I\!I_i} v_\lambda\,{}^i\psi_\lambda, \tag{3.3}$$

and its expansion coefficients satisfy the *norm equivalence*

$$\|v\|_{H_i} \;\sim\; \|{}^i\mathbf{D}\mathbf{v}\|_{\ell_2(I\!I_i)} \tag{3.4}$$

where ${}^i\mathbf{D}$ is some diagonal matrix. In other words, the *scaled* collection ${}^i\mathbf{D}^{-1}({}^i\Psi)$ constitutes a *Riesz basis* for H_i. We always use the convention that the basis is normalized in L_2, that is, when H_i agrees with L_2 it follows that ${}^i\mathbf{D}=\mathbf{I}$.

(II) *Locality:* The wavelets ${}^i\psi_\lambda$ are compactly supported with support

$$\operatorname{diam}(\operatorname{supp}{}^i\psi_\lambda) \;\sim\; 2^{-|\lambda|}. \tag{3.5}$$

With the aid of Riesz' representation theorem, one can conclude that for the dual pairing $\langle \cdot, \cdot \rangle$ for H_i and its dual H_i' there exists a collection

$$ {}^i\tilde{\Psi} := \{ {}^i\tilde{\psi}_\lambda : \lambda \in I\!I_i \} \subset H_i' \tag{3.6} $$

such that

$$ \langle {}^i\psi_\lambda, {}^i\tilde{\psi}_\mu \rangle = \delta_{\lambda\mu}, \qquad \lambda, \mu \in I\!I_i, \tag{3.7} $$

and ${}^i\mathbf{D}({}^i\tilde{\Psi})$ is a Riesz basis for H_i'. Here $\delta_{\lambda\mu}$ is the Kronecker delta. In fact, by a duality argument one concludes from (3.4) that the corresponding norm equivalence

$$ \|\tilde{v}\|_{H_i'} \sim \|{}^i\mathbf{D}^{-1}\tilde{\mathbf{v}}\|_{\ell_2(I\!I_i)} \tag{3.8} $$

holds for any $\tilde{v} = \tilde{\mathbf{v}}^T \, {}^i\tilde{\Psi} \in H_i'$ [D1]. The coefficients v_λ in the expansion (3.3) can then be expressed in terms of the dual basis as $v_\lambda = \langle v, {}^i\tilde{\psi}_\lambda \rangle$.

The pair $({}^i\Psi, {}^i\tilde{\Psi})$ is called *biorthogonal wavelets*. Of particular interest are the cases when the dual wavelets ${}^i\tilde{\Psi}$ also have compact support (3.5).

Remark 3.1 *Biorthogonal wavelets are particularly appropriate for the present applications since the primal functions can be arranged to be B–splines. The corresponding primal wavelets are then consequently also piecewise polynomials.*

We will always take in the sequel such a pair of compactly supported biorthogonal wavelets $(\Psi, \tilde{\Psi})$.

It will be convenient to use the following shorthand notation. We will both view Ψ as in (3.1) as a *collection* of functions and as a (possibly infinite) (column) *vector* containing all functions always assembled in some fixed unspecified order. For a countable collection of functions Θ and some single function σ, the quantity $\langle \Theta, \sigma \rangle$ is to be understood as the column vector with entries $\langle \theta, \sigma \rangle$, $\theta \in \Theta$, and correspondingly for rows with switched roles. For two collections Θ, Σ, the term $\langle \Theta, \Sigma \rangle$ is then a (possibly infinite) matrix with entries $(\langle \theta, \sigma \rangle)_{\theta \in \Theta, \, \sigma \in \Sigma}$ for which $\langle \Theta, \Sigma \rangle = \langle \Sigma, \Theta \rangle^T$. This also implies for a (possibly infinite) matrix $\mathbf{C}$ that $\langle \mathbf{C}\Theta, \Sigma \rangle = \mathbf{C}\langle \Theta, \Sigma \rangle$ and $\langle \Theta, \mathbf{C}\Sigma \rangle = \langle \Theta, \Sigma \rangle \mathbf{C}^T$.

In this notation, the expansion coefficients in (3.3) and (3.8) can explicitly be expressed as

$$ \mathbf{v}^T = \langle v, {}^i\tilde{\Psi} \rangle, \qquad \tilde{\mathbf{v}}^T = \langle \tilde{v}, {}^i\Psi \rangle. \tag{3.9} $$

Note that the *biorthogonality* or *duality conditions* (3.7) can now be written in terms of an infinite matrix,

$$ \langle {}^i\Psi, {}^i\tilde{\Psi} \rangle = \mathbf{I}, \tag{3.10} $$

where $\mathbf{I}$ is the identity matrix.

By now there are several constructions of wavelets available satisfying properties (I), (II) for function spaces on different domains or manifolds, including L–shaped domains and their boundaries.

Starting from biorthogonal wavelets on the interval [DKU2] which are constructed from the ones on all of $I\!R$ from [CDF], one can construct corresponding wavelets on the domain or its boundary that satisfy the norm equivalence (3.4). Depending on the smoothness of $\mathcal{H}$, one can either use domain decomposition ideas [CTU, DS1] for smoothness order $|s| < 1/2$ or more sophisticated function space characterizations for higher order smoothness [DS2]. Usually the set of *primal* wavelets Ψ consisting in these constructions of *piecewise polynomials* will be used for computation while the *dual* ones $\tilde{\Psi}$ are mainly needed for analysis purposes.

3.2 Discretization and Preconditioning

Having assured that a wavelet basis with the above properties exist for the space $\mathcal{H}$, one can transform a continuous operator equation satisfying (1.2) into a discrete one in terms of the scaled wavelet basis $\mathbf{\Psi}$ for $\mathcal{H}$ as follows. Expansion of the solution

$$U = (u_1, \ldots, u_M)^T = \left(\mathbf{u}_1^T \left({}^1\mathbf{D}^{-1}\right) {}^1\Psi, \ldots, \mathbf{u}_M^T \left({}^M\mathbf{D}^{-1}\right) {}^M\Psi\right)^T =: \mathbf{U}^T\mathbf{D}^{-1}\mathbf{\Psi} \quad \in \mathcal{H} \tag{3.11}$$

and the right hand side $F = \langle \mathbf{\Psi}, F\rangle^T \tilde{\mathbf{\Psi}} \in \mathcal{H}'$ yields the system of equations

$$\langle \mathbf{\Psi}, \mathcal{N}\mathbf{\Psi}\rangle \, \mathbf{D}^{-1} \, \mathbf{U} = \langle \mathbf{\Psi}, F\rangle. \tag{3.12}$$

Multiplying the system (3.12) by the block diagonal matrix $\mathbf{D}^{-1}$,

$$\mathbf{D}^{-1} \langle \mathbf{\Psi}, \mathcal{N}\mathbf{\Psi}\rangle \, \mathbf{D}^{-1} \, \mathbf{U} = \mathbf{D}^{-1} \langle \mathbf{\Psi}, F\rangle, \tag{3.13}$$

and recalling the norm equivalences (3.4) and (3.8), we obtain

$$\mathbf{N}\,\mathbf{U} = \mathbf{F} \tag{3.14}$$

where $\mathbf{N} : \ell_2(I\!I) \to \ell_2(I\!I)$ is an isomorphism. Here we have used the abbreviations

$$\mathbf{N} := \mathbf{D}^{-1} \langle \mathbf{\Psi}, \mathcal{N}\mathbf{\Psi}\rangle \, \mathbf{D}^{-1}, \quad \mathbf{F} := \mathbf{D}^{-1} \langle \mathbf{\Psi}, F\rangle. \tag{3.15}$$

Thus, (3.14) is well–posed in the Euclidean metric $\ell_2(I\!I)$ and the spectral condition number of $\mathbf{N}$ satisfies

$$\mathrm{cond}(\mathbf{N}) := \|\mathbf{N}\|_2 \, \|\mathbf{N}^{-1}\|_2 \quad \lesssim \quad 1. \tag{3.16}$$

Specifically the system (2.2) reads in this form (where now $M = 2$, ${}^1\Psi = \Psi_Y$, ${}^2\Psi = \Psi_Q$)

$$\mathbf{L}\begin{pmatrix}\mathbf{y}\\ \mathbf{p}\end{pmatrix} \;:=\; \begin{pmatrix}\mathbf{A} & \mathbf{B}^T\\ \mathbf{B} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{y}\\ \mathbf{p}\end{pmatrix} \;=\; \begin{pmatrix}\boldsymbol{f}\\ \mathbf{u}\end{pmatrix}. \tag{3.17}$$

Here we have used

$$\begin{aligned} \mathbf{A} &= \mathbf{D}_Y^{-1} \langle \Psi_Y, A\Psi_Y\rangle \, \mathbf{D}_Y^{-1}, & \boldsymbol{f} &= \mathbf{D}_Y^{-1} \langle \Psi_Y, f\rangle, \\ \mathbf{B} &= \mathbf{D}_Q^{-1} \langle \Psi_Q, B\Psi_Y\rangle \, \mathbf{D}_Y^{-1}, & \mathbf{u} &= \mathbf{D}_Q^{-1} \langle \Psi_Q, u\rangle. \end{aligned} \tag{3.18}$$

Together with Theorem 2.1, we have derived the following result.

Corollary 3.2 *The operator* $\mathbf{L}$ *defined in (3.17) is an* ℓ_2*–automorphism, i.e., for every* $(\mathbf{v}, \mathbf{q}) \in \ell_2 = \ell_2(I\!I_Y \times I\!I_Q)$ *one has*

$$\left\| \begin{pmatrix}\mathbf{v}\\ \mathbf{q}\end{pmatrix} \right\|_{\ell_2} \;\sim\; \left\| \mathbf{L}\begin{pmatrix}\mathbf{v}\\ \mathbf{q}\end{pmatrix} \right\|_{\ell_2}. \tag{3.19}$$

3.3 A Representer for Quadratic Cost Functionals

The *qualitative* behaviour of the cost functional (2.9) involving the general norms $\|\cdot\|_Z, \|\cdot\|_W$ can be captured by introducing *Riesz maps* $\mathcal{R}_1 : Z \to Z'$, $\mathcal{R}_2 : W \to W'$. They provide shifts between Sobolev spaces and induce norms that are equivalent to $\|\cdot\|_Z$ and $\|\cdot\|_W$, respectively. In terms of wavelets, one can derive like in [DKS] their representation as

$$\begin{aligned} \mathbf{R}_1 &:= \langle \mathcal{R}_1\Psi_Z, \Psi_Z\rangle &&= \langle \Psi_Z, \Psi_Z\rangle \, \mathbf{D}_Z^2 \, \langle \Psi_Z, \Psi_Z\rangle \\ \mathbf{R}_2 &:= \langle \mathcal{R}_2\tilde{\Psi}_Q, \tilde{\Psi}_Q\rangle &&= \langle \tilde{\Psi}_Q, \tilde{\Psi}_Q\rangle \, \mathbf{D}_Q^{-2} \, \langle \tilde{\Psi}_Q, \tilde{\Psi}_Q\rangle. \end{aligned} \tag{3.20}$$

Note that the order of the Sobolev spaces is encoded in the scaling of the entries in the diagonal matrices $\mathbf{D}_Z, \mathbf{D}_Q$.

Working in wavelet coordinates suggests to formulate the optimization problem already in terms of the discrete ℓ_2 norms to extract the main features of the approach. This means that instead of the minimization functional (2.9) we define the cost functional as

$$\mathbf{J}(\mathbf{y},\mathbf{u}) := \tfrac{\omega}{2}\|\mathbf{R}_1^{1/2}\mathbf{T}\mathbf{y} - \mathbf{y}_\Omega\|_{\ell_2}^2 + \tfrac{1}{2}\|\mathbf{R}_2^{1/2}\mathbf{u}\|_{\ell_2}^2, \tag{3.21}$$

where $\mathbf{T}$ is the discrete ℓ_2–automorphism representing the (identity or trace) operator T.

Thus, the minimization problem that is qualitatively *equivalent* to formulating a problem with cost functional (2.9) can be formulated in the following *discrete form*:

$$\begin{array}{l}\text{Find } (\mathbf{y},\mathbf{u}) \in \ell_2(I\!I_Y \times I\!I_Q) \text{ such that } \mathbf{J}(\mathbf{y},\mathbf{u}) \text{ defined in}\\ (3.21) \text{ is minimized subject to } (3.17).\end{array} \tag{3.22}$$

In the following, we will exclusively treat the new minimization problem (3.22). Following standard techniques from e.g. [Li, Z], we derive existence and uniqueness results for (3.22). To this end, defining

$$\begin{array}{lllllll} \hat{\mathbf{I}} & := \binom{\mathbf{0}}{\mathbf{I}}, & \mathbf{G} & := \binom{\mathbf{f}}{\mathbf{0}}, & \mathbf{S} & := (\mathbf{R}_1^{1/2}\mathbf{T},\ \mathbf{0}), \\ \mathbf{W} & := \binom{\mathbf{y}}{\mathbf{p}}, & \mathbf{W}_\Omega & := \binom{\mathbf{y}_\Omega}{\mathbf{0}}, & \mathbf{R} & := \operatorname{diag}(\mathbf{R}_1,\mathbf{R}_2), \end{array} \tag{3.23}$$

problem (3.22) can be reformulated as follows: Find $(\mathbf{W},\mathbf{u}) \in \ell_2$ which minimize the cost functional

$$\mathbf{J}(\mathbf{W},\mathbf{u}) = \tfrac{\omega}{2}\|\mathbf{S}\mathbf{W} - \mathbf{W}_\Omega\|_{\ell_2}^2 + \tfrac{1}{2}\|\mathbf{u}\|_{\ell_2}^2 \tag{3.24}$$

$$\text{subject to} \qquad \mathbf{L}\mathbf{W} = \hat{\mathbf{I}}\mathbf{u} + \mathbf{G}. \tag{3.25}$$

Since $\mathbf{L}$ is by Corollary 3.2 invertible, we can eliminate $\mathbf{W}$ from (3.25),

$$\mathbf{W} = \mathbf{L}^{-1}\hat{\mathbf{I}}\mathbf{u} + \mathbf{L}^{-1}\mathbf{G}, \tag{3.26}$$

and insert it into (3.24) to obtain a quadratic functional in terms of only the control,

$$\mathbf{J}(\mathbf{u}) := \tfrac{\omega}{2}\|\mathbf{S}\mathbf{L}^{-1}\hat{\mathbf{I}}\mathbf{u} + \mathbf{S}\mathbf{L}^{-1}\mathbf{G} - \mathbf{W}_\Omega\|_{\ell_2}^2 + \tfrac{1}{2}\|\mathbf{R}_2^{1/2}\mathbf{u}\|_{\ell_2}^2. \tag{3.27}$$

Denoting by $D^s\mathbf{J}(\mathbf{u};\mathbf{v}_1,\ldots,\mathbf{v}_s)$ the s-th variation of $\mathbf{J}$ at $\mathbf{u}$ in directions $\mathbf{v}_1,\ldots,\mathbf{v}_s$, where in particular

$$D\mathbf{J}(\mathbf{u};\mathbf{v}) = \langle \delta\mathbf{J}(\mathbf{u}),\mathbf{v}\rangle = \lim_{t\to 0}\frac{\mathbf{J}(\mathbf{u}+t\mathbf{v}) - \mathbf{J}(\mathbf{u})}{t}, \tag{3.28}$$

we can now collect a number of properties of $\mathbf{J}$ for later purposes.

Proposition 3.3 *The functional* $\mathbf{J}$ *defined in (3.27) is twice differentiable on* ℓ_2 *with derivative*

$$\begin{aligned} D\mathbf{J}(\mathbf{u};\mathbf{v}) &= \omega\langle \mathbf{S}\mathbf{W}(\mathbf{u}) - \mathbf{W}_\Omega,\ \mathbf{S}\mathbf{L}^{-1}\hat{\mathbf{I}}\mathbf{v}\rangle + \langle \mathbf{R}_2^{1/2}\mathbf{u},\mathbf{v}\rangle \\ &= \omega\langle \mathbf{S}\mathbf{L}^{-1}\hat{\mathbf{I}}\mathbf{u} + \mathbf{S}\mathbf{L}^{-1}\mathbf{G} - \mathbf{W}_\Omega,\ \mathbf{S}\mathbf{L}^{-1}\hat{\mathbf{I}}\mathbf{v}\rangle + \langle \mathbf{R}_2^{1/2}\mathbf{u},\mathbf{v}\rangle \end{aligned} \tag{3.29}$$

for all $\mathbf{v} \in \ell_2$. *We infer explicitly*

$$\delta\mathbf{J}(\mathbf{u}) = \omega\hat{\mathbf{I}}^T\mathbf{L}^{-T}\mathbf{S}^T(\mathbf{S}\mathbf{W}(\mathbf{u}) - \mathbf{W}_\Omega) + \mathbf{R}_2^{1/2}\mathbf{u} =: \mathbf{Q}\mathbf{u} + \mathbf{e}. \tag{3.30}$$

The second derivative of $\mathbf{J}$ *is for all* $\mathbf{v}, \mathbf{w} \in \ell_2$

$$D^2\mathbf{J}(\mathbf{u};\mathbf{v},\mathbf{w}) \;=\; \omega\langle \mathbf{S}\mathbf{L}^{-1}\hat{\mathbf{I}}\mathbf{v},\, \mathbf{S}\mathbf{L}^{-1}\hat{\mathbf{I}}\mathbf{w}\rangle + \langle \mathbf{R}_2^{1/2}\mathbf{v}, \mathbf{w}\rangle \tag{3.31}$$

or equivalently

$$D^2\mathbf{J}(\mathbf{u};\mathbf{v},\mathbf{w}) \;=\; \omega\,\langle \mathbf{S}\tilde{\mathbf{W}}, \mathbf{S}\bar{\mathbf{W}}\rangle + \langle \mathbf{R}_2^{1/2}\mathbf{v}, \mathbf{w}\rangle, \tag{3.32}$$

where $\tilde{\mathbf{W}} = \binom{\tilde{\mathbf{y}}}{\tilde{\mathbf{p}}}$ *and* $\bar{\mathbf{W}} = \binom{\bar{\mathbf{y}}}{\bar{\mathbf{p}}}$ *solve*

$$\mathbf{L}\,\tilde{\mathbf{W}} = \begin{pmatrix}\mathbf{0}\\ \mathbf{v}\end{pmatrix} \quad \textit{and} \quad \mathbf{L}\,\bar{\mathbf{W}} = \begin{pmatrix}\mathbf{0}\\ \mathbf{w}\end{pmatrix}, \tag{3.33}$$

respectively. Moreover, $D^2\mathbf{J}$ *satisfies for all* $\mathbf{v}, \mathbf{w} \in \ell_2$ *the estimates*

$$D^2\mathbf{J}(\mathbf{u};\mathbf{v},\mathbf{w}) \;\le\; C_*\,\|\mathbf{v}\|_{\ell_2}\|\mathbf{w}\|_{\ell_2} \tag{3.34}$$

and

$$D^2\mathbf{J}(\mathbf{u};\mathbf{v},\mathbf{v}) \;\ge\; c_*\,\|\mathbf{v}\|_{\ell_2}^2 \tag{3.35}$$

with constants

$$C_* := \omega(\mathbf{c}_{T_1,y}\,\mathbf{c}_{\mathbf{L}}^{-1})^2 + 1 \quad \textit{and} \quad c_* := 1. \tag{3.36}$$

Thus, $\mathbf{J}$ *is strictly convex on* ℓ_2*, implying that* $\mathbf{Q}$ *in (3.30) is symmetric positive definite.*

The constant $\mathbf{c}_{T_1,y}$ descends from the trace estimate (2.7) and the constants in the norm equivalence (3.4); $\mathbf{c}_{\mathbf{L}}$ is the constant for the lower estimate in (3.19).

Since the arguments are standard and since the assertions have been explicitly derived in [K4] for the special case $\mathbf{R}_1 = \mathbf{I}$, $\mathbf{R}_2 = \mathbf{I}$, we dispense with the proof here.

The following generalized Weierstrass theorem is a special case of Theorem 43.D from [Z], providing necessary and sufficient conditions in terms of derivatives of the *Lagrangian functional*

$$\textsc{Lagr}(\hat{\mathbf{y}},\hat{\mathbf{p}},\hat{\mathbf{u}},\hat{\mathbf{z}},\hat{\boldsymbol{\mu}}) \;:=\; \mathbf{J}(\hat{\mathbf{u}}) + \left\langle (\hat{\mathbf{z}},\hat{\boldsymbol{\mu}}),\, \mathbf{L}\begin{pmatrix}\hat{\mathbf{y}}\\ \hat{\mathbf{p}}\end{pmatrix} - \begin{pmatrix}\boldsymbol{f}\\ \hat{\mathbf{u}}\end{pmatrix}\right\rangle. \tag{3.37}$$

It is formed as usual by appending the conditions (3.17) by means of additional Lagrange multipliers $(\hat{\mathbf{z}},\hat{\boldsymbol{\mu}}) \in \ell_2$ to the minimization functional (3.27). Here $\hat{\mathbf{W}} = \binom{\hat{\mathbf{y}}}{\hat{\mathbf{p}}}$ is written explicitly to bring out the roles of $\hat{\mathbf{y}}, \hat{\mathbf{p}}$ again.

Theorem 3.4 *Let* $\mathbf{L}$ *be the* ℓ_2*–automorphism from Corollary 3.2 and let* $\mathbf{J}$ *be the quadratic functional defined in (3.27). Then the unique solution* $(\mathbf{y},\mathbf{p},\mathbf{u})$ *of (3.22) is determined by the necessary conditions*

$$\delta\textsc{Lagr}(\mathbf{y},\mathbf{p},\mathbf{u},\mathbf{z},\boldsymbol{\mu};\mathbf{V}) = 0 \qquad \textit{for } \mathbf{V} = \mathbf{z},\boldsymbol{\mu},\mathbf{u},\mathbf{y},\mathbf{p}. \tag{3.38}$$

The Euler equations are explicitly for $\mathbf{V} = \mathbf{z},\boldsymbol{\mu},\mathbf{u},\mathbf{y},\mathbf{p}$ *given by (3.17) and additional conditions for* $\mathbf{u}$ *and* $\mathbf{z},\boldsymbol{\mu}$*,*

$$\begin{aligned} \mathbf{L}\begin{pmatrix}\mathbf{y}\\ \mathbf{p}\end{pmatrix} &= \begin{pmatrix}\boldsymbol{f}\\ \mathbf{u}\end{pmatrix} \\ \mathbf{R}_2^{1/2}\mathbf{u} &= \boldsymbol{\mu} \end{aligned} \tag{3.39}$$

$$\widehat{\mathbf{L}}^T\begin{pmatrix}\mathbf{z}\\ \boldsymbol{\mu}\end{pmatrix} := \begin{pmatrix}\mathbf{A} & (\mathbf{R}_2^{1/2}\mathbf{B})^T\\ \mathbf{R}_2^{1/2}\mathbf{B} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{z}\\ \mathbf{u}\end{pmatrix} = \begin{pmatrix}-\omega\mathbf{T}^T(\mathbf{R}_1\mathbf{T}\mathbf{y} - \mathbf{R}_1^{1/2}\mathbf{y}_\Omega)\\ \mathbf{0}\end{pmatrix}. \tag{3.40}$$

Remark 3.5 *(i) In the following, we denote the constraints (3.17) or (3.25) as* primal system *while the system (3.40) will be called* dual system.

(ii) Control problems with distributed *control could in the present framework be formulated in terms of a cost functional of the form (3.21) with control* $\boldsymbol{f}$ *and with constraints (3.40). The corresponding Euler equations would be similar to (3.17) with (3.39), (3.40). Thus, aside from the Riesz operators problems with boundary and distributed control are equivalent.*

Equation (3.39) will be used in Section 4 to construct an inexact gradient method involving the iterative solution of (3.17) and (3.40). In view of this, it is useful to rewrite $\delta\mathbf{J}$ in (3.30) yet in another way.

Proposition 3.6 *In view of (3.17), (3.39) and (3.40), the first variation of* $\mathbf{J}$ *is*

$$\delta\mathbf{J}(\mathbf{u}) \;=\; \mathbf{R}_2^{1/2}\mathbf{u} - \boldsymbol{\mu}. \tag{3.41}$$

This entails that the evaluation of $\delta\mathbf{J}(\mathbf{u})$ is equivalent to solving first (3.25) and then (3.40).
Proof:
Recall from (3.30) that

$$\delta\mathbf{J}(\mathbf{u}) \;=\; \omega\mathbf{L}^{-T}\mathbf{S}^T\mathbf{R}^{1/2}(\mathbf{S}\mathbf{W}(\mathbf{u}) - \mathbf{W}_\Omega) + \mathbf{R}_2^{1/2}\mathbf{u}$$

where $\mathbf{W}(\mathbf{u})$ is the solution of (3.25). Writing the solution $\mathbf{X} = \mathbf{X}(\mathbf{u}) = (\mathbf{z}, \boldsymbol{\mu})^T$ of (3.40) explicitly,

$$\mathbf{X} = -\omega\mathbf{L}^{-T}\begin{pmatrix}\mathbf{T}^T(\mathbf{R}_1\mathbf{T}\,\mathbf{y}(\mathbf{u}) - \mathbf{R}_1^{1/2}\mathbf{y}_\Omega)\\ \mathbf{0}\end{pmatrix},$$

and applying $\hat{\mathbf{I}}^T$ from the left we infer

$$\boldsymbol{\mu} = -\omega\hat{\mathbf{I}}^T\,\mathbf{L}^{-T}\begin{pmatrix}\mathbf{T}^T(\mathbf{R}_1\mathbf{T}\,\mathbf{y}(\mathbf{u}) - \mathbf{R}_1^{1/2}\mathbf{y}_\Omega)\\ \mathbf{0}\end{pmatrix}$$

which in view of (3.30) confirms (3.41). ∎

At this point it is perhaps useful to summarize the results obtained in view of the general concept described in the introduction. Eliminating $\boldsymbol{\mu}$ from (3.39), we can write (3.17) together with (3.40) as a *weakly coupled* system of saddle point problems. Corollary 3.2 together with Theorem 3.4 then assures that the resulting operator is an ℓ_2–automorphism in the sense of 2.)

Corollary 3.7 *The operator* $\mathbf{N}$ *defined by*

$$\mathbf{N}\mathbf{U} := \begin{pmatrix}\mathbf{L} & \mathbf{E}\\ \widehat{\mathbf{E}} & \widehat{\mathbf{L}}^T\end{pmatrix}\begin{pmatrix}\mathbf{y}\\ \mathbf{p}\\ \mathbf{z}\\ \mathbf{u}\end{pmatrix} := \left(\begin{array}{cc|cc}\mathbf{A} & \mathbf{B}^T & \mathbf{0} & \mathbf{0}\\ \mathbf{B} & \mathbf{0} & \mathbf{0} & -\mathbf{I}\\ \hline \omega\mathbf{T}^T\mathbf{R}_1\mathbf{T} & \mathbf{0} & \mathbf{A}^T & \mathbf{B}^T\mathbf{R}_2^{1/2}\\ \mathbf{0} & \mathbf{0} & \mathbf{R}_2^{1/2}\mathbf{B} & \mathbf{0}\end{array}\right)\begin{pmatrix}\mathbf{y}\\ \mathbf{p}\\ \mathbf{z}\\ \mathbf{u}\end{pmatrix} \tag{3.42}$$

$$= \begin{pmatrix}\boldsymbol{f}\\ \mathbf{0}\\ -\omega\,\mathbf{T}^T\mathbf{R}_1^{1/2}\mathbf{y}_\Omega\\ \mathbf{0}\end{pmatrix} =: \mathbf{F}$$

is an ℓ_2–automorphism, $\ell_2 = \ell_2(I\!I) := \ell_2(I\!I_Y \times I\!I_{Q'} \times I\!I_Y \times I\!I_{Q'})$, i.e., for any $\mathbf{V} \in \ell_2$ the equivalence

$$\|\mathbf{N}\mathbf{V}\|_{\ell_2} \;\sim\; \|\mathbf{V}\|_{\ell_2} \tag{3.43}$$

holds.

The constants in (3.43) depend on the weight ω used in (3.27) to balance the two norms, on the constants $\mathbf{c_L}$ and $\mathbf{C_L}$ from (3.19) and on c_*, C_* defined in (3.36).

3.4 Stability of the Finite–Dimensional Systems

Next we need to ensure 3.) of the concept from Section 1, stability of the resulting *finite–dimensional* discrete system derived from (3.42),

$$\begin{aligned}
\mathbf{N}_\Lambda \mathbf{U}_\Lambda \;&\equiv\; \begin{pmatrix} \mathbf{L}_\Lambda & \mathbf{E}_\Lambda \\ \hat{\mathbf{E}}_\Lambda & \widehat{\mathbf{L}}_\Lambda^T \end{pmatrix} \begin{pmatrix} \mathbf{y}_\Lambda \\ \mathbf{p}_\Lambda \\ \mathbf{z}_\Lambda \\ \mathbf{u}_\Lambda \end{pmatrix} \\
&=\; \left(\begin{array}{cc|cc} \mathbf{A}_\Lambda & \mathbf{B}_\Lambda^T & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_\Lambda & \mathbf{0} & \mathbf{0} & -\mathbf{I} \\ \hline \omega \mathbf{T}_\Lambda^T (\mathbf{R}_1)_\Lambda \mathbf{T}_\Lambda & \mathbf{0} & \mathbf{A}_\Lambda^T & \mathbf{B}_\Lambda^T (\mathbf{R}_2)_\Lambda^{1/2} \\ \mathbf{0} & \mathbf{0} & (\mathbf{R}_2)_\Lambda^{1/2} \mathbf{B}_\Lambda & \mathbf{0} \end{array}\right) \begin{pmatrix} \mathbf{y}_\Lambda \\ \mathbf{p}_\Lambda \\ \mathbf{z}_\Lambda \\ \mathbf{u}_\Lambda \end{pmatrix} \\
&=\; \begin{pmatrix} \boldsymbol{f}_\Lambda \\ \mathbf{0} \\ -\omega\, \mathbf{T}_\Lambda^T (\mathbf{R}_1)_\Lambda^{1/2} (\mathbf{y}_\Omega)_\Lambda \\ \mathbf{0} \end{pmatrix} \equiv \mathbf{F}_\Lambda.
\end{aligned} \tag{3.44}$$

Here we use for convenience the index Λ for all quantities. In fact, Λ stands for possibly different index sets which are chosen according to conditions like (2.4) for the corresponding finite–dimensional spaces.

In view of Proposition 3.3 and Theorem 3.4, the derivation of the system (3.42) reveals that it suffices to ensure that the second derivation $D^2\mathbf{J}$ is *uniformly bounded* from above and below independent of the discretization and that the discrete stability of the finite–dimensional analog of (3.17) is satisfied. These conditions in turn essentially reduce for the situation at hand to the estimate in Remark 2.2 for the finite–dimensional spaces, the inf–sup condition (2.4) called the *Ladyšenskaja–Babuška–Brezzi (LBB) condition* and the continuity estimate (2.7) which is trivially satisfied for subspaces of $H^1(\Omega)$ and $H^{1/2}(\Gamma_y)$. In the context of finite elements, there has been an extensive discussion of the LBB condition, starting with the original work [Ba]. Corresponding conditions on the discretizations for finite–dimensional subspaces of $H^1(\Omega)$ and $(H^{1/2}(\Gamma))'$ for ensuring discrete stability of the finite–dimensional analog of (3.17) have been investigated in a general context of multiscale schemes and wavelet techniques in [DK] and [K2] for arbitrary dimensions using techniques from approximation theory. Roughly speaking, the LBB condition is satisfied if the discretization on Ω is sufficiently fine relative to the discretization of the boundary. Nevertheless, as the numerical tests in [DK] reveal, the theoretical estimates are much too pessimistic. From a numerical point of view, one often still obtains acceptable results even if the sufficient conditions for the LBB condition derived in [DK] are violated, see also Section 5. This subject is extensively discussed in [K2] so that we dispense here with further details.

Proposition 3.8 *Let finite–dimensional subspaces of $H^1(\Omega)$, $(H^{1/2}(\Gamma))'$, Z and W be chosen such that the conditions (2.3), (2.4) and the bounds (3.34), (3.35) are satisfied independent of the discretization. Then the resulting operator derived from (3.42) is an automorphism and satisfies (1.11), (1.12) with uniformly bounded constants.*

An alternative is to formulate (3.42) like in [DKS] as a least squares problem and apply the truncation techniques derived there to achieve stable finite discretizations.

4 An Inexact Gradient Method

Here we briefly recall the mayor results from [K4]. In [GL1] a coupled system of the form (3.44) has been solved in a semi–iterative way. Namely, the solution is achieved by solving each of the two discrete finite–dimensional systems derived from (3.17) and (3.40) by using a direct solver. One alternative would be to apply an iterative method to the whole system (3.44). However, since this system is neither symmetric, nor positive definite, nor stemming from a saddle point problem, it is not clear that a convergent iterative solver for the whole system can be found. Here we propose a fully iterative method for (3.44) which is based on solving the two finite–dimensional saddle point systems alternately. That is, the starting point is to recall that the finite–dimensional system (3.44) is the same as

$$\mathbf{L}_\Lambda \begin{pmatrix} \mathbf{y}_\Lambda \\ \mathbf{p}_\Lambda \end{pmatrix} = \begin{pmatrix} \mathbf{A}_\Lambda & \mathbf{B}_\Lambda^T \\ \mathbf{B}_\Lambda & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{y}_\Lambda \\ \mathbf{p}_\Lambda \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_\Lambda \\ \mathbf{u}_\Lambda \end{pmatrix} \tag{4.1}$$

$$\begin{aligned} \widehat{\mathbf{L}}_\Lambda^T \begin{pmatrix} \mathbf{z}_\Lambda \\ \mathbf{u}_\Lambda \end{pmatrix} &= \begin{pmatrix} \mathbf{A}_\Lambda^T & \mathbf{B}_\Lambda^T (\mathbf{R}_2)_\Lambda^{1/2} \\ (\mathbf{R}_2)_\Lambda^{1/2} \mathbf{B}_\Lambda & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{z}_\Lambda \\ \mathbf{u}_\Lambda \end{pmatrix} \\ &= -\omega \begin{pmatrix} \mathbf{T}_\Lambda^T ((\mathbf{R}_1)_\Lambda \mathbf{T}_\Lambda \mathbf{y}_\Lambda - (\mathbf{R}_1)_\Lambda^{1/2} (\mathbf{y}_\Omega)_\Lambda) \\ \mathbf{0} \end{pmatrix}. \end{aligned} \tag{4.2}$$

The convergence analysis is based on a gradient method for the functional $\mathbf{J}$.

We formulate the main algorithm first without specifying the type of iterative method which is used for the solution of the saddle point problems (4.1) and (4.2). We only assume that it converges and denote it the 'inner iteration', abbreviated as ALGORITHM INNIT, i.e., in some way new iterations

$$\mathbf{L}_\Lambda \begin{pmatrix} \mathbf{y}_\Lambda^{(i+1)} \\ \mathbf{p}_\Lambda^{(i+1)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_\Lambda \\ \mathbf{u}_\Lambda^{(i)} \end{pmatrix} \tag{4.3}$$

and

$$\widehat{\mathbf{L}}_\Lambda^T \begin{pmatrix} \mathbf{z}_\Lambda^{(i+1)} \\ \boldsymbol{\mu}_\Lambda^{(i+1)} \end{pmatrix} = -\omega \begin{pmatrix} \mathbf{T}_\Lambda^T ((\mathbf{R}_1)_\Lambda \mathbf{T}_\Lambda \mathbf{y}_\Lambda^{(i+1)} - (\mathbf{R}_1)_\Lambda^{1/2} (\mathbf{y}_\Omega)_\Lambda) \\ \mathbf{0} \end{pmatrix} \tag{4.4}$$

are determined. Since the overall algorithm involves the iterative solution of both systems, it is termed the 'outer iteration'.

ALGORITHM OUTIT

STEP 1: Choose $\mathbf{u}_\Lambda^{(0)}$, $\widetilde{\mathbf{y}}_\Lambda^{(0)}$, $\widetilde{\mathbf{p}}_\Lambda^{(0)}$, $\widetilde{\mathbf{z}}_\Lambda^{(0)}$, $\widetilde{\boldsymbol{\mu}}_\Lambda^{(0)}$, set up the blocks $\mathbf{L}_\Lambda$ in (4.1), $\hat{\mathbf{E}}_\Lambda$ in (4.2), and $\boldsymbol{f}_\Lambda$, $(\mathbf{y}_\Omega)_\Lambda$; let $\varepsilon_y(i)$, $\varepsilon_\mu(i)$ be some suitable stage–dependent tolerances to be determined later; set $i = 0$;

STEP 2: apply ALGORITHM INNIT to compute an approximate solution $(\widetilde{\mathbf{y}}_\Lambda^{(i+1)}, \widetilde{\mathbf{p}}_\Lambda^{(i+1)})^T$ of (4.3) with right hand side $(\boldsymbol{f}_\Lambda, \mathbf{u}_\Lambda^{(i)})^T$ and initial guesses $(\widetilde{\mathbf{y}}_\Lambda^{(i)}, \widetilde{\mathbf{p}}_\Lambda^{(i)})^T$ that satisfies

$$\left\| \mathbf{L}_\Lambda \begin{pmatrix} \widetilde{\mathbf{y}}_\Lambda^{(i+1)} \\ \widetilde{\mathbf{p}}_\Lambda^{(i+1)} \end{pmatrix} - \begin{pmatrix} \boldsymbol{f}_\Lambda \\ \mathbf{u}_\Lambda^{(i)} \end{pmatrix} \right\|_{\ell_2} \;<\; \varepsilon_y(i+1); \tag{4.5}$$

STEP 3: apply ALGORITHM INNIT to compute a solution $(\widetilde{\mathbf{z}}_\Lambda^{(i+1)}, \widetilde{\boldsymbol{\mu}}_\Lambda^{(i+1)})^T$ of (4.2) with right hand side $-\omega(\mathbf{T}_\Lambda^T((\mathbf{R}_1)_\Lambda \mathbf{T}_\Lambda \widetilde{\mathbf{y}}_\Lambda^{(i+1)} - (\mathbf{R}_1)_\Lambda^{1/2}(\mathbf{y}_\Omega)_\Lambda), \mathbf{0})^T$ and initial guesses $(\widetilde{\mathbf{z}}_\Lambda^{(i)}, \widetilde{\boldsymbol{\mu}}_\Lambda^{(i)})^T$ until

$$\left\| \mathbf{L}_\Lambda^T \begin{pmatrix} \widetilde{\mathbf{z}}_\Lambda^{(i+1)} \\ \widetilde{\boldsymbol{\mu}}_\Lambda^{(i+1)} \end{pmatrix} + \omega \begin{pmatrix} \mathbf{T}_\Lambda^T((\mathbf{R}_1)_\Lambda \mathbf{T}_\Lambda \widetilde{\mathbf{y}}_\Lambda^{(i+1)} - (\mathbf{R}_1)_\Lambda^{1/2}(\mathbf{y}_\Omega)_\Lambda) \\ \mathbf{0} \end{pmatrix} \right\|_{\ell_2} \;<\; \varepsilon_\mu(i+1); \tag{4.6}$$

STEP 4: update $\mathbf{u}_\Lambda^{(i)}$ by

$$\mathbf{u}_\Lambda^{(i+1)} = \mathbf{u}_\Lambda^{(i)} - \rho_i\,(\mathbf{u}_\Lambda^{(i)} - \widetilde{\boldsymbol{\mu}}_\Lambda^{(i+1)}); \tag{4.7}$$

STEP 5: set $i = i+1$ and repeat Step 2, 3, 4 until prescribed tolerance for $\mathbf{u}_\Lambda$ is reached.

The form of the right hand side in (4.7) stems from the fact that it just involves the first variation of $\mathbf{J}$ defined in (3.21) written in terms of $\mathbf{u}$ alone as in (3.27),

$$\delta\mathbf{J}(\mathbf{u}_\Lambda^{(i)}) \;=\; \mathbf{u}_\Lambda^{(i)} - \boldsymbol{\mu}_\Lambda^{(i+1)}, \tag{4.8}$$

see [K4]. The choice of the tolerances $\varepsilon_y(i+1), \varepsilon_\mu(i+1)$ as well as of the step size parameter ρ_i will be discussed below. The convergence of ALGORITHM OUTIT is not apparent beforehand since the iterative solution of (4.1) and (4.2) produces an additional error that appears again in the right hand side of the corresponding adjoint system. Thus, in the convergence analysis one needs to assure that the errors produced in the inner iterations do not accumulate and can be fully controlled.

It has been shown first in [K4] that ALGORITHM OUTIT converges for a certain range of parameters ρ_i if the systems (4.1) and (4.2) are solved *exactly* (e.g. by a direct solver, compare also [GL1]). This follows from the fact that $\mathbf{J}$ defined in (3.21), which can be expressed in terms of $\mathbf{u}_\Lambda$ alone, is a quadratic functional. The analysis in [K4] then yields the following result.

Denote by $\tau(\mathbf{u}_\Lambda)$ an estimate for the error between the exact solution of (4.2) and its initial guess,

$$\|\mathbf{u}_\Lambda^{(0)} - \mathbf{u}_\Lambda\|_{\ell_2} \le \tau(\mathbf{u}_\Lambda). \tag{4.9}$$

Furthermore, recall from e.g. [Bs, Br] that the convergence speed θ_{grad} of a gradient method for (3.27) is governed by the spectral condition number of $\mathbf{Q}_\Lambda$ defined as in (3.30) for the

finite–dimensional case,

$$\theta_{\text{grad}} = \frac{\kappa(\mathbf{Q}_\Lambda) - 1}{\kappa(\mathbf{Q}_\Lambda) + 1}. \tag{4.10}$$

Due to the preconditioning and scaling of the ingredients of $\mathbf{L}_\Lambda$, $\mathbf{Q}_\Lambda$ has a uniformly bounded condition number such that

$$\theta \leq \theta_{\text{grad}} \ < \ 1 \tag{4.11}$$

holds *independent* of the discretization. This means that in each iteration of the gradient method the error will be reduced by a fixed fraction θ, i.e.,

$$\|\mathbf{u}_\Lambda^{(i+1)} - \mathbf{u}_\Lambda\|_{\ell_2} \ \leq \ \theta \, \|\mathbf{u}_\Lambda^{(i)} - \mathbf{u}_\Lambda\|_{\ell_2} \tag{4.12}$$

where $\mathbf{u}_\Lambda$ is the exact solution of minimizing $\mathbf{J}(\mathbf{u}_\Lambda)$ from (3.27) over (4.1).

If for instance at the $(i+1)$th stage the tolerances $\varepsilon_y(i+1)$, $\varepsilon_\mu(i+1)$ are chosen as

$$\begin{aligned} \varepsilon_y(i+1) \ &:= \ \frac{1}{2} \frac{\mathbf{c_L}}{\mathbf{C_L}\,\rho_i} \frac{\theta^i}{(1+i)^2} \tau(\mathbf{u}_\Lambda), \\ \varepsilon_\mu(i+1) \ &:= \ \frac{1}{2} \frac{1}{\mathbf{C_L}\,\rho_i} \frac{\theta^i}{(1+i)^2} \tau(\mathbf{u}_\Lambda) \end{aligned} \tag{4.13}$$

where $\mathbf{c_L}, \mathbf{C_L}$ are the constants in the isomorphism relation (3.19), then one can prove the following.

Theorem 4.1 *If the tolerances $\varepsilon_y(i+1)$ and $\varepsilon_\mu(i+1)$ in (4.5) and (4.6) are selected at each stage according to (4.13) then* Algorithm OutIt *converges for ρ_i satisfying*

$$0 < \rho_* \ \leq \ \rho_i \ \leq \ \rho^* < 2\frac{c_*}{{C_*}^2},$$

where c_* and C_* are constants estimating the second variation of $\mathbf{J}$ from below and above.

In [K4] also a detailed complexity analysis has been given for this scheme, showing that this *basic iterative scheme* leads, in combination with a nested iteration strategy, to an *asymptotically optimal* method. Here we recall the main facts.

Up to this point, we have not specified the particular iterative method Algorithm InnIt by which (4.1) and (4.2) are solved. A simple iterative method for saddle point problems for symmetric $\mathbf{A}_\Lambda$ is the *Uzawa algorithm*. For a system of the form

$$\mathbf{L}_\Lambda \begin{pmatrix} \mathbf{y}_\Lambda \\ \mathbf{p}_\Lambda \end{pmatrix} \equiv \begin{pmatrix} \mathbf{A}_\Lambda & \mathbf{B}_\Lambda^T \\ \mathbf{B}_\Lambda & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{y}_\Lambda \\ \mathbf{p}_\Lambda \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_\Lambda \\ \mathbf{g}_\Lambda \end{pmatrix}, \tag{4.14}$$

the Uzawa algorithm in its simplest form reads for $i = 0, 1, \ldots$ when $\mathbf{y}_\Lambda^{(i)}, \mathbf{p}_\Lambda^{(i)}$ are chosen,

$$\begin{aligned} \mathbf{y}_\Lambda^{(i+1)} \ &= \mathbf{A}_\Lambda^{-1}(\boldsymbol{f}_\Lambda - \mathbf{B}_\Lambda^T \mathbf{p}_\Lambda^{(i)}) \\ &= \mathbf{y}_\Lambda^{(i)} + \mathbf{A}_\Lambda^{-1}(\boldsymbol{f}_\Lambda - \mathbf{A}_\Lambda \mathbf{y}_\Lambda^{(i)} - \mathbf{B}_\Lambda^T \mathbf{p}_\Lambda^{(i)}) \\ \mathbf{p}_\Lambda^{(i+1)} \ &= \mathbf{p}_\Lambda^{(i)} + \gamma\,(\mathbf{B}_\Lambda \mathbf{y}_\Lambda^{(i+1)} - \mathbf{g}_\Lambda). \end{aligned} \tag{4.15}$$

Here γ is some sufficiently small, fixed step size parameter. The first system in (4.15) is not solved exactly. Its iterative solution (by e.g. the conjugate gradient method) corresponds

to applying some approximation $(\mathbf{A}_\Lambda)_0^{-1}$ of $\mathbf{A}_\Lambda^{-1}$ which can be viewed as a preconditioner for $\mathbf{A}_\Lambda$. One usually also includes a preconditioner $(\mathbf{S}_\Lambda)_0$ for the second equation,

$$\begin{aligned} \mathbf{y}_\Lambda^{(i+1)} &= \mathbf{y}_\Lambda^{(i)} + (\mathbf{A}_\Lambda)_0^{-1}(\boldsymbol{f}_\Lambda - \mathbf{A}_\Lambda \mathbf{y}_\Lambda^{(i)} - \mathbf{B}_\Lambda^T \mathbf{p}_\Lambda^{(i)}) \\ \mathbf{p}_\Lambda^{(i+1)} &= \mathbf{p}_\Lambda^{(i)} + \gamma\,(\mathbf{S}_\Lambda)_0^{-1}(\mathbf{B}_\Lambda \mathbf{y}_\Lambda^{(i+1)} - \mathbf{g}_\Lambda). \end{aligned} \tag{4.16}$$

The role of $(\mathbf{S}_\Lambda)_0$ is explained below. Algorithm (4.16) is often called the *incomplete Uzawa algorithm* since the iterative method for the first equation corresponds to multiplying by an approximation $(\mathbf{A}_\Lambda)_0^{-1}$ of $\mathbf{A}_\Lambda^{-1}$, see [BPV].

For discussing the convergence properties of (4.15), one considers the *reduced equation*

$$\mathbf{B}_\Lambda \mathbf{A}_\Lambda^{-1} \mathbf{B}_\Lambda^T \mathbf{p}_\Lambda = \mathbf{B}_\Lambda \mathbf{A}_\Lambda^{-1} \boldsymbol{f}_\Lambda - \mathbf{g}_\Lambda \tag{4.17}$$

involving the *Schur complement* of (4.14). For symmetric and positive definite $\mathbf{A}_\Lambda$, the Uzawa method (4.15) is known to converge if $\mathbf{B}_\Lambda \mathbf{A}_\Lambda^{-1} \mathbf{B}_\Lambda^T$ is symmetric positive definite and if e.g. the step size parameter γ satisfies

$$\gamma \;<\; 2\,\|\mathbf{B}_\Lambda \mathbf{A}_\Lambda^{-1} \mathbf{B}_\Lambda^T\|^{-1}. \tag{4.18}$$

In fact, an iteration for (4.17) reads

$$\mathbf{p}_\Lambda^{(i+1)} = (\mathbf{I} - \gamma\,\mathbf{B}_\Lambda \mathbf{A}_\Lambda^{-1} \mathbf{B}_\Lambda^T)\,\mathbf{p}_\Lambda^{(i)} + \mathbf{B}_\Lambda \mathbf{A}_\Lambda^{-1} \boldsymbol{f}_\Lambda - \mathbf{g}_\Lambda \tag{4.19}$$

which converges if

$$\|\mathbf{I} - \gamma\,\mathbf{B}_\Lambda \mathbf{A}_\Lambda^{-1} \mathbf{B}_\Lambda^T\| < 1 \tag{4.20}$$

This inequality, in turn, follows from (4.18). In the present situation, the preconditioner $\mathbf{S}_0$ is actually only needed for a possibly diagonal scaling since the Schur complement already has a uniformly bounded condition number.

For the systems (4.3) and (4.4) which satisfy Corollary 3.7, we can prove the following.

Remark 4.2 *[K4] The convergence rate of solving (4.3) or (4.4) by the incomplete Uzawa algorithm (4.16) is for suitable choices of* $(\mathbf{A}_\Lambda)_0, (\mathbf{S}_\Lambda)_0$ *independent of the discretization.*

This result follows from the fact that only ℓ_2 operators are involved here. Consequently, in choosing the incomplete Uzawa method (4.16) as the inner iteration ALGORITHM INNIT in ALGORITHM OUTIT, in both STEP 2 and STEP 3 for any size of the systems (4.5), (4.6) only a *fixed* number of iterations is needed to reduce the error by a fixed fraction. Recall also that each iteration can be applied in an amount of work proportional to the size of the system since all operators in $\mathbf{L}_\Lambda$ can be realized by successively applying sparse matrices in terms of the Fast Wavelet Transform [K4]. Furthermore, a combination of the basic iterative method ALGORITHM OUTIT and the analysis can be exploited to prove the asymptotic optimality of the method as follows.

The systems in (4.5) and (4.6) can be solved by the incomplete Uzawa method (4.16) with a convergent rate independent of the discretizations according to Remark 4.2. Thus, taking as initial guess the solution from the previous level, only a uniformly bounded number of Uzawa steps is required to reduce the error by a fixed fraction. This is all that is needed in order to achieve discretization error accuracy on each level. On the lowest level the system is solved exactly. The operators in (4.5) and (4.6) can be applied at an expense that remains proportional to the number of unknowns on that level. Thus, with a geometric

series argument it follows that the overall work stays proportional to the computational work required by a matrix/vector multiplication on the highest level J. That is, the total work is proportional to $\mathcal{O}(N_J)$ where N_J is the number of unknowns on the highest level. In summary, one has the following result.

Theorem 4.3 *If in each iteration of* Algorithm OutIt *the systems (4.5) and (4.6) are solved up to tolerances corresponding to (4.13) and these solutions are taken as initial guesses for the next higher level, then* Algorithm OutIt *is an asymptotically optimal method in the sense that it provides the solution up to the discretization error on level J in an overall amount of $\mathcal{O}(N_J)$ operations where N_J is the number of unknowns in (4.5), (4.6) and (4.7).*

5 Numerical Example

We consider here the following numerical example similar to that from [K4]. The elliptic boundary value problem that plays the role of the constraints for the control problem (3.22) is

$$\begin{aligned} -\Delta y + y &= 1 \quad \text{in } \Omega, \\ y &= u \quad \text{on } \Gamma, \end{aligned}$$

where $\Gamma = \partial\Omega$. Here Ω is the disc with radius $R = 0.5$ around the mid point $(0.5, 0.5)$, $\Omega = \{x \in I\!R^2 : \|x - \binom{1/2}{1/2}\|_{\ell_2}^2 < R\}$. In order to treat this problem, the domain Ω is embedded into a fictitious domain $\square = (0,1)^2$ like in [K3]. This still yields a saddle point problem of the form (2.2) where, however, the operator A now lives on the extended space $H^1(\square)$. In the minimization functional we have chosen the first norm in (3.21) to be equivalent to $Z = Y = H^1(\square)$ which corresponds to the situation considered in the functional (1.17) in Example 1.1. Moreover, we have taken $y_\Omega \equiv 0$ and $\omega = 1$.

Thus, the discretized primal and dual systems read according to (4.1) and (4.2)

$$\mathbf{L}_\Lambda \begin{pmatrix} \mathbf{y}_\Lambda \\ \mathbf{p}_\Lambda \end{pmatrix} \equiv \begin{pmatrix} \mathbf{A}_\Lambda & \mathbf{B}_\Lambda^T \\ \mathbf{B}_\Lambda & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{y}_\Lambda \\ \mathbf{p}_\Lambda \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_\Lambda \\ \mathbf{u}_\Lambda \end{pmatrix}, \qquad \mathbf{L}_\Lambda \begin{pmatrix} \mathbf{z}_\Lambda \\ \mathbf{u}_\Lambda \end{pmatrix} = -\begin{pmatrix} \mathbf{y}_\Lambda \\ \mathbf{0} \end{pmatrix}. \tag{5.1}$$

The complete system is solved by applying the following variant of Algorithm OutIt. In Step 3 the system (4.6) is solved for $\mathbf{z}_\Lambda^{(i+1)}$ and $\mathbf{u}_\Lambda^{(i+1)}$ instead of $\boldsymbol{\mu}_\Lambda^{(i+1)}$. Step 4 is discarded since the update of $\mathbf{u}_\Lambda^{(i+1)}$ is already performed in Step 3. As inner iteration the Uzawa algorithm is used with a preconditioned conjugate gradient method to solve the first equation in (4.15) iteratively. That is, we have applied a CG–Uzawa algorithm. Our stopping criterion is based on the ℓ_2 norm of the residual for system (4.1),

$$\mathbf{r}_\Lambda := \begin{pmatrix} \mathbf{A}_\Lambda \mathbf{y}_\Lambda + \mathbf{B}_\Lambda^T \mathbf{p}_\Lambda - \boldsymbol{f}_\Lambda \\ \mathbf{B}_\Lambda \mathbf{y}_\Lambda - \mathbf{u}_\Lambda \end{pmatrix}$$

which by (3.19) is proportional to the error of (y_Λ, p_Λ) in $H^1(\square)$ and $H^{-1/2}(\Gamma)$. The first numbers in the third column of Table 1 shows the total number of CG iterations necessary to force the ℓ_2 error of the residual of the primal system to be smaller than $\texttt{tol} = \min\{2^{-j}, 2^{-\ell}\}$. The inner iterations are terminated when the residual is smaller than $0.01 * \texttt{tol}$. The numbers in parentheses show the number of Uzawa iterations. The increase of the iteration numbers when ℓ grows relative to j is caused by a violation of the sufficient conditions for the LBB

condition. In STEP 3 the adjoint system is solved for $\mathbf{z}_\Lambda^{(i+1)}$ and $\mathbf{u}_\Lambda^{(i+1)}$ up to the same tolerance `tol` for the corresponding residual of that system.

It is interesting to observe that for these tolerances the variant of ALGORITHM OUTIT *always terminates* after 1 cycle, that is, system (4.1) is solved in STEP 2 up to the requested tolerance, followed by the solution of the system (4.2) in STEP 3. Returning only once more to STEP 2 is in *all* cases sufficient to meet the required overall tolerance. For this reason, the cycle STEP 2 — STEP 3 — STEP 2 is called a *solution cycle* for the coupled saddle point problem (3.44). In Table 1 the iteration numbers for each step of the solution cycle are listed in columns 3, 4 and 5 and are termed *1st it.*, *2nd it.* and *3rd it.* The number #it means the total number of pcg–iterations, while the number of Uzawa steps is written behind in parentheses. We observe that the iteration numbers in the 3rd iteration are always smaller than the ones from the 1st iteration. This is a consequence of the fact that the solutions from the 1st iteration are taken as initial guesses to start the 3rd iteration.

Different variants on the tolerance for the inner iterations have been tested which suggests that ALGORITHM OUTIT turns out to be relatively robust in the following sense. Taking as inner tolerance a value up to `tol` itself instead of $0.01*$`tol` is sufficient for convergence. Of course, relaxing the tolerance can be carried out at the expense of increasing the number of PCG iterations and Uzawa steps. Nevertheless, one solution cycle was still sufficient.

Remark 5.1 *In summary, it seems to be sufficient that the adjoint system to determine the control is only solved once, sandwiched between two iterative solutions of the primal system.*

There are many variants one can think of to balance the amount of iterations needed in each step of the cycle with the necessary amount of iterations.

References

[Ad] R.A. Adams, *Sobolev Spaces*, Academic Press, 1978.

[Ba] I. Babuška, *The finite element method with Lagrange multipliers*, Numer. Math. 20, 1973, 179–192.

[Bs] D.P. Bertsekas, *Nonlinear Programming*, 2nd ed., Athena Scientific, 1999.

[Br] D. Braess, *Finite Elements: Theory, Fast Solvers and Applications in Solid Mechanics*, 2nd ed., Cambridge University Press, 2001.

[BPV] J.H. Bramble, J. Pasciak, A.T. Vassilev, *Analysis of the inexact Uzawa algorithm for saddle point problems*, SIAM J. Num. Anal. 34, 1997, 1072–1092.

[BPX] J.H. Bramble, J.E. Pasciak, J. Xu, *Parallel multilevel preconditioners*, Math. Comp. 55, 1990, 1–22.

[BF] F. Brezzi, M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer, 1991.

[BHM] W.L. Briggs, V.E. Henson, S.F. McCormick, *A Multigrid Tutorial*, SIAM, 2nd ed., 2000.

[CTU] C. Canuto, A. Tabacco, K. Urban, *The wavelet element method, part I: Construction and analysis*, Appl. Comput. Harm. Anal. 6, 1999, 1–52.

[CDF] A. Cohen, I. Daubechies, J.-C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, Comm. Pure Appl. Math. 45, 1992, 485–560.

[DHU] S. Dahlke, R. Hochmuth, K. Urban, *Adaptive wavelet methods for saddle point problems*, M^2AN 34, 2000, 1003–1022.

j	ℓ	1st it.	2nd it.	3rd it.	residual
Ω	Γ	STEP 2	STEP 3	STEP 2	$\|\mathbf{r}_\Lambda\|_{\ell_2}$
3	3	50(5)	21(1)	11(1)	$1.22007090e-1$
3	4	63(4)	22(1)	22(2)	$3.03893551e-2$
3	5	64(4)	23(1)	12(1)	$2.55219641e-2$
3	6	100(7)	25(1)	38(3)	$1.40530157e-2$
4	3	77(4)	28(1)	14(1)	$5.64009761e-2$
4	4	64(3)	28(1)	14(1)	$4.89317933e-2$
4	5	80(4)	30(1)	43(3)	$1.88224372e-2$
4	6	313(21)	33(1)	31(2)	$1.17742201e-2$
4	7	412(28)	55(2)	46(4)	$5.25188732e-3$
5	3	90(4)	33(1)	30(2)	$1.07381850e-2$
5	4	91(4)	34(1)	31(2)	$1.73216501e-2$
5	5	90(4)	35(1)	17(1)	$2.43685913e-2$
5	6	94(4)	38(1)	49(3)	$8.63815998e-3$
5	7	470(29)	64(2)	70(4)	$5.10956365e-3$
6	3	119(5)	58(2)	20(1)	$1.43552113e-2$
6	4	119(5)	39(1)	35(2)	$1.24542308e-2$
6	5	131(6)	40(1)	36(2)	$1.13766598e-2$
6	6	102(4)	42(1)	68(4)	$1.26067723e-2$
6	7	119(5)	70(2)	21(1)	$5.97800253e-3$

Table 1: Iteration numbers for the coupled saddle point problem (5.1) solved by ALGORITHM OUTIT using in each step the (preconditioned) CG–Uzawa method; #it: total number of PCG iterations with number of Uzawa steps in parentheses until the respective residual satisfies $\|\mathbf{r}_\Lambda\|_{\ell_2} \leq \mathtt{tol}$.

[D1] W. Dahmen, *Stability of multiscale transformations*, J. Four. Anal. Appl. 2, 1996, 341–361.

[D2] W. Dahmen, *Wavelet and multiscale methods for operator equations*, Acta Numerica 1997, 55–228.

[D3] W. Dahmen, *Wavelet methods for PDEs — Some recent developments*, IGPM–Preprint #183, RWTH Aachen, December 1999.

[DDU] S. Dahlke, W. Dahmen, K. Urban, *Adaptive wavelet methods for saddle point problems — Convergence rates*, Preprint, April 2001.

[DK1] W. Dahmen, A. Kunoth, *Multilevel preconditioning*, Numer. Math. 63, 1992, 315–344.

[DK] W. Dahmen, A. Kunoth, *Appending boundary conditions by Lagrange multipliers: Analysis the LBB condition*, Numer. Math. 88, 2001, 9–42.

[DKS] W. Dahmen, A. Kunoth, R. Schneider, *Wavelet least square methods for boundary value problems*, IGPM–Preprint #175, RWTH Aachen, September 1999, submitted for publication, revised, October 2000.

[DKU1] W. Dahmen, A. Kunoth, K. Urban, *A wavelet Galerkin method for the Stokes problem*, Computing 56, 1996, 259–302.

[DKU2] W. Dahmen, A. Kunoth, K. Urban, *Biorthogonal spline-wavelets on the interval – Stability and moment conditions*, Appl. Comput. Harm. Anal. 6, 1999, 132–196.

[DPS] W. Dahmen, S. Prößdorf, R. Schneider, Wavelet approximation methods for pseudodifferential equations I: Stability and convergence, Math. Z., 215, 1994, 583–620.

[DS1] W. Dahmen, R. Schneider, *Composite wavelet bases for operator equations*, Math. Comp. 68, 1999, 1533–1567.

[DS2] W. Dahmen, R. Schneider, *Wavelets on manifolds I: Construction and domain decomposition*, SIAM J. Math. Anal. 31, 1999, 184–230.

[DSX] W. Dahmen, R. Schneider, Y. Xu, *Nonlinear functionals of wavelet expansions — Adaptive reconstruction and fast evaluation*, Numer. Math. 86, 2000, 49–101.

[DH] J. Dankova, R. Haslinger, *Fictitious domain approach used in shape optimization: Neumann boundary condition*, in: Control of Partial Differential Equations and Applications, Casas, Eduardo (ed.), Dekker, Lect. Notes Pure Appl. Math. 174, 1996, 43–49.

[GJL] Ph.E. Gill, L.O. Jay, M.W. Leonard, L.R. Petzold, V. Sharma, *An SQP method for the optimal control of large-scale dynamical systems*, J. Comput. Appl. Math. 120, 2000, 197–213.

[GMPS] Ph.E. Gill, W. Murray, D.B. Ponecon, M. Saunders, *Preconditioners for indefinite systems arising in optimization*, SIAM J. Matrix Anal. Appl. 13, 1992, 292–311.

[GR] V. Girault, P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer, 1986.

[Glo] R. Glowinski, *Finite element methods for the numerical simulation of incompressible viscous flow. Introduction to the control of the Navier-Stokes equations*, Lect. Appl. Math. 28, 1991, 219–301.

[GPP] R. Glowinski, T.W. Pan, J. Periaux, *A fictitious domain method for Dirichlet problem and application*, Comp. Meth. Appl. Mechs. Eng. 111, 1994, 282–303.

[GoY] G.H. Golub, Q. Ye, *Inexact preconditioned conjugate gradient method with inner–outer iteration*, SIAM J. Sci. Comput. 21, 1999, 1305–1320.

[Gr] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, 1985.

[Gu] M.D. Gunzburger (ed.), *Flow Control*, The IMA Volumes in Mathematics and its Applications 68, Springer, 1995.

[GHS1] M.D. Gunzburger, L.S. Hou, Th. P. Svobodny, *Analysis and finite element approximation of optimal control problems for the stationary Navier–Stokes equations with Dirichlet controls*, M^2AN 25, 1991, 711–748.

[GHS2] M.D. Gunzburger, L.S. Hou, Th. P. Svobodny, *Analysis and finite element approximation of optimal control problems for the stationary Navier-Stokes equations with distributed and Neumann controls*, Math. Comp. 57, 1991, 123–151.

[GL1] M. D. Gunzburger, H. C. Lee, *Analysis, approximation, and computation of a coupled solid/fluid temperature control problem*, Comp. Meth. Appl. Mech. Engrg. 118, 1994, 133–152.

[GL2] M. D. Gunzburger, H. C. Lee, *Analysis and approximation of optimal control problems for first-order elliptic systems in three dimensions*, Appl. Math. Comput. 100, 1999, 49–70.

[Ha1] W. Hackbusch, *Fast solution of elliptic control problems*, J. Optim. Theory Appl. 31, 1980, 565–581.

[Ha2] W. Hackbusch, *Multigrid Methods and Applications*, Springer, New York, 1985.

[Has] J. Haslinger, *Fictitious domain approaches in shape optimization*, in: Borggard, Jeff (ed.) et al., Computational Methods for Optimal Design and Control, Birkhaeuser, Prog. Syst. Control Theory 24, 1998, 237–248.

[HHK] J. Haslinger, K.–H. Hoffmann, M. Kočvara, *Control fictitious domain method for solving optimal shape design problems*, RAIRO, Modelisation Math. Anal. Numer. 27, 1993, 157–182.

[HasN] P. Neittaanmäki, J. Haslinger, *Finite Element Approximation for Optimal Shape, Material and Topology Design*, Wiley, 2nd ed., 1996.

[Hei1] M. Heinkenschloss, *Formulation and analysis of a sequential quadratic programming method for the optimal Dirichlet boundary control of Navier-Stokes flow*, in: Optimal Control: Theory, Algorithms, and Applications H.W. Hager et al.(eds.), Kluwer, Appl. Optim. 15, 1998, 178–203.

[Hei2] M. Heinkenschloss, *The numerical solution of a control problem governed by a phase field model*, Optim. Methods Softw. 7, 1997, 211–263.

[HK1] M. Hinze, K. Kunisch, *On suboptimal control strategies for the Navier-Stokes equations*, ESAIM, Proc. 4, 1998, 181–198.

[HK2] M. Hinze, K. Kunisch, *Newton's method for tracking type control of the instationary Navier–Stokes equations*, Preprint, October 1999.

[IK] K. Ito, K. Kunisch, *Augmented Lagrangian-SQP methods for nonlinear optimal control problems of tracking type*, SIAM J. Contr. Optim. 34, 1996, 874–891.

[Ja] S. Jaffard, *Wavelet methods for fast resolution of elliptic equations*, SIAM J. Numer. Anal. 29, 1992, 965–986.

[K1] A. Kunoth, *Multilevel preconditioning — Appending boundary conditions by Lagrange multipliers*, Adv. Comp. Maths. 4, 1995, 145–170.

[K2] A. Kunoth, *Wavelet Methods — Boundary Value Problems And Control Problems*, Advances in Numerical Mathematics, Teubner, 2001.

[K3] A. Kunoth, *Wavelet techniques for the fictitious domain—Lagrange multiplier approach*, Manuscript, August 2000, to appear in: Numerical Algorithms.

[K4] A. Kunoth, *Fast iterative solution of saddle point problems in optimal control based on wavelets*, preprint, May 2000, submitted for publication, revised, March 2001, SFB 256 Preprint.

[Li] J.L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin, 1971.

[MM] H. Mauer, H.D. Mittelmann, *Optimization techniques for solving elliptic control problems with control and state constraints: Part 1. Boundary Control*, Comput. Optim. Appl. 16, 29–55, 2000.

[NT] P. Neittaanmäki, D. Tiba, *An embedding of domains approach in free boundary problems and optimal design*, SIAM J. Contr. Optim. 33, 1995, 1587–1602.

[OC] J.T. Oden, G.C. Carey, *Finite Elements – Mathematical Aspects*, Vol. IV, Prentice–Hall, New Jersey, 1983.

[O1] P. Oswald, *On discrete norm estimates related to multilevel preconditioners in the finite element method*, in: Constructive Theory of Functions, K.G. Ivanov, P. Petrushev, B. Sendov, (eds.), Proc. Int. Conf. Varna 1991, Bulg. Acad. Sci., Sofia, 1992, 203–214.

[PRG] L. Petzold, J.B. Rosen, Ph.E. Gill, L.O. Jay, K. Park, *Numerical optimal control of parabolic PDEs using DASOPT* L.T. Biegler et al. (eds.), Large-Scale Optimization with Applications. Part 2: Optimal Design and Control, Springer, IMA Vol. Math. Appl. 93, 1997, 271–299.

[Pi] J. Pitkäranta, *Boundary subspaces for the finite element method with Lagrange multipliers*, Numer. Math. 33, 1979, 273–289.

[Ri] A. Rieder, *A domain embedding method for Dirichlet problems in arbitrary space dimensions*, RAIRO, Modelisation Math. Anal. Numer. 32, 1998, pp. 405–431.

[Tr] F. Tröltzsch, *On the Lagrange–Newton–SQP method for the optimal control of semilinear parabolic equations*, SIAM J. Cont. Optim. 38, 1999, 294–312.

[Y] K, Yosida, *Functional Analysis*, Springer, 6th ed., 1994.

[Z] E. Zeidler, *Nonlinear Functional Analysis and its Applications; III: Variational Methods and Optimization*, Springer, 1985.