ENERGY ERROR ESTIMATES AND DOMAIN DECOMPOSITION METHODOLOGY IN FINITE ELEMENT METHOD

Thesis submitted in partial fulfillment of the requirements for
the award of degree of
Masters of Science
in
Mathematics and Computing

Submitted by
Archna Kumari
Reg. No.-301403002

Under
the guidance of
Dr. Vivek Sangwan

School of Mathematics
Thapar University, Patiala-147004
PUNJAB, INDIA
July, 2016

CERTIFICATE

I hereby certify that the work which is being presented in the thesis entitled **"ENERGY ERROR ESTIMATES AND DOMAIN DECOMPOSITION METHODOLOGY IN FINITE ELEMENT METHOD"** in partial fulfillment of the requirements for the award of degree of Master of Science in "Mathematics and Computing" to the School of Mathematics, Thapar University, Patiala is an authentic record of my own work studied under the supervision of Dr. Vivek Sangwan.

The matter presented in this thesis has not been submitted for the award of any other degree of this or any other university.

Archna

(*Archna Kumari*)

Reg. No. 301403002

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

Dr. Vivek Sangwan
Assistant Professor
School of Mathematics
Thapar University, Patiala

Countersigned by:

Dr. A.K. Lal
Associate Professor and Head, SOM
Thapar University, Patiala.

Prof. S.S. Bhatia
Dean of Academic Affairs
Thapar University, Patiala.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, **Dr. Vivek Sangwan**, Assistant Professor, School of Mathematics, Thapar University, Patiala, for the patient guidance, encouragement and advice he has provided throughout my time as his student. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly.

I would like to express my sincere thanks to **Dr. A.K. Lal**, Head, School of Mathematics, Thapar University, Patiala, for providing the necessary facilities needed for this work and directly or indirectly encouraged me to do work. I would also thank all academic and administrative staff of School of Mathematics, Thapar University, Patiala who have contributed directly or indirectly to this work.

I would also like to thank my parents for their continuous support which inspired me and encouraged me, without them I would not be able to complete my work.

Life at Thapar University, Patiala has been enjoyable with friends who have been always there for me, listening to me, rejoicing me, complaining and pondering my way throughout my study. I thank them all for their great company and finally to God who made all the things possible.

(ARCHNA KUMARI)

ABSTRACT

Differential equations arise in almost every area of science and engineering for example chemical engineering, aerodynamics, fluid dynamics, material science, astrophysics, economics etc. Day by day since the systems are becoming more and more complicated, the differential equations governing the respective models are also becoming more and more complex and nonlinear in nature. Because of this complexity and nonlinearity, it is very difficult to find the exact solution and we need numerical methods in order to solve the differential equations. In the present work, we will use finite element method to find the approximate solutions of the differential equations. Also, while applying any numerical technique, it becomes very important to estimate the error present in the numerical solution. Therefore, in the present work, due importance has been given to both the a priori and a posteriori error estimates. Since many differential equations arising in the fields like oceanography, elasticity etc. have very large domains and at the same time at many instances, because of the need of immediate solution requirements, it becomes very important to use parallel computing, in order to get an immediate solution. Domain decomposition techniques are highly used for incorporating parallelism. In the present work, a domain decomposition algorithm based on monotone Schwarz iterative process has been discussed in the finite element framework. The whole work has been divided into three chapters.

Chapter 1 introduces the basic concepts of differential equations and the solution methodology. A brief history of finite element methods has been presented and Galerkin finite element method has been explained with the help of examples.

Chapter 2 contains elementary estimates on a priori and a posteriori errors present in the finite element solution in the energy norm. These estimates have been derived for Galerkin

finite element discretizations of general linear elliptic operators in the research paper entitled “A tutorial in elementary finite element error analysis: A systematic presentation of a priori and a posteriori error estimates” by J. Stewart and T.J.R. Hughes. In the present work, these estimates have been derived by taking a particular linear elliptic differential equations.

In **Chapter 3**, a domain decomposition algorithm based on monotone Schwarz iterative process has been discussed using finite element method for a nonlinear problem. The monotone Schwarz iterative domain decomposition algorithm has been presented in the research paper entitled “A block monotone domain decomposition algorithm for a semilinear convection-diffusion problem” by Boglaev in the finite difference framework. In the presented work, the method has been discussed in the finite element framework. Bibliography has been presented in the last.

Contents

Chapter 1

Differential Equations and Finite Element Method - A Brief Overview

1.1 Introduction

Differential equations arise in almost every area of science and engineering for example chemical engineering, aerodynamics, fluid dynamics, electromagnetism, material science, astrophysics, economics etc. Day by day since the systems are becoming more and more complicated, the differential equations governing the respective models are also becoming more and more complex and nonlinear in nature. Because of this complexity and nonlinearity, it is very difficult to find the exact solution and we need numerical methods in order to solve the differential equations. In the present study, we will use finite element method to find the approximate solutions of the differential equations. In the starting, we will present the definitions of some of the concepts which have been used in the report.

1.2 Differential Equations

A differential equation is a relation between dependent variables, independent variables and the derivatives of the dependent variables with respect to the independent variables.

For example,

$$\begin{aligned} \frac{dy}{dx} + \sin(y) &= x; \\ \frac{d^2y}{dx^2} + 2\frac{dy}{dx} + y &= 0; \\ \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} &= xy. \end{aligned}$$

Differential equations can be categorized in two categories:

(1) **Ordinary differential equations**

(2) **Partial differential equations**

1.2.1 Ordinary Differential Equations

An ordinary differential equation (ODE) is a differential equation in which the unknown function is a function of a single independent variable. For example,

$$\begin{aligned} \frac{dy}{dx} + \sin(y) &= x; \\ \frac{d^2y}{dx^2} + 2\frac{dy}{dx} + y &= 0. \end{aligned}$$

1.2.2 Partial Differential Equations

A partial differential equation (PDE) is a differential equation in which the unknown function is a function of more than one independent variables and their partial derivatives. For example,

$$\begin{aligned} \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} &= xy; \\ \frac{\partial^2 z}{\partial x^2} + z\frac{\partial z}{\partial y} &= 0. \end{aligned}$$

Further, ordinary differential equations can be classified into two types:

(i) **Linear ordinary differential equations and**

(ii) Nonlinear ordinary differential equations

1.2.3 Linear Ordinary Differential Equations

An ordinary differential equation is said to be linear if the differential equation is linear in the dependent variable and its derivatives. For example,

$$\frac{dy}{dx} + \sin(x)y = x;$$
$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + y = 0.$$

1.2.4 Nonlinear Ordinary Differential Equations

An ordinary differential equation is said to be nonlinear if the differential equation is not linear in the dependent variable and its derivatives. For example,

$$\frac{dy}{dx} + \sin(y) = x;$$
$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + y^2 = 0.$$

Furthermore, partial differential equations can be classified into four types:

(i) **Linear partial differential equations;**

(ii) Semilinear partial differential equations;

(iii) Quasilinear partial differential equations; and

(iv) Nonlinear partial differential equations.

1.2.5 Linear Partial Differential Equations

A first order partial differential equation $f(x, y, z, p, q) = 0$ is said to be linear if it can be written in the form as

$P(x, y)p + Q(x, y)q = R(x, y)z + S(x, y)$, where P, Q and R are functions of x and y only and therefore, the partial differential equation is linear in z, and where $p = \frac{\partial z}{\partial x}$ and $q = \frac{\partial z}{\partial y}$. For example,

$$\frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} = xy.$$

1.2.6 Semilinear Partial Differential Equations

A first order partial differential equation $f(x, y, z, p, q) = 0$ is said to be semi-linear partial differential equation if it is linear in p and q. A semilinear first order partial differential equations can be written as

$P(x, y)p + Q(x, y)q = R(x, y, z)$, where P and Q are functions of x and y only and R is non-linear in z. For example,

$$\frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} = xyz^2.$$

1.2.7 Quasilinear Partial Differential Equations

A first order partial differential equation $f(x, y, z, p, q) = 0$ is said to be quasi-linear partial differential equation if it is linear in p and q and it can be written in the form

$P(x, y, z)p + Q(x, y, z)q = R(x, y, z)$, where at least one of P or Q is non-linear in z. For example,

$$\frac{\partial z}{\partial x} + z^3\frac{\partial z}{\partial y} = xyz^2.$$

1.2.8 Nonlinear Partial Differential Equations

A first order partial differential equation $f(x, y, z, p, q) = 0$ which does not come under the above three types is known as non-linear or fully non-linear partial differential equation. For example,

$$p^2 + q^2 = 1.$$

1.3 Solution Methodology

As stated earlier differential equations form the basis of many mathematical models of physical, chemical and biological phenomenon and more recently their use has spread into economics, financial forecasting, image processing and many other fields. Also these differential equations or models are nonlinear, complex and transcendental in nature. Most of the differential equations governing these model problems can not be solved analytically. Therefore, in order to investigate the behavior of these model problems, we need to depend on the numerical methods to find the solution. There are many numerical techniques available in the literature for example, finite difference methods, finite element methods, finite volume methods, collocation methods etc. In the present study, we will use finite element technique to solve the differential equations.

1.4 Brief History of Finite Element Methods

Though there are many numerical methods to solve differential equations, among these techniques, a particular class of numerical techniques namely "Finite Element Methods" is widely used for for approximating exact solution of differential equations. Finite element method represents a powerful and general class of techniques for approximating the exact solution of the differential equations. Finite element method was originally suggested by the famous mathematician Courant in 1943. Later on, finite element method(FEM) was developed in 1956 for the analysis of aircraft structural problems. Thereafter, within a decade, the potentialities of the method for finding the solution of problems in stress analysis, fluid flow, heat transfer, and in other areas were recognized. Over the years, the finite element technique has been so well established that today it is considered to be one of the best method for solving a wide variety of practical problems efficiently. In fact, the method has become one of the active research area for applied mathematicians. The systematic procedure of the FEM makes it possible to construct a general purpose software to solve a wide range of problems. Based on the variational principle, the basic procedures of FEM include:
domain discretization, the weak formulation of the problem, formation of element equations, assembly of element equations and then finally solving the global system of equations. Most commercial FEM software packages originated in the 1970's like Abaqus, Adina, Ansys, etc.

1.5 Finite Element Methods

The basic steps of finite element methods in computing an approximate solution of any problem consists :

1. Finite element discretization

2. Element equations
3. Assembly of element equations and solution
4. Convergence and error estimate

The basic ideas underlying the finite element method are introduced with the help of one example [12]:

1.5.1 Example 1

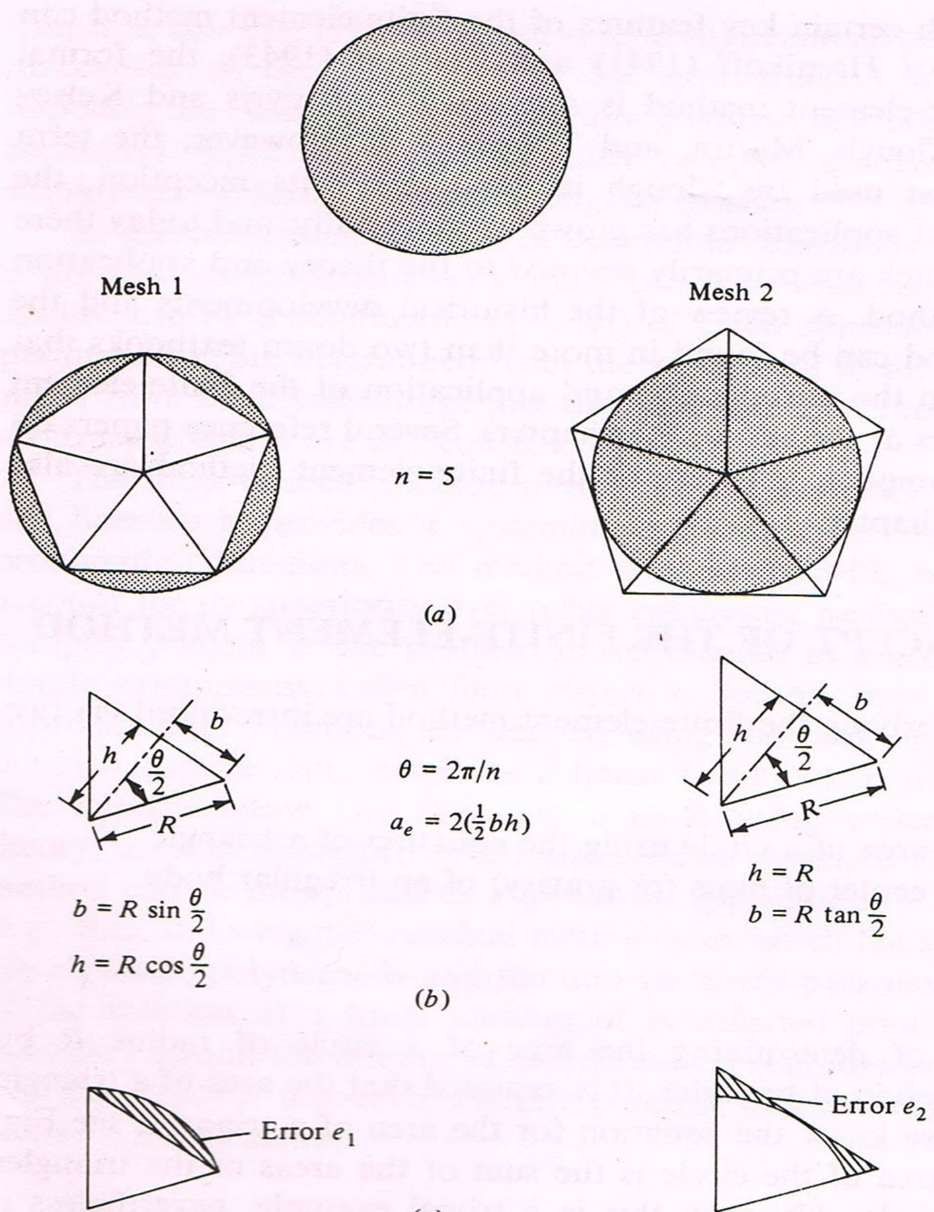
Consider the problem of determining the area of a circle of radius R by discretizing it as a collection of triangles. It is assumed that the area of the triangle can be calculated. The approximate area of the circle is the sum of the areas of the triangles used to represent the circle. Here we discuss the basic steps involved while finding finite element solution of any problem.

1. **Finite element discretization**: First, the circle is represented as a collection of n subregions, say triangles. This is called discretization of the domain. Each triangle is called an element. The collection of elements is called the finite element mesh. In our case, we discretize the circle into a mesh of five triangles. Since all the elements are of same size, the mesh is said to be uniform. We consider two discretizations of the circle say mesh 1 and mesh 2.

2. **Element equations**: A typical element i.e triangle from the discretized mess is selected and its properties (i.e. area in the present case) or equations are computed. It is here that we bring in the governing equation (i.e., the equation for computing the area) of the element to calculate the required property. Let a_e be the area of element e in mesh 1 and $\overline{a}_e$ be the area of element e in mesh 2. Thus, for the for element e, we have

$$a_e = \frac{R^2}{2} \sin\left(\frac{2\pi}{n}\right)$$
$$\overline{a}_e = R^2 \tan\left(\frac{\pi}{n}\right)$$

   where R is the radius of the circle. The above equations are called element equations.


Mesh 1
Mesh 2
$n = 5$
(a)
h
b
$\frac{\theta}{2}$
R
$b = R \sin \frac{\theta}{2}$
$h = R \cos \frac{\theta}{2}$
$\theta = 2\pi/n$
$a_e = 2(\frac{1}{2}bh)$
$h = R$
$b = R \tan \frac{\theta}{2}$
(b)
Error e_1
Error e_2
(c)

3. **Assembly of element equations and solutions**: The approximate area of the circle is obtained by putting together the element property of all the elements in the mesh, this is called assembly of the element equations. In the present case, the assembly means the sum of the areas of the triangles.

$$A_1 = \sum_{e=1}^{n} a_e$$
$$A_2 = \sum_{e=1}^{n} \overline{a}_e$$

   Since the mesh is uniform, a_e *or* $\overline{a}_e$ is the same for each of the elements in the mesh, and we have

$$A_1^{(n)} = \left\{ n\frac{R^2}{2} \ \sin\left(\frac{2\pi}{n}\right)\right\}$$
$$A_2^{(n)} = nR^2 \ \tan\left(\frac{\pi}{n}\right)$$

4. **Convergence and error estimate**: For the present problem the exact solution is the area of the circle i.e.: $A_0 = \pi R^2$. We can estimate the error in the finite element solution and show that the approximate solution converges to the exact solution as $n \to \infty$. Consider the typical element e. The error in the approximation is equal to the difference between the area of the sector and that of the triangle (see fig 1).

$$e_1 = |S_e - a_e|$$
$$e_2 = |S_e - \overline{a}_e|$$

   where $S_e = \frac{1}{2}R^2\theta$ is the area of the sector. Thus, the error estimates for an element in the meshes 1 and 2 are given by

$$e_1 = R^2\left(\frac{\pi}{n} - \frac{1}{2}\sin\left(\frac{2\pi}{n}\right)\right)$$
$$e_2 = R^2\left(\tan\left(\frac{\pi}{n}\right) - \frac{\pi}{n}\right)$$

The total error (called global error) is given by multiplying the e_i $'s$ by n:

$$\begin{aligned}
E_1^{(n)} &= R^2\big(\pi - \frac{n}{2}\sin\frac{2\pi}{n}\big) \\
&= \pi R^2 - A_1^{(n)} \\
E_2^{(n)} &= R^2\left(n\tan(\frac{\pi}{n} - \pi\right) \\
&= A_2^{(n)} - \pi R^2
\end{aligned}$$

We now show that E_1 and E_2 go to zero as $n \to \infty$.

Letting $x = 2/n$, we have

$$\begin{aligned}
A_1^{(n)} &= \frac{R^2 n}{2}\sin\left(\frac{2\pi}{n}\right) \\
&= \frac{R^2\pi}{\pi x}\sin\left(\pi x\right) \\
\lim_{n\to\infty} A_1^{(n)} &= \lim_{x\to 0} R^2\pi\left(\sin\left(\pi x\right)/\pi x\right) \\
&= \pi R^2
\end{aligned}$$

Similarly, letting $y = 1/n$, we have

$$\begin{aligned}
\lim_{n\to\infty} A_2^{(n)} &= \lim_{y\to 0} R^2\frac{\tan(\pi y)}{y} \\
&= \lim_{y\to 0} \pi R^2\sec^2(\pi y) \\
&= \pi R^2
\end{aligned}$$

Hence, $E_1^{(n)}$ and $E_2^{(n)}$ goes to 0 as $n \to \infty$. This completes the proof the convergence.

Also, one can use either rectangles or any other polygons as elements in order to discretize the area of the circle. The approximation error in each case is different and hence the finite element solution is also different in each case.

1.5.2 Example 2

Next, we will solve a second order linear differential equation using standard Galerkin finite element method [12].

Consider a second order linear differential equation

$$\frac{d^2u}{dx^2} - u + x = 0 \qquad\qquad 0 \leq x \leq 1.0$$

with boundary conditions; at $x = 0, u = 1$ and

at $x = 1, \frac{du}{dx} = 1$.

We take the nodes as $x = 0, 0.2, 0.5, 0.8, 1.0$.

Therefore, we have the following elements:

$$[0, 0.2], \quad [0.2, 0.5], \quad [0.5, 0.8], \quad [0.8, 1.0]$$

.

Now, at each node x_i, we define the basis functions $\phi_i(x)$ as

$$\phi_i(x) = \begin{cases} \dfrac{x - x_{i-1}}{x_i - x_{i-1}}, & x \in [x_{i-1}, x_i] \\ \dfrac{x_{i+1} - x}{x_{i+1} - x_i}, & x \in [x_i, x_{i+1}] \end{cases}$$

Let the finite element approximate solution be

$$\bar{u} = \sum_{j=1}^{5} \Phi_j(x) u_j. \tag{1.5.1}$$

Now, the weak formulation is given by

$$\int_0^1 \left(\frac{d^2\bar{u}}{dx^2} - \bar{u} + x \right) \Phi_i(x) dx = 0, \quad i = 1(1)5.$$

Integrating by parts and using $Q = -du/dx$, we get

$$\int_0^1 \frac{d\Phi_i}{dx} \frac{d\bar{u}}{dx} dx + \int_0^1 \Phi_i(x) \bar{u} dx = \int_0^1 x \Phi_i(x) dx - \Phi_i(x_5) Q_5 + \Phi_i(x_1) Q_1. \tag{1.5.2}$$

Substituting the value of $\bar{u}$ from (1.5.1), equation (1.5.2) becomes

$$\sum_{j=1}^{5} u_j \int_0^1 \frac{d\Phi_i}{dx} \frac{d\Phi_j}{dx} dx + \sum_{j=1}^{5} u_j \int_0^1 \Phi_i(x) \Phi_j(x) dx$$
$$= \int_0^1 x \Phi_i(x) dx - [\Phi_i(x_5) Q_5 - \Phi_i(x_1) Q_1], i = 1(1)5.$$

Rewriting the above integral element wise, we get

$$\sum_{r=1}^{4} \sum_{j=1}^{5} u_j \int_{e_r} \frac{d\Phi_i}{dx} \frac{d\Phi_j}{dx} dx + \sum_{r=1}^{4} \sum_{j=1}^{5} u_j \int_{e_r} \Phi_i(x) \Phi_j(x) dx = R - T, \quad i = 1(1)5 \tag{1.5.3}$$

where

$$R = \sum_{r=1}^{4} \int_{e_r} x\Phi_i(x)dx \tag{1.5.4}$$

and

$$T = [\Phi_i(x_5)Q_5 - \Phi_i(x_1)Q_1]. \tag{1.5.5}$$

Writing the first term on the left hand side of equation (1.5.3) in matrix form, we get

$$AU = \sum_{r=1}^{4} \sum_{j=1}^{5} u_j \int_{e_r} \frac{d\Phi_i}{dx}\frac{d\Phi_j}{dx}dx, \quad i = 1(1)5$$

Evaluating the integral over an element e having nodes 1 and 2, we get

$$A^e = \begin{bmatrix} 1/L & -1/L \\ -1/L & 1/L \end{bmatrix} \tag{1.5.6}$$

where L is defined as distance between the nodes.

Consider the second term on the left hand side of equation (1.5.3), and denote it in matrix form as

$$BU = \sum_{r=1}^{4} \sum_{j=1}^{5} u_j \int_{e_r} \Phi_i(x)\Phi_j(x)dx, \quad i = 1(1)5$$

Again, evaluating the integral over element e, we get

$$B^e = \begin{bmatrix} L/3 & L/6 \\ L/6 & L/3 \end{bmatrix} \tag{1.5.7}$$

Adding (1.5.6) and (1.5.7), we get

$$P^e = \begin{bmatrix} 3 + L^2/3L & 6 - L^2/6L \\ -(6 - L^2)/6L & 3 + L^2/3L \end{bmatrix}$$

Substituting the values $L_1 = 0.2, L_2 = 0.3, L_3 = 0.3, L_4 = 0.2$, we get

$$P_1 = \begin{bmatrix} 15.2/3 & -14.9/3 \\ -14.5/3 & 15.2/3 \end{bmatrix}$$

$$P_2 = \begin{bmatrix} 15.2/3 & -14.9/3 \\ -14.5/3 & 15.2/3 \end{bmatrix}$$

Therefore, we have $P_3 = P_2$ and $P_4 = P_1$ since $L_3 = L_2$ and $L_4 = L_1$.

These values of P_1, P_2, P_3, P_4 will be inserted in appropriate rows and columns of the global matrix.

The $i-$th term of the column vector R of equation (1.5.4) is given by

$$R_i = \int_0^1 x\phi_i(x)dx, \quad i = 1(1)5.$$

$$R_i^e = \sum_{r=1}^{4} \int_{e_r} x\Phi_i(x)dx, \quad i = 1(1)5.$$

Evaluating the above integral over the element e having nodes 1 and 2.

$$R_1^e = \frac{1}{2}x_1 L + \frac{L^2}{6}$$

$$R_2^e = \frac{1}{2}x_1 L - \frac{L^2}{6}$$

Similarly, we can get the expression for R over other elements.

$$R = \begin{bmatrix} \frac{1}{2}x_1L_1 + \frac{L_1^2}{6} \\ \frac{1}{2}x_2L_1 + \frac{L_1^2}{6} + \frac{1}{2}x_2L_2 - \frac{L_2^2}{6} \\ \frac{1}{2}x_3L_2 + \frac{L_2^2}{6} + \frac{1}{2}x_3L_3 - \frac{L_3^2}{6} \\ \frac{1}{2}x_4L_3 + \frac{L_3^2}{6} + \frac{1}{2}x_4L_4 - \frac{L_4^2}{6} \\ \frac{1}{2}x_5L_4 + \frac{L_4^2}{6} \end{bmatrix}$$

$$R = \begin{bmatrix} 0.00667 \\ 0.05833 \\ 0.15 \\ 0.19167 \\ 0.09333 \end{bmatrix}$$

Given that, $Q_5 = -1$,

Using nodal values for $x_1 = 0, x_2 = 0.2, x_3 = 0.5, x_4 = 0.8, x_5 = 1.0$, in

$T_i = [-\Phi_i(x_5)Q_5 + \Phi_i(x_1)Q_1]$, we get

$$T = \begin{bmatrix} Q_1 \\ 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

Using the values of $P_1, P_2, P_3, P_4, P_5, R$ and T, the system (1.5.3) becomes

$$\begin{aligned} 5.0667u_1 - 4.9667u_2 &= 0.00667 - Q_1 \\ -4.9667u_1 + 8.5u_2 - 3.2833u_3 &= 0.05833 \\ -3.2833u_2 + 6.8667u_3 - 3.28337u_4 &= 0.15 \\ -3.2833u_3 + 8.5u_4 - 4.9669u_5 &= 0.191667 \\ -4.9667u_4 + 5.0667u_5 &= 1.0933. \end{aligned}$$

Since Q_1 is not known, therefore we neglect first equation and consider the above remaining equations. Since $u_1 = 1$ is given, solving these linear equations, we get

$$\begin{aligned} u_2 &= 1.0667 \quad & u_3 &= 1.2307 \\ u_4 &= 1.4610 \quad & u_5 &= 1.6480 \end{aligned}$$

1.6 Basic results for Standard Finite Element Methods

Let V be a given Hilbert Space with norm $||.||$ and inner product $(.,.)$. Now, we present below some basic results for Ritz-Galerkin Method.

Consider a second order differential equation:

$$-\frac{d^2u}{dx^2} + b(x)\frac{du}{dx} + c(x)u = f(x),$$

with boundary conditions defined by $u(0) = u(1) = 0$.
We define the corresponding variational problem as follows:
Find $u \in V$ such that

$$a(u, v) = f(v), \qquad v \in V \tag{1.6.1}$$

where $a(.,.)$ is a bilinear form on $V \times V$ and $f(.)$ is a given continuous function on V defined as

$$a(u,v) = \int_0^1 (v^{'}w^{'} + (bv^{'} + cv)w)dx,$$

$$f(v) = \int_0^1 fvdx.$$

For the given problem, the solution will belong to the $V = H_0^1(0,1)$-space.

Now the Lax-Milgram Lemma gives the sufficient conditions for the existence and uniqueness of solutions of the variational problem (1.6.1). The Lax-Milgram Lemma stats that

Lemma 1.6.1 Lax-Milgram Lemma:

Let V be a Hilbert Space and $a(.,.)$ be a continuous and $V-$elliptic (or coercive) bilinear form. Then for any continuous linear functional $f \in V^{'}$, there exists a unique $u \in V$ such that $a(u,v) = (f,v), \quad \forall \ \ v \in V$.

In general, the solution space V is infinite-dimensional. But it will not be possible to find the solution in the infinite dimensional space. Therefore, in order to approximate the solution, we need to

1. approximate V by means of finite-dimensional space V_h, and
2. pose the variational problem in V_h.

Then solving this problem in V_h is equivalent to solving a finite-dimensional system of linear equations.

Assume that the method is confirming i.e. $V_h \subset V$.

Then the discrete problem corresponding to (1.6.1) is given by

Find $u_h \in V_h$ such that

$$a(u_h, v_h) = f(v_h) \qquad \forall v_h \in V_h. \tag{1.6.2}$$

This method is known as *Ritz-Galerkin* method.

Now, the Lax-Milgram Lemma implies that

1. the discrete problem has a unique solution

2. the discrete problem is stable.

Let dimension of the finite dimensional space V_h be N and let $\{w_i : i = 1, 2, ..., N\}$ be a basis for V_h. Therefore,

$$u_h = \sum_{i=1}^{N} u_i w_i \tag{1.6.3}$$

where the unknowns u_i are to be determined. Using this approximation (1.6.3) in (1.6.2), we get

$$\begin{aligned} & a\left(\sum_{i=1}^{N} u_i w_i, v_h\right) = f(v_h) \\ \Rightarrow \qquad & \sum_{i=1}^{N} u_i a(w_i, v_h) = f(v_h) \end{aligned} \tag{1.6.4}$$

Now, taking $v_h = w_1, w_2, ..., w_N$, we get a system of N equations, which can be written as

$$Au = b,$$

where $A_{i,j} = a(w_i, w_j), \quad U_i = u_i$ and $b_i = f(w_i)$.

Solving this system for $u_i's$, we get the approximate solution u_h.

Now, the Cea's lemma provides error estimates in the discretized solution u_h.

Lemma 1.6.2 **Cea's Lemma:**

Assume that the hypothesis of the Lax-Milgram lemma are satisfied. Let u and u_h denote the solutions of the continuous problem (1.6.1) *and the discrete problem* (1.6.2) *respectively. Then*

$$||u - u_h|| \leq \frac{\beta}{\alpha} \inf_{v_h \in V_h} ||u - v_h|| \tag{1.6.5}$$

Proof: Since u and u_h are solutions of (1.6.1) and (1.6.2) respectively, i.e.

$$a(u, v_h) = f(v_h)$$

$$\textit{and} \quad a(u_h, v_h) = f(v_h), \qquad \forall v_h \in V_h$$

Subtracting these equations, we get

$$a(u - u_h, v_h) = 0, \qquad \forall v_h \in V. \tag{1.6.6}$$

Now, using coercivity and continuity of $a(.,.)$,, we get

$$\begin{aligned}\alpha||e||^2 &\leq a(u-u_h, u-u_h) \qquad (using \;\; coercivity)\\ &= a(u-u_h, u) - a(u-u_h, u_h)\\ &= a(u-u_h, u) - a(u-u_h, v_h) \qquad (from \;\; (1.6.6))\\ &= a(u-u_h, u-v_h)\\ &= a(e, u-v_h)\\ &\leq \beta||e||||u-v_h|| \qquad (using \;\; continuity)\end{aligned}$$

Thus,

$$||u-u_h|| \leq \frac{\beta}{\alpha}||u-v_h|| \qquad \forall v_h \in V_h. \tag{1.6.7}$$

This proves the result.

1.6.1 Energy Error Estimates

Suppose $a(.,.)$ is coercive, symmetric and bilinear form, then define the energy product and the related energy norm as

$$(v,w)_E = a(v,w) \qquad and \qquad ||v||_E^2 = (v,v)_E.$$

Continuing in the same way as above in the Cea's lemma, we get

$$||u-u_h||_E = \inf_{v_h \in V_h} ||u-v_h||_E.$$

Chapter 2

ELEMENTARY A PRIORI AND A POSTERIORI FINITE ELEMENT ERROR ESTIMATES IN ENERGY NORM

In finite element error estimates, we determine the bound on error $\|u - u_h\|$ where u is exact solution and u_h is finite element solution of any given differential equation. In general, finite element error estimates can be classified into two categories:

1) A Priori Error Estimates

2) A Posteriori Error Estimates

In a priori error estimates, the estimates are obtained before the solution is known. These type of estimates are useful for the design of finite element methods and are valuable for determining the asymptotic rate of convergence of the method as well as for finding dependencies on the problem parameters. A priori estimates are not computable because they are given in term of exact solution which is generally not known. But these estimates require knowledge about the derivatives of the (unknown) exact solution expected.

The second kind of estimates are a posteriori estimates. In a posteriori error estimates, bounds are obtained on the error in the finite element approximation in terms of the residual error. These error bounds can be evaluated once the finite element solution is computed and are used to estimate the error. Through the a posteriori error estimates, it is possible to estimate and adaptively control the finite element error to a desired tolerance level by suitably refining the mesh i.e. the error in the finite element approximation can be made arbitrarily small by refining the mesh provided that the solution u is sufficiently smooth to allow the interpolation error bounds not to depend on the approximate solution to be computed. These type of estimates are computable since they are expressed in the term of the known finite element solution.

In the present study, the a priori estimates have been derived in the energy norm for both the symmetric and non-symmetric operators. The a posteriori estimates have also been derived in the energy norm. Before proceeding to a priori estimates, first we define some norms.

2.1 Some Important Norms

$$\begin{aligned} L_p - Norm \qquad \|v\|_p &= \Big(\int_0^L |v(x)|^p \Big)^{1/p}, \quad 0 < p < 1 \\ L_2 - Norm \qquad \|v\|_2 &= \Big(\int_0^L |v(x)|^2 \Big)^{1/2}, \quad 0 < p < 1 \\ Energy - Norm \qquad \|v\|_E &= \Big(\int_0^L a(x)|v(x)|^2 \Big)^{1/2} \end{aligned}$$

2.2 A Priori Error Estimates in Energy Norm

In this Section, the a priori bounds on the error in computed finite element solution has been derived in the energy norm. The objective has been achieved by categorizing the operator as symmetric and nonsymmetric operators.

2.2.1 Symmetric Operator Case

Consider a differential equation:

$$-\epsilon \frac{d^2u}{dx^2} + u = 0 \qquad on \quad \bar{\Omega} = [0, 1]; \tag{2.2.1}$$

with boundary conditions $u(0) = u_0$ and $u(1) = u_1$.

In order to derive the a priori error estimates, we first find the finite element solution. For this first discretize the domain $[0,1]$ into N elements.

Let $0 = x_0 < x_1 < x_2 < ... < x_N = 1$ be any discretization of the domain.

Now, at each node x_j, we define a basis function $\phi_j(x)$ as

$$\phi_j(x) = \begin{cases} \dfrac{x - x_{j-1}}{x_j - x_{j-1}}, & x \in [x_{j-1}, x_j] \\ \dfrac{x_{j+1} - x}{x_{j+1} - x_j}, & x \in [x_j, x_{j+1}] \end{cases}$$

In general, the solution space, say, S is infinite dimensional. But since it is not possible to find the solution in the infinite dimensional space, we approximate the solution space in the finite dimension solution space S^h. In S^h, let the finite element solution is given by

$$\bar{u} = \sum_{j=1}^{n} \phi_j(x) u_j, \tag{2.2.2}$$

where $\bar{u} \in S^h$.

Multiply equation (2.2.1) by weight function say $w(x)$ and then integrating over the domain $[0,1]$, we get

$$\int_0^1 -\epsilon \frac{d^2u(x)}{dx^2} w(x)dx + \int_0^1 u(x)w(x)dx = 0 \tag{2.2.3}$$

Integrating by parts, the first term in (2.2.3) becomes

$$\int_0^1 -\epsilon \frac{d^2u(x)}{dx^2} w(x)dx = -\epsilon \Big[w(x) \frac{du(x)}{dx} \Big|_0^1 - \int_0^1 \frac{dw(x)}{dx} \frac{du(x)}{dx} dx \Big]$$

Substituting this value in equation (2.2.3), we get

$$\begin{aligned} & -\epsilon w(x) \frac{du(x)}{dx} \Big|_0^1 + \epsilon \int_0^1 \frac{dw(x)}{dx} \frac{du(x)}{dx} dx + \int_0^1 u(x)w(x)dx = 0 \\ & \epsilon \int_0^1 \frac{dw(x)}{dx} \frac{du(x)}{dx} dx + \int_0^1 u(x)w(x)dx = \epsilon w(x) \frac{du(x=1)}{dx} - \epsilon w(x) \frac{du(x=0)}{dx} \end{aligned}$$

Putting

$$Q = \frac{du}{dx},$$

we get,

$$\epsilon \int_0^1 \frac{dw(x)}{dx} \frac{du(x)}{dx} dx + \int_0^1 u(x)w(x)dx = \epsilon w(x_N) Q_N - \epsilon w(x_1) Q_1 \tag{2.2.4}$$

Writing the above weak formulation in the finite dimensional space S^h, we get

$$\epsilon\int_0^1 \frac{dw(x)}{dx}\frac{d\bar{u}(x)}{dx}dx + \int_0^1 \bar{u}(x)w(x)dx = \epsilon w(x_N)Q_N - \epsilon w(x_1)Q_1 \tag{2.2.5}$$

Substitute equation (2.2.2) in equation (2.2.5), we get

$$\epsilon\int_0^1 \frac{dw(x)}{dx}\frac{d}{dx}\Big(\sum_{j=1}^{N} u_j\phi_j\Big)dx + \int_0^1 w(x)\Big(\sum_{j=1}^{N} u_j\phi_j\Big)dx = \epsilon w(x_N)Q_N - \epsilon w(x_1)Q_1$$

$$\sum_{e=1}^{N}\Big(\sum_{j=1}^{N} u_j\epsilon\int_{x_{e-1}}^{x_e} \frac{dw(x)}{dx}\frac{d\phi_j}{dx}dx + \sum_{j=1}^{N} u_j\int_{x_{e-1}}^{x_e} w(x)\phi_j dx\Big) = \epsilon w(x_N)Q_N - \epsilon w(x_1)Q_1,$$

$$\Rightarrow \sum_{e=1}^{N}\Big(\sum_{j=1}^{N} u_j B_e(w,\phi_j)\Big) = \epsilon w(x_N)Q_N - \epsilon w(x_1)Q_1. \tag{2.2.6}$$

where

$$B_e\Big(w,\phi_j\Big) = \int_{x_{e-1}}^{x_e}\Big(\epsilon\frac{dw(x)}{dx}\frac{d\phi_j}{dx} + w(x)\phi_j(x)\Big)dx.$$

Let

$$B\Big(w,\phi_j\Big) = \sum_{e=1}^{N} B_e\Big(w,\phi_j\Big)$$

$$\Rightarrow B(u,v) = \epsilon\int_0^1 \frac{du}{dx}\frac{dv}{dx}dx + \int_0^1 uv dx. \tag{2.2.7}$$

The above equation (2.2.6) is called finite element weak formulation. By calculating the equation (2.2.6) for different test functions $w = \phi_i; i = 1, 2, ..., N$, we get a linear system of equations which can be written in the form

$$AU = B$$

Simplifying this linear algebraic system, we get the finite element solution for problem (2.2.1)

Now

$$B(u,v) = \epsilon\int_0^1 \frac{du}{dx}\frac{dv}{dx}dx + \int_0^1 uv dx.$$

$$\Rightarrow B(u,v) = \epsilon(u_x, v_x) + (u,v) \tag{2.2.8}$$

where

$$
\begin{aligned}
\epsilon(u_x, v_x) &= \epsilon \int_0^1 \frac{du}{dx}\frac{dv}{dx} dx \\
(u, v) &= \int_0^1 uv dx
\end{aligned}
$$

and $B(.,.)$ is a bilinear form on $V \times S$.

S and V are the trial solution space and weighting function space respectively. The solution space S consists of $H^1(\Omega)$ functions which satisfy the given boundary conditions and the test space V consists of test functions which belong to the space $H^1(\Omega)$. It can easily be seen that the bilinear form $B(u, v)$ defined by (2.2.8) is symmetric. Now the energy norm is defined by

$$\|u\|_E^2 = B(u, u) \tag{2.2.9}$$

The Cauchy Schwarz inequality holds true with respect to this norm and is given by

$$|B(u, v)| \leq \|u\|_E \|v\|_E$$

Remark: The Cauchy Schwarz inequality is valid in this form since the bilinear form $B(u, v)$ is symmetric.

Introduce the following splitting of the error which will be useful for the proof. Let $\tilde{u}_h$ be nodal interpolation of the exact solution u. Letting e denote the finite element solution of error, we have

$$
\begin{aligned}
e &= u_h - u \\
&= u_h - \tilde{u}_h + \tilde{u}_h - u \\
&= e_h + \eta
\end{aligned} \tag{2.2.10}
$$

where $e_h = u_h - \tilde{u}_h$ is the portion of the error in V^h, and

$$\eta = \tilde{u}_h - u = e - e_h$$

is the interpolation error and is a member of V.

Now the weak formulation of the given problem can be written as

$$B(u, v) = 0. \tag{2.2.11}$$

Since $v_h \in V$, so (2.2.8) will also hold for all $v_h \in V_h$, a finite dimensional subspace of V,

$$\Rightarrow B(u, v_h) = 0 \tag{2.2.12}$$

Also, from the finite element weak formulation, we have

$$B(u_h, v_h) = 0. \tag{2.2.13}$$

Subtracting (2.2.12) from (2.2.13), we get

$$\begin{aligned} B(u, v_h) - B(u_h, v_h) &= 0 \\ \Rightarrow B(u - u_h, v_h) &= 0 \end{aligned}$$

Thus, we get Galerkin orthogonality property

$$B(e, e_h) = 0,$$

where $e = u - u_h$.

Using the energy norm $\|.\|_E$ given by (2.2.9), for $e \in V$, we get

$$\begin{aligned} \|e\|_E^2 &= |B(e, e)| \\ &= |B(e, e_h + \eta)| \quad (using \;\; (2.2.10)) \\ &= |B(e, e_h) + B(e, \eta)| \end{aligned}$$

Using Galerkin orthogonality property i.e. $B(e, e_h) = 0$,

$$\begin{aligned} \|e\|_E^2 &= |B(e, \eta)| \\ &\leq \|e\|_E \|\eta\|_E \qquad (Using \;\; Cauchy - Schwarz \; inequality) \end{aligned} \tag{2.2.14}$$

Therefore,

$$\|e\|_E \leq \|\eta\|_E \tag{2.2.15}$$

which states the best approximation property with respect to the energy norm.

Again,

$$\begin{aligned} \|e\|_E &\leq (|B(\eta, \eta)|)^{1/2} \qquad\qquad (\because \|\eta\|^2 = B(\eta, \eta)) \\ &= (|\epsilon(\eta_x, \eta_x) + (\eta, \eta)|)^{1/2} \qquad (by \;\; equation \;\; ((2.2.8)) \\ &= (\epsilon\|\eta_x\|^2 + \|\eta\|^2)^{1/2} \\ &= (\frac{\epsilon}{L^2}|\eta|_1^2 + \|\eta\|^2)^{1/2} \qquad (\text{where } \|\eta_x\| = |\eta|_1) \\ \Rightarrow \|e\|_E &\leq \bar{C}\|\eta\|_1 \end{aligned} \tag{2.2.16}$$

Expanding $H^1(\Omega)$ norm for η, we get

$$\begin{aligned}\|\eta\|_1^2 &= \int_0^1 \eta^2 dx + \int_0^1 \Big(\frac{d\eta}{dx}\Big)^2 dx\\ &= \|\eta\|^2 + \|\eta_x\|^2\\ &= |\eta|_0^2 + |\eta|_1^2,\end{aligned}$$

where $\|.\|$ denotes the $L_2(\Omega)$ norm, and $\|.\|_1$ and $|.|_1$ denote the $H^1(\Omega)$ norm and $H^1(\Omega)$ semi-norm respectively.

We can now apply the standard interpolation estimates [7, 14] of the form

$$\|\eta\|_1 \leq C_i(p,\Omega)h^p\|\phi\|_{p+1}, \tag{2.2.17}$$

where it is assumed that ϕ has regularity $r \geq p+1$. Substituting (2.2.17) into (2.2.16), we get

$$\|e\|_E \leq Ch^p\|\phi\|_{p+1}, \tag{2.2.18}$$

where $C = \bar{C}C_i(p,\Omega)$.

2.2.2 Non-symmetric Operator Case

Consider the differential equation

$$-\epsilon u^{''}(x) - bu^{'}(x) + cu(x) = f(x) \quad in \quad \Omega = (0,1) \tag{2.2.19}$$

with the boundary conditions

$$u(0) = u(1) = 0.$$

Before proceeding to estimate the a priori error, we first briefly explain the finite element formulation in order to get the bilinear form for problem (2.2.19). We consider the domain discretization as in the symmetric operator case. Similarly, at each x_j, we define a basis function $\phi_j(x)$ as

$$\phi_j(x) = \begin{cases} \dfrac{x - x_{j-1}}{x_j - x_{j-1}}, & x \in [x_{j-1}, x_j] \\ \dfrac{x_{j+1} - x}{x_{j+1} - x_j}, & x \in [x_j, x_{j+1}] \end{cases}$$

Then the finite element solution is given by:

$$\bar{u} = \sum_{j=1}^{N} \phi_j(x)u_j \tag{2.2.20}$$

Multiplying equation (2.2.19) by weight function say $w(x)$ and then integrating over the domain Ω, we get

$$\int_0^1 \Big(-\epsilon u^{''}(x) - bu^{'}(x) + cu(x)\Big) w(x)dx = \int_0^1 f(x)w(x)dx.$$

Integrating by parts, the first term in of above equation becomes

$$-\epsilon \int_0^1 u^{''}(x)w(x)dx = -\epsilon w(x)u^{'}(x)\Big|_0^1 + \epsilon \int_0^1 u^{'}(x)w^{'}dx$$

Using this result, the above equation becomes

$$-\epsilon w(x)u^{'}(x)\Big|_0^1 + \epsilon \int_0^1 u^{'}(x)w^{'}(x)dx - b\int_0^1 u^{'}(x)w(x) + c\int_0^1 u(x)w(x)dx = \int_0^1 f(x)w(x)dx$$

$$\epsilon \int_0^1 u^{'}(x)w^{'}(x)dx - b\int_0^1 u^{'}(x)w(x)dx + c\int_0^1 u(x)w(x)dx = \int_0^1 f(x)w(x)dx + \epsilon w(x)u^{'}(x)\Big|_0^1$$

Put $Q = u^{'}$

$$\begin{aligned}\epsilon \int_0^1 u^{'}(x)w^{'}(x)dx - b\int_0^1 u^{'}(x)w(x)dx + c\int_0^1 u(x))w(x)dx &= \int_0^1 f(x)w(x)dx \\ &+ \epsilon(w(x_N)Q_N - w(x_1)Q_1)\end{aligned}$$

Now the finite element weak formulation is obtained by using the approximation $\bar{u}$ of u in the above equation. Substituting the value of $\bar{u}$ in the finite element weak formulation, we get

$$\begin{aligned}\sum_{e=1}^{N}\Big(\sum_{j=1}^{N} u_j\epsilon\int_{x_{e-1}}^{x_e} w^{'}(x)\phi_j^{'}(x)dx - \sum_{j=1}^{N} u_j b\int_{x_{e-1}}^{x_e} w(x)\phi_j^{'}(x)dx + \sum_{j=1}^{N} u_j c\int_{x_{e-1}}^{x_e} w(x)\phi_j(x)dx\Big) \\ = \int_0^1 f(x)w(x)dx + \epsilon(w(x_N)Q_N - w(x_1)Q_1) \\ \Rightarrow \sum_{e=1}^{N}\Big(\sum_{j=1}^{N} u_j B_e(\phi_j, w)\Big) = (f(x), w) \quad (\because\ w(x_N) = 0 = w_(x_1))\end{aligned} \tag{2.2.21}$$

where

$$B_e(\phi_j, w) = \int_{x_{e-1}}^{x_e} \Big(\epsilon w^{'}\phi_j^{'} - bw\phi_j^{'} + cw\phi_j\Big)dx$$

Take

$$B(\phi_j, w) = \sum_{e=1}^{N} B_e(\phi_j, w)$$

$$\Rightarrow \sum_{j=1}^{N} B(\phi_j, w) = L(w)$$

$$\Rightarrow B(u, w) = L(w)$$

where

$$B(u,v) = \epsilon \int_0^1 u^{'} v' dx - b \int_0^1 u^{'} v dx + c \int_0^1 uv dx \tag{2.2.22}$$

$$L(v) = \int_0^1 f(x) v dx. \tag{2.2.23}$$

$B(u,v)$ is the bilinear form and $L(v)$ is the linear functional corresponding to the problem (2.2.19).

The above equation is called weak formulation. Here, in the present case, we can easily see that the bilinear form $B(u,v)$ is not symmetric.

Calculating equation (2.2.21) element wise, we get

$$AU = B$$

Solving this algebraic system we get the finite element solution. Next, we will estimate the a priori error in the finite element formulation.

In order to derive the a priori error estimates, the non symmetric operator can be systematically decomposed into symmetric part and skew symmetric part i.e.

$$B(u,v) = B_{symm}(u,v) + B_{skew}(u,v), \tag{2.2.24}$$

where

$$B_{symm}(u,v) = 1/2(B(u,v) + B(v,u)) \tag{2.2.25}$$

$$B_{skew}(u,v) = 1/2(B(u,v) - B(v,u)) \tag{2.2.26}$$

Substituting (2.2.22) into (2.2.25) and (2.2.26), we get

$$\begin{aligned} B_{symm}(u,v) = 1/2\Big[\epsilon \int_0^1 u^{'} v^{'} dx - b \int_0^1 u^{'} v dx + c \int_0^1 uv dx + \epsilon \int_0^1 v^{'} u^{'} dx \\ - b \int_0^1 v^{'} u dx + c \int_0^1 vu dx\Big] \end{aligned} \tag{2.2.27}$$

Now using integration by parts

$$\begin{aligned} b \int_0^1 v^{'} u dx &= buv \mid_0^1 - b \int_0^1 u^{'} v dx \\ &= -b \int_0^1 u^{'} v dx. \end{aligned}$$

Substituting this value in equation (2.2.27), we get

$$\begin{aligned} B_{symm}(u,v) &= 1/2\Big[\epsilon\int_0^1 u^{'}v^{'}dx - b\int_0^1 u^{'}vdx + c\int_0^1 uvdx + \epsilon\int_0^1 v^{'}u^{'}dx \\ &\qquad + b\int_0^1 u^{'}vdx + c\int_0^1 vudx\Big] \\ &= \epsilon\int_0^1 u^{'}v^{'}dx + c\int_0^1 uvdx' \end{aligned}$$

$$B_{skew}(u,v) = 1/2\Big[\epsilon\int_0^1 u^{'}v^{'}dx - b\int_0^1 u^{'}vdx + c\int_0^1 uvdx - \epsilon\int_0^1 v^{'}u^{'}dx \tag{2.2.28}$$

$$+ b\int_0^1 v^{'}udx - c\int_0^1 vudx\Big] \tag{2.2.29}$$

Again using integration by parts, we get

$$\begin{aligned} b\int_0^1 v^{'}udx &= buv\mid_0^1 - b\int_0^1 u^{'}vdx \\ &= -b\int_0^1 u^{'}vdx. \end{aligned}$$

Substituting this value in equation (2.2.28), we have

$$\begin{aligned} B_{skew}(u,v) &= 1/2\Big[\epsilon\int_0^1 u^{'}v^{'}dx - b\int_0^1 u^{'}vdx + c\int_0^1 uvdx - \epsilon\int_0^1 v^{'}u^{'}dx \\ &\qquad - b\int_0^1 u^{'}vdx - c\int_0^1 vudx\Big] \\ &= -b\int_0^1 u^{'}vdx \end{aligned}$$

The proof of the previous Section break down here for the non-symmetric operator case since the Cauchy-Schwarz inequality does not hold for non-symmetric operators. This is because the skew symmetric operator does not contribute to the energy norm, as if $B(.,.)$ is skew-symmetric, then

$$\begin{aligned} \|u\|_E^2 &= |B(u,u)| \\ &= |B_{skew}(u,u)| \\ &= |1/2(B(u,u) - B(u,u))| \\ &= 0. \end{aligned}$$

The approach taken here is to define a dual norm to accommodate the skew symmetric part. This definition is canonical and allow the proof to proceed in a natural way, along the similar lines of the preceding derivation for symmetric operators. The 'skew-norm' is defined by

$$\|u\|_{skew} = \sup_{v \in V} \frac{|B_{skew}(u,v)|}{\|v\|_E}. \tag{2.2.30}$$

Clearly, from (2.2.30),

$$|B_{skew}(u,v)| \leq \|u\|_E \|v\|_E \tag{2.2.31}$$

This inequality is derived directly from the definition of the skew norm. It is similar in the form to the Cauchy-Schwarz inequality, however, note the appearance of the symmetric norm (i.e. the energy norm) on the right-hand side.

The proof of a priori error estimate may now be proceeded. We begin by applying the definition of $\|.\|_E$ on e, as in the previous section, to get

$$\begin{aligned}
\|e\|_E^2 &= |B(e,e)| \\
&= |B(e^h + \eta, e)| \\
&= |B(e^h, e) + B(\eta, e)| \\
&= |B(\eta, e)| \qquad (\because B(e, e^h) = 0) \\
&= |B_{symm}(\eta, e) + B_{skew}(\eta, e)| \qquad (using \quad (2.2.24)) \\
&\leq |B_{symm}(\eta, e)| + |B_{skew}(\eta, e)| \\
&\leq \|\eta\|_E \|e\|_E + \|\eta\|_{skew} \|e\|_E
\end{aligned}$$

Therefore,

$$\|e\|_E \ \leq \|\eta\|_E + \|\eta\|_{skew}. \tag{2.2.32}$$

Expending $\|\eta\|_{skew}$ and noting that $|B(\eta,v)| = |B(v,\eta)|$, we get

$$
\begin{aligned}
\|\eta\|_{skew} &= \sup_{v\in V} \frac{|B_{skew}(\eta,v)|}{\|v\|_E} \\
&= \sup_{v\in V} \frac{|B_{skew}(v,\eta)|}{\|v\|_E} \\
&= \sup_{v\in V} \frac{|b(v,\eta^{'})|}{(\epsilon\int_0^1 v'v'dx + c\int_0^1 vvdx)^{1/2}} \\
&\le \sup_{v\in V} \frac{b\|v\|\|\eta^{'}\|}{\epsilon\|v^{'}\|^2 + c\|v\|^2} \qquad (using \ \ Cauchy-Schwarz \ \ inequality) \\
&\le \sup_{v\in V} \frac{b\|v\|\|\eta^{'}\|}{(\epsilon C_{PF}^{-2}+c)^{1/2}\|v\|} \\
&= \bar{C}_n\|\eta^{'}\| \\
&= \frac{\bar{C}_n}{L}|\eta|_1 \\
&\le C_n\|\eta\|_1
\end{aligned} \tag{2.2.33}
$$

where C_{PF} is the domain-dependent constant arising from the Poincare-Friedrich inequality.

$$\|v\| \le C_{PF}\|v^{'}\| \tag{2.2.34}$$

which hold for all functions in $V = H^1(\Omega)$. The constant in (2.2.33) is taken as

$$C_n = \frac{b}{L(\epsilon C_{PF}^{-2}+c)^{1/2}} \tag{2.2.35}$$

Applying the interpolation estimates (2.2.17) to (2.2.33),

$$\|\eta\|_{skew} \le C_n C_i(p,\Omega)h^p\|\phi\|_{p+1} \tag{2.2.36}$$

Finally, substituting (2.2.36) into (2.2.32) and using the estimates for $\|\eta\|_E$ derived in (2.2.18), we get

$$\|e\|_E \le C_N h^p\|\phi\|_{p+1} \tag{2.2.37}$$

where

$$C_N = C + C_n c_i$$

Comparing (2.2.37) to (2.2.18), it is seen that the skew symmetric part does not destroy the error estimate unless C_n is large.

2.3 A Posteriori Estimates in Energy Norm

In this Section, we consider here the same differential equation as considered in the previous Section. Consider the equation

$$-\epsilon u_x - bu_x + cu = f(x) \qquad in \;\; \Omega = (0,1) \tag{2.3.1}$$

with boundary conditions

$$u(0) = u(1) = 0.$$

Multiplying above equation by a test function $w(x)$ and then integrating over the domain Ω, we get

$$\begin{aligned} \epsilon \int_0^1 u_x w_x dx - b \int_0^1 u_x w(x) dx + c \int_0^1 uw(x) dx &= \int_0^1 f(x) w(x) dx \\ \epsilon(u_x, w_x) - b(u_x, w) + c(u, w) &= (f, w) \\ B(w, u) &= L(w) \end{aligned}$$

where $B(w,u)$ is the bilinear form and $L(w)$ is the linear functional corresponding to the problem (2.3.1).

Replacing the test function w by v, we get

$$B(v, u) = L(v) \tag{2.3.2}$$

where

$$B(v, u) = \epsilon(u_x, v_x) - b(u_x, v) + c(u, v) \tag{2.3.3}$$

and

$$L(v) = (f, v). \tag{2.3.4}$$

The methodology for deriving a posteriori error estimates is same for both the symmetric and nonsymmetric operators, since the Cauchy-Schwarz inequality in the energy norm given by

$$\|e\|_E \leq \|e\|_E \|\eta\|_E$$

is not needed.

The derivation begins the same way as for the a priori estimates.

We know that $\|w\|_E = |B(w,w)|$ regardless of the symmetric properties of $B(.,.)$.

$$\begin{aligned}\|e\|_E^2 &= |B(e,e)| \\ &= |B(e_h + \eta, e)|\end{aligned}$$

where

$$\begin{aligned} e &= e_h + \eta \\ \|e\|_E^2 &= |B(e_h, e) + B(\eta, e)| \\ \|e\|_E^2 &= |B(\eta, e)| \qquad (Using\ \ Galerkin\ \ consistancy) \end{aligned} \tag{2.3.5}$$

which is same as in the case of a priori estimates. Here, we diverge from the a priori estimates proof. For the a priori estimates, the objective is to cancel the error term on the right with one on the left, and then apply interpolation estimates to bound the remaining η term as a function of the exact solution u. For the a posteriori estimates, the goal is to form residual, and then apply interpolation estimates to bound the remaining η term as a function of the error. This error term then cancel with one on the left, leaving a bound in term of only residuals.

Now, continuing from (2.3.5), we have,

$$\begin{aligned} \|e\|_E^2 &= |B(\eta, u^h - u)| \\ &= |B(\eta, u^h) - B(\eta, u)| \\ \|e\|_E^2 &= |B(\eta, u^h) - L(\eta)| \qquad (using\ \ B(v,v) = L(v)) \end{aligned} \tag{2.3.6}$$

The effect of the third line is to remove the exact solution u from the expression, leaving only Galerkin solution u^h. Also the right hand side of equation (2.3.6) is not zero since $\eta \notin V^h$. The next step is to form residuals, which is always done through integration by parts. Furthermore, the integration by parts must be performed separately on each element K, since u^h is only C^0 continuous across element interfaces.

Let any K-th element be defined by

$$\Omega_K = (x_{K-1}, x_K) \qquad where \quad K = 1, ..., N$$

and N is the total number of elements, $x_0 = 0$ and $x_N = 1$. Thus, the elements are numbered from left to right. In addition, let

$$\begin{aligned} x_K^+ &= lim_{\epsilon \to 0}(x_K + \epsilon) \\ x_K^- &= lim_{\epsilon \to 0}(x_K - \epsilon) \end{aligned}$$

where $\epsilon > 0$. Subsituting (2.3.3) and (2.3.4) into (2.3.6) and integrating by parts, we get

$$\|e\|_E^2 = \Big|\epsilon(\eta_x, u_x^h) - b(\eta_x, u^h) + c(\eta, u^h) - (\eta, f)\Big|. \tag{2.3.7}$$

Now

$$\begin{aligned} \epsilon(\eta_x, u_x^h) &= \epsilon \int_0^1 \eta_x u_x^h dx \\ &= \epsilon \sum_{K=1}^{N} \int_{x_{K-1}^+}^{x_K^-} \eta_x u_x^h dx \\ &= \epsilon \sum_{K=1}^{N} \Big(u_x^h \eta \mid_{x_{K-1}^+}^{x_K^-} - \epsilon \int_{x_{K-1}^+}^{x_K^-} \eta u_{xx}^h dx \Big) \\ &= \sum_{K=1}^{N} \Big[\epsilon u_x^h \eta \mid_{x_{K-1}^+}^{x_K^-} - \epsilon(\eta, u_{xx}^h) L_{2(\Omega_K)} \Big] \end{aligned} \tag{2.3.8}$$

$$\begin{aligned} \Rightarrow b(\eta_x, u^h) &= b \int_0^1 \eta_x u^h dx \\ &= \sum_{K=1}^{N} b \int_{x_{K-1}^+}^{x_K^-} \eta_x u^h dx \\ &= \sum_{K=1}^{N} \Big(b u^h \eta \mid_{x_{K-1}^+}^{x_K^-} - b \int_{x_{K-1}^+}^{x_K^-} \eta u_x^h dx \Big) \\ &= \sum_{K=1}^{N} \Big[b u^h \eta \mid_{x_{K-1}^+}^{x_K^-} - b(\eta, u_x^h)_{L_{2(\Omega_K)}} \Big] \end{aligned} \tag{2.3.9}$$

Substituting (2.3.8) and (2.3.9) in (2.3.7), we get

$$\begin{aligned} \|e\|_E^2 = \Big| \sum_{K=1}^{N} [\epsilon u_x^h \eta \mid_{x_{K-1}^+}^{x_K^-} - \epsilon(\eta, u_{xx}^h) L_{2(\Omega_K)} - b u^h \eta \mid_{x_{K-1}^+}^{x_K^-} + b(\eta, u_x^h)_{L_{2(\Omega_K)}} \\ + c(\eta, u^h) - (\eta, f)_{L_{2(\Omega_K)}}] \Big| \end{aligned}$$

$$\Rightarrow \|e\|_E^2 = \Big| \sum_{K=1}^{N} [(\eta, -\epsilon u_{xx}^h + b u_x^h + c u^h - f) + \epsilon u_x^h \eta \mid_{x_{K-1}^+}^{x_K^-} - b u^h \eta \mid_{x_{K-1}^+}^{x_K^-}] \Big|$$

$$= \Big| \sum_{K=1}^{N} [(\eta, r^h)] + \sum_{K=1}^{N-1} [\eta(x_K)(\epsilon u_x^h(x_K^+) - \epsilon u_x^h(x_K^-))] \Big|$$

where r^h is the residual on element interior, given by

$$r^h = -\epsilon u_{xx}^h + b u_x^h + c u^h - f,$$

and

$$\sum_{K=1}^{N} -b u^h \eta \mid_{x_{K-1}^+}^{x_K^-} = 0$$

as

$$\begin{aligned}\sum_{K=1}^{N} -b u^h \eta|_{x_{K-1}^+}^{x_K^-} &= -b u^h \eta|_{x_0^+}^{x_1^-} ... - b u^h \eta|_{x_{K-1}^+}^{x_K^-} - b u^h \eta|_{x_K^+}^{x_{K+1}^-} - ... - b u^h \eta|_{x_{N-1}^+}^{x_N^-} \\ &= - b\eta(x_1^-) u^h(x_1^-) + b\eta(x_0^+) u^h(x_0^+) ... - b\eta(x_K^-) u^h(x_K^-) \\ &\quad + b\eta(x_{K-1}^+) u^h(x_{K-1}^+) - b\eta(x_{K+1}^-) u^h(x_{K+1}^-) + b\eta(x_K^+) u^h(x_K^+) - ... \\ &\quad - b\eta(x_N^-) u^h(x_N^-) + b\eta(x_{N-1}^+) u^h(x_{N-1}^+)\end{aligned}$$

Here, inbetween terms cancel with each other as η and u^h are continuous functions and hence

$$b\eta(x_K^-) u^h(x_K^-) = b\eta(x_K^+) u^h(x_K^+).$$

Therefore, we get

$$\sum_{K=1}^{N} -b u^h \eta \Big|_{x_{K-1}^+}^{x_K^-} = b\eta(x_0) u^h(x_0) - b\eta(x_N) u^h(x_N)$$

Again, since u^h vanishes on the boundary because $u^h \in S$ and

$$S = \{u \in H_1(\Omega) \mid u(x_0) = 0 \quad and \;\; u(x_N) = 0\},$$

thus

$$\sum_{K=1}^{N} -b u^h \eta \mid_{x_{K-1}^+}^{x_K^-} = 0.$$

But since u_h is only C^0 continuous function i.e. u^h may not be differentiable at the inter-element nodal points, hence,

$$\eta(x_K)(\epsilon u_x^h(x_K^+) - \epsilon u_x^h(x_K^-)) \neq 0$$

Therefore, at last we are left only with

$$\begin{aligned}\|e\|_E^2 &= \Big|\sum_{K=1}^{N}[(\eta, r^h)] + \sum_{K=1}^{N-1}[\eta(x_K)(\epsilon u_x^h(x_K^+) - \epsilon u_x^h(x_K^-))]\Big| \\ &\leq \Big|\sum_{K=1}^{N}[(\eta, r^h)]\Big| + \Big|\sum_{K=1}^{N-1}[\eta(x_K)(\epsilon u_x^h(x_K^+) - \epsilon u_x^h(x_K^-))]\Big|\end{aligned}$$

which gives us the required a posteriori error estimates in the energy norm.

2.4 Conclusion

In the present work, elementary estimates for both the a priori and a posteriori errors present in the finite element solution have been derived in the energy norm. For a priori error estimates, two different problems have been considered which results in symmetric and non-symmetric bilinear forms. Then, the a priori estimates have been derived separately for both these symmetric and non-symmetric bilinear forms. These estimates are generally not computable since they include the unknown exact solution. For a posteriori error estimates, a linear convection-diffusion problem has been considered. After obtaining the corresponding bilinear form, a posteriori error estimates have been derived. On contrary to the a priori estimates, these estimates are computable since they are expressed in the terms of residual error which uses only the finite element solution.

Chapter 3

A SCHWARZ ITERATIVE MONOTONE DOMAIN DECOMPOSITION ALGORITHM FOR A NONLINEAR CONVECTION-DIFFUSION PROBLEM

3.1 Introduction

The traditional numerical methods for solving singular perturbation problem require a fine mesh covering the whole domain in order to resolve local details. These methods are inefficient since the fine mesh is not needed in those part of the domain where the solution has a moderate variation.

In the present work we are interested in applying an iterative domain decomposition method for solving a nonlinear convection-diffusion problem which possesses boundary layers. The governing differential equation is given by

$$-\epsilon\frac{d^2u}{dx^2}+b\frac{du}{dx}+f(x,u)=0,\qquad \Omega=(0,1)$$

with boundary condition

$$u=g \quad \text{on} \quad \partial\Omega,$$

We assume that $b \geq \beta > 0 \quad \text{on} \quad \bar{\Omega}, \quad \text{where} \quad \bar{\Omega}=[0,1] \quad \text{and}$

$$0<c_*<f_u(x,u)\leq c^*, \qquad (x,u)\in\bar{\Omega}\times(-\infty,\infty) \quad \text{and} \quad f_u=\frac{\partial f}{\partial u} \tag{3.1.1}$$

ϵ is a small positive parameter, c^* and c_* are constants, $\partial\Omega$ denotes the boundary of Ω. If $f(x,u)$ is sufficiently smooth, then under suitable continuity and compatibility conditions on the data, a unique solution $u(x)$ of (3.1.1) exists (see [10] for details). For $\epsilon \ll 1$, problem (3.1.1) is singularly perturbed and the boundary layer occurs at $x=1$.

In the study of numerical solution of nonlinear singularly perturbed problems by the finite element method, the corresponding weak problem is formulated into finite dimensional weak problem which is further written as a system of nonlinear algebraic equations. A major point about this system is to obtain reliable and efficient computational algorithms for computing the solution.

In the present work, we are interested in applying a finite element method to the given nonlinear problem (3.1.1) by using monotone iterative (known as a method of lower and upper solutions). This method leads to iterative algorithms which converge globally and solve only linear discrete system at each iterative step.

In [2], for solving nonlinear reaction-diffusion problems, Boglaev proposed the discrete iterative algorithm which combines the monotone approach and the iterative domain decomposition method based on the Schwarz alternating procedure from [1]. In [3], for solving problem (3.1.1), Boglaev proposed the monotone domain decomposition algorithm based on the Schwarz alternating procedure from [8].

In the present work, the computational domain is partitioned into overlapping subdomains. The domain decomposition algorithm consists of the two iterative processes: outer iterations and inner iterations. One outer iteration step represents computing finite element subproblems on the subdomains in serial, starting from the left subdomain. Thus, the multiplicative

Schwarz method is the outer part of the algorithm. At the level of inner iterations, each common subdomain is discretized into overlapping nodes and the natural parallelism can be invoked, since each of box-subdomains can be treated by its own processor. For inner iterations, one can use a version of the Schwarz method. In [3], Boglaev investigated only the outer part of domain decomposition algorithm from [8].

The purpose of the present work is to apply the monotone domain decomposition algorithm in the finite element frame work. The nonlinear finite element weak formulation has been approximated by the linear finite element weak formulation by making use of the monotone Schwarz iterative algorithm. The same strategy has already been applied for solving the semilinear (singular perturbed) differential equations in the finite difference frame work by the Boglaev [4]. The same methodology [4] has been discussed in the present work but in the finite element framework for one-dimensional nonlinear problem. A numerical experiment has been presented from [4] using finite difference method.

To be brief in Section 2, for solving the nonlinear finite element scheme, a Schwarz iterative method which possesses the monotone convergence has been discussed. In Section 3, a monotone domain decomposition algorithm has been constructed and the monotone Schwarz iterative strategy has been employed in each of these subdomains.

3.2 Monotone Iterative Method

On $\bar{\Omega}$, introduce a non-uniform mesh $\bar{\Omega}^h$:

$$\bar{\Omega}^h = \{x_i, \quad 0 \le i \le N; \quad x_0 = 0, \quad x_N = 1; \quad h_i = x_{i+1} - x_i\},$$

In order to find the solution $U(x),\ x \in \bar{\Omega}^h$, we use the traditional finite element scheme which is thoroughly explained in Chapter 1. The nonlinear finite element weak formulation for problem (3.1.1) is defined as follows:

Find $U(x) \in H^1(\Omega)$ such that

$$\begin{aligned} \epsilon(U_x, V_x) + b\ (U_x, V) + (f(x,U), V) &= 0 \\ \Rightarrow \qquad B(U,V) + (f(x,U), V) &= 0 \end{aligned}$$

where the bilinear form is given by

$$\begin{aligned} B(U,V) &= \epsilon(U_x,V_x) + b\ (U_x,V) \\ \Rightarrow L^hU(x) + (f(x,U),V) &= 0, \quad x \in \Omega^h, \end{aligned} \tag{3.2.1}$$

where U satisfies the boundary condition $u = g\ on\ \ \partial\Omega$,
and $L^hU(x)$ is defined by

$$L^hU(x) = B(U,V)$$

Now, we construct the Schwarz iterative algorithm for solving the nonlinear finite element formulation (3.2.1) which possesses the monotone convergence. This method is based on the approach from [5].
Also, we assume that $f(x,U)$ from (3.1.1) satisfies bounds

$$0 < c_* \leq f_u \leq c^*, \tag{3.2.2}$$

where c_* and c^* are fixed constants. We define $\bar{U}(x)$ as an upper solution of (3.2.1) if it satisfies the following inequalities

$$\begin{aligned} L^h\bar{U}(x) + (f(x,\bar{U}),V) \geq 0, &\quad x \in \Omega^h, \\ \bar{U} \geq g &\quad \text{on} \quad \partial\Omega^h, \end{aligned}$$

where the operator L_h is defined as above.
Similarly, $\underline{U}(x)$ is called a lower solution of (3.2.1) if it satisfies the following inequalities

$$\begin{aligned} L^h\underline{U}(x) + (f(x,\underline{U}),V) \leq 0, &\quad x \in \Omega^h, \\ \underline{U} \leq g &\quad \text{on} \quad \partial\Omega^h. \end{aligned}$$

The iterative sequence $U^{(n)}(x)$ is constructed using the following recurrence formulas:
Let $U^{(0)}(x)$ be any fixed mesh function which satisfies the given boundary condition i.e.

$$U^{(0)}(x) = g(x), \qquad x \in \partial\Omega^h$$

Now, let the solution at first iteration be denoted by $U^{(1)}$ and is given by

$$U^{(1)} = U^{(0)} + Z^{(1)}, \tag{3.2.3}$$

where $Z^{(1)}$ is the error in $U^{(0)}$.

Now $U^{(1)}$ must satisfies the discrete equation (3.2.1) i.e.

$$L^h U^{(1)}(x) + (f(x, U^{(1)}(x)), V) = 0; \quad x \in \Omega^h. \tag{3.2.4}$$

Substituting equation (3.2.3) in (3.2.4), we get

$$L^h(U^{(0)} + Z^{(1)}) + (f(x, U^{(0)} + Z^{(1)}), V) = 0.$$

Since L^h is a linear operator, hence

$$L^h(U^{(0)} + Z^{(1)}) = L^h(U^{(0)}) + L^h(Z(1)).$$

Expanding the second term in above equation by using Taylor's series and simplifying, we get

$$L^h(U^{(0)}(x)) + L^h(Z^{(1)}(x)) + (f(x, U^{(0)}(x)), V) + (Z^{(1)}(x) f_u(x, U^{(0)}(x)), V) = 0, \quad x \in \Omega^h.$$

$$\Rightarrow L^h Z^{(1)}(x) + (c^* Z^{(1)}(x), V) = -[L^h U^{(0)}(x) + (f(x, U^{(0)}(x)), V)]$$

$$Z^{(1)}(x) = 0, \quad x \in \partial\Omega^h$$

Generalizing the iterative process, we get

$$U^{(n)}(x) = U^{(n-1)}(x) + Z^{(n)}(x) \tag{3.2.5}$$

$$L^h Z^{(n)}(x) + (c^* Z^{(n)}(x), V) = -[L^h U^{(n-1)}(x) + (f(x, U^{(n-1)}(x)), V)] \tag{3.2.6}$$

$$Z^{(n)}(x) = 0, \quad x \in \partial\Omega^h.$$

Hence, we get the Schwarz iterative solution $U^{(n)}(x)$ for the given problem (3.1.1) using the finite element methodology.

3.2.1 Remark 1

Let $\bar{U}^{(0)}$, $\underline{U}^{(0)}$ be upper and lower solution of (3.2.1), and let $f(x, U)$ satisfies (3.2.2). Then the upper sequence $\bar{U}^{(n)}(x)$ generated by (3.2.5) converges monotonically from above to the unique solution U of (3.2.1) [4] and the lower sequence $\underline{U}^{(n)}(x)$ generated by (3.2.5) converges monotonically from below to U

$$\underline{U}^{(0)} \le \underline{U}^{(1)} ... \le \underline{U}^{(n)} \le \underline{U}^{(n+1)} \le U \le \bar{U}^{(n+1)} \le \bar{U}^{(n)} ... \le \bar{U}^{(1)} \le \bar{U}^{(0)} \qquad \text{on } \bar{\Omega}^h.$$

3.2.2 Remark 2

We can consider the following approach for constructing initial upper and lower solutions $\bar{U}^{(0)}$ and $\underline{U}^{(0)}$. Consider a mesh function $W(x)$ which is defined on $\bar{\Omega}^h$ and satisfies the boundary condition $W(x) = g(x)$ on $\partial\Omega^h$. Introduce the following finite element problem

$$L^h Z_\mu^{(0)}(x) + (c_* Z_\mu^{(0)}(x), V) = \mu |L^h W(x) + (f(x, W(x)), V)|, \quad x \in \Omega^h, \tag{3.2.7}$$

$$Z_\mu^{(0)}(x) = 0, \quad x \in \partial\Omega^h, \quad \mu = 1, -1. \tag{3.2.8}$$

Then the functions $\bar{U}^{(0)} = W + Z_1^{(0)}$, $\underline{U}^{(0)} = W + Z_{-1}^{(0)}$ can be taken as the upper and lower solutions, respectively.

Theorem: Prove that $\bar{U}^{(0)}$ as defined above forms an upper solution.

Proof: We will prove the result by making use of discrete maximum principle. Here we have $\mu = 1$. Hence from (3.2.7), we get

$$\begin{aligned} L^h Z_1^{(0)}(x) + (c_* Z_1^{(0)}(x), V) &= |L^h W(x) + (f(x, W(x)), V)|, \\ &\geq 0 \quad \forall \;\; x \in \Omega^h. \end{aligned}$$

Also, since on the boundary $\partial\bar{\Omega}^h$, from (3.2.8) we have $Z_1^{(0)}(x) = 0$.
From discrete maximum principle, we get $Z_1^{(0)}(x) \geq 0$ on $\bar{\Omega}^h$.
Again, using the equation (3.2.7) for $Z_1^{(0)}(x)$ and using Mean Value Theorem, we have

$$\begin{aligned} L^h(W + Z_1^{(0)}) + (f(x, W + Z_1^{(0)}), V) = {} & [L^h(W) + (f(x, W), V)] \\ & + |[L^h(W) + (f(x, W), V)]| + ((f_u^{(0)} - c_*) Z_1^{(0)}, V). \end{aligned}$$

Since $f_u^{(0)} \geq c_*$ and $Z_1^{(0)}$ is non-negative, we conclude that $\bar{U}^{(0)}$ is an upper solution.

3.3 Monotone Domain Decomposition Strategy

In this Section, we consider a monotone domain decomposition algorithm based on the Schwarz alternating algorithm from [8] which has been explained using upwind finite difference framework for a two dimensional problem. In [8], domain decomposition algorithm consists of the two iterative processes: outer iterations and inner iterations. One outer iterative step represents computing M difference subproblems on overlapping vertical strips

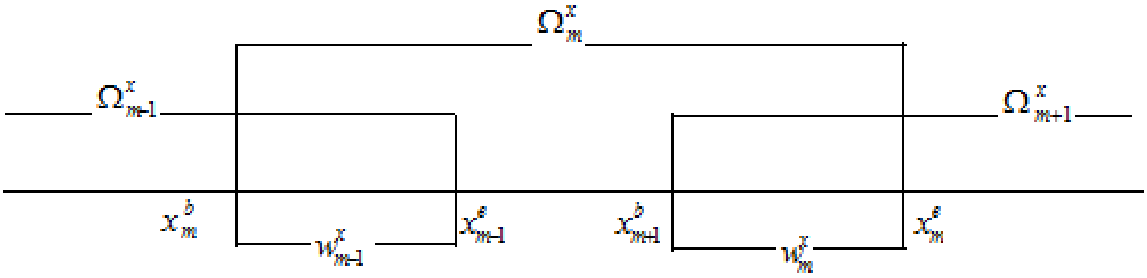


Fig. 1

$\bar{\Omega}_m^{hx}$, $m = 1, ..., M$ in serial, starting from $\bar{\Omega}_1^{hx}$ and finishing off on $\bar{\Omega}_M^{hx}$. The multiplicative Schwarz method is the outer part of the algorithm from [8]. At the level of inner iterations, each vertical strips $\bar{\omega}_m^h$ is split into overlapping boxes, and the natural parallelism can be used since each of the subdomains can be treated by its own processor. In the present work, we shall consider only the outer part of domain decomposition algorithm from [8] but in the finite element framework for one dimensional nonlinear problem. We introduce the set of the overlapping subdomains as follows:

$$\Omega_m = (x_m^b, x_m^e), \quad m = 1, ..., M, \quad x_1^b = 0, \quad x_M^e = 1,$$
$$\partial\Omega_m = \{x_m^b, x_m^e\}, \quad \Gamma_m^b = \{x_m^b\}, \quad \Gamma_m^e = \{x_m^e\}.$$

Thus,

$$\bar{\Omega}_m \cap \bar{\Omega}_{m+1} = \bar{w}_m, \quad m = 1, ..., M-1, \quad \text{where} \quad w_m = (x_{m+1}^b, x_m^e),$$

where $\bar{w}_m, m = 1, ..., M$ are the overlap domains between two subdomains $\bar{\Omega}_m$ and $\bar{\Omega}_{m+1}$.

$$\bar{\Omega}_m^h = \bar{\Omega}_m \cap \bar{\Omega}^h, \qquad \{x_m^b, x_m^e\}_{m=1}^M = \bar{\Omega}^h.$$

The domain decomposition of the domain $\bar{\Omega} = [0, 1]$ is illustrated in Fig.1

3.3.1 Monotone Domain Decomposition Algorithm

In this Section, we explain the monotone domain decomposition algorithm from [2], where one complete iterative step includes solving a sequence of M problems on $\bar{\Omega}^h_m, m = 1, ..., M$ in serial. For computing the problem on subdomains $\bar{\Omega}_m, 1 < m \leq M$, Dirichlet boundary condition on the left boundary Γ^b_m is updated by using the solution of the problem on the previous subdomains $\bar{\Omega}_{m-1}$ (previous step of the current iterative step). As for the monotone iterative method (3.2.5), we assume that $f(x, U)$ from (3.1.1) satisfies (3.2.2).

Step 0: Initialization: On the whole mesh $\bar{\Omega}^h$, choose an upper or lower solution $U^{(0)}(x), x \in \bar{\Omega}^h$ of (3.2.1) satisfying the boundary condition $U^{(0)}(x) = g(x)$ on $\partial\bar{\Omega}^h$.

Step 1: On the subdomains $\bar{\Omega}^h_m,\ m = 1, ..., M$, compute, in serial, mesh function $Z^{(n)}_m(x),\ m = 1, ..., M$, satisfying the following finite element scheme:

$$L^h Z^{(n)}_m(x) + (c^* Z^{(n)}_m(x), V) = -[L^h U^{(n-1)}(x) + (f(x, U^{(n-1)}(x)), V)], \quad x \in \Omega^h_m,$$

$$Z^{(n)}_m(x) = \begin{cases} 0, & x \in \Gamma^b_1,\ m = 1, \\ Z^{(n)}_{m-1}(x), & x \in \Gamma^b_m,\ m = 2, ..., M, \\ 0, & x \in \Gamma^e_m \end{cases} \tag{3.3.1}$$

and denote

$$W^{(n)}_m(x) = U^{(n-1)}(x) + Z^{(n)}_m(x), \qquad x \in \bar{\Omega}^h_m.$$

Step 2: Compute the solution $U^{(n)}(x),\ x \in \bar{\Omega}^h$ by piecing the solution on the subdomains

$$U^{(n)}(x) = \begin{cases} W^{(n)}_m(x), & x \in \bar{\Omega}^h_m \setminus w^h_m,\ \ m = 1, ..., M, \\ W^{(n)}_M(x), & x \in \bar{\Omega}^h_M. \end{cases} \tag{3.3.2}$$

Step 3: Stopping criterion: As soon as, the iterative solution $U^{(n)}(x);\ x \in \bar{\Omega}^h$ attains any given accuracy, then stop; otherwise go to Step 1.

One of the possible approaches for constructing initial upper and lower solution for the finite element formulation (3.2.1) has been suggested in Remark 2 of the previous Section.

Similar to Remark 1, we get the following convergence property of algorithm (3.3.1)-(3.3.2).

3.3.2 Remark 3

Let $\bar{U}^{(0)}$, $\underline{U}^{(0)}$ be upper and lower solutions of (3.2.1), and let $f(x, U)$ satisfy (3.2.2). Then the upper solution sequence $\bar{U}^{(n)}(x)$ generated by (3.3.1)-(3.3.2) converges monotonically from above to the unique solution $U(x)$ of (3.2.1) and the lower sequence $\underline{U}^{(n)}(x)$ generated by (3.3.1)-(3.3.2) converges monotonically from below to $U(x)$ and satisfies

$$\underline{U}^{(0)} \leq \underline{U}^{(1)} \leq ...\underline{U}^{(n)} \leq \underline{U}^{(n+1)} \leq U \leq \bar{U}^{(n+1)} \leq \bar{U}^{(n)} \leq ...\bar{U}^{(1)} \leq \bar{U}^{(0)} \qquad \text{on} \ \ \bar{\Omega}^h$$

3.4 Numerical Result

In this Section, a nonlinear singularly perturbed problem has been considered and its numerical results have been obtained presented which have been obtained using monotone Schwarz iterative domain decomposition algorithm in [6] and the finite difference scheme has been employed. Also, a grid refinement technique which solves the problem in interfacial subdomains has been employed in [6]:

3.4.1 Example

Consider the problem

$$L_\epsilon u(x) \equiv \epsilon u^{''} - f(x, u) = 0, \qquad x \in (0, 1),$$
$$u(0) = 0, \quad u(1) = 0$$

$$f_u \geq \beta_0^2, \quad \beta_0 = const > 0, \quad (x, u) \in (0, 1) \times \Re. \tag{3.4.1}$$

where

$$f(x, u) = exp\Big[- (1 - x)^2 \Big] - exp(-u)$$

Clearly, we can check that $f_u \geq \beta_0^2, \quad \beta_0 = const > 0, \quad (x, u) \in (0, 1) \times \Re$

Take $\mu = \sqrt{\epsilon}$. Boundary layer width h_ϵ is defined as

$$h_\epsilon = \mu |\ln(\mu)|/\beta_0.$$

It is easy to see that this test problem has the boundary layer at $x = 0$. Introduce a non-equidistant mesh $\{x_i, i = 0, 1, ..., N\}$. In the boundary layer region $[0, h_\epsilon]$, the mesh generating function is of logarithmic type function from [1].

$$x_i \in [0, h_\epsilon]: \quad x_i = -\mu ln[1 - (1-\mu)i/n_\epsilon]/\beta_0, \qquad i = 0, 1, ..., n_\epsilon = N/2;$$

$$x_i \in (h_\epsilon, 1]: \quad x_i = x_{i-1} + (1 - h_\epsilon)/n_\epsilon, \qquad i = n_\epsilon + 1, ..., N = 2n_\epsilon.$$

If $h_\epsilon > 0.5$, then we choose the following mesh:

$$x_0 = 0, \qquad x_i = x_{i-1} + h_0\gamma^{i-1}, \quad i = 1, 2, ..., n_\epsilon,$$

$$h_0 = \min\{0.5/n_\epsilon; \ -\mu ln[1 - (1-\mu)/n_\epsilon]/\beta_0\},$$

where γ satisfies the condition

$$h_0 \sum_{i=0}^{n_\epsilon - 1} \gamma^i = 0.5,$$

i.e.,

$$x_{n_\epsilon} = 0.5,$$

and

$$x_i = x_{i-1} + 0.5/n_\epsilon, \qquad i = n_\epsilon + 1, ..., N.$$

The differential equation from (3.4.1) is approximated by a simple variable-mesh formula. The nonlinear algebraic system are solved by the Newton iterative method upto an accuracy of 10^{-5}. The iterative algorithm are finished to achieve an accuracy of 10^{-5}.

In Table 1.1 and Table 1.2, we give the numerical results of the Schwarz alternative algorithm and interfacial procedure for the various value of μ and h (the number of the mesh points $n_\epsilon = 100$) have been presented [6]. K_{A1} and K_{A2} denote number of iterations for Schwarz alternative algorithm and interfacial procedure, respectively, to the achieve an accuracy of 10^{-5} (in the table $*$ denotes numbers K_{A1},$\mathrm{K}_{A2} > 200$). The table contains the value of K_{A1}, K_{A2} for the two cases: $h_{out} = h/16$ and $h_{out} = h$. One can see that, in the first case, the inequality $\mathrm{K}_{A1} \geq \mathrm{K}_{A2}$ is fulfilled if $h/\mu \leq 1$. In the case $h_{out} = h$, we have $K_{A1} \simeq K_{A2}$ for all values of μ and h.

Table 3.1:

h	K_{A1};K_{A2} h_{out} =h/16				
2^{-2}	7;11	5;6	5;6	3;3	3;3
2^{-4}	21;35	8;13	6;6	4;5	3;3
2^{-6}	65;111	23;39	8;13	6;6	4;4
2^{-8}	*;*	74;126	23;38	8;12	6;6
2^{-10}	*;*	*;*	72;123	19;32	7;10
μ	2^{-2}	2^{-4}	2^{-6}	2^{-8}	2^{-10}

Table 3.2:

h	K_{A1};K_{A2} h_{out} =h				
2^{-2}	7;6	5;5	5;5	3;3	3;3
2^{-4}	21;19	8;7	6;6	4;4	3;3
2^{-6}	65;64	23;22	8;7	6;6	4;4
2^{-8}	*;*	74;72	23;21	8;7	6;6
2^{-10}	*;*	*;*	72;70	19;18	7;7
μ	2^{-2}	2^{-4}	2^{-6}	2^{-8}	2^{-10}

3.5 Conclusion

In the present study, we have considered a nonlinear convection-diffusion problem. Firstly, monotone Schwarz iterative methodology has been discussed under finite element framework for solving the problem under consideration. Since the problem is non-linear in nature, if we directly apply the finite element method, the method results in a system of nonlinear algebraic equations. A major point about this system is to obtain reliable and efficient computational algorithms for computing the solution. A major benefit of the monotone Schwarz iterative strategy is that it leads to iterative algorithm which converge globally and solve only linear discrete system at each iterative step. Further, the iterative scheme has been incorporated in the domain decomposition algorithm in the finite element framework. The algorithm has the benefit that the problem can be solved on subdomains in parallel.

Bibliography

[1] I. Boglaev, V. Duoba, Domain decomposition for an advection-diffusion problem with parabolic layers, Appl.Math.Comput. 146 (2003) 27-53.

[2] I. Boglaev, On monotone iterative method for the a nonlinear singularly perturbed reaction-diffusion problem, J.Comput.Appl.Math. 162 (2004) 445-466.

[3] I. Boglaev, A monotone Schwarz algorithm for a semilinear convection-diffusion problem, J.Numer.Math. Accepted for publication.

[4] I. Boglaev, A block monotone domain decomposition algoritm for a semilinaer convection-diffusion problem, J.Comput.Appl.Math 173 (2005) 259-277.

[5] I. Boglaev, V. Sirotkin, The integral-difference method for quasi-linear singular perturbation problems,in: J.J.H.Miller (Ed.), Computational methods for Boundary and Interior Layers in Several Dimension, Boole Press, Dublin, 1991, pp. 1-26.

[6] I. Boglaev, V.V. Sirotkin, Computational method for a singular perturbation problem via domain decomposition and its parallel implementation, Appl.Math.Comput 56 (1993) 71-95.

[7] P.G. Ciarlet, The Finite Element Method for Elliptic Problems (North-Holland, amsterdam, 1978).

[8] M. Garbey, Yu.A. Kuznetsov, Yu.V. Vassilevski, A parallel Schwarz method for a convection-diffusion problem, SIAM J.Sci.Comput. 22 (2000) 891-916.

[9] R.S. Gupta, Elements of numerical analysis, Macmillan, 2008.

[10] O.A. Ladyzenskaja, N.N. Ural'ceva, Linear and Quasi-linear elliptic equations, Academic Press, New York, 1968.

[11] M.D. Raisinghania, Ordinary and partial differential equations, S. Chand, 2008.

[12] J.N. Reddy, An Introduction to finite element methods, McGraw-Hill, New York, 2006.

[13] H.G. Roos, M. Stynes, L. Tobiska, Numerical methods for singularly perturbed Differential Equations: Convection-Diffusion and Flow problems, Springer-Verlag, Berlin, 1996.

[14] G. Strang and G. Fix, An analysis of the finite element (Prentice-Hall, Englewood Cliffs) 1973.