JOHANNES KEPLER
UNIVERSITÄT LINZ
Netzwerk für Forschung, Lehre und Praxis

Multigrid Method for Elliptic Control Problems

MASTERARBEIT

zur Erlangung des akademischen Grades

MASTER OF SCIENCE

in der Studienrichtung

INDUSTRIAL MATHEMATICS

Angefertigt am *Institut für Numerische Mathematik*

Betreuung:

O. Univ. Prof. Dipl. Ing. Dr. Helmut Gfrerer

Eingereicht von:

Muzhinji Kizito

Linz, August 2008

This dissertation is dedicated to my wife Wachenuka, our two girls Nyashadzashe and Tapiwanashe, my parents Mr and Mrs Muzhinji, brothers and Sisters, Dr and Mrs Unganai for their support and encouragement. A special dedication goes to my daughter, Tapiwanashe, she was born during the beginning of my two year study and I will see her for the first time when she is already two years old.

Contents

Acknowledgments

This work presented so many new challenges during the inceptional and implementational phases. Without the support, patience and guidance of the following people, it would have taken years to be completed. I owe my profound gratitude to the following:

1. Professor Helmut Gfrerer who undertook to act as my supervisor despite his various commitments. Without his support, patience and guidance this work could have been hardly a success.

2. My wife, Wachenuka and our two daughters, Nyashadzashe and Tapiwanashe for allowing me to be away from them for two years of study. My parents Mr and Mrs Muzhinji who always supported, encouraged and prayed for my studies. Above all, a Big Thanks to Dr and Mrs Unganai, without them I could have failed to make it for my studies.

3. My Lecturers both at TU Kaiserslautern(Germany) and Johannes Kepler Linz University(Austria) for providing all the necessary knowledge during the course of my studies.

4. All my friends who were always with and inspired me in this programme.

Finally, I would like acknowledge the European Union for financing my studies through the Erasmus Mundus Scholarship.

Muzhinji Kizito
Linz, August 2008

Abstract

In this work we studied elliptic control problem. It is an optimization problem that consists of the cost functional to be minimized subject to constraints governed by elliptic differential equations with a Neumann boundary condition. We transformed the optimization problem to the optimality system which was characterised in the form of the integral equation on which the multigrid method is applied to find an optimal control. The major goal is to find the optimal control. We achieve this by computing the distributed control problem using finite element method with piecewise linear functions. The domain was partitioned by regular triangulation. The existence and uniqueness of the solutions to the optimal control and the discrete optimal problem is studied and error estimates obtained. Different examples were considered. The convergence of the multigrid method is analysed and the numerical results agree with the theoretical claims.

List of Figures

List of Tables

List of Notation and Symbols

Symbol	Description
Ω	Solution Domain, $\Omega \subset \mathbb{R}^2$
Ω_l	Discrete domain $\Omega_l \subset \Omega$
$\Gamma = \partial\Omega$	Boundary of Ω
$H^m(\Omega)$	$\left\{v \in L^2(\Omega) : \forall\alpha, \quad \lvert\alpha\rvert \leq m, \quad \partial^{\lvert\alpha\rvert} v \in L^2(\Omega)\right\}$
$y, \quad y_l$	Continuous and discrete state solution
p	Adjoint variable
$u, \quad u_l$	Continuous and discrete control solutions
y_d	Desired/target state
$A, \quad A^\star$	Elliptic differential operators and adjoints
$A_l, \quad N_l$	Stiffness and mass matrices at grid level l
$K, \quad K_l$	Continuous and discrete integral operators
$r_{l,l-l}, \quad p_{l-1,l}$	The restriction and the prolongation operators
MGM	Multigrid Method

Chapter 1

Introduction

An optimal control problem consists of a governing system, a description of the control mechanism, and a criterion defining the cost functional, that models the purpose of the control and describes the cost of its action. The system is governed by the elliptic partial differential equations in our case. The formulation of an optimal control problem involves the cost functional to minimize under the constraint given by the modeling equations. The necessary conditions for such a minimum result in a set of coupled equations called the optimality system. In this and in the next section we describe a control- and a state-constrained optimal control problems.

The main thrust of this work is to apply the multigrid method to solve for the constrained optimization problem governed by the partial differential equations. The multigrid method is applied to find the optimal control which is the minimiser of the cost functional. The multigrid method has been shown to be very efficient and successful in solving elliptic control problems(Hachbusch, Borzi). The first step is to transform the optimization problem to the first order optimality system then characterize the first order necessary optimal condition by an integral equation on which the multigrid method is developed, analyzed and finally numerically implemented.

The key features of the multigrid method are smoothing and coarse grid correction that involves the inter-grid transfers and a solution correction step. The main results of the work are the convergence of the multigrid method in calculating the control variable. We make numerical experiments for the elliptic control problem

$$min_{(y,u)\in(Y\times U)}\ J(y,u)=\frac{1}{2}\|y-y_d\|^2_{L^2(\Omega)}+\frac{\delta}{2}\|u\|^2_{L^2(\Omega)} \tag{1.1}$$

subject to

$$\begin{aligned} -\Delta y+y &= f+u \quad in \quad \Omega \\ \frac{\partial y}{\partial n} &= g \qquad\quad on \ \ \partial\Omega \end{aligned}$$

The Lagrangian Principle is applied to the elliptic control problem to get the optimal control system(2.13). With u we denote the control function belonging to a set of admissible controls $\mathcal{U}_{ad} \subset \mathcal{U}$, where $\mathcal{U}$ is a real Hilbert space. The state of the system is a function of the control denoted by $y(u) \in \mathcal{Y}$, where Y is a real Hilbert space.

The optimal control u together with the corresponding state variable y and the co-state/adjoint variable p is the solution of the system

$$\begin{aligned} -\Delta y + y &= f + u && in \ \ \Omega \\ \frac{\partial y}{\partial n} &= g && on \ \partial\Omega \\ -\Delta p + p &= y - y_d && in \ \ \Omega \\ \frac{\partial p}{\partial n} &= 0 && on \ \partial\Omega \\ u &= -\frac{1}{\delta} \cdot p && in \ \ \Omega \end{aligned}$$

Where

- $y(u)$ is the solution of the state elliptic partial differential equation corresponding to the control u.
- $p(u)$ is the solution of the adjoint elliptic partial differential equation corresponding to the state $y(u)$.
- y_d is the desired/target state.

The state and adjoint solutions depend on the control function u. The control may be involved by the differential equation as above or by the boundary condition. The coupled system is decoupled into an integral equation form(chapter 2 and chapter 3). The decoupling results into a single equation which is not more expensive to solve than the system of elliptic partial differential equations if the fast solver is applied Hachbusch([6], [7]). The optimal control is the solution of the equation of the form

$$(I - K)u = q \tag{1.2}$$

where

- K is the integral operator(3.34)
- I is the identity operator
- q is the right hand side involving the $y_d, \ \ f \ \ and \ \ g$.

The idea of solving the integral equation characterisation of the optimality system using the multigrid method was used by Hachbusch ([6] and [7]). The optimal system is discretized using the finite element method. The integral equation replaces the system of elliptic equations on which the multigrid method is applied. The solution of the resulting discrete integral equation is presented in (chapter 5). The two grid method which is the basis of the multigrid is elaborated. The multigrid method algorithm is formulated based on the application of the two grid recursively. The convergence of the multigrid is also analyzed with the main results.

In chapter 2 the notion, notations and the several examples of control problems are presented. The finite element discretization which is vital for the assembling of matrices is reviewed in chapter 3 and the characterization of the optimal control by the integral equation including the discretization of the operator. Chapter 4 contains the description of the multigrid method and its properties. The discussion of the numerical solution is presented in chapter 5.

Chapter 2

ELLIPTIC CONTROL PROBLEMS

The formulation of optimal control of systems governed by elliptic partial differential equations requires the following terms:

- The definition of a control function u that represents the driving influence of the environment on the system.
- The elliptic partial differential equations modeling the controlled system, represented by the state function $y(u)$.
- The cost functional which models the purpose of the control on the system.

This work is done on an open domain $\Omega \subset \mathbb{R}^2$. The domain $\Omega = (0,1)^2$ with boundary $\partial\Omega = \Gamma$ as shown on the figure below.

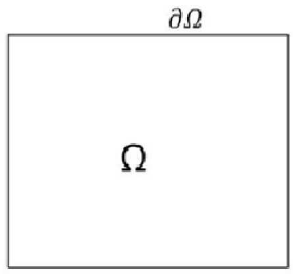

Figure 2.1: Model Domain $\Omega \in \mathbb{R}^2$ *with boundary* $\partial\Omega = \Gamma$.

The control can be defined in the following ways:

Definition 2.1. *With reference to the domain figure 2.1 define*

1. ***Distributed Control***- *if the control is defined on the whole or in some parts of the inside of the domain.*

2. ***Boundary Control***- *if the control is defined on the whole or parts of the boundary.*

The examples of the elliptic differential equations for the distributed and boundary control problems are as follows:

Distributed Control: Find the state $y(u)$ so that for the given control $u \in \mathcal{U}$ the following equation is satisfied:

$$\begin{aligned} Ay(u) &= f + u \quad \text{in} \quad \Omega \\ By(u) &= g \quad \text{on} \quad \partial\Omega \end{aligned} \tag{2.1}$$

Boundary Control: Find the state $y(u)$ so that for the given control $u \in \mathcal{U}$ the following equation is satisfied:

$$\begin{aligned} Ay(u) &= f \quad \text{in} \quad \Omega \\ By(u) &= u + g \quad \text{on} \quad \partial\Omega \end{aligned} \tag{2.2}$$

This means that $y(u)$ is the solution of (2.1) or (2.2) for a given control u either in the whole domain or on the boundary. A full description of the examples of the distributed and boundary control is given below. Some important notations that are used in the descriptions are given below.

- Let A be a differential operator of the second order for example $A = (-\Delta + I)$ and A^* is the adjoint of the operator A.
- Let B and C be boundary operators of the first order with smooth coefficients such that the Green's formula holds

$$\langle Ay, p\rangle_{L^2(\Omega)} - \langle y, A^\star p\rangle_{L^2(\Omega)} = \langle y, Cp\rangle_{L^2(\Gamma)} - \langle By, p\rangle_{L^2(\Gamma)} \tag{2.3}$$

.

- Let $\mathcal{U}$ be the linear space of control functions that are either distributed(defined on Ω) or boundary(defined on Γ).

If $\mathcal{U}$ is defined on Γ we have boundary control problem for example the optimal temperature distribution otherwise we have a distributed control problem like the optimal heat source distributed on the whole domain. Define the set of admissible controls $\mathcal{U}_{ad} \subset \mathcal{U}$. In our case we assume that there are no constraints on the control, then $\mathcal{U}_{ad} = \mathcal{U}$. The space of the state, controls and the adjoint are defined on the Sobolev spaces order k, H^k, H_0^k which is the closure of all smooth functions with compact support in Ω(chapter 3)

2.1 Distributed Control Problems

This section involves the study of the examples of the distributed control problems with the boundary condition either Neumann or Dirichlet type. In this case the control is distributed on the whole or on parts of the whole domain Ω. The first order necessary conditions will explicitly be given for the distributed control problems.

2.1.1 Model Problem

The goal is to find the control functions in the domain such that the objective functional is minimized. For this the optimisation problem is defined as:

$$J(y, u) := \frac{1}{2} \parallel y - y_d \parallel^2_{L^2(\Omega)} + \frac{\delta}{2} \parallel u \parallel^2_{L^2(\Omega)} \tag{2.4}$$

which has to be minimized

subject to the constraints

$$\begin{aligned} Ay(u) &= f + u \quad \text{in} \quad \Omega \\ By(u) &= g \quad\quad\; \text{on} \quad \partial\Omega \end{aligned} \tag{2.5}$$

with

- $f \in L^2(\Omega)$, $g \in H^{\frac{1}{2}}$ *fixed and* $u \in \mathcal{U}_{ad} = L^2(\Omega)$ varies.
- The goal is to get $y(u) \simeq y_d$ for a given function $y_d \in L^2(\Omega)$ with the control as small as possible.
- The equation (2.5) is called the state equation with the variable y(u) called the state and the y_d is the desired/target state.
- $\delta > 0$ is the weighting parameter of the cost functional.

Since for a given control u we can find the corresponding state $y(u)$, then we can define the control to state mapping

$$S : U \to Y, \quad Su := y(u)$$

exists and is continuously differentiable such that the new cost functional will be defined as

$$\begin{aligned} F(u) &= J(y(u), u) && (2.6) \\ &= J(Su, u) && (2.7) \end{aligned}$$

This means that S is the solution operator. The new cost functional for the optimization problem is now defined as

$$F(u) := \frac{1}{2} \parallel Su - y_d \parallel^2_{L^2(\Omega)} + \frac{\delta}{2} \parallel u \parallel^2_{L^2(\Omega)} \tag{2.8}$$

which has to be minimized

2.1.2 First Order Optimality Condition

Consider the objective function (2.8). The first order optimality condition is given by the theorem below

Theorem 2.2. *The control* $u \in \mathcal{U}_{ad}$ *is optimal if and only if* $u \in \mathcal{U}_{ad}$ *and* v *satisfies the variational inequality*
$\langle F'(u), v - u \rangle \geq 0 \quad \forall v \in \mathcal{U}$

PROOF:
Let $u \in \mathcal{U}$ be the optimal solution and choose $v \in \mathcal{U}$
By the convexity of $\mathcal{U}_{ad}$ we have

$$w_t = u + t(v - u) \in \mathcal{U}_{ad} \quad \forall t \in [0, 1].$$

Now by the optimality of u yields

$$F(w_t) - F(u) \geq 0 \text{ for all } t \in [0,1].$$

Then

$$\begin{aligned} \langle F'(u), v-u\rangle_{\mathcal{U}^\star,\mathcal{U}} &= \lim_{t\to 0} \tfrac{F(u+t(v-u))-F(u)}{t} \\ &\geq 0 \end{aligned}$$

On the other hand, using the convexity of F and for all $u \in \mathcal{U}_{ad}$ we have

$$\begin{aligned} 0 &\leq \langle F'(u), v-u\rangle_{\mathcal{U}^\star,\mathcal{U}} \\ &= \lim_{t\to 0} \tfrac{F(u+t(v-u))-F(u)}{t} \\ &\leq \lim_{t\to 0} \frac{tF(v)-(1-t-1)F(u)}{t} \\ &= F(v) - F(u). \end{aligned}$$

Hence u is optimal control.

The multigrid algorithm is applied to find the optimal control of the optimality system(chapter 4). To apply this algorithm we need the adjoint/costate variable which can be found by solving the adjoint elliptic differential equation of the given state partial differential equation. To find the adjoint elliptic partial differential equation, the Lagrangian Principle is used.

2.1.3 Lagrange Principle

In this section we apply the Lagrange Principle to derive the optimality system for the distributed control problem

$$J(y(u),u) := \frac{1}{2} \parallel y(u) - y_d \parallel^2_{L^2(\Omega)} + \frac{\delta}{2} \parallel u \parallel^2_{L^2(\Omega)} \quad \text{which has to be minimized}$$

subject to the constraints

$$\begin{aligned} Ay(u) &= f+u \quad \text{in} \quad \Omega \\ By(u) &= g \qquad \text{on} \quad \partial\Omega \end{aligned} \tag{2.9}$$

Introducing the Lagrange multiplier p results in the Lagrange function:

$$\begin{aligned} \mathcal{L}\big(y(u),u,p\big) &= J(y(u),u) - \langle Ay(u)-f-u, p\rangle_{L^2(\Omega)} - \langle By(u)-g,p\rangle_{L^2(\Gamma)} \\ &= J(y(u),u) - \langle Ay(u),p\rangle_{L^2(\Omega)} + \langle f+u,p\rangle_{L^2(\Omega)} - \langle By(u),p\rangle_{L^2(\Gamma)} \\ &\quad + \langle g,p\rangle_{L^2(\Gamma)} \end{aligned}$$

From the Green's formula(2.3) we get

$$\langle Ay, p\rangle_{L^2(\Omega)} = \langle y, A^\star p\rangle_{L^2(\Omega)} + \langle y, Cp\rangle_{L^2(\Gamma)} - \langle By, p\rangle_{L^2(\Gamma)} \tag{2.10}$$

This results in the formulation

$$\begin{aligned}\mathcal{L}\big(y(u), u, p\big) &= J(y(u), u) - \langle y(u), A^\star p\rangle_{L^2(\Omega)} - \langle y(u), Cp\rangle_{L^2(\Gamma)} + \langle By(u), p\rangle_{L^2(\Gamma)} \\ &\quad + \langle f + u, p\rangle_{L^2(\Omega)} - \langle By(u), p\rangle_{L^2(\Gamma)} + \langle g, p\rangle_{L^2(\Gamma)} \\ &= J(y(u), u) - \langle y(u), A^\star p\rangle_{L^2(\Omega)} - \langle y(u), Cp\rangle_{L^2(\Gamma)} + \langle f + u, p\rangle_{L^2(\Omega)} + \langle g, p\rangle_{L^2(\Gamma)}\end{aligned}$$

This implies that we are looking for the optimality conditions of the linear optimization problem under the assumptions

1. for all $u \in \mathcal{U}$ there exists a unique $y = y(u)$ such that
   $Ay(u) = f + u \ \ in \ \ \Omega, \quad By(u) = g \ \ on \ \ \partial\Omega$
2. $\mathcal{U}_{ad}$ is convex, bounded and closed.

Theorem 2.3. *Let $(y(u), u)$ be an optimal solution then under the assumptions(1-2) there exists a Lagrangian multiplier p such that the optimality system holds*

- $Ay(u) = f + u$ *state equation.*
- $\mathcal{L}_y\big(y(u), u, p\big) = 0$ *adjoint equation*
- $\langle \mathcal{L}_u\big(y(u), u, p\big), v - u\rangle_{\mathcal{U}^\star, \mathcal{U}} \geq 0$ *for all* $v \in \mathcal{U}_{ad}$

From the theorem(2.2), we can observe that $u \in U_{ad}$ is an optimal control iff

$$\langle F'(u), v - u\rangle \geq 0 \quad \forall\ v \in \mathcal{U}_{ad}$$

This means that

$$\langle y(u) - y_d, y(v) - y(u)\rangle_{L^2(\Omega)} + \delta\langle u, v - u\rangle_{L^2(\Omega)} \geq 0 \quad \forall\ v \in \mathcal{U}_{ad}$$

Now let $p(u)$ be chosen such that

$$\begin{aligned} A^\star p(u) &= y(u) - y_d \quad \text{in} \quad \Omega \\ Cp(u) &= 0 \quad\quad\quad\quad \text{on} \quad \partial\Omega \end{aligned} \tag{2.11}$$

and we characterize the optimal control u by

$$u = -\delta^{-1} p(u) \tag{2.12}$$

From the relations (2.9), (2.11) and the Green formula(2.3) we get

$$\begin{aligned}\langle y(u)-y_d, y(v)-y(u)\rangle_{L^2(\Omega)} &= \langle A^\star p(u), y(v)-y(u)\rangle_{L^2(\Omega)} \\ &= \langle p(u), A\big(y(v)-y(u)\big)\rangle_{L^2(\Omega)} + \langle p(u), B\big(y(v)-y(u)\big)\rangle_{L^2(\Gamma)} \\ &\quad -\langle Cp(u), y(v)-y(u)\rangle_{L^2(\Gamma)} \\ &= \langle p(u), v-u\rangle_{L^2(\Omega)}\end{aligned}$$

Since $\mathcal{U}_{ad}$ is a linear space,

$$\langle p(u)+\delta u, v-u\big\rangle_{L^2(\Omega)} \geq 0, \quad \forall\, v \in \mathcal{U}_{ad}$$

We obtain (2.12).
Now eliminating u in (2.9) by (2.12) , we get a coupled system of two elliptic partial differential equations.

$$\begin{aligned} Ay(u) &= f-\delta^{-1}p(u) \quad \text{in} \quad \Omega \\ By(u) &= g \quad \text{on} \quad \partial\Omega \\ A^\star p(u) &= y(u)-y_d \quad \text{in} \quad \Omega \\ Cp(u) &= 0 \quad \text{on} \quad \partial\Omega \end{aligned} \tag{2.13}$$

A multigrid method is applied to a coupled system(chapter5). The boundary conditions are Neumann type. By and Cp can be replaced by $y(u)|_\Gamma$ and $p(u)|_\Gamma$ to get the analogous results for the Dirichlet problem. The optimality system can be represented in different ways. The next section deals with the representation of the control on the boundary that is the boundary control.

2.2 Boundary Control Problem

This section involves the study of the examples of the boundary control problems with the boundary condition either Neumann or Dirichlet type. In this case the control is restricted on the whole or on some parts of the boundary of the domain Ω. The first order necessary conditions will explicitly be given for the two boundary control problems.

2.2.1 Neumann Boundary Control Problem

The boundary control is an example of the control problem involving an elliptic differential equation with Neumann boundary conditions where the control in defined on $\partial\Omega$. Let $y(u)$ be defined by

$$\begin{aligned} Ay(u) &= f \quad \text{in} \quad \Omega \\ By(u) &= u+g \quad \text{on} \quad \partial\Omega \end{aligned} \tag{2.14}$$

for $u \in \mathcal{U}_{ad} = \mathcal{U} = L^2(\Gamma)$.
The cost functional

$$J(y(u),u) := \frac{1}{2} \parallel y(u)-y_d \parallel^2_{L^2(\Omega)} + \frac{\delta}{2} \parallel u \parallel^2_{L^2(\Gamma)}$$

which has to be minimized where $y_d \in L^2(\Omega)$ is the target state.

The optimal control is defined on the boundary and is given by

$$u = -\delta^{-1} p(u)|_\Gamma \tag{2.15}$$

where $p(u)$ is the solution of the adjoint elliptic differential equation(2.11).
The resulting optimality system

$$\begin{array}{rcll} Ay(u) & = & f & \text{in} \quad \Omega \\ By(u)|_\Gamma & = & g - \delta^{-1} p(u)|_\Gamma & \text{on} \quad \Gamma \\ A^\star p(u) & = & y(u) - y_d & \text{in} \quad \Omega \\ Cp(u) & = & 0 & \text{on} \quad \Gamma \end{array} \tag{2.16}$$

The system above describes the behaviour on the boundary. The other way in which the boundary control problem with Neumann boundary condition can be represented is:

Consider the cost functional which describes the behaviour on the boundary

$$J(y(u), u) := \frac{1}{2} \parallel y(u) - y_d \parallel^2_{L^2(\Gamma)} + \frac{\delta}{2} \parallel u \parallel^2_{L^2(\Gamma)}$$

which has to be minimized where $y_d \in L^2(\Gamma)$ is the target state.

Let $y(u)$ fulfils the state equation (2.14) with (2.15) yields the optimal control provided $p(u)$ solves the adjoint partial differential equation

$$\begin{array}{rcll} A^\star p(u) & = & 0 & \text{in} \quad \Omega \\ Cp(u)|_\Gamma & = & y(u)|_\Gamma - y_d & \text{on} \quad \Gamma \end{array} \tag{2.17}$$

Then the resulting optimality system becomes

$$\begin{array}{rcll} Ay(u) & = & f & \text{in} \quad \Omega \\ By(u)|_\Gamma & = & g - \delta^{-1} p(u)|_\Gamma & \text{on} \quad \Gamma \\ A^\star p(u) & = & 0 & \text{in} \quad \Omega \\ Cp(u)|_\Gamma & = & y(u)|_\Gamma - y_d & \text{on} \quad \Gamma \end{array} \tag{2.18}$$

which is a coupled by the boundary condition.

2.2.2 Dirichlet Boundary Control Problem

Let the target/desired state $y_d \in L^2(\Omega)$ and the set of admissible controls defined as $\mathcal{U}_{ad} = \mathcal{U} = L^2(\Gamma)$. Let $y(u)$ be defined by

$$\begin{array}{rcll} Ay(u) & = & f & \text{in} \quad \Omega \\ y(u) & = & u + g & \text{on} \quad \partial\Omega \end{array} \tag{2.19}$$

The cost functional

$$J(y(u), u) := \frac{1}{2} \parallel y(u) - y_d \parallel^2_{L^2(\Omega)} + \frac{\delta}{2} \parallel u \parallel^2_{L^2(\Gamma)}$$

which has to be minimized.

The optimal control on the boundary is given $u = -\delta^{-1}p(u)|_\Gamma$ where $p(u)$ is the solution of the adjoint elliptic differential equation.

$$\begin{aligned} A^\star p(u) &= y(u)|_\Gamma - y_d \quad \text{in} \quad \Omega \\ p(u) &= 0 \qquad\qquad\quad \text{on} \quad \Gamma \end{aligned} \tag{2.20}$$

Then the resulting optimal system becomes

$$\begin{aligned} Ay(u)) &= f \qquad\qquad\qquad \text{in} \quad \Omega \\ y(u)|_\Gamma &= g - \delta^{-1}p(u)|_\Gamma \quad \text{on} \quad \Gamma \\ A^\star p(u) &= y(u)|_\Gamma - y_d \qquad \text{in} \quad \Omega \\ p(u) &= 0 \qquad\qquad\qquad \text{on} \quad \Gamma \end{aligned} \tag{2.21}$$

which is a coupled by the boundary condition.

2.2.3 Existence and Uniqueness of the Function Minimum

In this section, the existence and uniqueness of the control function as the minimizer of the cost functional is analyzed. We have the cost functional $J(y(u), u)$ with $J : \mathcal{Y} \times \mathcal{U} \longrightarrow \mathbb{R}$ is investigated with $y(u)$ and u related by the operator

$$S : \mathcal{U} \longrightarrow \mathcal{Y}.$$

With the new cost functional defined by (2.6, 2.7, 2.8).

Theorem 2.4. ***(Existence of the minimiser)***
Let $\mathcal{U}$ be a Hilbert space, $\mathcal{U}_{ad} \subset \mathcal{U}$ non-empty, bounded, closed and convex and $F : U \longrightarrow \mathbb{R}$ weakly lower semi-continuous and $\mathcal{U}$ radially unbounded, that is

1. *$u_n \rightharpoonup u^\star$ follows that $F(u^\star) \leq \liminf_{n\to\infty} F(u_n)$(Weakly lower semi-continuous).*
2. *$\| u \| \longrightarrow \infty$ follows that $F(u) \longrightarrow \infty$.(radially unbounded)*

With linear and bounded operator S, assume F is bounded from below by a constant $C \in \mathbb{R}$ with $C \leq F(u) \leq \infty$ for $u \in \mathcal{U}$. Then the minimization problem $min_{u^\star \in \mathcal{U}_{ad}} F(u)$ has a solution $u^\star \in \mathcal{U}_{ad}$.

PROOF.
Let F be bounded from below, it implies the existence of the infimum. This means that there is a sequence $(u_n) \subset \mathcal{U}_{ad}$ with

$$F^\star = \lim_{n\to\infty} F(u_n).$$

By the radial unboundedness of F, (u_n) is bounded, and there exists a weakly converging subsequence (u_{n_k}) of u_n such that

$$u_{n_k} \rightharpoonup u^\star \ as \ k \to \infty.$$

Since $\mathcal{U}_{ad}$ is closed and convex subset of the Hilbert space it is weakly closed and hence $u^\star \in \mathcal{U}_{ad}$. Since S is linear and bounded, it follows weak convergence in $\mathcal{Y}$;

$$Su_{n_k} \rightharpoonup Su^\star \ as \ k \to \infty.$$

Due to weakly lower semi continuity of F we obtain

$$F(u^\star) \leq \liminf_{n\to\infty} F(u_{n_k}) = F^\star \ for \ k \to \infty.$$

Since $F^\star$ is the infimum if and only if $F(u^\star) = F^\star$ Hence the minimum is attained at $u^\star$.

Theorem 2.5. ***(Uniqueness)***
Let the conditions of theorem(2.4) holds. If additionally F is strictly convex, then there exists at most one optimal control.

Chapter 3

DISCRETIZATION

This chapter deals with the discretization of the optimization model(optimality system(2.13)) problem, distributed control by the finite element method. The ingredients of the finite element discretization are the variational formulation where the test spaces are introduced, the existence and uniqueness of the solution. The finite element method described here is based on the references (Braess([3]), Knabner and Angermann([11])). The convergence, stability and the approximation property of the solution are also going to briefly discussed. The discretization of the optimality system and finally the discretization of the operator K of the integral equation characterisation of the optimality system are also going to be demonstrated. In this chapter the Galerkin method is used to discretize the distributed control problem (2.4, 2.5) with Neumann boundary condition. From now on we assume that $A = -\Delta + I$ and $B = \frac{\partial}{\partial n}$.

3.1 Variational Formulation

Assuming the existence of a classical solution, the following steps are performed in general:

- **Step 1:** Multiplication of the differential equation by test functions that are chosen compatible with the type of boundary condition and subsequent integration over the domain Ω.
- **Step 2:** Integration by parts and incorporation of the boundary conditions in order to derive a suitable bilinear form.
- **Step 3:** Verification of the required properties like ellipticity and continuity(Knaber and Angermann([11])). The solution of the elliptic differential equation is based on the weak formulation in the Sobolev spaces.

We demonstrate the variational formulation of the state partial differential equation and that for the adjoint follows immediately.

We choose the test function $w \in W = H^1$, multiply by the test function and integrate by parts

$$\begin{aligned}
\int_\Omega (-\Delta y + y)\cdot w\mathbf{d}x &= \int_\Omega (f+u)\cdot w\mathbf{d}x \\
\int_\Omega \nabla y\cdot\nabla w\mathbf{d}x - \int_\Gamma (\nabla y\cdot\mathbf{n})w\mathbf{d}s + \int_\Omega y\cdot w\mathbf{d}x &= \int_\Omega (f+u)\cdot w\mathbf{d}x \\
\int_\Omega \nabla y\cdot\nabla w\mathbf{d}x + \int_\Omega y\cdot w\mathbf{d}x &= \int_\Omega (f+u)\cdot w\mathbf{d}x + \int_\Gamma g\cdot w\mathbf{d}s
\end{aligned}$$

Define

- the bilinear form

$$a : H^1 \times H^1 \longrightarrow \mathbb{R}$$
$$a(y,w) = \int_\Omega \nabla y\cdot\nabla w\mathbf{d}x + \int_\Omega y\cdot w\mathbf{d}x \tag{3.1}$$

- the linear form

$$F : H^1 \longrightarrow \mathbb{R}$$
$$F(w) = \int_\Omega (f+u)\cdot w\mathbf{d}x + \int_\Gamma g\cdot w\mathbf{d}s \tag{3.2}$$

- The variational formulation
  Find $y \in H^1$ such that

$$a(y,w) = F(w) \quad \forall w \in W = H^1 \tag{3.3}$$

Definition 3.1. *Let W be a Hilbert space. A bilinear form $a : W \times W \longrightarrow \mathbb{R}$ is called:*

- *symmetric if $a(y,w) = a(w,y) \qquad \forall w,\ y \in W$.*
- *continuous if $|a(y,w)| \leq C \parallel y \parallel_W \cdot \parallel w \parallel_W$.*
- *coercive if $a(w,w) \geq \gamma \parallel w \parallel_W^2$.*

The variational formulation for the adjoint and the state elliptic partial differential equations for the distributed optimal control system(2.15, 2.16), for the given control u

$$\begin{aligned}
a(y(u),w) &= \langle f+u, w\rangle_{L^2(\Omega)} + \langle g, w\rangle_{L^2(\Gamma)} \qquad \text{for all } w \in W = H^1 \\
a(w,p(u)) &= \langle w, y(u)+y_d\rangle_{L^2(\Omega)} \qquad \text{for all } w \in W = H^1
\end{aligned}$$

In this work we consider $a(y(u),w) = a(w,p(u))$, so it will be enough to assemble $a(y(u),w)$.

The central theorem the ensure the unique solvability of the variational problem is the Lax Milgram theorem which holds for convex sets.

Theorem 3.2. ***The Lax Milgram Theorem***
Let W be a Hilbert space. Let $a : W \times W \longrightarrow \mathbb{R}$ be a continuous, coercive bilinear form and $F : W \longrightarrow \mathbb{R}$ a linear functional. Then the variational equation

$$a(y, w) = F(w) \quad \textit{for all } w \in W \tag{3.4}$$

has a unique solution $y \in W$. Moreover, the solution satisfies the estimate

$$\| y \|_W \leq \frac{1}{\gamma} \| F \|_{W^\star} \tag{3.5}$$

3.2 Finite Element Method

In this section we discuss the finite element discretization of our domain $\Omega = (0, 1)^2$. This can be achieved with the following general steps.

1. Discretize the domain $\overline{\Omega}$, that is to divide the solution region into finite elements(subdomains). The solution domain is divided into several simpler finite elements, where each element has a simple geometry, so appropriate assumed solutions can easily be written for the element. In this thesis the sub-domains are the triangles.

2. Establish the matrix equation for the finite element which relates the nodal values of the unknown function to other parameters. In this work we use the Galerkin method.

3. To find the global equation system for the whole solution, all element equations must be assembled. This involves the combination local element equations for all elements used for discretization. The Neumann boundary condition is incorporated on the right hand side

4. Solving the global equation system. Since the finite element global equation system is typically sparse, symmetric and positive definite, direct and iterative methods can be used for the solution. In this work the multigrid method which is an iterative solver is used. The nodal values of the sought function are produced as a result of the solution, since approximating functions are determined in terms of nodal values of as physical field which is sought.

3.2.1 Formulation of the Finite Element Method

As we have mentioned in (1-4) above, several approaches can be used to transform the physical formulation of the problem to its finite element discrete state. Since the formulation of the problem is described by an elliptic differential equation then the most popular method of its finite element formulation is the Galerkin method.

The Galerkin Method

The Variational Problem of the model problem is given by

Find $y \in W: \quad a(y, w) = F(w), \qquad \forall\, w \in W$
where

- W is the solution space with $W = H^1$
- $a(.,.)$ is a continuous bilinear form on $W \times W$.
- $F(.)$ is a continuous linear form on W.

The idea is to replace the infinite dimensional space W by a finite dimensional space $W_h \subset W$ which consists of a fixed degree associated with the subdivisions of the computational domain. Let $\mathcal{H}_l$ for $l \in \mathbb{N}_0$ be a sequences of subspaces of the finite dimensional subspace W_h defined on each level of refinement l.

$$\mathcal{H}_l \subset \mathcal{H}_{l+1} \subset W_h \subset W = H^1$$

where $\mathcal{H}_{l+1}$ subspace which correspond to Ω_{l+1} is the refinement of Ω_l with subspace $\mathcal{H}_l$ such that $\Omega_l \subset \Omega_{l+1} \subset \Omega$. An example is illustrated in the figure below where Ω_0 is the initial mesh $l = 0$ and Ω_1 is a refinement of Ω_0 as shown on the figure below

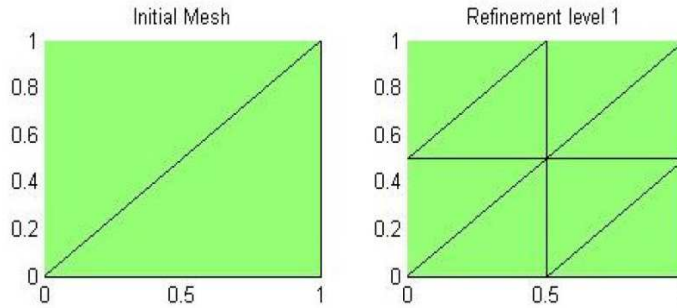

Figure 3.1: **Left:** Initial Mesh Ω_0, **Right:** One Refinement Ω_1

Now consider the variational problem in the finite dimensional subspace

$$\text{Find } y_l \in \mathcal{H}_l : \quad a(y_l, w_l) = F(w_l), \quad \forall w_l \in \mathcal{H}_l$$

Suppose that the $dim \mathcal{H}_l = N_l$ and to calculate the solution we choose the basis functions which are linearly independent with a small support

$$\mathcal{H}_l = span\Big\{\phi_1, \phi_2, ..., \phi_{N_l}\Big\}$$

Now expressing the approximate solution y_l in terms of the basis functions

$$y_l = \Sigma_{i=1}^{N_l} \xi_i \phi_i \quad\quad \xi_i \in \mathbb{R} \;\; i = 1, 2, ..., N_l$$

Then the new variational problem reads

$$\text{Find } \xi = (\xi_1, \xi_2, ..., \xi_{N_l}) \in \mathbb{R}^{N_l} :$$

$$\Sigma_{i=1}^{N_l} a(\phi_j, \phi_i)\xi_i = F(\phi_j) \quad j = 1, ..., N_l \tag{3.6}$$

This is a linear system of equations for $\xi = (\xi_1, \xi_2, ..., \xi_{N_l})^T$ with matrix

$$A_l = a(\phi_j, \phi_i) \in \mathbb{R}^{N_l \times N_l} \text{ and } b_j = F(\phi_j) \in \mathbb{R}^{N_l \times 1}$$

Since ϕ_i have small support, $a(\phi_j, \phi_i) = 0$ for most i and j the matrix A_l is sparse(most of the entries are zeros).

Lemma 3.3. *Let a be a coercive, bilinear form then the matrix A_l is positive definite. A symmetric bilinear form implies symmetry of matrix A_l.*

3.2.2 Finite Elements

In this section the construction of the Finite Element method is described. Let $\mathcal{H}_l \subset W_h$ consists of continuous piecewise linear functions. Let $\Omega = (0, 1)^2$ be a bounded domain, with boundary $\partial\Omega = \Gamma$. The domain Ω can be covered with a finite number of triangles as shown on the figure below

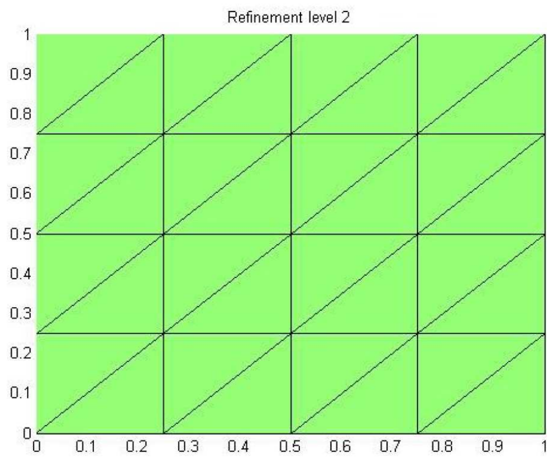

Figure 3.2: An example of a triangulation

The subdivision of the domain into triangles is called triangulation.

Definition 3.4. ***(Triangulation)***
Let $\Omega \subset \mathbb{R}^2$. A partition $\mathcal{T}_h = \left\{T_1, T_2, ..., T_N\right\}$ of $\overline{\Omega}$ into triangular elements is called admissible provided the following properties hold

1. $\overline{\Omega} = \cup_{T \in \mathcal{T}_h} T$.
2. *if $T_i \cap T_j$ consists of exactly one point, then it is a common vertex of T_i and T_j.*
3. *if for $i \neq j$, $T_i \cap T_j$ consists of more than one point, then $T_i \cap T_j$ is a common edge of T_i and T_j.*

Where $h := max_{1 \leq i \leq N}\ diam(T_i)$ denotes the maximum diameter of all $T \in \mathcal{T}_h$. The items 2 and 3 implies that T_i and T_j are adjacent. With each node we associate the basis function ϕ which is equal to 1 at the node and 0 at all other nodes. For example at node x_j, we have that

$$\phi_i(x_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

ϕ is assumed to be a continuous function on Ω and linear on each of the triangles.

Suppose that the nodes are labeled $1, ..., N_l$ and let $\phi_1(x, y), ..., \phi_{N_l}(x, y)$ be the corresponding basis functions. The functions $\phi_1, ..., \phi_{N_l}$ are linearly independent and span a N_l dimensional linear subspace $\mathcal{H}_l$ at level l.
For our the model problem the finite element can be restated as

$$\text{Find } \xi = (\xi_1, \xi_2, ..., \xi_{N_l}) \in \mathbb{R}^{N_l} :$$

$$\Sigma_{i=1}^{N_l} \int_\Omega \left(\frac{\partial \phi_i}{\partial x}\frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y}\frac{\partial \phi_j}{\partial y} + \phi_i\phi_j\right)\xi_i dxdy = \int_\Omega (f+u)\phi_j dxdy + \int_{\partial\Omega} g\phi_j ds \qquad (3.7)$$

for $j = 1, ..., N_l$.

Letting $A_l = (a_{ij})$ *and* $b_l = (F_1, ..., F_{N_l})^T$ where

- $a_{ij} = a_{ji} = \int_\Omega \left(\frac{\partial \phi_i}{\partial x}\frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y}\frac{\partial \phi_j}{\partial y} + \phi_i\phi_j\right) dxdy$
- $b_j = \int_\Omega \left(f + u\right)\phi_j dxdy + \int_{\partial\Omega} g\phi_j ds$

The finite element approximation can be written as a system of linear equation
$A_l \xi = b_l$

3.2.3 Assembling of Matrices

In subsection(3.2.2), we have developed a system of equations. We now want to calculate each of these terms involved in these systems. It is important to note that each ϕ_i defined in the previous subsection has support over at most two elements thus when a regular triangulation $\mathcal{T}_h$ has been generated for the domain, we can calculate the stiffness matrix A_l, the mass matrix N_l and the right hand side b_l which is the sum of f_l and g_l at each level l.

For the stiffness matrix and the mass matrix we have

$$A_{ij} = \sum_{T\in\mathcal{T}_h} \int_T \nabla\phi_i \nabla\phi_j dx, \quad N_{ij} = \sum_{T\in\mathcal{T}_h} \int_T \phi_i\phi_j dx$$

and for the right hand side we consider the construction of f and g.

$$f_j = \sum_{T\in\mathcal{T}_h} \int_T f\phi_j dx \quad \textit{and } g_j = \sum_{T\in\mathcal{T}_h} \int_E g\phi_j ds$$

Definition 3.5. *The local stiffness and mass matrices* $A_l^T,\ N_l^T \in \mathbb{R}^{3\times 3}$ *are defined by*

$$(A_l^{(T)})_{ij} = \int_T \nabla\phi_i \nabla\phi_j dx \quad \textit{for } i,j = 1,2,3 \tag{3.8}$$

$$(N_l^{(T)})_{ij} = \int_T \phi_i\phi_j dx \quad \textit{for } i,j = 1,2,3 \tag{3.9}$$

For a triangular element $T \in \mathcal{T}_h$, with vertices $(x_1, y_1),\ (x_2,\ y_2), (x_3, y_3)$ and $\phi_1,\ \phi_2,\ \phi_3$ be the corresponding basis functions in $\mathcal{H}_l$. We denote the area of the triangle by $|T|$ with

$$|T| = \begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} \tag{3.10}$$

Since

$$\begin{pmatrix} \phi_1(x,y) \\ \phi_2(x,y) \\ \phi_2(x,y) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} \tag{3.11}$$

It can easily be computed that

$$\nabla\phi_i(x,y) = \frac{1}{2}\begin{pmatrix} y_{i+1} - y_{i+2} \\ x_{i+2} - x_{i+1} \end{pmatrix} \qquad \text{indices are modulo 3}$$

Then

$$\int_T \nabla\phi_i\nabla\phi_j dx = \frac{1}{4|T|}\left(y_{i+1}-y_{i+2} \quad x_{i+2}-x_{i+1} \right)\left(\begin{array}{c} y_{j+1}-y_{j+2} \\ x_{j+2}-x_{j+1} \end{array} \right)$$

This means that we get the local stiffness matrix for the term $\int_T \nabla\phi_i\nabla\phi_j dx$. This can be expressed as

$A_l^{(T)} = \frac{|T|}{2} \cdot Grad \cdot Grad^T$ where the

$$Grad = \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.12}$$

Now for the mass matrix which is from the term $\int_T \phi_i\phi_j dx$. Using the quadrature rule we get

$$N_l^{(T)} = \frac{|T|}{12}(1+\delta_{ij}) \quad i,j = 1,2,3$$

which is

$$N_l^{(T)} = \frac{|T|}{12}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \tag{3.13}$$

After computing the element stiffness and mass matrices we sum over all the elements to obtain the global matrices. As mentioned already above the stiffness and the mass matrix for the state and the adjoint equations take the same form.

The Right Hand Side

Since we are going to apply the multigrid method to a system of elliptic partial differential equations, we need only to assemble the desired state. The solution procedure will require us to input the initial control into the state equation, solve for the state, input the state into the adjoint equation and finally find the optimal control. To assemble for the desired state, we use the quadrature rule. To assemble the term involving $\int_\Omega y_{d,l}\phi_j dx$. This can be written as

$$\int_\Omega y_{d,l}\phi_j dx = \sum_{T\in\mathcal{T}_h}\int_T y_{d,l}\phi_j dx \tag{3.14}$$

We approximate the integral $\int_T y_{d,l}\phi_j dx$ using the midpoint rule such that (x_s, y_s) is the centroid of each element. Numerical realization of this term also in the simplest case involves one point numerical quadrature. The integral can be evaluated as

$$\int_T y_{d,l}\phi_j dx = \frac{|T|}{3} y_{d,l}(x_s, y_s) \tag{3.15}$$

And this is the same as multiplying the element mass matrix by the value of $y_{d,l}$ at the nodes.

$$(y_{d,l})^{(T)} = N_l^{(T)} \cdot y_d(x_i, y_i) \quad i = 1,2,3 \tag{3.16}$$

To assemble the term involving $\int_\Omega f_l \phi_j dx$. This can be written as

$$\int_\Omega f_l \phi_j dx = \sum_{T \in \mathcal{T}_h} \int_T f_l \phi_j dx \tag{3.17}$$

We approximate the integral $\int_T f_l \phi_j dx$ using the midpoint rule such that (x_s, y_s) is the centroid of each element. This can be written as

$$\int_\Omega f_l \phi_j dx = \sum_{T \in \mathcal{T}_h} \int_T f(x_s, y_s) \phi_j dx \tag{3.18}$$

Numerical realization of this term also in the simplest case involves one point numerical quadrature. The integral can be evaluated as

$$\int_T f_l \phi_j dx = \frac{|T|}{3} f_l(x_s, y_s) \tag{3.19}$$

Now for the boundary term

$$g_j = \sum_{T \in \mathcal{T}} \int_E g_l \phi_j ds \tag{3.20}$$

where E is the edge of one of the element. Let (x_m, y_m) be the midpoint of the edge. Then the integration over the edge yields

$$\int_E g_l \phi_j ds \simeq g_l(x_m, y_m)|E| \tag{3.21}$$

where $|E|$ is the length of the edge.

3.2.4 Approximation Properties

In this section, having established the construction of the finite element for the state and the adjoint elliptic equation, now we establish approximation properties of our solutions.

We have the variational formulation for the state and the adjoint elliptic partial differential equations

$$\begin{aligned} a(y(u), w) &= \langle f + u, w \rangle_{L^2(\Omega)} \qquad \forall w \in H^1 \\ a(w, p(u)) &= \langle w, y(u) - y_d \rangle_{L^2(\Omega)} \qquad \forall w \in H^1 \end{aligned}$$

We keep in mind that the two equations have the same bilinear $a(.,.)$ form which is $W - elliptic$, continuous and symmetric.
Let $\mathcal{H}_l$ for $l \in \mathbb{N}_0$ be a sequence of subspaces with

$$\mathcal{H}_l \subset \mathcal{H}_{l+1} \subset H^s \qquad s \geq 1.$$

In our case $s = 1$.
Then the discrete solutions at each level of discretization is defined by

Definition 3.6. *Find discrete solutions $y_l(v), p_l(v) \in \mathcal{H}_l$ such that*

$$a(y_l(v), w_l) = \langle f_l + v, w_l \rangle_{L^2(\Omega)} \qquad \forall w_l \in \mathcal{H}_l \tag{3.22}$$

$$a(w_l, p_l(v)) = \langle w_l, y_l(v) - y_{d,l} \rangle_{L^2(\Omega)} \qquad \forall w_l \in \mathcal{H}_l \tag{3.23}$$

This means that at each level, we can compute the discrete solutions for the state, adjoint depending on the discrete control v_l. The subspaces $\mathcal{H}_l$ are connected by the step size h_l and the approximation property.
To begin the approximation property is the *Céa*'s lemma in the general subspace..

Lemma 3.7. *Let the bilinear form $a : W \times W \rightarrow \mathbb{R}$ be continuous and W-elliptic. Suppose y and y_h are the solutions of variational formulation in W and $W_h \subset W$ respectively then*

$$\| y - y_h \|_W \leq \frac{C}{\gamma} inf_{w_h \in W_h} \| y - w_h \|_W \tag{3.24}$$

PROOF:
by the definition of y *and* y_h

$$a(y, w) = F(w) \qquad \forall w \in W$$

$$a(y_h, w) = F(w) \qquad \forall w \in W_h$$

Since $W_h \subset W$, by subtraction

$$a(y - y_h, w) = 0 \qquad \forall w \in W_h \qquad Galerkin \;\; Orthogonality$$

Let $w_h \in W_h$ with $w = w_h - y_h \in W_h \subset W$, then

$$a(y - y_h, w_h - y_h) = 0$$

and by coercivity and continuity

$$\begin{aligned} \gamma \| y - y_h \|_W^2 &\leq a(y - y_h, y - y_h) \\ &= a(y - y_h, y - w_h) + a(y - y_h, w_h - y_h) \\ &\leq C \| y - y_h \|_W \cdot \| y - w_h \|_W \qquad \forall w_h \in W_h \end{aligned}$$

hence dividing by $\| y - y_h \|_W$ the result follows.

- The *Céa*'s lemma says that the accuracy of the numerical solution depends on the choice of the function spaces which are capable of approximating the solution y.
- The choice of W_h is important. Let the space W_h be a space of linear piecewise finite elements

$$W_h = \left\{ y \in C^0(\Omega) : y|_{T_i} \in P_1, \quad T_i \in \mathcal{T} \right\}$$

- The error estimates are primarily based on the estimating the approximation error

$$inf_{w_h \in W_h} \parallel y - w_h \parallel_W .$$

- From the *Céa*'s lemma the discretization error is estimated by the approximation error which is estimated by the interpolation error

Let the continuous solution be regular enough, that is $y \in H^2(\Omega)$. Now for arbitrary $w \in H^2(\Omega)$ define the interpolation operator

$$\begin{aligned} \Pi_h : H^2(\Omega) &\rightarrow W_h \\ w &\mapsto \Pi_h w \end{aligned}$$

Hence it follows that

$$inf_{w_h \in W_h} \parallel y - w_h \parallel_{H^1} \leq \parallel y - \Pi_h y \parallel_{H^1} \tag{3.25}$$

As a consequence, it is enough to deal with the interpolation error $\parallel y - \Pi_h y \parallel_{H^1}$ for convergence results. The main ideas are

1. localise the error on the triangles(elements)
2. transformation of the triangles on the reference triangles.
3. compute the local interpolation error.
4. inverse transformation back to the triangles(elements).

Following the ideas 1-4 above, the interpolation error in our case is given by the theorem

Theorem 3.8. *Let $y \in H^2(\Omega)$ with $\Omega = (0,1)^2$. Then the interpolation error satisfies*

$$\parallel y - \Pi_h y \parallel_{H^1} \leq Ch \parallel y \parallel_{H^2}$$

Then the convergence follows from the $C\grave{e}a$'s Lemma.

Theorem 3.9. *Let $\mathcal{T}$ a quasi-uniform triangulation of Ω. Let y and y_h be the solutions of the continuous and discrete variational equations respectively. Then for $y \in H^2(\Omega)$.*

$$\| y - y_h \|_{H^1} \leq Ch \| y \|_{H^2} \tag{3.26}$$

Where C is independent of the level l

This gives the convergence in H^1-norm. This means that there is a linear convergence for finite element error in H^1-norm, that is $\| y - y_h \|_{H^1} = O(h)$

Similarly, the convergence in L^2-norm follows for the Aubin-Nitsche theorem on shape regular triangulation.

Theorem 3.10. ***(Aubin-Nitsche)***
Suppose that Ω is a convex polygon and that $\mathcal{T}_h$ is regular family of meshes on Ω. Then

$$\| y - y_h \|_{L^2(\Omega)} \leq Ch^2 \| y \|_{H^2} \tag{3.27}$$

for some constant $C > 0$

This implies that we have a quadratic convergence for the finite element in L^2-norm, that is $\| y - y_h \|_{L^2(\Omega)} = O(h^2)$. Hence for our case the following stability estimates are valid for $f = y_d = 0$ and $g = 0$

$$\begin{aligned} \| y_l(u) \|_{H^1(\Omega)} &\leq C \| u \|_{L^2(\Omega)} \\ \| p_l(u) \|_{H^4(\Omega)} &\leq C \| y_l(u) \|_{H^1(\Omega)} \end{aligned}$$

3.3 Discrete Optimality System

For the optimization problem there are two ways: first optimize, then discretize or first discretize, then optimize(Hinze[10]). In this work the approach is optimize the optimization problem then discretize. The continuous optimal system(2.17) was transformed into the discrete optimal system by discretization using finite element method. Consider a sequence of discretizations with different step sizes h and level of refinement l. Fix the coarsest grid size h_0 and define

$$h_l = 2^{-l} h_0 \quad l \in \mathbb{N}_0 = 0, 1, 2, ...$$

where l is the level number. The discrete optimal system resulting from the Galerkin method is

$$min_{(y_l,u_l)\in L^2\times L^2} J(y_l,u_l) = \frac{1}{2} \parallel y_l - y_{d,l} \parallel_{L^2(\Omega)} + \frac{\delta}{2} \parallel u_l \parallel^2_{L^2(\Omega)} \quad (3.28)$$

subject to the constraints

$$A_l y_l + N_l y_l = N_l f_l + N_l u_l \quad (3.29)$$

$$A_l p_l + N_l p_l = N_l y_l - N_l y_{d,l} \quad (3.30)$$

$$N_l u_l = -\frac{1}{\delta} N_l p_l \quad (3.31)$$

where A_l and N_l are the stiffness and mass matrices at level l respectively.

Remark 3.11. *The state and adjoint elliptic partial differential equations have the same stiffness and the mass matrices at each level*

The idea is to solve a set of constraints for the values of the state and the control that minimise the objective function. Then the solution procedure is defined as follows(Hackbusch[6])

1. choose the initial control u_0
2. find $y_l = Su_0$, solve the elliptic state equation, where S involves A_l *and* N_l.
3. find $p_l = S^\star(y_l - y_{d,l})$
4. find $u_l = -\frac{1}{\delta} \cdot p_l$
5. where S is the solution operator with $S = S^\star$
6. set $u_0 = u_l$ and go to (1) with the multigrid method coming into play.

From (1-6) above it means that we can calculate the discrete state $y_l(v_l)$ from the discrete control v_l and the discrete adjoint $p_l(v_l)$ from the discrete state $y_l(v_l)$. Then we can define the optimal discrete control u_l for the distributed control problem as

$$u_l = -\delta^{-1} \cdot p_l(v_l) \quad (3.32)$$

The whole set of constraints of our discrete optimal control system (3.28 - 3.30)can be expressed as a matrix system

$$\begin{pmatrix} A_l + N_l & O & -N_l \\ -N_l & A_l + N_l & O \\ O & N_l & \delta N_l \end{pmatrix} \begin{pmatrix} y_l \\ p_l \\ u_l \end{pmatrix} = \begin{pmatrix} N_l f_l \\ -N_l y_{d,l} \\ O \end{pmatrix} \quad (3.33)$$

This defines the matrix equation of our solution procedure for the optimal control. The formulation(3.33) can be characterized by an integral equation.

3.3.1 Integral Equation Characterizing the Optimal Control

The Operator K

The mapping $u \mapsto y(u) \mapsto p(u) \mapsto -\delta^{-1}p(u)$ is affine and defined on a linear operator K such that the optimal control(2.15) can have the representation

$$-\delta^{-1}p(u) = Ku + q \tag{3.34}$$

This defines the solution procedure(Hackbusch[6]). The operator K is a linear operator. The operator K or the powers of K map into the space with finer topology. The powers of K signifies the number of times the operator K is applied in the solution process. Let B_0 and $B_1 \subset B_0$ be two Banach spaces, where B_1 is finer than B_0. K^m has to satisfy

$$\| K^m \|_{B_0 \to B_1} \leq C, \quad m \geq 1 \ fixed$$

In Ref.(5), for the choices of the two Banach spaces B_0 and B_1 for the distributed control equation(2.15, 2.16). Let $u \in L^2(\Omega)$ implies that $y(u) \in H^2(\Omega)$ and $p(u) \in H^4(\Omega)$, if Γ and the coefficients are smooth. If $m = 1, \quad B_0 = L^2(\Omega), \quad B_1 = H^4(\Omega)$. This means that more generally,

$$K^m : L^2(\Omega) \to H^{4m}(\Omega) \ \ continuous \ \ for \ \ all \ \ m \geq 1$$

For the other examples

1. Considering the Distributed control with Dirichlet boundary condition

$$K^m : L^2(\Omega) \to H^m(\Omega) \quad continuous$$

2. the Neumann boundary control problem(2.19)

$$K^m : L^2(\Gamma) \to H^{3m}(\Gamma) \quad continuous$$

3. Considering the Boundary control with Dirichlet boundary condition(2.24)

$$K^m : L^2(\Gamma) \to H^m(\Gamma) \quad \text{continuous}$$

The discrete optimality system (3.28 - 3.31) is going to be solved using the multigrid method. Let the discrete control v_l, discrete state and the adjoint $y_l(v_l),\ p_l(v_l)$ and the desired control u_l. The discrete optimality system(3.33) simplifies to

$$N_l u_l = -\frac{1}{\delta} N_l A_l^{-1} N_l \left[A_l^{-1} N_l (u_l + f_l) - y_{d,l} \right] \tag{3.35}$$

Definition 3.12. ***(Discrete Integral Equation)*** *The discrete optimality system(3.28 - 3.30) is equivalent to the discrete integral equation*

$$(I_l - K_l)u_l = q_l \tag{3.36}$$

with

$$\begin{aligned} K_l &= -\frac{1}{\delta} A_l^{-1} N_l (A_l^{-1} N_l) \\ q_l &= \frac{1}{\delta} A_l^{-1} N_l \left[A_l^{-1} N_l f_l - y_{d,l} \right] \end{aligned}$$

K_l is the discrete operator of K. The knowledge of the entries of the matrix K is not necessary except at the coarsest level($l = 0$). The similar approach can be used to derive the analogous discrete integral equation for the boundary control problem. In Hachbusch([6]), the most important requirement for K_l is that

$$\| K_l^m \|_{B_0^l \to B_1^l} \leq C, \quad l \in \mathbb{N}_0.$$

This will be applied in the next in chapter 4, on the convergence analysis of the multigrid method. The solution of the discrete integral equation is clearly the optimal control. To solve the discrete system the multigrid method(chapter 4) is applied.

Chapter 4

Solution of the Discretized Optimal Control System

In this chapter we want to develop a multigrid algorithm for solving the discretized optimal control system. The main goal being to find the pair (y_l, u_l) of the discrete control and the discrete state variables at the finest level l. To calculate this, an multigrid algorithm is developed over the discrete integral equation that characterizes the discrete optimal control. As has been already been highlighted in section(3.3.1), the discrete optimal constraints are reduced to one system of elliptic equations. So the multigrid algorithm will require only the numerical solution of a sequence of single elliptic equations. The optimal control can be obtained by solving one system of two elliptic equations. In this work we will illustrate the numerical treatment of a discrete optimal control problem by applying the multigrid method(MGM).

4.1 Multigrid Algorithm

In this section the multigrid algorithm is developed for the distributed control problem with the Neumann boundary condition(2.9). In this case we take $m = 1$ for the power of integral operator K. The multigrid algorithm adopted in this work is based on the work by Hachbusch([6]). The main ingredients of the multigrid iteration are the smoothing and coarse grid correction. The coarse grid correction process is carried out by a restriction, coarse grid solve, interpolation. Let $l \in \mathbb{N}_0$ be the refinement levels.

- For $l = 0$ the equation $u_l = K_l u_l + q_l$ where u_l is the desired control, is solved exactly by LU-decomposition of $I_0 - K_0$. At this level the entries matrix K_0 are known by evaluation of $K_0 v_0 + q_0$ for $q_0 = 0$ and all unit vectors v_0.
- **Smoothing:** for $l > 0$, that is if the level of refinement is not the coarsest, firstly the initial control u_l^ν at the level l(finest level) is smoothed by

$$u_l^{\nu+\frac{1}{2}} = K_l u_l^\nu + q_l \tag{4.1}$$

  where q_l and K_l are defined in (3.35, 3.36). For q_l is involved by $f_l, g_l, y_{d,l}$ and if we consider $f,\ g = 0$, then q_l will be involved by the desired state $y_{d,l}$ only. So we need to solve $K_l u_l$ separately. The explicit illustration follows

1. We need to solve the equation $u_l^{\nu_1} = K_l u_l^{\nu}$
   - choose the initial value $u_l^{\nu_0}$
   - With the initial control, solve for the state variable using the equation $A_l y_l = N_l u_l^{\nu_0}$ where A_l *and* N_l are the stiffness and mass matrices respectively at level l.
   - With y_l solve for the adjoint variables using the equation $A_l p_l = N_l y_l$ with same matrices defined above.
   - With p_l find the new control using the relation $u_l^{\nu_1} = -\delta^{-1} p_l$.
2. finally, $u_l^{\nu+\frac{1}{2}} = u_l^{\nu_1} + q_l$
3. repeat the smoothing process(1) with the initial $u_l^{\nu+\frac{1}{2}}$ to the equation, to get $u_l^{\nu_2} = K_l u_l^{\nu+\frac{1}{2}} + q_l$
4. result(2) is a smoother control than the initial one.

- **Calculating the defect:** After the smoothing process we calculate the corresponding defect

$$\begin{aligned} d_l &= (I_l - K_l) u_l^{\nu+\frac{1}{2}} - q_l & (4.2)\\ &= u_l^{\nu+\frac{1}{2}} - K_l u_l^{\nu+\frac{1}{2}} - q_l & (4.3)\\ &= K_l \big[u_l^{\nu} - u_l^{\nu+\frac{1}{2}} \big] & (4.4)\\ &= u_l^{\nu+\frac{1}{2}} - u_l^{\nu_2} & (4.5) \end{aligned}$$

- From the two smoothing processes above the defect can be expressed as $d_l = u_l^{\nu+\frac{1}{2}} - u_l^{\nu_2}$

- **Restrict the defect:** The restriction is an inter-grid transfer process. The process transfers the defect from the finer grid to a coarser grid. By a suitable restriction $r_{l,l-1} : B_0^l \to B_0^{l-1}$ to a coarser grid, we obtain the result

$$d_{l-1} = r_{l,l-1} d_l \in B_0^{l-1} \tag{4.6}$$

- Approximate on the coarser grid that is on the level $l-1$ by

$$w_{l-1} = (I_{l-1} - K_{l-1})^{-1} d_{l-1} \tag{4.7}$$

  two iterations of the multigrid method on the level $l-1$.

- **Prolongate the Approximate:** The prolongation/interpolation is an inter-grid transfer process. The process transfers the smooth error from the coarser grid to a finer grid. It is a linear mapping. By a suitable prolongation $p_{l-1,l} : B_0^{l-1} \to B_0^l$ to a finer grid and coarse grid correction, we obtain the result

$$u_l^{\nu+1} = u_l^{\nu+\frac{1}{2}} - p_{l-1,l} w_{l-1} \tag{4.8}$$

The above description gives the two grid algorithm. Applying the two grid recursively results in multigrid algorithm. Defining the multigrid method(MGM) recursively. We define the algorithm MGM_l at level $l > 0$ by means of the algorithm MGM_{l-1} corresponding to the coarser grid. Now we define the multigrid algorithm.

Multigrid Algorithm
We define the multigrid algorithm at level l as $MGM_l(u_l^{new}, u_l^{old}, q_l)$ where

- u_l^{new} is the output of one step of the multigrid algorithm at level l.
- u_l^{old} is the input at level l.
- q_l is defined implicitly by $f_l, g_l, y_{d,l}$ at level l.
- $u_l^{\nu} := u_l^{old} \mapsto u_l^{\nu+1} =: u_l^{new}$

Algorithm $MGM_l(u_l^{new}, u_l^{old}, q_l)$
if $l = 0$(coarsest grid)

$$\begin{aligned} u_0 &= (I_0 - K_0)^{-1} q_0 \\ &= MGM_0(u_0, q_0) \end{aligned}$$

else $l > 0$ define $MGM_l(u_l^{new}, u_l^{old}, q_l)$

1. **Smoothing**

$$\widetilde{u}_l = K_l u_l^{old} + q_l$$

   - **defect computation**

$$d_l = \widetilde{u}_l - K_l \widetilde{u}_l - q_l$$

2. **restrict the defect**

$$d_{l-1} = r_{l,l-1} d_l$$

3. **approximate solution**

$$v_{l-1} = K_{l-1} v_{l-1} + d_{l-1}$$

4. Applying two iterations of MGM_{l-1} at the recursive call:
   - Set $v_{l-1}^{(0)} = 0$
   - compute

$$v_{l-1}^{(1)} = MGM_{l-1}(v_{l-1}^{(1)}, v_{l-1}^{(0)}, d_{l-1})$$

   - compute

$$v_{l-1}^{(2)} = MGM_{l-1}(v_{l-1}^{(2)}, v_{l-1}^{(1)}, d_{l-1})$$

   - If $l = 1$, one call of $MGM_0(v_0^2, d_0)$ is sufficient.

5. **Correction Step**
   Define the new iterate by

$$u_l^{new} := \widetilde{u}_l - p_{l-1,l} v_{l-1}^{(2)}$$

Remark 4.1. *From the above multigrid algorithm*

- *The entries of the matrix K_l may be unknown*
- *We need the performance of mapping $v_l \mapsto K_l v_l$ which involves $v_l^{old} \mapsto y_l(v) \mapsto p_l(v) \mapsto -\delta^{-1} p_l(v) = v_l^{new}$*
- *The whole smoothing process involves the mapping $v_l \mapsto K_l v_l + q_l$ where q_l involves f_l, g_l, $y_{d,l}$*
- ***restriction:*** $r_{l,l-1}$ *is a restriction operator that takes the fine mesh function d_l to the coarse grid function d_{l-1}. In this work, restriction was be chosen by simply taking the fine-grid values at coarse-grid points (injection) of the parent nodes of the triangle elements.*
- ***prolongation:*** $p_{l-1,l}$ *is the prolongation operator of the coarse grid function $v_{l-1}^{(2)}$ to the fine mesh function w_l. This involves the parent grid points(coarse grid points) and their intermediate values obtained by averaging*

In this work we use a multigrid w-cycle which starts at a finest level l. The fine level solution is then transferred to next coarser level, (restriction). After some relaxation(smoothing) cycles on the coarse level, the solution is then restricted to next coarser level until the coarsest level is reached. The solution obtained at the coarsest level is than interpolated back to the finer level(prolongation). The solution from this finer level is interpolated to next finer level after some relaxation iterations. The solution is prolongated till the finest level is reached. The whole process is repeated until satisfactory convergence is reached.

4.2 Convergence of the Multigrid Method

In this section we look at the convergence analysis of the multigrid method in finding the optimal control with an aim of establishing the convergence rates. From the chapter 3 we realize that:

- Integral characterisation of the continuous optimality system $u = Ku + q$ and its corresponding discrete system $u_l = K_l u_l + q_l$ which represent the desired control.
- K is a linear operator such that K and powers of K map into a finer topology with $B_1 \subset B_0$ such that

$$K^m : B_0 \to B_1,$$

$$\| K^m \|_{B_0 \to B_1} \leq C_1, \qquad \| \cdot \|_0 \leq \| \cdot \|_1 \tag{4.9}$$

- Let $B_1^l \subset B_0^l$ be discrete vector spaces of the discrete control v_l

Assumption 4.2. *Let $l, \quad l-1$ be the two levels with l finer and $l-1$ coarser*

- *The norms of discrete spaces B_0^l and B_1^l with norms $| \cdot |_{0,l}$ and $| \cdot |_{1,l}$ and the continuous spaces B_0 and B_1 are connected by the continuous restriction and prolongation operators $R_l : B_i \to B_i^l, \ i = 0, 1$ and $P_l : B_0^l \to B_0$ respectively.*
- *The corresponding discrete restriction and prolongation operators $r_{l,l-1} : B_i^l \to B_i^{l-1}$ and $p_{l-1,l} : B_i^{l-1} \to B_i^l$ for $i = 0, 1$ respectively.*
- *The continuous and discrete integral equation matrices is invertible and bounded*

$$\| (I - K)^{-1} \|_{B_0 \to B_0} \leq C_2, \quad \| (I_l - K_l)^{-1} \|_{B_0^l \to B_0^l} \leq C_2, \quad \forall l. \tag{4.10}$$

  This is the stability condition.

- *The continuous and the discrete operators bounded*

$$\| K \|_{B_0 \to B_0} \leq C_3, \qquad \| K_l \|_{B_0^l \to B_0^l} \leq C_3, \quad \forall l. \tag{4.11}$$

$$\| K \|_{B_1 \to B_1} \leq C_3, \qquad \| K_l \|_{B_1^l \to B_1^l} \leq C_3, \quad \forall l. \tag{4.12}$$

- *Analogous to (4.9) is*

$$\| K_l^m \|_{B_l^0 \to B_1^l} \leq C_1', \qquad \| \cdot \|_{0,l} \leq \| \cdot \|_{1,l} \tag{4.13}$$

- *The discrete restriction and prolongation operators are bounded*

$$\| r_{l,l-1} \|_{B_0^l \to B_0^{l-1}} \leq C_4, \quad \| p_{l-1,l} \|_{B_0^{l-1} \to B_0^l} \leq C_4, \quad \forall l. \tag{4.14}$$

and also

$$\| r_{l,l-1} \|_{B_1^l \to B_1^{l-1}} \leq C_4, \quad \| R_l \|_{B_1 \to B_1^l} \leq C_4, \quad \forall l. \tag{4.15}$$

Then there exists

$$\widehat{P}_l : B_1^l \to B_1 \text{ with } R_l \widehat{P}_l = I_l, \quad \| \widehat{P}_l \|_{B_1^l \to B_1} \leq C_4 \tag{4.16}$$

- *The condition*

$$\| I_l - p_{l-1,l} r_{l,l-1} \|_{B_1^l \to B_0^l} \leq C_5 (n_{l-1})^{-\alpha} \tag{4.17}$$

$\forall n_{l-1} \in \mathbb{N}$ *dimension of coarser grid* , $\alpha \geq 0$ *means that the smooth functions from* B_1^l *to* B_0^{l-1} *can be approximated.*

-

Remark 4.3. *From (4.9), (4.10) and (4.12) it can be concluded that The continuous and discrete integral equation matrices is invertible and bounded*

$$\| (I-K)^{-1} \|_{B_1 \to B_1} \leq C_2' := C_1 C_2 + 1 + C_3 + ... + C_3^{m-1} \tag{4.18}$$

PROOF
Let $q \in B_1, \quad u := (I-K)^{-1} q$*, with the repeated application of* K *from the beginning* $u := Ku + q$*, we have*

$$\begin{aligned} u &= K^2 u + Kq + q \\ &= K^m u + K^{m-1} q + ... + Kq + q \end{aligned}$$

Then from (4.9) we get

$$\begin{aligned} \| K^m u \|_1 &\leq C_1 \| u \|_0 \\ &= C_1 \| (I-K)^{-1} q \|_0 \\ &\leq_{(4.10)} C_1 C_2 \| q \|_0 \\ &\leq_{(4.9)} C_1 C_2 \| q \|_1 \end{aligned}$$

and the term $K^v q$ *with* $0 \leq v \leq m-1$ *using (4.13) gives*

$$\| K^v q \|_1 \leq C^v \| q \|_1$$

Hence the result.

- *The consistence condition is formulated*

$$\begin{aligned} |\big[(I_l - K_l)R_l - R_l(I - K)\big]u|_{0,l} &= |(K_l R_l - R_l K)u|_{0,l} \\ &= |(I_l - K_l)R_l u - R_l q|_{0,l} \\ &\leq C_6 n_l^{-\beta} \| q \|_1 \end{aligned} \tag{4.19}$$

$\forall n_l \in \mathbb{N}$ *dimension of finer grid and* $\beta > 0$

- *Stability(4.10) and Consistency(4.19) implies convergence*

$$|u_l - R_l u|_{0,l} \leq C_7 n_l^{-\beta} \| q \|_1 \tag{4.20}$$

ANALYSIS OF THE MULTIGRID ALGORITHM

An iteration of single multigrid step consists of a combination of smoothing step and a coarse grid correction step. The following relations result from the application of each step

- The exact solution is given by the relation

$$u_l = K_l u_l + q_l \tag{4.21}$$

- **Smoothing:** is done by the application of the relation

$$u_l^\nu \mapsto \widetilde{u}_l := K_l u_l^\nu + q_l \tag{4.22}$$

apply K m-times $m > 1$ we have the error

$$\begin{aligned} \widetilde{u}_l - u_l &= K^m(u_l^\nu - u_l) &(4.23) \\ &= K^m \triangle u_l^\nu &(4.24) \end{aligned}$$

- **Calculating the defect**

$$d_l = (I_l - K_l)\widetilde{u}_l - q_l \tag{4.25}$$

- From the exact solution and the defect relations we get

$$u_l = \widetilde{u}_l - (I_l - K_l)^{-1} d_l \tag{4.26}$$

From this relation we have

$$d_l = (I_l - K_l)(\widetilde{u}_l - u_l) \tag{4.27}$$

- **Coarse grid correction**: Produces the relation for the new iterate

$$u_l^{\nu+1} = \widetilde{u}_l - p_{l-1,l}(I_{l-1} - K_{l-1})^{-1} r_{l,l-1} d_l \tag{4.28}$$

$$\underbrace{=}_{4.27} \widetilde{u}_l - p_{l-1,l}(I_{l-1} - K_{l-1})^{-1} r_{l,l-1}(I_l - K_l)(\widetilde{u}_l - u_l) \tag{4.29}$$

- The error of the new iterate, subtract the exact solution from both sides

$$\begin{aligned} u_l^{\nu+1} - u_l &= \widetilde{u}_l - p_{l-1,l}(I_{l-1} - K_{l-1})^{-1} r_{l,l-1}(I_l - K_l)(\widetilde{u}_l - u_l) - u_l \\ \triangle u_l^{\nu+1} &= \widetilde{u}_l - u_l - p_{l-1,l}(I_{l-1} - K_{l-1})^{-1} r_{l,l-1}(I_l - K_l)(\widetilde{u}_l - u_l) \\ &\underbrace{=}_{4.23} K^m \triangle u_l^\nu - p_{l-1,l}(I_{l-1} - K_{l-1})^{-1} r_{l,l-1}(I_l - K_l) K^m \triangle u_l^\nu \\ &= \left[I_l - p_{l-1,l}(I_{l-1} - K_{l-1})^{-1} r_{l,l-1}(I_l - K_l)\right] K^m \triangle u_l^\nu \end{aligned}$$

- The above relation for the error can be expressed as

$$\triangle u_l^{\nu+1} = M_l \triangle u_l^\nu \tag{4.30}$$

where M_l is the iteration matrix given by

$$M_l = \left[I_l - p_{l-1,l}(I_{l-1} - K_{l-1})^{-1} r_{l,l-1}(I_l - K_l)\right] K^m \tag{4.31}$$

- The idea is to establish the convergence rates for our multigrid algorithm which are defined by the relation

$$\frac{\| \triangle u_l^{\nu+1} \|_0}{\| \triangle u_l^\nu \|_0} = \| M_l \|_{B_0 \to B_0} \tag{4.32}$$

The convergence property of the iterative process depends on $\| M_l \|_{B_0 \to B_0}$

Theorem 4.4. *Let $l, l-l \in \mathbb{N}$ and $0 < \sigma \leq \frac{n_{l-1}}{n_l} < 1$, then from the conditions (4.9), (4.10), (4.11), (4.12), (4.13), (4.14),(4.15), (4.16), (4.17) and (4.20) it holds*

$$\| M_l \|_{B_0 \to B_0} \leq C_{10} n_l^{-\sigma} + C_{11} n_l^{-\beta}$$

The method converges for sufficiently large n_l.

PROOF

Let $w_l \in B_0^l$ with norm $\| w_l \|_{0,l} = 1.$ and $v_l = K_l^m w_l$ (by smoothing).

Applying (4.13) we get

$$|v_l|_{1,l} \leq C_1^{'} \tag{4.33}$$

Since the right hand side cancels out, set

- defect: $d_l = v_l - K_l v_l$
- restriction: $d_{l-1} = r_{l,l-1} d_l$
- approximation: $v_{l-l} = (I_{l-1} - K_{l-1})^{-1} d_{l-1}$

With the continuous solution $v \in B_1$ gives defect $v - Kv = d := \widehat{P} d_l$
The new iterate applying the iteration matrix to initial value is given by

$$M_l w_l = v_l - p_{l,l-1} v_{l-1} \tag{4.34}$$

The equation (4.34) can be divided into three parts

1. error after smoothing $v_l - R_l v$ where R_l is a continuous restriction operator
2. prolongation of the error $p_{l-1,l}(R_{l-1} v - v_{l-1})$
3. prolongation of the continuous value $R_l v - p_{l-1,l} R_{l-1} v$

Now analysing the items (1-3)

- For (1) show that it is bounded

$$\begin{aligned} |v_l - R_l v|_{0,l} &\underbrace{\leq}_{4.20} C_7 n_l^{-\beta} \| d \|_1 \\ &\underbrace{\leq}_{4.16} C_4 C_7 |d_l|_{1,l} n_l^{-\beta} \\ &\underbrace{\leq}_{4.12} (1 + C_3) C_4 C_7 n_l^{-\beta} |v_l|_{1,l} \\ &\underbrace{\leq}_{4.33} C_1^{'} (1 + C_3) C_4 C_7 n_l^{-\beta} \end{aligned}$$

- For (2) we follow the same process

$$\begin{aligned}
|p_{l,l-1}(R_{l-1}v - v_{l-1})|_{0,l} &\underbrace{\leq}_{4.14} C_4|R_{l-1}v - v_{l-1}|_{0,l} \\
&\underbrace{\leq}_{as\ above} C_4^2 C_7 (n_{l-1})^{-\beta}|d_{l-1}|_{1,l-1} \\
&\underbrace{\leq}_{4.15} C_4^3 C_7 (n_{l-1})^{-\beta}|d_l|_{1,l} \\
&\underbrace{\leq}_{as\ in\ 1} C_1^{'}(1+C_3)C_4^3 C_7 \sigma^{-\beta} n_l^{-\beta}
\end{aligned}$$

since $0 < \sigma \leq \dfrac{n_{l-1}}{n_l} < 1$

- For the third we use the expression (4.17) and the remark (4.3)

$$\begin{aligned}
|R_l v - p_{l-1,l}R_{l-1}v|_{0,l} &\underbrace{\leq}_{r_{l,l-1}R_l = R_{l-1}} |(I_l - p_{l-1,l}r_{l,l-1})R_l v|_{0,l} \\
&\underbrace{\leq}_{4.17} C_5(n_{l-1})^{-\alpha}|R_l v|_{1,l} \\
&\underbrace{\leq}_{4.15} C_4 C_5 (n_{l-1})^{-\alpha} \parallel v \parallel_1 \\
&\underbrace{\leq}_{remark(4.3)} C_2^{'} C_4 C_5 (n_{l-1})^{-\alpha} \parallel d \parallel_1 \\
&\underbrace{\leq}_{4.16} C_2^{'} C_4^2 C_5 (n_{l-1})^{-\alpha}|d_l|_{1,l} \\
&\underbrace{\leq}_{as\ above} C_1^{'} C_2^{'}(1+C_3)C_4^2 C_5 \sigma^{-\alpha} n_l^{-\alpha}
\end{aligned}$$

since $0 < \sigma \leq \dfrac{n_{l-1}}{n_l} < 1$

Collecting things together , the result follows with

$$C_{10} = C_1^{'} C_2^{'}(1+C_3)C_4^2 C_5 \sigma^{-\alpha} \quad \text{and} \quad C_{11} = C_1^{'}(1+C_3)C_4 C_7(1+C_4^2)\sigma^{-\beta}.$$

Finally, from the above proof we can deduce that $\parallel M_l \parallel_{B_0 \to B_0} \leq C_{10} n_l^{-\sigma} + C_{11} n_l^{-\beta}$

$$\begin{aligned}
\parallel M_l \parallel_{B_0 \to B_0} &\leq C_{10} n_l^{-\sigma} + C_{11} n_l^{-\beta} &(4.35)\\
&\leq C n_l^{-\beta} &(4.36)
\end{aligned}$$

where $C = C_{10} + C_{11}$ and $max(\sigma, \beta) = \beta$

Conclusion 4.5. *The rate of convergence of the multigrid method on the level $l \in \mathbb{N}_0$ is proportional to h_l^{β} for some $\beta > 0$. This means that the estimate*

$$\| u_l^{\nu+1} - u_l \|_{B_0^l} = C h_l^{\beta} \| u_l^{\nu} - u_l \|_{B_0^l} \tag{4.37}$$

where u_l is the discrete exact solution, holds for two consecutive iterates. Since u_l is unknown we use the continuous control u to get thet convergence rates of the multigrid algorithm. This is achieved by integrating over each triangle element(chapter5, p.44).

Hachbusch([5]) concluded that rate of convergence is proportional to by a factor h_l^{β} which means $\beta > 0$

Chapter 5

Numerical Results

In this chapter we present the result of a distributed optimal control problem with Neumann boundary condition. We pay particular attention to the computational performance of the proposed multigrid scheme as a solver of the distributed control problems. In chapter 2 we considered the distributed control problem(2.4, 2.5), transformed it into an optimality system(2.17) and finally the characterization by integral equation(3.34, 3.35) as applied by Hachbusch([6]). We consider examples where $f = g = 0$ and the right hand side q will only depend on the target state y_d. The numerical treatment is given to the integral equation which characterises the optimality system.

Implementations were performed on a windows XP platform with 1.6 GHz speed intel(R) processor by using Matlab 7 programming language.

For all the examples, we approximate $\overline{\Omega} = (0,1)^2$ by a triangular mesh and use the Matlab pdetool for the mesh generation. The mesh data describes the triangulation and it consists of a list of coordinates for nodes **p**(array of coordinates), geometry **g**, a list of triangles **t**(array of the vertices of **p**), edge connections **e** and a list of all edges that describe the Neumann boundary(has no contribution in this work since the Neumann condition is zero). The mesh is stored for each refinement and the assembled matrices are also stored for each refinement level.

The Mesh

In this work we use the structured mesh and regular refinements. The meshes are generated by the matlab pdetool. Since the multigrid method requires a hierarchy of grids which are produced by successive refinements, we need to choose the coarse mesh(should be as coarse as possible), finest mesh which corresponds to the maximum level of refinement. The figures below show an example of the refinement levels(the examples below we use the coarse level to have 25 nodes).

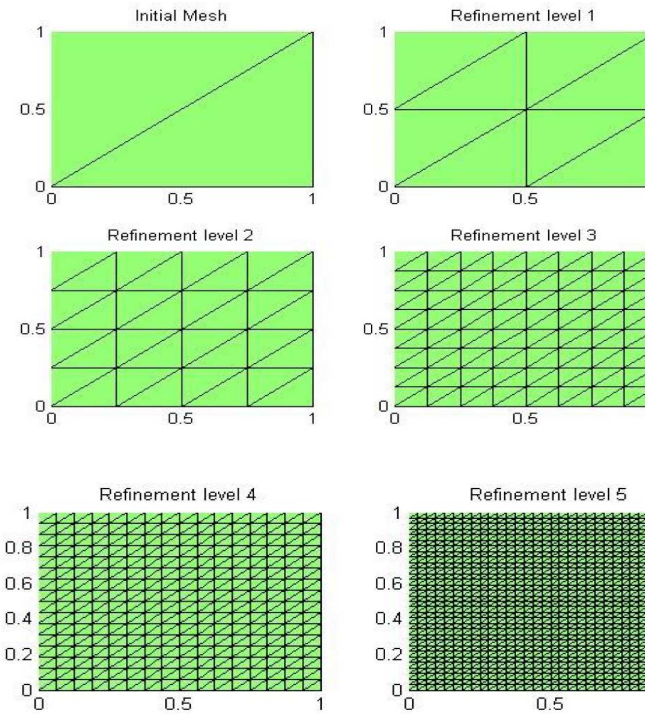

Figure 5.1: Levels of refinement

The table below shows the refinement levels and the number of grid points for each level.

Refinement Level(l)	mesh size(h_l)	Nodes(number of grid points)
0	$\frac{1}{4}$	25
1	$\frac{1}{8}$	81
2	$\frac{1}{16}$	289
3	$\frac{1}{32}$	1089
4	$\frac{1}{64}$	4225
5	$\frac{1}{128}$	16641

Table 5.1: Refinement Levels and Number of Nodes

Since we have the Neumann condition, the number of nodes is the same as the degrees of freedom. In this work we consider the coarse level to have 25 nodes. The following cases for the examples which we tackle numerically in this chapter are

- when the exact solutions for u, p, y are known.
- testing with different target state.
- testing with different initial control u_0 at the finest level that is
  - the initial control $u_0 = 0$.
  - the initial control $u_0 = exact \;\; solution$

We have established from literature that accuracy of the approximation by applying piecewise linear functions is $O(h^2)$ in L^2-norm. To compute the L^2-error, we approximate the element with both the numerical solution u_l and the exact solution u at the centroid $(x(s), y(s))$ of each of the triangles T_i. Then the L^2-error is calculated according to $\| u - u_l \|_{L^2} = \sqrt{\sum area(T_i) \cdot (u(x(s), y(s)) - u_l(x(s), y(s)))^2}$. We have also established that the convergence rate of the multigrid algorithm is proportional to $h_l^\beta, \;\; \beta > 0$.

5.1 Test Example 1

In this example we consider the target state depending on the exact solutions of the state, control and the adjoint. The exact solutions for the distributed optimality system(2.17) are

$$y = \cos(\pi x1)\cos(\pi x2) \tag{5.1}$$
$$u = (2\pi^2 + 1)\cos(\pi x1)\cos(\pi x2) \tag{5.2}$$
$$p = -\delta(2\pi^2 + 1)\cos(\pi x1)\cos(\pi x2) \tag{5.3}$$

We get the corresponding desired state as

$$y_d = (\delta(2\pi^2+1)^2+1)\cos(\pi x1)\cos(\pi x2)$$

The graphs of the target state and the exact solution for the control are given below

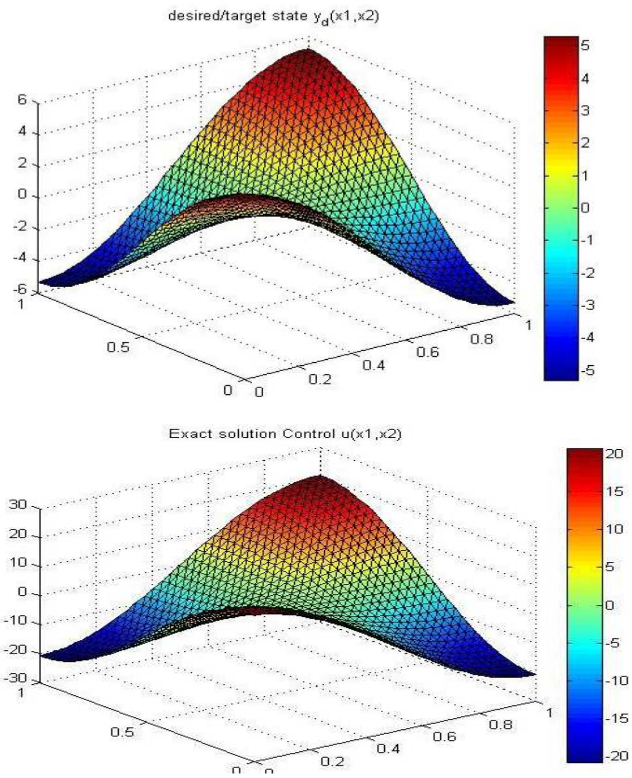

Figure 5.2: **Above:** Target state. **Below:** Exact solution for the control at level 4

We will consider two cases for the initial control

- example 1: we take the initial control $u_0 = 0$.
- example 2: we take the initial control $u_0 = exact \ \ solution$

The goal is to find the optimal control so that the target state can be achieved by solving the integral equation(3.36) as outlined in chapter 4. In the tables (5.2) and (5.3) below, we present the numerical results of the multigrid method with $l = 4$ and $l = 5$ grid levels with the weighting parameter $\alpha = 1e - 2$ chosen. On the finest grid $l = 5$, it is important to note that in both cases the multigrid method converges very fast in 3 iterations. This means that in successive iterations the distance between the iterates of the control becomes continuously small. The stopping criteria(tolerance) was chosen to be $tolerance = 10^{-8}$. That is $\| u_l^{k+1} - u_l^k \|_{L^2(\Omega)} \leq 10^{-8}$. The results on the tables also confirms a well known result that $\| u - u_l^k \|_{L^2(\Omega)} = O(h_l^2)$. This means that the multigrid method converges very fast. This feature is also reflected in the error.

iterations	Control	tolerance	error
k	$\| u_l^k \|_{L^2(\Omega)}$	$\| u_l^{k+1} - u_l^k \|_{L^2(\Omega)}$	$\| u - u_l^k \|_{L^2(\Omega)}$
1	10.36169892	10.36169892	0.008463995
2	10.36035543	3.446123e-6	0.009776728
3	10.36035560	8.1541817e-14	0.009767091

Table 5.2: $l = 4, \quad nodes = 4225, \quad \delta = 1e - 2, \quad h_l = \frac{1}{64}, \quad tolerance = 1e - 8$

iterations	Control	tolerance	error
k	$\| u_l^k \|_{L^2(\Omega)}$	$\| u_l^{k+1} - u_l^k \|_{L^2(\Omega)}$	$\| u - u_l^k \|_{L^2(\Omega)}$
1	10.36062264	10.36762264	0.002118022
2	10.36029109	2.1515235e-7	0.002118023
3	10.36029109	3.17562278e-16	0.002443111

Table 5.3: $l = 5, \ nodes = 16641, \quad \delta = 1e - 2, \ h_l = \frac{1}{128}, \ tolerance = 1e - 8$

In table (5.4), a detailed consideration of the iteration errors for all the levels. The table also shows that the iteration errors are reduced from the coarse level to finest level by $\frac{1}{4}$ with the fineness of the grids. This means that further refinement reduces the error and as the step size becomes smaller and the iteration error approaches zero. All these confirms what the theory says.

level	iterations	Control	error
l	k	$\parallel u_l^k \parallel_{L^2(\Omega)}$	$\parallel u - u_l^k \parallel_{L^2(\Omega)}$
0	1	8.27853197	2.178158212
1	5	9.79716843	0.602113822
2	3	10.22288193	0.154778976
3	3	10.33267581	0.038988659
4	3	10.36035560	0.009767709
5	3	10.36029109	0.002443111

Table 5.4: $Convergence\ \ results\ \ \delta = 1e-2,\ \ tolerance = 1e-8$

The behaviour of the error from the coarse level to the finest level is represented in the figure below. The figure gives the visual representation of the behaviour of the error in table(5.4). It depicts a rapid convergence.

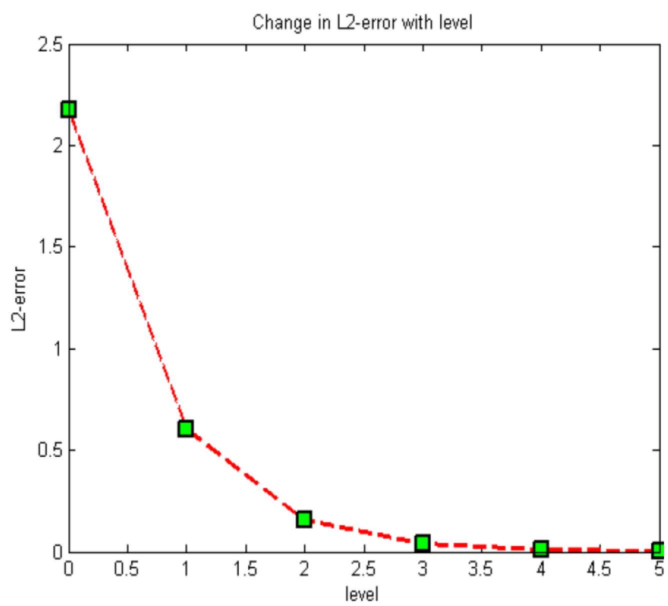

Figure 5.3: Behaviour of L2-error

The weighting factor of the control has also an impact in the results. The table below shows the effects of the weighting factor on the performance of the method. The results in the table show that the change weighting factor lead the change in the optimal control and the error. The increase in error with increase in the weighting factor means that the weighting factor of the cost functional should be chosen reasonably small. If we take large values of δ the error increases. For this work the value $\delta = 1e-2$ produce optimal results. The table shows that for large values of weighting factor the discretization error grows. It has been noted in this work that at any grid level the method diverges for the values $\delta \leq 1e-3$. The results in the table demonstrates the effect at finest level.

w. parameter	iterations	Control	error
δ	k	$\parallel u_l^k \parallel_{L^2(\Omega)}$	$\parallel u - u_l^k \parallel_{L^2(\Omega)}$
1e-2	3	10.36029109	0.00244311
5e-2	3	10.36686262	0.00290681
1e-1	3	10.36679819	0.00306427
5e-1	2	10.36674483	0.00428668
7.5e-1	2	10.36674001	0.00490374

Table 5.5: *Changes in* δ, $l = 5$, $nodes = 16641$, $h_l = \frac{1}{128}$

Snapshot of the approximate optimal control at level 4

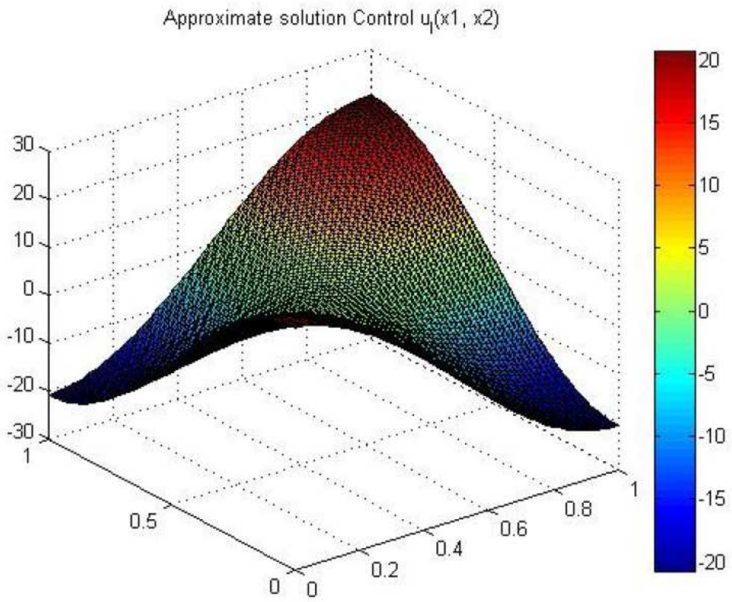

Figure 5.4: Snapshot of the optimal control at l=4

5.2 Example 2

In this section we focus on the same distributed optimality system in example 1 and solve it with the initial control different from zero. For example, we take $u_0 = (2\pi^2 + 1)\cos(\pi x1)\cos(\pi x2)$. That is taking the exact solution as the initial control. We observe that we achieve the same convergence results and that it converges fast in 2 iterations. The figures in example 1 are the same for this case. We present the results at levels 4 and 5. From the calculations on the finest level we achieve the following results

iterations	Control	tolerance	error
k	$\parallel u_l^k \parallel_{L^2(\Omega)}$	$\parallel u_l^k - u_l^{k+1} \parallel_{L^2(\Omega)}$	$\parallel u - u_l^k \parallel_{L^2(\Omega)}$
1	10.36035509	1.83907065e-5	0.00976765
2	10.36035560	1.35312409e-13	0.00976709

Table 5.6: $l = 4, \quad nodes = 4225, \quad \delta = 1e-2, \quad h_l = \frac{1}{64}, \quad tolerance = 1e-8$

iterations	Control	tolerance	error
k	$\parallel u_l^k \parallel_{L^2(\Omega)}$	$\parallel u_l^k - u_l^{k+1} \parallel_{L^2(\Omega)}$	$\parallel u - u_l^k \parallel_{L^2(\Omega)}$
1	10.36029106	1.15190037e-6	0.002444314
2	10.36029109	2.0581872e-15	0.002443111

Table 5.7: $l = 5, \quad nodes = 16641, \quad \delta = 1e-2, \quad h_l = \frac{1}{128}, \quad tolerance = 1e-8$

5.3 Example 3

In this section we consider the case where the exact solutions are not known. We choose the target state not depending on the exact solutions. We choose the target state as $y_d = x_1 + x_2$ and the using a zero initial control.

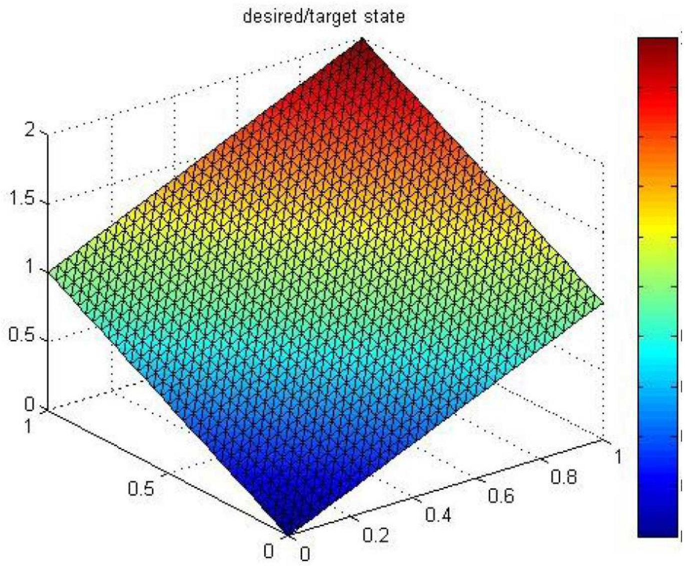

Figure 5.5: Desired/Target State at l=4

The behaviour of the control and the stopping criteria on tables(5.8, 5.9, 5.10) from the coarse grid to the finest grid reflect that the method converges to the optimal control.

level	iterations	Control
l	k	$\| u_l^k \|_{L^2(\Omega)}$
1	4	2.23636568
2	3	2.24634503
3	3	2.24879971
4	3	2.24941154
5	3	2.24946390

Table 5.8: $Control, \quad \delta = 1e-2, \quad tolerance = 1e-8$

In this case we illustrate the results at levels 4 and 5. The results for all the levels are presented in the tables(5.9) and (5.10).

no. of iterations	Control	tolerance
k	$\parallel u_l^k \parallel_{L^2(\Omega)}$	$\parallel u_l^k - u_l^{k+1} \parallel_{L^2(\Omega)}$
1	2.24956732	5.06055475
2	2.249411145	2.680565e-7
3	2.249411154	5.1360268e-15

Table 5.9: $l = 4, \quad nodes = 4225, \quad \delta = 1e-2, \quad h_l = \frac{1}{64}, \quad tolarence = 1e-8$

The results for the finest level

no. of iterations	Control	error
k	$\parallel u_l^k \parallel_{L^2(\Omega)}$	$\parallel u_l^k - u_l^{k+1} \parallel_{L^2(\Omega)}$
1	2.24960274	5.0607125
2	2.24995638	1.674958e-8
3	2.24946390	2.0000481e-17

Table 5.10: $l = 5, \quad nodes = 16641, \quad \delta = 1e-2, \quad h_l = \frac{1}{128}, \quad tolerance = 1e-8$

The figure below represents the optimal control for this case. The result was achieved in 5 iterations of the multigrid method,

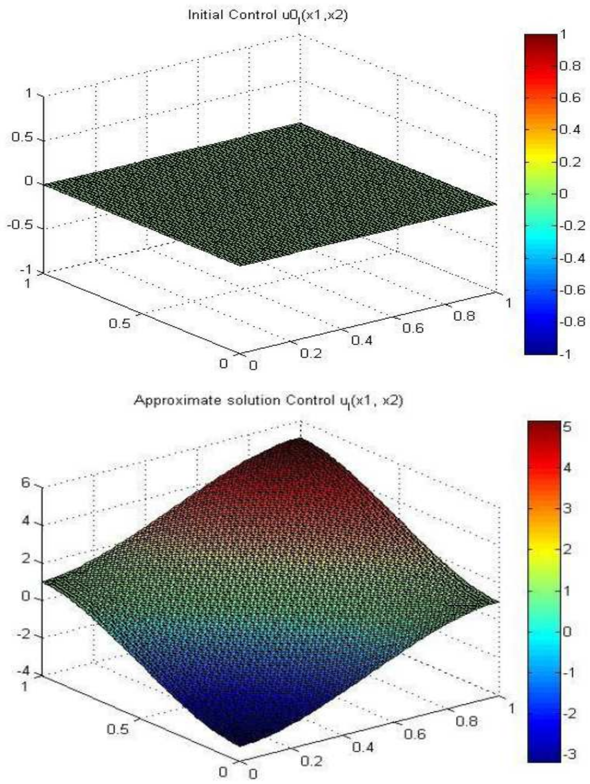

Figure 5.6: **Above:** Initial Control. **Below:** Approximate solution for the control at level 4

5.4 Conclusion

In this work we considered the mathematical model of the optimal control problems governed by the elliptic partial differential equations. We have highlighted that the model can be described by the distributed control or boundary control with Neumann or Dirichlet boundary conditions. However, in this work we have given the mathematical treatment to the distributed control model governed by the elliptic partial differential equation with a zero Neumann boundary data. The same mathematical treatment can be applied to control problems the with Dirichlet, non-zero Neumann data. Also the same treatment can be applied to the models where the control is on the boundary, boundary control optimal problems. We described the existence of minimizer of the cost functional which is the control variable. For the distributed control we have shown that the optimal control problem can be expressed as a integral equation where the multigrid method is applied to compute for the optimal control. We have established that multigrid method converges by a factor proportional to the refinement step h between consecutive multigrid iterations. In this work we have used the W-cycle of the multigrid method. We started with an initial guess for initial control, smooth, calculate the defect, restrict, solve exactly on the coarse grid then prolongate and coarse grid correction. For the numerical implementation we use the finite element with piecewise linear elements.

We test our method by an example with different cases namely, non-zero initial control, zero initial control and a case where the exact solutions are unknown where the target state is prescribed. We observe that the iteration for the multigrid method decreases rapidly with the fineness of the grids.

Bibliography

[1] Borzi, A. and Kunisch, K. A Multigrid Scheme for Elliptic Constrained Optimal Control Prpblems, Computational Optimaization and Applications, 31(2005), 309-333.

[2] Borzi, A . Smoother for Control-and State- constrained Optimal Control Problems, Computational Science, 11(2008), 59-66.

[3] Braess, D. Finite Elements, Cambridge University press, 2007.

[4] Brenner , C. S and Scott, R. L. The Mathematical Theory of Finite Element Methods, Springer, New York, 2008.

[5] Ciarlet, Philippe G. The finite element methods for elliptic problems, North- Holland publishing company, 1978

[6] Hackbusch, W. Fast solution of elliptic control problems, Journal of Optimization Theory and Appli- cation, 31 (1980), 565581.

[7] Hackbusch, W. On the fast solving of parabolic boundary control problems, SIAM J. Control and Optim., 17 (1979), pp. 231244.

[8] Hackbusch, W. Multi-grid Methods and Applications, Springer-Verlag, New York, 1985.

[9] Hackbusch, W. Die schnelle Auflsung der Fredholmschen Integralgleichung zweiter Art, Beitrige Numer. Math 9.(1981), 47-62.

[10] Hinze, M.(etal). Modelling and Optimization of PDE-constrained problems, Short Summary of Autumn School 2005, 26.-30. September 2005, Hamburg.

[11] Knabner, P. and Angermann, L. Numerical Methods for Elliptic and Parabolic Differential Equations, Springer, Berlin, 2003

Eidesstattliche Erklärung

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

Linz, August 2008
Muzhinji Kizito