A POSTERIORI ERROR ESTIMATES FOR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS IN THE FINITE ELEMENT METHOD

By
Manar Jafar Wazwaz

Supervisor
Dr. Hasan Almanasreh

Submitted in partial fulfillment of the requirements for the degree of Master of Science, Faculty of Graduate Studies, Hebron University, Hebron, Palestine.

2017

A POSTERIORI ERROR ESTIMATES FOR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS IN THE FINITE ELEMENT METHOD

By
Manar Jafar Wazwaz

This thesis was defended successfully on 20/4/2017 and approved by:

Committee Members:		**Signature**
• Dr. Hasan Almanasreh	Supervisor	....
• Dr. Tareq Amro	Internal Examiner	....
• Dr. Ali Zein	External Examiner	....

Dedication

I dedicate my thesis to my parents, brothers, sisters and friends who supported me each step of the way.

Acknowledgments

I am heartily thankful to my supervisor, Dr. Hasan Al manasreh, for providing me with unfailing support and continuous encouragement from the initial to the final level enabled me to develop and understand the subject.

I acknowledge Hebron University for supporting this work, and I wish to pay my great appreciation to all respected staff in the department of mathematics.

I must express my very profound gratitude to my family whose encouragement, guidance and support throughout my years of study and through the process of researching and writing this thesis.

Finally, I offer my regards and thanks to all of those who supported me during the completion of this thesis.

الإقرار

أنا الموقع أدناه مقدم الرسالة التي تحمل العنوان :

الأخطاء البعدية للمعادلات التفاضلية الجزئية القطعية الناقصة باستخدام طريقة العناصر المحدودة

A POSTERIORI ERROR ESTIMATES FOR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS IN THE FINITE ELEMENT METHOD

أقر بأن ما اشتملت عليه هذه الرسالة إنما هو نتاج جهدي الخاص،
باستثناء ما تمت الإشارة إليه حيثما ورد، و أن هذه الرسالة ككل لم تقدم
من قبل لنيل أية درجه علمية أو بحث علمي أو بحثي لدى أية مؤسسة تعليمية او بحثية أخرى.

Declaration

The work provided in this thesis, unless otherwise referenced, is the result of the researcher's work, and has not been submitted elsewhere for any other degree or qualification.

Abstract

In the present thesis we study the a posteriori error estimates for elliptic partial differential equations in the finite element method to assess the error of the approximate solution in the employed triangulation. For this purpose, we firstly discuss the general theory of the finite element method in one and two dimensions: the variational formulation for general differential equation is obtained, also, the discretization system of the variational form is discussed for general boundary conditions. In this essence, numerical solutions for some differential equations are provided using the Matlab software.

A posteriori error estimator is a quantity which bounds or approximates the error and can be computed from the knowledge of numerical solution and input data. The advantage of any a posteriori error estimator is to provide an estimate and ideally bounds for the solution error in a specified norm or in a functional of interest if the problem data and the finite element solution are available. To achieve our goal, required basic concepts to obtain a posteriori error estimates for the finite element solution of an elliptic linear differential equation are reviewed. We give the basic ideas to establish global error estimates for the energy norm as well as goal-oriented error estimates. Moreover, a posteriori error estimates for general differential equations are obtained in more details.

الملخص

في هذه الرسالة سوف ندرس الأخطاء البعدية الناتجة من استخدام طريقة العناصر المحدودة للحلول العددية القطعية الناقصة وذلك لتقييم مقدار الخطأ في الحل التقريبي باستخدام هذه الطريقة والمطبق على تجزئة المنطقة (المجال) إلى مثلثات. لهذا الغرض، سوف نناقش أولاً طريقة العناصر المحدودة للحلول العددية للمعادلات التفاضلية أحادية وثنائية البُعد. في هذا السياق، سوف نستعرض صيغة الحل التبايني ونظام التجزئة للمعادلات التفاضلية بمختلف الشروط على حدود مجال التعريف. أيضاً أمثلة توضيحية سوف يتم دراستها باستخدام لغة البرمجة الرياضية الماتلاب.

الأخطاء البعدية التي نتناولها في هذه الدراسة هي كميات تحد وتقرب الأخطاء الناتجة عن الحلول العددية للمعادلات التفاضلية باستخدام طريقة العناصر المحدودة والتي يمكن تقديرها بالاعتماد على الحل التقريبي والمعلومات المتوفرة عن المعادلة التفاضلية. فائدة الأخطاء البعدية للحلول العددية للمعادلات التفاضلية يتلخص في معرفة مقدار الخطأ في الحل التقريبي وذلك من أجل جعله أقل ما يمكن وبالتالي الحصول على أفضل تقريب للحل. ولمناقشة هذه الأخطاء البعدية، مفاهيم أساسية سوف يتم تناولها أولاً لتوضيح هذه الأخطاء بعد ذلك سوف يتم معالجة الأخطاء البعدية بالتفصيل لعدد من المعادلات التفاضلية الجزئية.

Contents

Introduction **vii**

1 *General theory of FEM* **4**
1.1 Sobolev spaces 4
1.2 Classification of partial differential equations (PDEs) 8
1.3 Finite element method (FEM) 10
1.4 Variational formulation and discretization 11
1.4.1 One-dimensional problem 11
1.4.2 Two-dimensional problem 17
1.5 Existence and uniqueness theorems 20

2 *FEM for general differential equations* **21**
2.1 One dimensional problem 21
2.1.1 Variational formulation and discretization 21
2.1.2 Numerical Examples 25
2.2 Two dimensional problem 29
2.2.1 Variational formulation and Discretization 29
2.2.2 Numerical Examples 32

3 *Error estimation* **34**
3.1 Introduction 34
3.2 A priori error estimates 38
3.3 A posteriori error estimates 43
3.3.1 A short history 43
3.3.2 A posteriori error estimate for Poisson equation 44

4 ***The a posteriori error estimator for poisson equation*** **46**
4.1 A posteriori error estimator for homogeneous boundary condition 47
4.2 A posteriori error estimator for mixed boundary condition 50

5 ***The a posteriori error estimator for reaction-diffusion and convection-diffusion problems*** **56**
5.1 Reaction-diffusion problems 56
5.1.1 A priori error estimation 57
5.1.2 A posteriori error estimation 58
5.2 *Convection-diffusion problems* 61
5.2.1 The convection-diffusion problems 62
5.2.2 A posteriori error estimation 63

Conclusion **72**

Appendix **73**

References **81**

Introduction

The finite element method (FEM) is a numerical technique for solving problems which are described by partial differential equations that arise in scientific and engineering applications. The FEM uses a variational form of the problem that involves an integral form of the differential equation over a given domain where this domain is divided into a number of subdomains called finite elements.

We are interested in the existence and a posteriori estimates of the weak solutions of linear elliptic differential equations. Such problems arise in a variety of situations in biology, chemistry, or physics,...etc. The purpose of this thesis is to study the finite element method for second order elliptic problems in one and two dimensions and find the a posteriori error estimates for Poisson, reaction-diffusion, and convection-diffusion problems.

In this thesis the Sobolve spaces that are used in the variational formulation of differential equations and some other required concepts are defined. The classifications of the differential equations according to the value of demonstrate to elliptic, hyperbolic and parabolic, and according to the boundary conditions, Dirichlet, Neumann, and Robin problems are explained. We formulate the general theorems for existence and uniqueness in Hilbert space framework and state the conditions that spaces and bilinear form should satisfy. These results are applied to investigate the solvability of particular partial differential equations.

The core of this work starts with the discussion of the variational formulation and the discretization of the problem with homogenous and mixed boundary conditions. We construct a variational (weak) formulation by multiplying both sides of the differential equation by a test function $v(x) \in V$ such that V is some Sobolve space and then integrate over the interval by parts. The aim of introducing the notation of weak formulation is to give access to the existence and uniqueness results for the solutions which is well suited for the numerical approximation of such problems. In the discretization we construct a finite element dimensional space V_h of continuous linear functions on the partition τ_h, and find $U(x) \in V_h$ such that the variational formulation holds. Then we discuss the error estimate which is the difference between the approximate solution u_h and the exact solution u. The both types of error are a priori and a posteriori error estimates. The first type is error bounds given by known information on the solution of the variational problem and the finite element function space, where the second type is error bounds given by information on the numerical solution obtained on the finite element function space.

In this thesis we will focus on three types of problems :Poisson, Reaction-Diffusion, and Convection-Diffusion Problems, where the main task is to discuss the a posteriori error estimates for these problems.

The Poisson equation as the model problem for elliptic partial differential equation. It arises, e.g., in structural mechanics, theoretical physics as gravitation, electromagnetism, elasticity and in many other areas of science and engineering. The poisson problem is defined as :

$$-\Delta u(x) = f(x), \quad x \in \Omega.$$

The reaction-diffusion problem arises naturally in systems consisting of many interacting components as chemical reactions, and are widely used to describe pattern-formation phenomena in variety of biological, chemical and physical systems. The typical form is as follows:

$$-\varepsilon \Delta u + cu = f, \quad \text{in } \Omega \ .$$

The convection-diffusion problems very often happen that the solution have a convective nature on most of the domain of the problem, and the diffusive part of the differential operator affect only in certain small subdomains. They usually have a degree of instability. Common sources of convection-diffusion problems are the Navier-Stokes equations, semiconductor device modeling, and from financial modeling,[6], the Black-sholes equation. The convection-diffusion is defined as :

$$-\varepsilon\Delta u + b.\nabla u + cu = f, \quad \text{in } \Omega.$$

This project consists of six chapters. Chapter One will be about the FEM in general. Chapter Two talks about the variational formulation and discretization of differential equation. The error estimation in its both types, a posteriori and a priori, will be explained in Chapter Three. The a posteriori error estimate for Poisson equation is discussed in Chapter Four. Chapters five and six will be about a posteriori error estimate for reaction-diffusion and convection-diffusion problems, respectively.

Chapter 1

General theory of FEM

1.1 Sobolev spaces

In this section we introduce a class of spaces called Sobolev spaces that are used in the variational formulation of differential equations. We begin the concept of a multi-index, let $\mathbb{N}$ denote the set of all non-negative integer. An n-tuple $\alpha = (\alpha_1\ , \alpha_2\ , \dots\ , \alpha_n) \in \mathbb{N}^n$ is called a **multi-index**. The non-negative integer $|\alpha| = \sum_{i=1}^n \alpha_i$ is the length of the multi-index.
Let

$$D^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \dots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} = \frac{\partial^{|\alpha|}}{\partial {x_1}^{\alpha_1} \dots \partial {x_n}^{\alpha_n}},$$

Ω be an open set in $\mathbb{R}^n$ and let $k \in \mathbb{N}$.

Definition 1. *Let Ω be an open set in $\mathbb{R}^n$, $u \in \mathbb{N}$. Define spaces $C^k(\Omega), C^k(\bar{\Omega}), C^\infty(\Omega)$ by*

$$\begin{aligned} C^k(\Omega) &= \{u : \Omega \to \mathbb{R} : D^\alpha u \textit{ is continuous in } \Omega, \quad \forall\ |\alpha| \le k\}, \\ C^k(\bar{\Omega}) &= \{u : \bar{\Omega} \to \mathbb{R} : D^\alpha u \textit{ is continuous in } \bar{\Omega}, \quad \forall\ |\alpha| \le k\}, \\ C^\infty(\Omega) &= \{u : \Omega \to \mathbb{R} : D^\alpha u \textit{ is continuous in } \Omega, \quad \forall\ \alpha \in \mathbb{N}^n\}, \end{aligned}$$

where $\bar{\Omega}$ is the closure of Ω. If Ω is bounded, $\bar{\Omega} = \Omega \cup \partial\Omega$, where $\partial\Omega$ is the boundary of Ω. Denote $C(\Omega) = C^0(\Omega)$ and $C(\bar{\Omega}) = C^0(\bar{\Omega})$.

Definition 2. *The **support** of a continuous function u defined on an open set $\Omega \subset \mathbb{R}^n$ is defined as the closure in Ω of the set $\{x \in \Omega : u(x) \neq 0\}$. We shall write supp u for the support of u. Thus, supp u is the smallest closed subset of Ω such that $u \neq 0$ in Ω.*

Let $C_0^\infty(\Omega) = \bigcap_{k \geqslant 0} C_0^k(\Omega)$, where $C_0^k(\Omega)$ is the set of all u contained in $C^k(\Omega)$ whose support is a bounded subset of Ω.

Definition 3. *A function $f : \Omega \to \mathbb{R}$ is **locally integrable** if $f \in L^1(K)$ (i.e., $\int_k |f| < \infty$) for every bounded open set K such that $\bar{K} \subset \Omega$. The space $L^1_{loc}(\Omega)$ consists of locally integrable functions.*

Suppose that u is a smooth function, say $u \in C^k(\Omega)$, with Ω an open subset of $\mathbb{R}^n$, and let $v \in C_0^\infty(\Omega)$, then

$$\int_\Omega D^\alpha u(x)v(x)\ dx = (-1)^{|\alpha|} \int_\Omega u(x) D^\alpha v(x)\ dx, \quad |\alpha| \leq k.$$

Note that this formula is just integration by parts and all terms involving over the boundary Ω have disappeared because v and all of its derivatives are zero on the boundary.

Definition 4. *Suppose that u and w_α are locally integrable function defined on Ω such that*

$$\int_\Omega w_\alpha(x)\phi(x)\ dx = (-1)^{|\alpha|} \int_\Omega u(x) D^\alpha \phi(x)\ dx, \quad \forall \phi \in C_0^\infty(\Omega),$$

*then the **weak derivative** of u of order α denoted by $D^\alpha u$ is defined by $w_\alpha = D^\alpha u$.*

Remark : If a locally integrable function has a weak derivative then it is unique.

Definition 5. *Let Ω denote an open subset of $\mathbb{R}^n$ and assume $p \in [1, \infty)$. The space $L_p(\Omega)$ of integrable functions is defined by*

$$L_p(\Omega) = \Big\{ v : \Omega \to \mathbb{R}; \int_\Omega |v(x)|^p\ dx < \infty \Big\}.$$

Definition 6. *A set M is called **complete** if every Cauchy sequence in M converges to an element in M.*

Note that a complete inner product space is called a **Hilbert space**.

Definition 7. *[12] Let k be a non-negative integer and suppose that $p \in [1, \infty]$,we define*

$$w_p^k(\Omega) = \{u \in L_p(\Omega) : D^\alpha u \in L_p(\Omega), |\alpha| \leq k\},$$

$w_p^k(\Omega)$ is called a ***Sobolev space*** *of order k.*

Define the Sobolev norm

$$||u||_{w_p^k(\Omega)} = \left(\sum_{|\alpha| \leq k} ||D^\alpha u||^p_{L_p(\Omega)} \right)^{1/p}, \text{when} \quad 1 \leq p < \infty,$$

$$\text{where } ||u||_{L_p(\Omega)} = \big(\int_\Omega |u(x)|^p dx \big)^{1/p}.$$

$$||u||_{w_\infty^k(\Omega)} = \sum_{|\alpha| \leq k} ||D^\alpha u||_{L_\infty(\Omega)}, \text{when} \quad p = \infty,$$

$$\text{where} \quad ||u||_{L_\infty(\Omega)} = sup_{x \in \Omega} |u(x)|.$$

An important special case corresponds to taking $p = 2$, the space $w_2^k(\Omega)$ is then a Hilbert space with the inner product

$$(u, v)_{w_2^k(\Omega)} = \sum_{|\alpha| \leq k} (D^\alpha u, D^\alpha v),$$

$$\text{where the inner product is defined as } (u, v) = \int_\Omega u(x)v(x)dx.$$

We usually write $H^k(\Omega)$ instead of $w_2^k(\Omega)$.

Some definitions of $w_p^k(\Omega)$, see [52]; and its norm, for $p = 2, k = 1$,

$$H^1(\Omega) = \left\{ u \in L_2(\Omega) : \frac{\partial u}{\partial x_j} \in L_2(\Omega), j = 1, ..., n \right\}.$$

$$||u||_{H^1(\Omega)} = \left(||u||^2_{L_2(\Omega)} + \sum_{j=1}^n || \frac{\partial u}{\partial x_j} ||^2_{L_2(\Omega)} \right)^{1/2}.$$

$$|u|_{H^1(\Omega)} = \left(\sum_{j=1}^n || \frac{\partial u}{\partial x_j} ||^2_{L_2(\Omega)} \right)^{1/2}.$$

For $p = 2, k = 2$,

$$H^2(\Omega) = \left\{u \in L_2(\Omega) : \frac{\partial u}{\partial x_j} \in L_2(\Omega), \frac{\partial^2 u}{\partial x_i \partial x_j} \in L_2(\Omega), i, j = 1, ..., n\right\}.$$

$$||u||_{H^2(\Omega)} = \left(||u||^2_{L_2(\Omega)} + \sum_{j=1}^{n} ||\frac{\partial u}{\partial x_j}||^2_{L_2(\Omega)} + \sum_{i,j=1, i\neq j}^{n} ||\frac{\partial^2 u}{\partial x_i \partial x_j}||^2_{L_2(\Omega)}\right)^{1/2}, i \neq j.$$

$$|u|_{H^2(\Omega)} = \left(\sum_{i,j=1}^{n} ||\frac{\partial^2 u}{\partial x_i \partial x_j}||^2_{L_2(\Omega)}\right)^{1/2}.$$

Finally, we define the special Sobolev space H_0^1 as the closure of C_0^∞ in the norm $||.||_{H^1(\Omega)}$.

$$H_0^1(\Omega) = \{u \in H^1(\Omega) : u = 0 \ on \ \partial\Omega\}.$$

Note that $H_0^1(\Omega)$ is a Hilbert space, with the same norm and inner product as $H^1(\Omega)$.

Lemma 1. ***Poincare Inequality***
Suppose that Ω is a bounded open set in $\mathbb{R}^n$ and let $u \in H_0^1$, then there exist a constant C_Ω , independent of u, such that

$$\int_\Omega |u(x)|^2 dx \leq C_\Omega \sum_{i=1}^{n} \int_\Omega |\frac{\partial u(x)}{\partial x_i}|^2 dx \quad ,$$

i.e.,

$$||u||_{L_p(\Omega)} \leq C_\Omega ||\nabla u||_{L_p(\Omega)}.$$

see [52] for the proof.
Young's Inequality.
When $1 < p < \infty$ and $a, b \geq 0$, Young's inequality is the expression

$$ab \leq \frac{p-1}{p} a^{\frac{p-1}{p}} + \frac{1}{p} b^p$$

In particular, if $p = 2$, we have Cauchy's inequality

$$ab \leq \frac{1}{2}(a^2 + b^2).$$

Definition 8. *Let V be a vector space, then*

1. *A* ***linear functional*** *on V is a function $\phi : V \to \mathbf{R}$ that is linear in the sense that*

$$\begin{aligned}\phi(v+w) &= \phi(v)+\phi(w)\\ \phi(\alpha v) &= \alpha\phi(v)\end{aligned}$$

   $\forall v,w \in V$, and $\forall\ \alpha \in \mathbf{R}$.

2. *The* ***dual space*** *V' of the vector space V is the set of all linear functional on V, or the space of linear functional on V.*

1.2 Classification of partial differential equations (PDEs)

In this section, we will consider the general second order linear PDE and will reduce it to one of three distinct types of equations.
The most general form of a linear second order PDE in two independent variables x, y and the dependent $U(x, y)$ is

$$AU_{xx} + BU_{xy} + CU_{yy} + DU_x + EU_y + FU + G = 0, \text{with } A, ..., G \text{ are given functions.}$$

Let $M = B^2 - 4AC$ denote the discriminant of the given equation, then this equation is called

1. Elliptic when $M < 0$.
2. Parabolic when $M = 0$.
3. Hyperbolic when $M > 0$.

Another form of a linear second order PDE can be described as follows

$$-\nabla.(A\nabla u) + b.\nabla u + cu = f, \tag{1.1}$$

where $A : \Omega \to \mathbb{R}^{n\times n}$ is matrix of real-valued functions, $b : \Omega \to \mathbb{R}^n$ is a vector of real-valued functions, and $c : \Omega \to \mathbb{R}$ is a real-valued function.

Definition 9. *An $n \times n$ matrix operator A is called* ***positive definite*** *if $v^T Av > 0$ for any $v \neq 0$.*

Definition 10. *The equation (1.1) is said to be, see [59],*

1. elliptic *in $x \in \mathbb{R}^n$, if $A(x)$ is positive definite or negative definite .*

2. hyperbolic *in $x \in \mathbb{R}^n$, if $n-1$ eigenvalues of $A(x)$ are of the same sign and one is of the opposite sign.*

3. parabolic *in $x \in \mathbb{R}^n$, if $n-1$ eigenvalues of $A(x)$ are of the same sign, one equals null and $rang(A, b) = n$.*

Note that we say equation (1.1) is elliptic in Ω if $A(x)$ for all $x \in \Omega$ is positive definite or if for all $x \in \Omega$ is negative definite.

Now we will classify the second order linear PDE by the boundary conditions [42], where there are three types of boundary conditions. Defining a domain Ω, and its boundary $\partial\Omega$, the three boundary conditions are :

1. *Dirichlet* boundary conditions with $u = g$ on $\partial\Omega$.

   (a) Homogeneous Dirichlet if $g = 0$.

   (b) Inhomogeneous Dirichlet if $g \neq 0$.

2. *Neumann* boundary conditions with $\frac{\partial u}{\partial n} = n.\nabla u = g$ on $\partial\Omega$, where n is outward normal to $\partial\Omega$.

   (a) Homogeneous Neumann if $g = 0$.

   (b) Inhomogeneous Neumann if $g \neq 0$.

3. Mixed *(Robin)* boundary conditions if $\frac{\partial u}{\partial n} + \gamma u = g$ on $\partial\Omega$, where n is outward normal to $\partial\Omega$ and γ is constant.

.

1.3 Finite element method (FEM)

The finite element method is a numerical technique for solving problems which are described by partial differential equations that arise in scientific and engineering applications. It uses a variational problem that include an integral of the differential equation over the problem domain. This domain is divided into a number of subdomains called finite elements.

The advantage of the FEM compared with finite difference method is that the FEM useful for problem with complicated geometry and material properties where analytical solutions can not be obtained, and general boundary conditions can be handled relatively easily. Also, the FEM has a solid theoretical foundation which gives added reliability and in many cases makes it possible to mathematically analyze and estimate the error in the approximation.

In this method we start from a reformulation of the given differential equation as an equivalent variational problem.

A short history

The finite element method, [10, 20, 28], was first proposed in 1909 by Ritz who developed an effective method [48] for the approximate solution of problems which contains an approximation of energy functional by the known functions with unknown coefficients. In 1943 the German mathematician Richard Courant [17] increased possibilities of the Ritz method by introduction the special linear functions defined over piecewise linear approximations on subregions, and he used a finite element type of procedure in a potential energy minimization of a functional for the torsion stress function using grid point values as the unknown parameters. Over the period $1950 - 1962$, Turner, Clough, Martin and Topp [53] generalized and perfected the Direct Stiffness Method and he oversaw the development of the first continuum based finite elements. An important contribution was brought into finite element method development by the papers of Argyris and Kelsey [3, 4] Clough [15, 16], Hrennikov [26]. Although the name finite element method was not introduced until 1960 when it was proposed by

Clough [15, 16] The first book on finite element method was published in 1967 by Zienkiewicz and Cheung [57] and called (The finite element method in structural and continuum mechanics). The mathematical analysis of these methods began in the 1960's. In 1962 Friedrichs [21] used piecewise linear function on triangles to derive a system of equations for solving problems on a general domain. In 1963 Oganesjan [40] proved the first a priori estimate for the error in H^1 norm for Laplace's and more second order elliptic equations, e.g., for plates. And in 1968 The a priori error estimate for quadratic elements on triangles was produced by Zlamal[58]. In the late 1970's, work on a posteriori error analysis began, the paper of Babuska and Rheinboldt [7, 8] published in 1978 is often cited as the first work aimed at developing rigorous global error bounds for finite element approximations of linear elliptic two-point boundary value problems.

1.4 Variational formulation and discretization

To make the concept of the FEM clear to the reader, we will study in this section the following two problems, consider the ODE and PDE problems.

$$\begin{aligned} \text{1-D} : -u''(x) &= f(x), \quad \alpha < x < \beta, \quad u(\alpha) = u(\beta) = 0. \\ \text{2-D} : -\Delta u &= f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega, \end{aligned}$$

where $\Delta u = \frac{\partial^2 u}{\partial {x_1}^2} + \frac{\partial^2 u}{\partial {x_2}^2}$, and Ω is a domain in the plane with boundary $\partial\Omega$.

1.4.1 One-dimensional problem

Consider the 1-D two-point boundary value problem

$$-u''(x) = f(x), \ \alpha < x < \beta,$$

$$u(\alpha) = u(\beta) = 0,$$

where f is a source function. Construct a variational (weak) formulation by multiplying both sides of the differential equation by a test function $v(x) \in V$, where $V = \{v : v$ is continuous function on $[a, b]$, v' is piecewise and bounded on $[\alpha, \beta]$, and $v(\alpha) = v(\beta) = 0\}$.

Now integrate over the interval $[\alpha, \beta]$

$$\int_{\alpha}^{\beta} -u''(x)v(x)dx = \int_{\alpha}^{\beta} f(x)v(x)dx$$

$$-u'(x)v(x)\Big|_{\alpha}^{\beta} + \int_{\alpha}^{\beta} u'(x)v'(x)dx = \int_{\alpha}^{\beta} f(x)v(x)dx,$$

but $v(\alpha) = v(\beta) = 0$,then

$$\int_{\alpha}^{\beta} u'(x)v'(x)dx = \int_{\alpha}^{\beta} f(x)v(x)dx, \quad v \in V. \tag{1.2}$$

The solution u is then known as a weak solution to the problem (1.2). Using the notations

$$\begin{aligned} a(u, v) &= \int_{\alpha}^{\beta} u'v'dx, \\ F(v) &= \int_{\alpha}^{\beta} f(x)v(x)dx, \end{aligned}$$

we then reformulate the variational formulation (1.2) in an abstract form as follows:
Find $u \in H_0^1(\Omega)$ such that

$$a(u, v) = F(v), \quad \forall v \in H_0^1(\Omega). \tag{1.3}$$

Discretization

Construct a finite element dimensional space V_h of continuous linear functions on the partition τ_h , $\alpha = x_0 < x_1 < ... < x_{m+1} = \beta$.
The problem is now to find $U(x) \in V_h$ such that

$$\int_{\alpha}^{\beta} U'(x)v'(x)dx = \int_{\alpha}^{\beta} f(x)v(x)dx, \quad \forall v \in V_h, \tag{1.4}$$

where $V_h = \{w \in V : w \text{ is linear on each subinterval } I_j \text{ , } w(\alpha) = w(\beta) = 0 \}$.
We will define $I_j = [x_{j-1}, x_j]$, $h_j = x_j - x_{j-1}$, $h = \max h_j$ for $j = 1, ..., m+1$. The

subspace V_h can be spanned by the following basis functions,

$$\phi_j(x) = \begin{cases} \frac{x-x_{j-1}}{h_j} & , x_{j-1} \le x \le x_j, \\ \frac{x_{j+1}-x}{h_{j+1}} & , x_j \le x \le x_{j+1}, \\ 0 & , \text{otherwise}, \end{cases}$$

which is called the *hat functions*. It is clear that

$$\phi_j(x_i) = \begin{cases} 1 & , i = j \\ 0 & , i \ne j \end{cases} = \delta_{ji}, \quad \text{the Kronecker delta functions.}$$

Since $U \in V_h$, then U can be written as a unique linear combination of ϕ_j's .

$$U(x) = \sum_{j=0}^{m+1} \xi_j \phi_j(x) \quad , \text{where } \xi_j = U(x_j).$$

Note that $\xi_0 = \xi_{m+1} = 0$ since $U = 0$ at $x_0 = \alpha, x_{m+1} = \beta$, thus, use $U(x) = \sum_{j=1}^{m} \xi_j \phi_j(x)$ in (1.4) to get

$$\sum_{j=1}^{m} \xi_j \int_\alpha^\beta \phi_j' v' dx = \int_\alpha^\beta f v dx, \quad \forall v \in V_h.$$

Since $\{\phi_i\}_{i=1}^m$ is a basis of V_h, then let $v = \phi_i$, hence

$$\sum_{j=1}^{m} \xi_j \int_a^b \phi_j' \phi_i' dx = \int_a^b f \phi_i dx, \quad i = 1, ..., m, \tag{1.5}$$

which is a quadratic system of m linear equation's and m unknowns. Use the notations

$$\begin{aligned} A &= (a_{ij}), \\ b &= (b_1, ..., b_m)^T, \\ \text{and } \xi &= (\xi_1, ..., \xi_m)^T, \text{ where} \end{aligned}$$

$$a_{ij} = \int_\alpha^\beta \phi_j' \phi_i' dx, \quad \text{Stiffness matrix,}$$

$$b_i = \int_\alpha^\beta f \phi_i dx, \quad \text{load vector,}$$

then (1.5) can be written in a matrix form as

$$A\xi = b.$$

Remarks:

1. $\int_\alpha^\beta \phi_j'\phi_i' dx = 0$ if $|i-j| > 1$, i.e., A is tridiagonal matrix.

2. $a_{ij} = \int_\alpha^\beta \phi_j'\phi_i' dx = \int_\alpha^\beta \phi_i'\phi_j' dx = a_{ji}$, hence A is symmetric.

Definition 11. *(Properties of bilinear forms), [27],*
Let $(V, ||.||_V)$ be a Banach space, $a(.,.)$ be a bilinear form on $V \times V$ to $\mathbf{R}$, and $F(.)$ be a linear form on V then

1. *$a(.,.)$ is positive if $a(v,v) \geq 0,\ \forall\ v \in V$.*

2. *$a(.,.)$ is symmetric if $a(u,v) = a(v,u),\ \forall\ u,v \in V$.*

3. *$a(.,.)$ is continuous if $\exists\ \gamma > 0$ such that $|a(u,v)| \leq \gamma||u||_V||v||_V,\ \forall\ u,v \in V$.*

4. *$a(.,.)$ is coercive (V-elliptic) if $\exists\ \alpha > 0$ such that $a(v,v) \geq \alpha||v||_V^2,\ \forall\ v \in V$.*

5. *$F(.)$ is continuous if $\exists\ \Gamma > 0$ such that $|F(v)| \leq \Gamma||v||_V,\ \forall\ v \in V$.*

Theorem 1. *The stiffness matrix A is positive definite.*
proof.
Write $v = (v_1, ..., v_n)$, then since $a(.,.)$ is a bilinear form, it follows that

$$\begin{aligned} v^T Av &= \sum_{i=1}^{n}\sum_{j=1}^{n} v_i a_{ij} v_j = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i a(\phi_i, \phi_j) v_j \\ &= a\Big(\sum_{i=1}^{n} v_i\phi_i, \sum_{j=1}^{n} v_j\phi_j\Big) = a(v,v) \geq \alpha||v||_V > 0, \end{aligned}$$

for any non zero v, i.e., the positive definiteness of the coefficient matrix comes form the V-ellipticity.

Using the linear basis functions defined before, for $j = 1, ..., m$, we have

$$(\phi_j', \phi_j') = \int_{x_{j-1}}^{x_j} (\phi_j')^2 dx + \int_{x_j}^{x_{j+1}} (\phi_j')^2 dx = \frac{1}{h_j} + \frac{1}{h_{j+1}}.$$

and

$$(\phi_j', \phi_{j-1}') = \int_{x_{j-1}}^{x_j} \frac{-1}{h_j^2} dx = \frac{-1}{h_j}.$$

and

$$(\phi_j', \phi_{j+1}') = \int_{x_j}^{x_{j+1}} \frac{-1}{h_j^2} dx = \frac{-1}{h_{j+1}}.$$

Also, in the case of uniform partition, $h_j = h = \frac{b-a}{m+1}$, then $A\xi = b$ becomes

$$\frac{1}{h} \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Example: Let u be the solution to

$$\begin{aligned} -u''(x) &= 1, \quad 0 < x < 1, \\ u(0) &= u(1) = 0, \end{aligned} \tag{1.6}$$

and let $I = (0, 1)$ be divided into a uniform mesh with $h = \frac{1}{m}$, calculate the finite element approximation U for $m = 3$.
solution:

$$\text{Let} \quad U = \sum_{j=0}^{3} \xi_j \phi_j = \xi_0\phi_0 + \xi_1\phi_1 + \xi_2\phi_2 + \xi_3\phi_3 = \xi_1\phi_1 + \xi_2\phi_2,$$

since $U = 0$ at x_0 and x_3, i.e., $\xi_0 = \xi_3 = 0$. The finite element formulation is: multiply (1.6) by a test function v such that $v(0) = v(1) = 0$ and integrate over $(0\ ,\ 1)$, we get

$$-u'v|_0^1 + \int_0^1 u'v'dx = \int_0^1 1vdx$$

$$\Longleftrightarrow \int_0^1 u'v'dx = \int_0^1 vdx.$$

Now, let $U = \xi_1\phi_1 + \xi_2\phi_2$, and $v = \phi_i$, i = 1, 2, then

$$\xi_1 \int_0^1 \phi_1'\phi_1' dx + \xi_2 \int_0^1 \phi_2'\phi_1' dx = \int_0^1 \phi_1 dx,$$

and

$$\xi_1 \int_0^1 \phi_1'\phi_2' dx + \xi_2 \int_0^1 \phi_2'\phi_2' dx = \int_0^1 \phi_2 dx.$$

In a matrix form, this is equivalent to

$$\frac{1}{h}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Note that $\frac{1}{h} = 3$, and $b_1 = b_2 = \frac{1}{3}$ since

$$b_1 = \int_0^1 \phi_1 dx = \frac{1}{2}.\frac{2}{3}.1 = \frac{1}{3},$$

$$b_2 = \int_0^1 \phi_2 dx = \frac{1}{2}.\frac{2}{3}.1 = \frac{1}{3},$$

which is the area of triangles. Thus, we get

$$\begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$$

$$\begin{aligned} 6\xi_1 - 3\xi_2 &= \frac{1}{3}. \\ -3\xi_1 + 6\xi_2 &= \frac{1}{3}. \end{aligned}$$

Solving this system of equations gives

$$\xi_1 = \xi_2 = \frac{1}{9}.$$

Hence,

$$U = \frac{1}{9}\phi_1(x) + \frac{1}{9}\phi_2(x).$$

To study the variational formulation for 2-D problems. Firstly we set some definitions and theorems.

Theorem 2. ***The Divergence Theorem***
Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open domain with smooth boundary $\partial\Omega$, and let $A : \bar{\Omega} \to \mathbb{R}^n$, $A = (A_1, A_2, ..., A_n)^T$ be a smooth vector-valued function, then we have

$$\int_\Omega \nabla.A \; dx = \int_{\partial\Omega} A.n \; ds \;,$$

where $\nabla.A(x) = div\ A = \frac{\partial A_1(x)}{\partial x_1} + . \; . \; . + \frac{\partial A_n(x)}{\partial x_n}$, $n(x) = (n_1(x), ..., n_n(x))^T$ is the outward unit normal to $\partial\Omega$, and ds is a curve element on $\partial\Omega$.

Definition 12. *Let v: $\bar{\Omega} \to \mathbb{R}$ be a smooth function, then*

1. *The **gradient** of v*
$$\nabla v = \left(\frac{\partial v}{\partial x_1}, \frac{\partial v}{\partial x_2}\;, . \; . \; . \;, \frac{\partial v}{\partial x_n}\right)^T.$$

2. *The **Laplacian** of v*
$$\Delta v = \frac{\partial^2 v}{\partial x_1^2} + \frac{\partial^2 v}{\partial x_2^2} + . \; . \; . + \frac{\partial^2 v}{\partial x_n^2}.$$

3. *The **normal derivative** of v*
$$\frac{\partial v}{\partial n} = \nabla v.n = \frac{\partial v}{\partial x_1} n_1 + \frac{\partial v}{\partial x_2} n_2 + . \; . \; . + \frac{\partial v}{\partial x_n} n_n, \text{ where } n \text{ is the outward unit normal.}$$

The Green's formula: Let $u, v \in H^2(\Omega)$ where Ω is a bounded domain, then

$$\int_\Omega \Delta u v \; dx = \int_{\partial\Omega} (\nabla u.n) v \; ds - \int_\Omega \nabla u.\nabla v \; dx.$$

Note that this formula is a simple consequence of the divergence theorem.

1.4.2 Two-dimensional problem

Consider the 2-D boundary value problem

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

Construct a variational formulation by multiplying both sides of the differential equation by a test function $v \in V$, where $V = \{v : v$ is continuous on $\bar{\Omega}$, $\frac{\partial v}{\partial x_1}, \frac{\partial v}{\partial x_2}$ are bounded and piecewise continuous on $\bar{\Omega}$, and $v = 0$ on $\partial\Omega$ $\}$. Now integrate over Ω,

$$\int_\Omega -\Delta uv \ dx = \int_\Omega fv \ dx, \quad \forall v \in V.$$

By the Green's theorem, we have

$$-\int_{\partial\Omega} (\nabla u.n)v \ ds + \int_\Omega \nabla u.\nabla v \ dx. = \int_\Omega fv \ dx.$$

But $v = 0$ on $\partial\Omega$, then

$$\int_\Omega \nabla u.\nabla v \ dx = \int_\Omega fv \ dx, \tag{1.7}$$

which called the weak formulation.

Discretization:

Construct a finite-dimensional subspace V_h of V; define $T_h = k_1, ..., k_m$ where k_i are non-overlapping triangles such that $\bar{\Omega} = \bigcup_{k_i \in T_h} k_i$, and let $h = \max_{k_i \in T_h}(diam(k_i))$, where $diam(k_i)$ is the largest side of k_i, so

$$V_h = \{U \in V : U \text{ is continuous on } \Omega \ ,U|_{k_i} \text{ is linear for } k_i \in T_h \ ,U = 0 \text{ on } \partial\Omega\}.$$

Find $U \in V_h$ such that

$$\int_\Omega \nabla U.\nabla v \ dx = \int_\Omega fv \ dx, \quad \forall v \in V_h. \tag{1.8}$$

Denote the internal nodes in the triangulation by $N_1, N_2, ..., N_m$, for $j = 1, 2, ..., m$ we define the linear function ϕ_j such that

$$\phi_j(N_i) = \delta_{ij} = \begin{cases} 1 & , i = j, \\ 0 & , i \neq j, \end{cases}$$

then $\{\phi_j\}_{j=1}^m$ is a basis of the finite-dimensional space V_h.
Since $U \in V_h$, then U can be written as a unique linear combination of ϕ_j's

$$U(x) = \sum_{j=1}^m \xi_j \phi_j(x),$$

where $U = 0$ on $\partial\Omega$. Now let $U(x) = \sum_{j=1}^{m} \xi_j \phi_j(x)$, and substitute in (1.8), then we get

$$\sum_{j=1}^{m} \xi_j \int_{\Omega} \nabla\phi_j.\nabla v dx = \int_{\Omega} f v dx, \quad \forall v \in V_h.$$

Since $\{\phi_j\}_{j=1}^{m}$ is a basis of V_h, assume $v = \phi_i$,

$$\sum_{j=1}^{m} \xi_j \int_{\Omega} \nabla\phi_j.\nabla\phi_i dx = \int_{\Omega} f\phi_i dx, \quad i = 1, ..., m, \tag{1.9}$$

which is a system of m linear equations and m unknowns. Let $A = (a_{ij})$,
$b = (b_1, ..., b_m)^T$, and $\xi = (\xi_1, ..., \xi_m)^T$, where

$$a_{ij} = \int_{\alpha}^{\beta} \nabla\phi_i.\nabla\phi_j dx, \quad \text{Stiffness matrix},$$

$$b_i = \int_{\alpha}^{\beta} f\phi_i dx, \quad \text{load vector},$$

then (1.9) is written in a matrix form as

$$A\xi = b.$$

Remark:

$$a_{ij} = \int_{\Omega} \nabla\phi_i.\nabla\phi_j dx = 0 \quad \text{unless } N_i \text{ and } N_j \text{ are nodes of the same triangle.}$$

1.5 Existence and uniqueness theorems

We discuss the general theorems for existence and uniqueness in Hilbert space and set the conditions that spaces and bilinear forms should satisfy. The existence and uniqueness of a solution to the weak formulation of the problem can be proved using the Lax-Milgram Theorem which states that the weak formulation has a unique solution.

Theorem 3. *(Lax-Milgram theorem)*
Let $a(.,.)$ be a bilinear form on $V \times V$, where V is a Hilbert space. Assume that $a(.,.)$ is continuous and coercive. Then for any continuous linear form $F(.)$ on V, there exists a unique $u \in V$ such that

$$a(u, v) = F(v), \quad \forall v \in V. \tag{1.10}$$

see, e.g., [12], for the proof.

The Galerkin method

The standard finite element method, which just replaces in the variational formulation (1.3) the space V by $V_h \subset V$, is called Galerkin method: Find $u_h \subset V_h$ such that $\forall$ $v_h \in V_h$

$$a(u_h, v_h) = F(v_h).$$

Chapter 2

FEM for general differential equations

In this chapter, we consider linear elliptic problems that are commonly found in mechanical and physical partial differential models. The aim is to introduce the notation of weak formulation that gives access to existence and uniqueness results for the solutions and that is well suited for the numerical approximation of such problems.

2.1 One dimensional problem

2.1.1 Variational formulation and discretization

Consider the two-point boundary value problem

$$\begin{aligned} -au''(x) + bu'(x) + cu(x) &= f(x), \quad \alpha < x < \beta, \\ au'(\alpha) &= \gamma(\alpha)[u(\alpha) - g_D(\alpha)] + g_N(\alpha), \\ -au'(\beta) &= \gamma(\beta)[u(\beta) - g_D(\beta)] + g_N(\beta), \end{aligned} \tag{2.1}$$

where $u(x)$, denoting the concentration of the substance, is unknown function that we wish to compute. The following are data to the problem:
$a(x)$ is diffusion coefficient, $b(x)$ is convection coefficient, $c(x)$ is rate coefficient, and $f(x)$ is source function. And

- $\gamma(\alpha), \gamma(\beta)$: permeability at the end-points.

- $g_D(\alpha), g_D(\beta)$: ambient concentration.
- $g_N(\alpha), g_N(\beta)$: externally induced flux through the boundary.

To derive the variational formulation, we multiply the differential equation by $v \in H^1$, and integrate over $[\alpha, \beta]$, we get

$$-\int_\alpha^\beta au''vdx + \int_\alpha^\beta bu'vdx + \int_\alpha^\beta cuvdx = \int_\alpha^\beta fvdx.$$

Integrating by parts gives

$$-au'v\Big|_\alpha^\beta + \int_\alpha^\beta au'v'dx + \int_\alpha^\beta bu'vdx + \int_\alpha^\beta cuvdx = \int_\alpha^\beta fvdx,$$

which is equivalent to

$$-au'(\beta)v(\beta) + au'(\alpha)v(\alpha) + \int_\alpha^\beta au'v'dx + \int_\alpha^\beta bu'vdx + \int_\alpha^\beta cuvdx = \int_\alpha^\beta fvdx.$$

Use the boundary condition in (2.1),

$$au'(\alpha) = \gamma(\alpha)[u(\alpha) - g_D(\alpha)] + g_N(\alpha),$$

$$-au'(\beta) = \gamma(\beta)[u(\beta) - g_D(\beta)] + g_N(\beta),$$

we obtain,

$$\Big(\gamma(\beta)\big(u(\beta) - g_D(\beta)\big) + g_N(\beta)\Big)v(\beta) + \Big(\gamma(\alpha)\big(u(\alpha) - g_D(\alpha)\big) + g_N(\alpha)\Big)v(\alpha) +$$

$$+ \int_\alpha^\beta au'v'dx + \int_\alpha^\beta bu'vdx + \int_\alpha^\beta cuvdx = \int_\alpha^\beta fvdx\ .$$

Rearrange the terms in the previous equation to obtain,

$$\gamma(\beta)u(\beta)v(\beta) + \gamma(\alpha)u(\alpha)v(\alpha) + \int_\alpha^\beta au'v'dx + \int_\alpha^\beta bu'vdx + \int_\alpha^\beta cuvdx = \tag{2.2}$$

$$\big(\gamma(\beta)g_D(\beta) - g_N(\beta)\big)v(\beta) + \big(\gamma(\alpha)g_D(\alpha) - g_N(\alpha)\big)v(\alpha) + \int_\alpha^\beta fvdx\ .$$

Thus, the variational formulation of the boundary value problem (2.1) is to find $u \in H^1$ such that (2.2) holds. For more details see e.g., [5, 14, 19, 20].

Discretization

Introducing the vector space V_h of continuous piecewise linear functions on a partition $\alpha = x_1 < x_2 < ... < x_{m-1} < x_m = \beta$ of $[\alpha, \beta]$. Find $U(x) \in V_h$ such that

$$\gamma(\beta)U(\beta)v(\beta) + \gamma(\alpha)U(\alpha)v(\alpha) + \int_\alpha^\beta aU'v'dx + \int_\alpha^\beta bU'vdx + \int_\alpha^\beta cUvdx = \tag{2.3}$$

$$\big(\gamma(\beta)g_D(\beta) - g_N(\beta)\big)v(\beta) + \big(\gamma(\alpha)g_D(\alpha) - g_N(\alpha)\big)v(\alpha) + \int_\alpha^\beta fvdx.$$

We construct a set of basis functions which called the hat functions $\{\phi_i\}_{i=1}^m \subset V_h$ such that

$$\phi_i(x_j) = \delta_{ij} = \begin{cases} 1 & , i = j, \\ 0 & , i \neq j, \end{cases}$$

for $i, j = 1, ..., m$.

Now, $\forall\ U \in V_h$, $U(x)$ can be written as a unique linear combination of ϕ_i's. i.e., $U(x) = \sum_{j=1}^m \xi_j \phi_j(x)$. To construct the discrete system of linear equations, we substitute $U(x)$ into (2.3), and take $\alpha = x_1$, $\beta = x_m$, then we get

$$\gamma(x_m)\xi_m v(x_m) + \gamma(x_1)\xi_1 v(x_1) + \sum_{j=1}^m \xi_j \left(\int_{x_1}^{x_m} a\phi_j' v' dx + \int_{x_1}^{x_m} b\phi_j' v dx + \int_{x_1}^{x_m} c\phi_j v dx \right) =$$

$$\Big(\gamma(x_m)g_D(x_m) - g_N(x_m)\Big)v(x_m) + \Big(\gamma(x_1)g_D(x_1) - g_N(x_1)\Big)v(x_1) + \int_{x_1}^{x_m} fvdx, \quad \forall\ v \in V_h.$$

Since $\{\phi_i\}_{i=1}^m \subset V_h$ is a basis of V_h, we can assume $v = \phi_i$, $i = 1, 2, ..., m$, to get

$$\gamma(x_m)\xi_m\phi_i(x_m) + \gamma(x_1)\xi_1\phi_i(x_1) + \sum_{j=1}^m \xi_j \left(\int_{x_1}^{x_m} a\phi_j'\phi_i' dx + \int_{x_1}^{x_m} b\phi_j'\phi_i dx + \int_{x_1}^{x_m} c\phi_j\phi_i dx \right) =$$

$$\Big(\gamma(x_m)g_D(x_m) - g_N(x_m)\Big)\phi_i(x_m) + \Big(\gamma(x_1)g_D(x_1) - g_N(x_1)\Big)\phi_i(x_1) + \int_{x_1}^{x_m} f\phi_i dx. \tag{2.4}$$

Which is a system of m linear equations and m unknowns. We will use the notations $A = (a_{ij})$, $C = (c_{ij})$, $M = (m_{ij})$, and $b = (b_1, ..., b_m)^T$, where

$$a_{ij} = \int_{x_1}^{x_m} a\phi_j'\phi_i' dx, \qquad \text{stiffness matrix,}$$

$$c_{ij} = \int_{x_1}^{x_m} b\phi_j'\phi_i dx, \qquad \text{convection matrix,}$$

$$m_{ij} = \int_{x_1}^{x_m} c\phi_j\phi_i dx, \qquad \text{mass matrix,}$$

$$b_i = \int_{x_1}^{x_m} f\phi_i dx, \qquad \text{load vector.}$$

Taking into account

$$\phi_i(x_1) = \begin{cases} 1 & , i = 1, \\ 0 & , i \neq 1, \end{cases} \qquad and \quad \phi_i(x_N) = \begin{cases} 1 & , i = N, \\ 0 & , i \neq N, \end{cases}$$

we can write the system of equations (2.4) as

$$\begin{pmatrix} (\gamma(x_1) + a_{11} + c_{11} + m_{11})\xi_1 + \ldots + (a_{1N} + c_{1N} + m_{1N})\xi_N = b_1 + \gamma(x_1)g_D(x_1) - g_N(x_1) \\ (a_{21} + c_{21} + m_{21})\xi_1 + \ldots + (a_{2N} + c_{2N} + m_{2N})\xi_N = b_2 \\ \vdots \\ (a_{N1} + c_{N1} + m_{N1})\xi_1 + \ldots + (a_{NN} + c_{NN} + m_{NN} + \gamma(x_N))\xi_N = b_N + \gamma(x_N)g_D(x_N) - g_N(x_N) \end{pmatrix}$$

In a matrix form, this is read as

$$\big(A + M + R + C\big)\xi = b + rv,$$

where

$$R = \begin{pmatrix} \gamma(x_1) & 0 & \ldots & 0 \\ 0 & \ldots & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & \ldots & \gamma(x_N) \end{pmatrix}$$, contains the boundary contributions to the system matrix.

$$rv = \begin{pmatrix} \gamma(x_1)g_D(x_1) - g_N(x_1) \\ 0 \\ \vdots \\ \gamma(x_N)g_D(x_N) - g_N(x_N) \end{pmatrix}$$, contains the boundary contributions to the right hand side.

2.1.2 Numerical Examples

Example 1. *Consider the problem*

$$\begin{cases} -u''(x) = 6x, \ 1 < x < 2, \\ \quad u(1) \ = \ u(2) \ = \ 0. \end{cases} \tag{2.5}$$

Let $I = (1,2)$ be divided into a uniform mesh $h = \frac{1}{m}$, below we calculate the finite element approximation U for $m = 5$, and $m = 10$.

Solution:

Firstly, find the variational formulation to the problem by multiplying (2.5) by a test function v, such that $v(1) = v(2) = 0$ and integrate over the interval $(1,2)$, we get

$$-u'v|_1^2 + \int_1^2 u'v'dx = \int_1^2 6xvdx$$

Find $U(x)$ such that

$$\int_1^2 U'v'dx = \int_1^2 6xvdx.$$

Using the linear basis functions in page (12), U, for $m = 5$, can be written as a unique linear combination of ϕ_j's as

$$U(x) = \sum_{j=0}^{5} \xi_j \phi_j, \ \text{where } \xi_j = U(x_j),$$

$$= \sum_{j=1}^{4} \xi_j \phi_j, \ \text{since } \xi_0 = \xi_5 = 0.$$

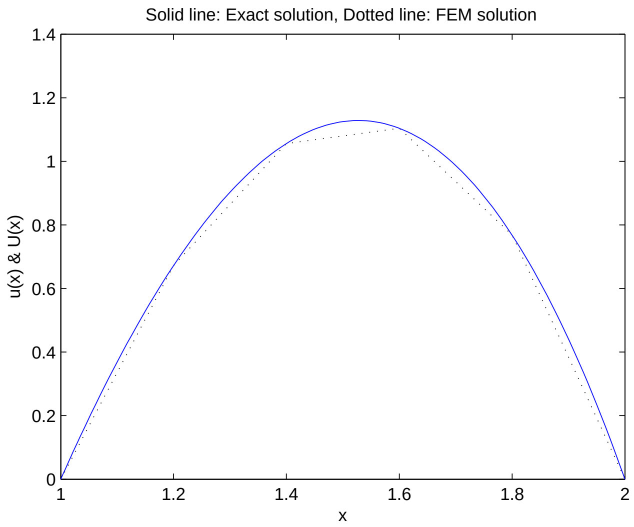
Thus

$$\sum_{j=1}^{4} \xi_j \int_1^2 \phi_j' v'dx \ = \ \int_1^2 6xvdx.$$

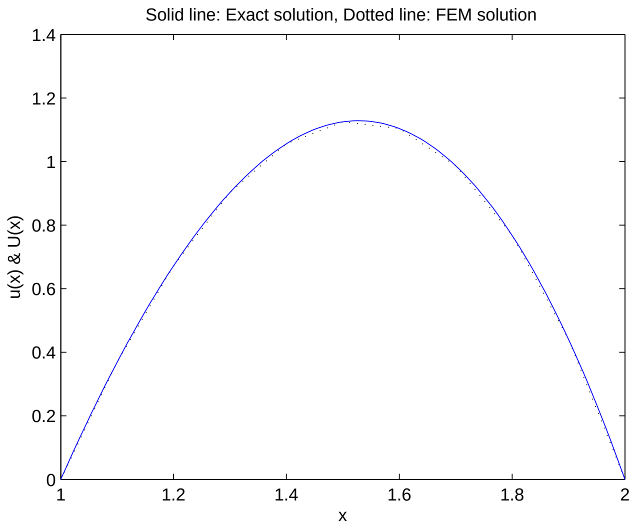
Since $\{\phi_i\}_{i=1}^m$ is a basis, then let $v = \phi_i$, so

$$\sum_{j=1}^{4} \xi_j \int_1^2 \phi_j' \phi_i' dx \ = \ \int_1^2 6x\phi_i dx, \ i = 1,2,3,4.$$

The following two figures are the approximations with $m = 5$ and $m = 10$ respectively, depicted with the exact solution $u(x) = -x^3 + 7x - 6$. The approximation is obtained using the Matlab software.



m=5



m=10

The figures show both the exact solution (the solid line) and the FEM solution (the dotted line). The first one is with $m = 5$, and the second is with $m = 10$.
Note that when m increases then the approximate solution becomes closed to the exact one.

Example 2. *Consider the problem*

$$\begin{cases} u''(x) - 5u'(x) + 6u(x) = 6x - 11, & 0 < x < 1, \\ u'(0) \;=\; u'(1) \;=\; 2. \end{cases} \tag{2.6}$$

Let $I = (0,1)$ be divided into a uniform mesh with $h = \frac{1}{m}$, below we calculate the finite element approximation U for $m = 10$, where the exact solution for this problem is $u(x) = ((e^3 - 1)/(2(e^3 - e^2)))e^{2x} + ((1 - e^2)/(3(e^3 - e^2)))e^{3x} + x - 1.$

Solution:

Firstly, find the variational formulation by multiplying (2.6) by a test function v, and integrate over $(0,1)$, we get

$$\int_0^1 u''v\,dx - \int_0^1 5u'v\,dx + \int_0^1 6uv\,dx = \int_0^1 (6x - 11)v\,dx$$

$$\Longleftrightarrow\; u'v|_0^1 - \int_0^1 u'v'\,dx - \int_0^1 5u'v\,dx + \int_0^1 6uv\,dx = \int_0^1 (6x - 11)v\,dx$$

$$\Longleftrightarrow\; u'(1)v(1) - u'(0)v(0) - \int_0^1 u'v'\,dx - \int_0^1 5u'v\,dx + \int_0^1 6uv\,dx = \int_0^1 (6x - 11)v\,dx$$

$$\Longleftrightarrow\; -\int_0^1 u'v'\,dx - \int_0^1 5u'v\,dx + \int_0^1 6uv\,dx = \int_0^1 (6x - 11)v\,dx - 2(v(1) - v(0)).$$

Find $U(x)$ such that

$$-\int_0^1 U'v'\,dx - \int_0^1 5U'v\,dx + \int_0^1 6Uv\,dx = \int_0^1 (6x - 11)v\,dx - 2(v(1) - v(0)).$$

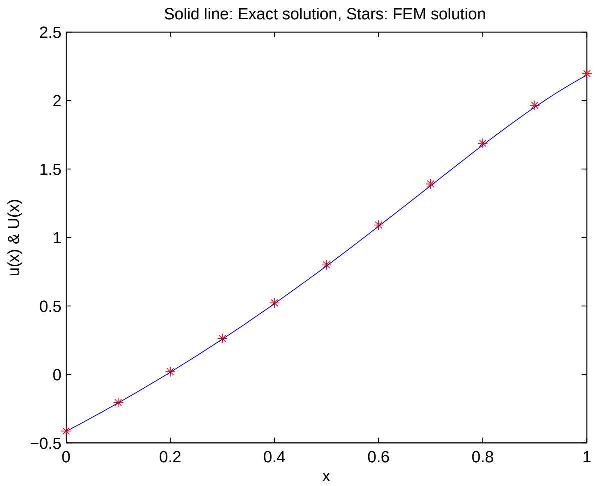
Now, using the basis functions such that

$$U = \sum_{j=0}^{10} \xi_j \phi_j, \text{ and let } v = \phi_i,$$

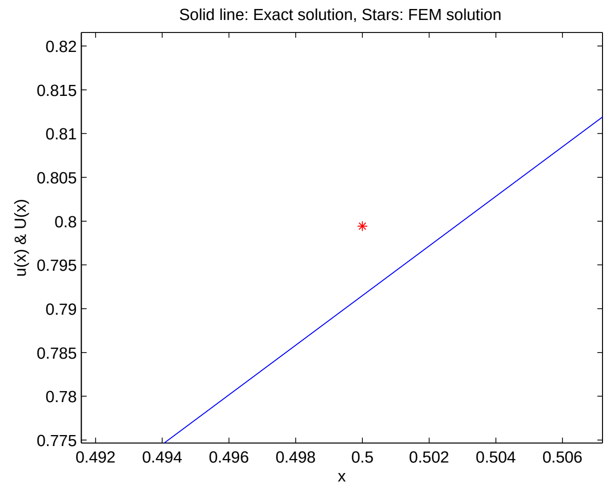
so,

$$-\sum_{j=0}^{10} \xi_j \int_0^1 \phi_j'\phi_i'\,dx - \sum_{j=0}^{10} \xi_j \int_0^1 5\phi_j'\phi_i\,dx + \sum_{j=0}^{10} \xi_j \int_0^1 6\phi_j\phi_i\,dx = \int_0^1 (6x - 11)\phi_i\,dx - 2(\phi_i(1) - \phi_i(0)),$$

where $i = 0, 1, ..., 10$. Now, using the Matlab software we get the below finite element approximation.



The exact solution (the solid line) and the FEM solution (the stars) are shown in the figure above.
From the figure, it seems that the approximation is the same as the exact solution at the nodal points, but when we zoom in, the error becomes clear, as it noted from the figure below.



For examples 1 and 2, the Matlab codes are appended to the thesis, see the appendix.

2.2 Two dimensional problem

2.2.1 Variational formulation and Discretization

Consider the boundary value problem

$$-\nabla.(a\nabla u) + b.\nabla u + cu = f, \quad x \in \Omega,$$

$$-n.(a\nabla u) = \gamma(u - g_D) + g_N, \quad x \in \partial\Omega,$$

where γ is permeability of the boundary, g_D ambient concentration, and g_N is externally induced flux through the boundary. To drive the variational formulation, we multiply the differential equation by a test function $v \in V$ such that $V = \{v : v$ is continuous on $\bar{\Omega}$, $\frac{\partial v}{\partial x_1}, \frac{\partial v}{\partial x_2}$ are bounded and piecewise continuous on $\bar{\Omega}$ $\}$. Now, integrate over Ω, to get

$$\int_\Omega -\nabla.(a\nabla u)vdx + \int_\Omega b.\nabla uvdx + \int_\Omega cuvdx = \int_\Omega fvdx,$$

then by the Green's formula we have

$$-\int_{\partial\Omega}(n.(a\nabla u))vds + \int_\Omega a\nabla u.\nabla vdx + \int_\Omega b.\nabla uvdx + \int_\Omega cuvdx = \int_\Omega fvdx.$$

Use the boundary condition above

$$-n.(a\nabla u) = \gamma(u - g_D) + g_N, \quad x \in \partial\Omega,$$

to obtain

$$\int_{\partial\Omega}\gamma uvds + \int_\Omega a\nabla u.\nabla vdx + \int_\Omega b.\nabla uvdx + \int_\Omega cuvdx = \int_\Omega fvdx + \int_{\partial\Omega}(\gamma g_D - g_N)ds. \quad (2.7)$$

We thus state the following variational formulation; find $u \in V$ such that (2.7) holds $\forall$ $v \in V$.

Discretization

Introducing the vector space V_h of continuous piecewise linear functions on a triangulation of Ω. Find $U \in V_h$ such that

$$\int_{\partial\Omega} \gamma U v ds + \int_{\Omega} a\nabla U.\nabla v dx + \int_{\Omega} b.\nabla U v dx + \int_{\Omega} c U v dx = \tag{2.8}$$

$$= \int_{\Omega} f v dx + \int_{\partial\Omega} (\gamma g_D - g_N) ds, \quad \forall v \in V_h.$$

We construct a set of basis which called the tent functions $\{\phi_i\}_{i=1}^{m} \subset V_h$, where

$$\phi_i(N_j) = \delta_{ij} = \begin{cases} 1 & , i = j, \\ 0 & , i \neq j, \end{cases}$$

for $i, j = 1, ..., m$, N_j are the nodes in the generated triangulation, refers to the number of nodes in the given triangulation.
Now, $\forall\ U \in V_h$, $U(x)$ can be written as a unique linear combination of ϕ_i's, i.e., $U(x) = \sum_{j=1}^{m} \xi_j \phi_j(x)$, see [5, 12, 19, 27, 52].

For the construction of the discrete system of linear equations, we substitute $U(x) = \sum_{j=1}^{m} \xi_j \phi_j(x)$ into (2.8), to get

$$\sum_{j=1}^{m} \xi_j \Big(\int_{\partial\Omega} \gamma \phi_j v ds + \int_{\Omega} a\nabla\phi_j.\nabla v dx + \int_{\Omega} b.\nabla\phi_j v dx + \int_{\Omega} c\phi_j v dx \Big) =$$

$$= \int_{\partial\Omega} (\gamma g_D - g_N) v ds + \int_{\Omega} f v dx, \quad \forall v \in V_h.$$

Since $\{\phi_i\}_{i=1}^{m} \subset V_h$ is a basis of V_h, then we can choose $V = \phi_i, i = 1, 2, ..., m$, thus

$$\sum_{j=1}^{m} \xi_j \Big[\int_{\partial\Omega} \gamma \phi_j \phi_i ds + \int_{\Omega} a\nabla\phi_j.\nabla\phi_i dx + \int_{\Omega} b.\nabla\phi_j \phi_i dx + \int_{\Omega} c\phi_j \phi_i dx \Big] =$$

$$\int_{\partial\Omega} (\gamma g_D - g_N) \phi_i ds + \int_{\Omega} f \phi_i dx, \quad i = 1, 2, ...m.$$

Which is a system of m linear equations and m unknowns. To simplify we introduce the following notations,

$$
\begin{aligned}
a_{ij} &= \int_\Omega a\nabla\phi_j.\nabla\phi_i dx, \\
m_{ij} &= \int_\Omega c\phi_j\phi_i dx, \\
c_{ij} &= \int_\Omega b\nabla\phi_j.\phi_i dx, \\
r_{ij} &= \int_{\partial\Omega} \gamma\phi_j\phi_i ds, \\
rv_i &= \int_{\partial\Omega} (\gamma g_D - g_N)\phi_i ds, \\
b_i &= \int_\Omega f\phi_i dx.
\end{aligned}
\tag{2.9}
$$

Thus, we can write the system of equations as :

$$
\big(A + C + M + R\big)\xi = rv + b, \tag{2.10}
$$

where

$$
\begin{aligned}
A &= (a_{ij}), && \text{Stiffness matrix,} \\
M &= (m_{ij}), && \text{Mass matrix,} \\
C &= (c_{ij}), && \text{Convection matrix,} \\
b &= (b_i), && \text{Load vector,} \\
R &= (r_{ij}), && \text{containes the boundary contributions to the system matrix,} \\
rv &= (rv_i), && \text{containes the boundary contributions to the right hand side .}
\end{aligned}
$$

2.2.2 Numerical Examples

Example 3. *Consider the problem*

$$\begin{cases} -\Delta u = 2\pi \sin(\pi x)\sin(\pi y), & x, y \in \Omega, \\ \quad\; u = 0, & on\ \partial\Omega, \end{cases} \tag{2.11}$$

where $\Omega = [0,1] \times [0,1]$.
Solution:
To find the variational formulation of the problem multiply (2.11) by a test function v, and integrate over Ω, we get

$$\int_\Omega -\Delta u v dx = \int_\Omega f v dx,$$

$$\Longleftrightarrow \; -\int_{\partial\Omega} (\nabla u.n) v ds + \int_\Omega \nabla u.\nabla v dx = \int_\Omega f v dx,$$

since $u = 0$ on $\partial\Omega$, then

$$\int_\Omega \nabla u.\nabla v dx = \int_\Omega f v dx.$$

Find U such that

$$\int_\Omega \nabla U.\nabla v dx = \int_\Omega f v dx,$$

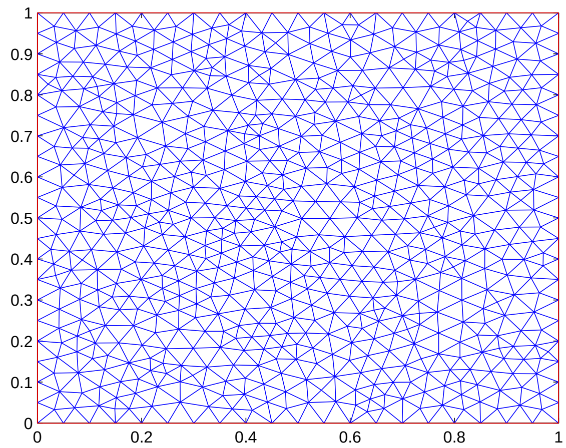
Use the basis function in page 29, and let

$$U = \sum_{j=1}^{n} \xi_j \phi_j, \;\; v = \phi_i, \;\; i = 1, 2, ..., n,$$

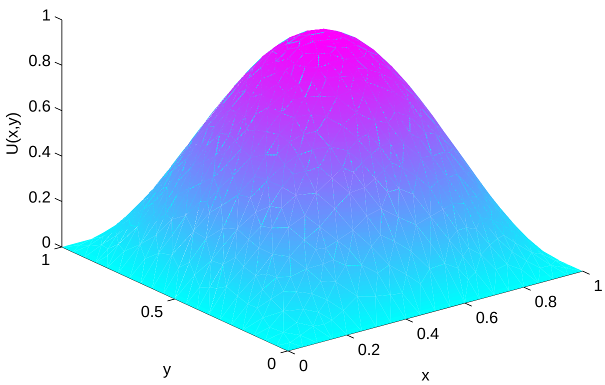
where n is the number of the internal nodes. Then,

$$\sum_{j=1}^{n} \xi_j \int_\Omega \nabla \phi_j.\nabla \phi_i dx = \int_\Omega f \phi_i dx.$$

Now, using the Matlab software we get the following two figures



The triangulation of randomly generated mesh for Ω.



The approximation.

The first figure is the mesh plot (triangulation) of the region of the example. The second figure is the solution of the given poisson equations on $\Omega = [0, 1] \times [0, 1]$.

Chapter 3

Error estimation

3.1 Introduction

Errors are considered so important in numerical analysis. There are many kinds of errors including roundoff error, truncation error, error in data, and uncertainty in the model. One of the most important issue concerning errors is to be able to control the error which will be so useful in solving problems. Estimating errors is so helpful as it can help us in evaluating the solution or the model itself.

One of the main and essential factor in computational sciences, is the mathematical theory of estimating discretization error. It was noticed that many of computational results that used the mathematical models include numerical errors. This actually can help us in assessing the reliability of the computation of the numerical process. The use of measures of error to control time steps in the numerical solution of ordinary differential equations probably represents the first use of a posteriori estimates to control discretization error in numerical solutions of initial- or boundary-value problems.

In fact, the error in the numerical solution is defined as the difference between the exact and approximation solutions. The purpose of error estimation is to avoid inaccuracy in the numerical solution, including the errors that come from inaccurate discretization of the solution domain and discretization errors. Also to bound the discretization error $e = u - u_h$ in a Sobolev space or Lebesgue norm, where u is the exact solution to the variational problem

$$a(u,v) = f(v), \quad \forall v \in V, \tag{3.1}$$

and u_h is the approximation solution to the variational problem

$$a(u_h, v_h) = f(v_h), \quad \forall v_h \in V_h.$$

The error estimate is the difference between approximate solution u_h and the exact solution u, i.e., we want the approximate solution converges to the exact solution as the discretization parameter goes to zero. Actually, estimating the error will be easier to understand if its dimensions are exactly similar to the solution variable. In general, this process has not been accepted due to its impracticality for singular problems. The kinds of the approximation solution rely on both the discretization parameters and the choice of the exact element space.

Here, it is of importance to shed the light on the difference between two concepts: The error estimates and the error bounds. The error estimate express an amount that is nearly almost the real unknown error. In contrast, upper and lower error bounds are amount that are smaller than the actual unknown error. Thus, error bounds are actually inevitable but still considered as incorrect, on the other hand, error estimate seem to be more accurate even though they include the original true error. Hence, error bounds can be guaranteed but still be inaccurate, it over or underestimate the true error, whereas an error estimate should be accurate, in general.

Error estimate typically proceeds in two steps, see [20]:

(i) Showing that u_h is optimal in the sense that the error $u - u_h$ satisfies

$$||u - u_h|| = \min_{v \in V_h} ||u - v|| \tag{3.2}$$

in an appropriate norm, and

(ii) finding an upper bound for the right-hand side of (3.2).

The appropriate norm to use with (3.2) for the model problem (3.1) is the strain energy norm

$$||v||_E = \sqrt{a(v,v)}.$$

The finite element solution might not satisfy (3.2) with other norms. For example, finite element solutions are not optimal in any norm for non-self-adjoint problems, [20]. In these cases, (3.2) is replaced by the weaker statement

$$||u - u_h|| \leq C \min_{v \in V_h} ||u - v||, \quad \text{where } C > 1.$$

Thus, the solution is closed to the best solution but it only differs by a constant from the best possible solution in the space.

Error estimators that are based directly on the finite element approximation and the data of the problem are usually referred to as explicit error estimators which involve a direct computation of the interior element residuals and the jumps at the element boundaries to find an estimate for the error in the energy norm. In contrast, implicit error estimators require the solution of auxiliary local boundary value problems and involve the solution of the auxiliary boundary value problems whose solution yields an approximation to the actual error, [23]. Hence, explicit error estimators in general require less computational effort than implicit schemes. A third class of error estimators is the recovery-based error estimators which make use of the fact that the gradient of the finite element solution is in general discontinuous across the interelement boundaries.

Error estimates for FEM for poisson equation

Consider the problem

$$\begin{cases} -\Delta u = f, & \text{in } \Omega, \\ \quad u \ = 0, & \text{on } \partial\Omega, \end{cases} \tag{3.3}$$

where $\Omega \subset \mathbf{R}^d$, $d = 1, 2, 3$. By the Green's theorem,

$$\int_\Omega \nabla u . \nabla v \; dx = \int_\Omega f v \; dx, \quad \forall v \in V_h, \tag{3.4}$$

with variational formulation: Find $U \in V_h$ such that

$$\int_\Omega \nabla U . \nabla v \; dx = \int_\Omega f v \; dx, \quad \forall v \in V_h. \tag{3.5}$$

For the error $e = u - U$ we have

$$\nabla e = \nabla u - \nabla U = \nabla(u - U).$$

Subtraction of (3.5) from the (3.4) yields the Galarkin Orthogonality

$$\int_\Omega (\nabla u - \nabla U).\nabla v \; dx = \int_\Omega \nabla e.\nabla v \; dx = 0, \forall v \in V_h. \tag{3.6}$$

On the other hand we may write

$$||\nabla e||^2 = \int_\Omega \nabla e.\nabla e \; dx = \int_\Omega \nabla e.\nabla u \; dx - \int_\Omega \nabla e.\nabla U \; dx,$$

now, using the Galarkin Orthogonality (3.6), and since $U \in V_h$, we have

$$\int_\Omega \nabla e.\nabla U \; dx = 0.$$

Thus, inserting $\int_\Omega \nabla e.\nabla v \; dx = 0$, $\forall v \in V_h$, to get

$$\begin{aligned}
||\nabla e||^2 &= \int_\Omega \nabla e.\nabla e \; dx \\
&= \int_\Omega \nabla e.\nabla (u - U) \; dx \\
&= \int_\Omega \nabla e.(\nabla u - \nabla U) \; dx \\
&= \int_\Omega \nabla e.\nabla u \; dx - \int_\Omega \nabla e.\nabla U \; dx \\
&= \int_\Omega \nabla e.\nabla u \; dx \\
&= \int_\Omega \nabla e.\nabla u \; dx - \int_\Omega \nabla e.\nabla v \; dx \\
&= \int_\Omega \nabla e.\nabla (u - v) \; dx
\end{aligned}$$

$$\leq ||\nabla e|| \; ||\nabla (u - v)||.$$

Hence,

$$||\nabla (u - U)|| \leq ||\nabla (u - v)||, \quad \forall v \in V_h. \tag{3.7}$$

This means that the finite element solution $U \in V_h$ is the best approximation of the solution u among functions in V_h, i.e., U is closer to u than any other $v \in V_h$.

Theorem 4. *Let $a = x_0 < x_1 < ... < x_n = b$ be a partition of $[a, b]$ and h be the step size and let $\pi_h v(x)$ be the piecewise linear interpolant of $v(x)$. Then there is an interpolant constant c_i such that, for $1 \leq p \leq \infty$, [19],*

$$||\pi_h v - v||_{L_p} \leq c_i ||h^2 v''||_{L_p},$$

$$||(\pi_h v)' - v'||_{L_p} \leq c_i ||hv''||_{L_p},$$

$$||\pi_h v - v||_{L_p} \leq c_i ||hv'||_{L_p}.$$

In this chapter, we shall focus on two types of error estimates for the finite element method, a priori and a posteriori estimates. A priori error estimates are error bounds that use information about the unknown solution u to estimate the error before we compute the approximate solution u_h. They tell us about the order of convergence of a given finite element method, that is, they tell us that the finite element error $||u - u_h||$ in some norm $||.||$ is $O(h^\alpha)$, where h is the maximum mesh size and α is a positive integer. Additionally, the a priori error estimates supply information on convergence rates but are difficult to use for quantitative error information. A posteriori error estimates, which use the computed solution, provide more practical accuracy appraisal, [20]. In contrast, a posteriori estimates use the computed solution u_h in order to give us an estimate of the form $||u - u_h|| \leq \epsilon$, where ϵ is a small number.

The main difference between a priori and a posteriori estimates is that a priori error is error bounds given by known information on the solution of the variational problem and the finite element function space. It gives us a reasonable measure of the efficiency of a given method by telling us how fast the error decreases as we decrease the mesh size. But a posteriori estimates are error bounds given by information on the numerical solution obtained on the finite element function space. The a posteriori estimate provides a much better idea of the actual error in a given finite element computation than a priori estimates and they can be used to perform adaptive mesh refinement.

3.2 A priori error estimates

A priori estimate (also called a priori bound) is a Latin expression which means from before and refers to the fact that the estimate for the solution is derived before the solution is known to exist. The a priori estimation of errors in numerical methods has long been a project of numerical analysis. Such estimates give information on the convergence and stability of various solvers.

A priori error estimators provide information on the asymptotic behavior of the discretization errors but are not designed to give an actual error estimate for a given mesh, and they play an important role in the proof of existence of the solution. These estimates also give us an excellent tool for dealing with a very practical problem.

There exist different methods which give a priori estimates of solutions of elliptic problems [31]: The first method called blow-up was first introduced by B. Gidas and J. Spruck in [22]. Another method is the method of Rellich-Pohozaev identities and moving planes which introduced by D. G. de Figueiredo, P.-L. Lions and R. D. Nussbaum [37]. Moreover, we have the method of Hardy-Sobolev inequalities by H. Brezis and R. E. L. Turner [13]. Finally, the bootstrap procedure by P. Quittner and Ph. Souplet [43].

Our goal is to find bounds for the error $u - u_h$ in the finite element approximation of the solution u to our general boundary value problem. The most important property of any conforming finite element formulation based on a symmetric bilinear form is the optimality condition, see [23],

$$||u - u_h||_E = \min ||u - v_h||_E, \quad \forall v_h \in V_h, \tag{3.8}$$

which states that u_h is the best approximation in the finite element space V_h, i.e., u_h is closer to the exact solution. In Cea's lemma, choosing $V \subset H^1$ and employing interpolation estimates, it turns out that the error measured in the H^1-norm is $O(h^p)$, see [23], that is

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

where c is a stability and interpolation constant which does not depend on the actual interpolation space, and h denotes the maximum of all element sizes.
Furthermore, we have for the error in the L^2-norm

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

which means that the convergence rate for the solution itself is $O(h^{p+1})$, [11].

These types of estimates, cf [31], date back to 1974 when R.E.L. Turner [54] studied problem (3.3), he proved that if a continuous function f defined on $\Omega \times [0, \infty)$ satisfies

$$C_1 u^p \leq f \leq C_2(1 + u^p), \quad p < 3,$$

for some constants $C_1, C_2 > 0$, then any nonnegative classical solution of (3.3) satisfies the a priori bound

$$||u||_\infty \leq C. \tag{3.9}$$

One year later, R. Nussbaum [37] proved that any positive classical solution of (3.3) with f satisfying

$$|f| \leq C(1 + |u|^p)$$

satisfies the a priori bound (3.9).
In 1981, B. Gidas and J. Spruck [22] derived an a priori estimate for positive solutions of (3.3) in the optimal range of exponents. In 2004, P. Quittner and Ph. Souplet [43] showed that any very weak solution of (3.3), is a classical solution if f is smooth enough.

In 2007, M. del Pino, M. Musso and F. Pacard [18] constructed positive very weak solutions of (3.3), which vanish in the sense of traces on $\partial\Omega$, but which are singular at prescribed points of $\partial\Omega$.

Theorem 5. *(A Priori Error Estimate), [12],*
Let $V_h \subset V$, and assume the conditions of the Lax-Milgram theorem. Then there is a unique solution of the problem to find $u_h \in V_h$ such that

$$a(u_h, v_h) = F(v_h), \quad \forall v_h \in V_h, \tag{3.10}$$

and it holds the error estimate

$$||u - u_h||_V \leq \frac{\Gamma}{\alpha} \inf_{v_h \in V_h} ||u - v_h||_V,$$

where u is the unique solution of the continuous problem (1.10) and the constants are defined in the definition (11).

proof.
The existence and uniqueness of u and u_h follows directly from the Lax-Milgram theorem, since the subspace of a Hilbert space is also a Hilbert space and the properties of the bilinear form carry over from V to V_h.

Computing the difference of the continuous equation (1.10) and the discrete equation (3.10) yields

$$a(u-u_h,v_h)=0,\quad \forall v_h\in V_h,$$

with $\alpha||v||_V^2\le a(v,v)$ and $|a(u,v)|\le\Gamma||u||_V||v||_V$. It follows for all $v_h\in V_h$ that

$$||u-u_h||_V^2\le\frac{1}{\alpha}a(u-u_h,u-u_h),$$

but since $a(u-u_h,v_h)=0$, so,

$$||u-u_h||_V^2\le\frac{1}{\alpha}a(u-u_h,u-v_h)$$

$$\le\frac{\Gamma}{\alpha}||u-u_h||_V||u-v_h||_V.\square$$

This inequality is equivalent to the statement of the theorem, which called Cea's lemma, [36]. By Cea's lemma, the discretization error is bounded by the best-approximation error. But the estimate in Cea's lemma is weaker than the corresponding estimate (3.8) for the model problem since the symmetry of the bilinear form allows characterizing solutions of (3.1) as minimizers of a functional.

A priori error estimate for poisson equation

Consider the problem

$$\begin{cases} -\Delta u=f, & \text{in } \Omega,\\ \quad u\ =0, & \text{on } \partial\Omega. \end{cases}\tag{3.11}$$

Theorem 6. *The finite element approximation U satisfies (3.7). In particular, there is a constant C_i such that*

$$||u-U||_E\le||\nabla(u-U)||\le C_i\ ||hD^2u||,$$

where C_i is an interpolation constant, and

$$D^2u=(u_{xx}^2+2u_{xy}^2+u_{yy}^2)^{1/2}.$$

Now, we will find a priori error estimate for the solution $e=u-U$. For a general mesh we have the following a priori error estimate for the solution of the Poisson equation (3.11).

Theorem 7.

$$||e|| \leq C^2 C_\Omega^2 h^2 ||f||.$$

proof.
Let ϕ be the solution of the dual problem

$$\begin{cases} -\Delta\phi = e, & \text{in } \Omega, \\ \quad \phi \ = 0, & \text{on } \partial\Omega. \end{cases}$$

Then,

$$\begin{aligned} ||e||^2 &= \int_\Omega e.e \ dx \\ &= \int_\Omega e(-\Delta\phi) \ dx \\ &= \int_\Omega \nabla e.\nabla\phi \ dx, \quad \text{by Green's formula} \\ &= \int_\Omega \nabla e.\nabla\phi \ dx - \int_\Omega \nabla e.\nabla v \ dx, \quad \text{by Galarkin Orthogonality} \\ &= \int_\Omega \nabla e.\nabla(\phi - v) \ dx. \end{aligned}$$

So,

$$||e||^2 \leq ||\nabla e|| \ ||\nabla(\phi - v)||, \quad \forall v \in V_h.$$

Let v be an interpolation of ϕ such that

$$||\nabla(\phi - v)|| \leq C||h \ D^2\phi||, \quad \text{by } (\mathbf{??}).$$

Hence,

$$\begin{aligned} ||e||^2 &\leq ||\nabla e||C||h \ D^2\phi|| \\ &\leq ||\nabla e|| \ C(\max_\Omega h) \ ||D^2\phi||. \end{aligned} \tag{3.12}$$

To complete the proof, we need the following lemma [5, 19].

Lemma 2. *(regularity lemma) Assume that Ω has no re-intrents. We have for $u \in H^2(\Omega)$; with $u = 0$ or $(\frac{\partial u}{\partial n} = 0)$ on $\partial\Omega$ that, ,*

$$||D^2 u|| \leq C_\Omega ||\Delta u||.$$

proof. *see [5].*

Now, applying this lemma to ϕ,

$$||D^2 \phi|| \leq C_\Omega . ||\Delta \phi|| = C_\Omega ||e||.$$

Then, (3.12) implies

$$||e||^2 \leq ||\nabla(u - U)|| C(\max_\Omega h) C_\Omega ||e||.$$

Thus, using Theorem 6 , the following a priori error estimate is obtained,

$$||e|| \leq C^2 C_\Omega (\max_\Omega h) ||hD^2 u||.$$

Which, using the lemma above, for a uniform (constant) h, can be written as an stability estimate, [19],

$$||e|| \leq C^2 C_\Omega^2 h^2 ||f||. \square$$

3.3 A posteriori error estimates

A posteriori estimates present a necessary tool in the adaptive procedures used in computer simulation and is known to be essential for reliable scientific computing. It used to control discretization error in numerical solutions of initial or boundary value problems.

3.3.1 A short history

The term a posteriori error estimator, was firstly used by Ostrowski [41] in 1940. To the authors knowledge, the first use of error estimates for adaptive meshing strategies in significant engineering problems was given in the work of Guerra [24] in 1977. The paper of Babuska and Rheinboldt [7] published in 1978 is often cited as the first work aimed at developing rigorous global error bounds for finite element approximations of linear elliptic two-point boundary value problems. In the period spanning over two decades since these works, significant advances have been made. A brief history of the subject is given in the book of Ainsworth and Oden [1, ?], see also the books and survey articles of Verfurth [56], Babuska and Strouboulis [9], Oden and Demkowicz [38] . It can be argued that until quite recently, the vast majority of the published work on a posteriori error estimation dealt with global estimates of errors in finite element approximations of linear elliptic problems, these estimates generally being in energy-type norms.

The aims of a posteriori error estimation is developing quantitative methods in which the error $e = u - u_h$ is estimated in post-processing procedures using the solution u_h as data for the error estimates. A posteriori error estimator is a quantity which bounds or approximates the error and can be computed from the knowledge of numerical solution and input data. The advantage of any a posteriori error estimator is to supply an estimate and ideally bounds for the solution error in a specified norm if the problem data and the finite element solution are available.

3.3.2 A posteriori error estimate for Poisson equation

To study a posteriori error analysis, where instead of the unknown value of $u(x)$, we use the known value of the approximate solution to estimate the error, [23]. This means that the error analysis performed after the computation is completed. We shall denote the error by $e(x)$, i.e., $e(x) = u(x) - U(x)$.

Theorem 8. *Let u be the solution of the Poisson equation (3.11) and U is the continuous piecewise linear finite element approximation. Then there is constant C, independent of u and h, such that*

$$||u - U|| \leq C||h^2 r||, \tag{3.13}$$

where $r = f + \Delta U$ is the residual.

proof.
Consider the following dual problem

$$\begin{cases} -\Delta\phi(x) = e(x), & x \in \Omega, \\ \quad \phi(x) \ = 0, & x \in \partial\Omega, \end{cases} \tag{3.14}$$

where it is clear that

$$e(x) = 0, \forall\ x \in \partial\Omega.$$

Using the Green's formula, the L_2 norm of the error can be written as

$$||e||^2 = \int_\Omega e^2\ dx = -\int_\Omega e(\Delta\phi)\ dx = \int_\Omega \nabla e.\nabla\phi\ dx.$$

Thus, by the Galerkin orthogonality and using the boundary condition, we get

$$\begin{aligned}
||e||^2 &= \int_\Omega \nabla e.\nabla\phi\ dx - \int_\Omega \nabla e.\nabla v\ dx \\
&= \int_\Omega \nabla e.\nabla(\phi - v)\ dx \\
&= \int_\Omega (-\Delta e)(\phi - v)\ dx.
\end{aligned}$$

But

$$-\Delta e = -\Delta u + \Delta U = f + \Delta U = r,$$

where r is the residual and v is an interpolant of ϕ, so

$$||e||^2 \leq ||h^2 r||\ ||h^{-2}(\phi - v)||.$$

Using the inequality

$$||(\phi - v)|| \leq C||h^2 D^2\phi|| \leq CC_\Omega||\Delta\phi||,$$

where C and C_Ω are constants, we get

$$||e||^2 \leq CC_\Omega||h^2 r||\ ||\Delta\phi||$$

$$\leq CC_\Omega||h^2 r||\ ||e||.$$

Thus, for this problem, the final a posteriori error estimate is

$$||u - U|| \leq c||h^2 r||.\square$$

In the following chapters we will talk about a posteriori error estimates for poisson, reaction-diffusion, and convection-diffusion problems. The finite element a posteriori estimate of such of which will be dwelled upon with more details.

Chapter 4

The a posteriori error estimator for poisson equation

The poisson equation is the model problem for elliptic partial differential equation, much like the heat and wave equations are for parabolic and hyperbolic PDE.

One of the most important encountered equations in many mathematical models of physical phenomena is the poisson equation. Just as an example, the solution of this equation gives the electrostatic potential for a given charge distribution. It also frequently appears in structural mechanics, theoretical physics as gravitation, electromagnetism, elasticity and many other areas of science and engineering.

The poisson equation is named after the French mathematician Siméen-Denis Poisson. The formulation of the poisson problem is

$$-\Delta u = f, \quad x \in \Omega,$$

where Ω is n-dimensional domain. The unknown function u, may be considered, e.g., as the electrostatic potential data, f, is the charge distribution. Note that if the charge distribution vanishes, this equation becomes Laplace's equation and the solution to the Laplace equation is called harmonic function, i.e., the equation $-\Delta u = 0$ is called the Laplace's equation.

4.1 A posteriori error estimator for homogeneous boundary condition

Consider the problem

$$-\Delta u = f, \quad \text{in } \Omega, \tag{4.1}$$

$$u = 0, \quad \text{on } \partial\Omega,$$

where $\Omega \subset \mathbb{R}^n$ is a polygonal domain with Lipschitz continuous boundary $\partial\Omega$ and $f \in L_2(\Omega)$. The variational formulation of (4.1) is find $u \in H_0^1(\Omega)$ such that

$$a(u, v) = F(v), \quad \forall\, v \in V,$$

where

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

The trial and test space V is the usual space of functions from $H_0^1(\Omega)$ which defined as

$$V = \{v : v \text{ is continuous on } \bar{\Omega} : v = 0 \text{ on } \partial\Omega\}.$$

The form $a(u, v)$ is assumed to be a V-elliptic bilinear from: $V \times V$ and the linear functional $F(v)$ is an element of the dual space V' (the dual space V' of the vector space V is the set of all linear functional on V). Associated with the bilinear form is the energy norm defined by $||v||_E = \sqrt{a(v, v)}$ and $||v||_{L_2} = \sqrt{(v, v)}$.
Note that the existence and uniqueness of the variational solution is provided by the Lax-Milgram theorem. The boundary of each element is also assumed to be Lipschitz continuous. The finite element approximation means to find a function $u_h \in V_h$ such that

$$a(u_h, v_h) = F(v_h), \quad \forall\, v_h \in V_h \subset V.$$

The error of the finite element approximation denoted by $e = u - u_h$ satisfies the error representation

$$\begin{aligned} a(e,v) &= a(u,v) - a(u_h,v) \\ &= F(v) - a(u_h,v) \\ &= R(v), \quad \forall v \in V. \end{aligned} \tag{4.2}$$

Here, $R(.)$ is called the residual functional or the weak residual.

If the choice of test functions is restricted to the finite element space, the fundamental Galerkin orthogonality condition follows,

$$R(v_h) = a(e, v_h) = 0, \quad \forall\ v_h \in V_h.$$

Assuming that the bilinear form is positive definite, it follows that the norm of the residual functional R is equal to the energy norm of the error [23, 55],

$$||R||_{V'} = \sup_{v\in V} \frac{|R(v)|}{||v||_E} = \sup_{v\in V} \frac{|a(e,v)|}{||v||_E} = ||e||_E$$

i.e., the error estimation in the energy norm is equivalent to the computation of the dual norm of the given residual. We break up the integrals over Ω into sums of integrals over the triangles $k \in T_h$ and integrate by parts over each triangle, i.e., if the integration is split into the contributions from each element, then $R(v)$ in (4.2) can be rewritten as

$$\begin{aligned} R(v) &= F(v) - a(u_h, v) \\ &= \sum_{k\in T_h} \int_k (fv - \nabla u_h.\nabla v)dx, \end{aligned}$$

where T_h is a family of triangulations, k denotes an element in T_h. Applying the Green's theorem and rearranging terms leads to, [23],

$$R(v) = \sum_{k\in T_h} \int_k (fv + \Delta u_h v)dx + \sum_{\gamma\in\partial T_h} \int_\gamma Jvds, \quad \forall\ v \in V,$$

where J is the jump of the gradient across the element edge γ

$$J = -n_\gamma.\nabla u_h, \quad \text{where } n_\gamma \text{ is the unit normal to the edge } \gamma.$$

So,

$$||e||_E^2 = a(e,e) = \sum_{k\in T_h} \int_k (f+\Delta u_h) e dx + \sum_{\gamma\in\partial T_h} \int_\gamma Jeds. \tag{4.3}$$

Now, we use the Galerkin orthogonality condition to introduce the interpolant $\pi_h e$ into (4.3), then we have

$$||e||_E^2 = a(e,e) = \sum_{k\in T_h} \int_k (f+\Delta u_h)(e-\pi_h e) dx + \sum_{\gamma\in\partial T_h} \int_\gamma J(e-\pi_h e)ds.$$

Let $r = f + \Delta u_h$ and applying the Cauchy-Schwarz inequality, we get

$$||e||_E^2 \leq \sum_{k\in T_h} ||r||_{L_2(k)} ||e-\pi_h e||_{L_2(k)} + \sum_{\gamma\in\partial T_h} ||J||_{L_2(\gamma)} ||e-\pi_h e||_{L_2(\gamma)}.$$

According to results of interpolation theory we have, see [23],

$$\begin{aligned} ||e-\pi_h e||_{L_2(k)} &\leq Ch||e||_{H^1(\bar{k})}, \\ ||e-\pi_h e||_{L_2(\gamma)} &\leq C\sqrt{h}||e||_{H^1(\bar{k})}, \end{aligned}$$

where h is the diameter of the element k and $\bar{k}$ denotes the subdomain of element sharing a common edge with k, and C is an interpolation constant which depends, for our model problem, on the shape of the element. Using these estimates in $||e||^2$ we get

$$||e||_E^2 \leq C||e||_{H^1(\bar{k})} \bigg(\sum_{k\in T_h} h||r||_{L_2(k)} + \sum_{\gamma\in\partial T_h} \sqrt{h}||J||_{L_2(\gamma)} \bigg).$$

Employing the inequality $||e||_{H^1(\bar{k})} \leq C||e||_E$, we get

$$||e||_E^2 \leq C||e||_E \bigg(\sum_{k\in T_h} h||r||_{L_2(k)} + \sum_{\gamma\in\partial T_h} \sqrt{h}||J||_{L_2(\gamma)} \bigg),$$

then

$$||e||_E \leq C \bigg(\sum_{k\in T_h} h||r||_{L_2(k)} + \sum_{\gamma\in\partial T_h} \sqrt{h}||J||_{L_2(\gamma)} \bigg).$$

Now, squaring both sides of this inequality

$$||e||_E^2 \leq C \bigg(\sum_{k\in T_h} h||r||_{L_2(k)} + \sum_{\gamma\in\partial T_h} \sqrt{h}||J||_{L_2(\gamma)} \bigg)^2.$$

By Young inequality, we have

$$||e||_E^2 \leq C\bigg(\sum_{k\in T_h} h^2||r||^2_{L_2(k)} + \sum_{\gamma\in\partial T_h} h||J||^2_{L_2(\gamma)}\bigg).$$

Now split the constant C into two contributions C_1 and C_2 corresponding to the element residual and the jump terms, respectively, then

$$||e||_E^2 \leq \sum_{k\in T_h} \big(C_1h^2||r||^2_{L_2(k)} + C_2h||J||^2_{L_2(\partial k)}\big).$$

Let

$$\mu_k^2 = C_1h^2||r||^2_{L_2(k)} + C_2h||J||^2_{L_2(\partial k)},$$

then,

$$||e||_E^2 \leq \sum_{k\in T_h} \mu_k^2.$$

Remark: The constants C that appear in the inequalities above are different from each other, where we assume only one notation just for simplicity.

4.2 A posteriori error estimator for mixed boundary condition

Let Ω be a bounded domain with Lipschitz continuous boundary Γ. Suppose that Γ consists of two measurable parts Γ_D and Γ_N such that $\Gamma = \Gamma_N \cup \Gamma_D$ where Γ_D and Γ_N are the Dirichlet and Neumann boundaries, respectively. Consider the mixed boundary value problem: Find a function u such that

$$\begin{aligned} -\Delta u &= f, \quad \text{in } \Omega, \\ u &= 0, \quad \text{on } \Gamma_D, \\ n.\nabla u &= g, \quad \text{on } \Gamma_N, \end{aligned} \tag{4.4}$$

where n is the outward normal to Γ. We assume that $f \in L_2(\Omega)$ and $g \in L_2(\Gamma_N)$.
A variational formulation of this problem is: Find $u \in V$ such that

$$\int_\Omega \nabla u.\nabla v \; dx = \int_\Omega fv \; dx + \int_{\Gamma_N} gv \; ds, \quad \forall v \in V,$$

where the test functions space V is defined as

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

This solution can be characterized equivalently as the minimizer of the following variational formulation: Find $u \in V$ such that $J(u) = \inf_{v \in V} J(v)$, where

$$J(v) = \frac{1}{2} \int_\Omega |\nabla v|^2 \, dx - \int_\Omega f v \, dx - \int_{\Gamma_N} g v \, ds.$$

To derive the dual variational formulation we employ the relation, [46, 47],

$$J(u) = \inf_{v \in V} \sup_{y^\star \in L^2(\Omega, \mathbf{R}^n)} \{ \int_\Omega (\nabla v . y^\star - \frac{1}{2} |y^\star|^2 - f v) \, dx - \int_{\Gamma_N} g v \, ds \}.$$

Define

$$Q^\star_{f,g} = \left\{ q^\star \in L^2(\Omega, \mathbf{R}^n); \int_\Omega \nabla . q^\star w \, dx = \int_\Omega -f w \, dx, \int_{\Gamma_N} (q^\star . n) w \, ds = \int_{\Gamma_N} g w \, ds, \ \forall \ w \in V \right\},$$

to find $p^\star \in Q^\star_{f,g}$ such that $I^\star(p^\star) = \sup_{q^\star \in Q^\star_{f,g}} I^\star(q^\star)$, where

$$I^\star(q^\star) = \int_\Omega (\nabla u . q^\star - \frac{1}{2} |q^\star|^2 - f u) \, dx - \int_{\Gamma_N} g u \, ds, \quad \text{is the dual variational functional.}$$

Let

$$\begin{aligned} J(u) &= I^\star(p^\star), \\ \nabla u &= p^\star, \end{aligned}$$

then we have the following theorem, [46].

Theorem 9. *For all $v \in V$ and $q^\star \in Q^\star_{fg}$, we have*

$$||\nabla(v - u)||^2 \leq ||\nabla v - q^\star||^2, \quad \forall v \in V, \quad \forall q^\star \in Q^\star_{f,g}.$$

proof.
We will begin as

$$
\begin{aligned}
J(v)-J(u) &= J(v)-I^{\star}(p^{\star}) \\
&= J(v)-I^{\star}(\nabla u) \\
&= \int_{\Omega}(\frac{1}{2}|\nabla v|^2-fv)\ dx-\int_{\Gamma_N} gv\ ds-\big(\int_{\Omega}(\nabla u.\nabla u-\frac{1}{2}|\nabla u|^2-fu)\ dx-\int_{\Gamma_N} gu\ ds\big) \\
&= \int_{\Omega}\big(\frac{1}{2}|\nabla(v-u)|^2+\nabla u.\nabla v-fv-\nabla u.\nabla u+fu\big)\ dx-\int_{\Gamma_N}(gv-gu)\ ds,
\end{aligned}
$$

but since

$$
\begin{aligned}
\int_{\Omega}\nabla u.\nabla v\ dx &= \int_{\Omega} fv\ dx+\int_{\Gamma_N} gv\ ds \\
\int_{\Omega}\nabla u.\nabla u\ dx &= \int_{\Omega} fu\ dx+\int_{\Gamma_N} gu\ ds.
\end{aligned}
$$

then we have,

$$
J(v)-J(u)=\frac{1}{2}||\nabla(v-u)||^2,\quad \forall v\in V.
$$

Hence, one can derive

$$
\begin{aligned}
\frac{1}{2}||\nabla(v-u)||^2 &= J(v)-J(u) \\
&= J(v)-I^{\star}(p^{\star}) \\
&= J(v)-\sup_{q^{\star}\in Q^{\star}_{f,g}}\{I^{\star}(q^{\star})\} \\
&= J(v)+\inf_{q^{\star}\in Q^{\star}_{f,g}}\{-I^{\star}(q^{\star})\} \\
&= \inf_{q^{\star}\in Q^{\star}_{f,g}}\{J(v)-I^{\star}(q^{\star})\}.
\end{aligned}
$$

For the term $J(v)-I^{\star}(q^{\star})$ we have

$$
\begin{aligned}
J(v)-I^{\star}(q^{\star}) &= \int_{\Omega}\left(\frac{1}{2}|\nabla v|^2-fv\right)\ dx-\int_{\Gamma_N} gv\ ds-\big(\int_{\Omega}\nabla u.q^{\star}-\frac{1}{2}|q^{\star}|^2-fu\big)\ dx-\int_{\Gamma_N} gu\ ds \\
&= \int_{\Omega}\big(\frac{1}{2}|\nabla v-q^{\star}|^2+q^{\star}.\nabla v-q^{\star}.\nabla u-fv+fu\big)\ dx+\int_{\Gamma_N}(gu-gv)\ ds,
\end{aligned}
$$

but

$$\begin{aligned}\int_\Omega q^\star.\nabla v\ dx &= -\int_\Omega \nabla.q^\star v\ dx + \int_{\Gamma_N} (q^\star.n)v\ ds\\ &= \int_\Omega fv\ dx + \int_{\Gamma_N} gv\ ds.\end{aligned}$$

Similarly,

$$\int_\Omega q^\star.\nabla u\ dx = \int_\Omega fu\ dx + \int_{\Gamma_N} gu\ ds.$$

So, we get that

$$J(v) - I^\star(q^\star) = \frac{1}{2}||\nabla v - q^\star||^2, \quad \forall v \in V, \quad q^\star \in Q^\star_{f,g},$$

and that,

$$||\nabla(v-u)||^2 = \inf_{q^\star \in Q^\star_{fg}} ||\nabla v - q^\star||^2.$$

We immediately deduce the estimate

$$||\nabla(v-u)||^2 \le ||\nabla v - q^\star||^2, \quad \forall v \in V, \forall q^\star \in Q^\star_{f,g}.\square \tag{4.5}$$

Now we will present a much simplified way of deriving functional type a posteriori estimates using a variant of the Helmholtz decomposition [44, 45] for the space $L_2(\Omega, \mathbf{R}^n)$. The Helmholtz decomposition of a vector field is the decomposition of the vector field into two vector fields, one a divergence-free and a curl-free fields. The space $L_2(\Omega, \mathbf{R}^n)$ is used for vector-valued functions with components in $L_2(\Omega)$. Here, we will

use the trace theorem, [29], that is

$$||u||_{0,\Gamma} \le C_\Gamma ||u||_{1,\Omega} \quad ,\forall\ v \in H^1(\Omega), \tag{4.6}$$

where C_Γ is positive constants depending only on Γ, and $||.||_{1,\Omega}$ stands for the standard norm in $H^1(\Omega)$, and the symbol $||.||_{0,\Gamma}$ means the norm is $L_2(\Gamma)$, see, e.g., [35].

Theorem 10.
Let $u \in V$ be the solution to the problem (4.4) and v be any function from V. Then, see [46, 47],

$$||\nabla(v-u)||^2 \leq (1+\beta)||\nabla v - y^\star||^2 + \left(1+\frac{1}{\beta}\right)\left(1+\frac{1}{\gamma}\right)C_{\Gamma_N}^2(1+C_\Omega^2)||y^\star.n-g||^2_{L_2(\Gamma_N)} + \tag{4.7}$$

$$+\left(1+\frac{1}{\beta}\right)(1+\gamma)C_\Omega^2||div\ y^\star + f||^2,$$

where β is an arbitrary positive number, $y^\star$ is any function from $\tilde{H}(\Omega, div) = \{y^\star \in L_2(\Omega, \mathbf{R}^n) : div\ y^\star \in L_2(\Omega), y^\star.n \in L_2(\Gamma_N)\}$,[47], C_Ω is the constant from Ponicare inequality, and C_{Γ_N} is the constant in the trace inequality for the domain Ω.

proof.
Consider

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \Gamma_D, \\ n.\nabla u &= g \quad \text{on } \Gamma_N, \end{aligned}$$

by (4.5) we have

$$||\nabla(v-u)||^2 \leq ||\nabla v - q^\star||^2, \quad \forall v \in V, \forall q^\star \in Q^\star_{f,g}.$$

To estimate the right-hand side for any $v \in V$, we take an arbitrary function $y^\star \in \tilde{H}(\Omega, div)$. Define the auxiliary function w as the solution to the problem

$$\begin{aligned} \Delta w &= div\ y^\star + f \quad \text{in } \Omega, \\ w &= 0 \quad \text{on } \Gamma_D, \\ n.\nabla w &= y^\star.n + g \quad \text{on } \Gamma_N. \end{aligned} \tag{4.8}$$

As $y^\star \in L_2(\Omega, \mathbf{R}^n)$, we have for $y^\star$ the Holmholtz decomposition $y^\star = q^\star + \nabla w$, where $q^\star \in Q^\star_{f,g}$ and $w \in V$.
Then, using Young's inequality, we obtain

$$||\nabla v - q^\star||^2 \leq (1+\beta)||\nabla v - y^\star||^2 + \left(1+\frac{1}{\beta}\right)||\nabla w||^2, \quad \forall\ \beta > 0. \tag{4.9}$$

Since $w \in V$ and $\Delta w \in L_2(\Omega)$, then by Poincare inequality we get

$$\begin{aligned} ||\nabla w||^2 &= \int_{\Gamma_N} \frac{\partial w}{\partial n} w \; ds - \int_{\Omega} (\Delta w) w \; dx \\ &\leq ||\frac{\partial w}{\partial n}||_{L_2(\Gamma_N)} C_{\Gamma_N} (1 + C_\Omega^2)^{\frac{1}{2}} ||\nabla w|| + C_\Omega ||\Delta w|| ||\nabla w||, \end{aligned}$$

that is,

$$||\nabla w|| \leq C_{\Gamma_N} (1 + C_\Omega^2)^{\frac{1}{2}} ||\frac{\partial w}{\partial n}||_{L_2(\Gamma_N)} + C_\Omega ||\Delta w||, \tag{4.10}$$

where C_Ω is the constant of Poincare inequality, and C_{Γ_N} is the constant of the trace inequality. Now by (4.5) we have

$$||\nabla(v - u)||^2 \leq ||\nabla v - q^\star||^2, \quad \forall v \in V, \forall q^\star \in Q^\star_{f,g}.$$

Using (4.9),(4.10), and Young's inequality to get

$$||\nabla(v - u)||^2 \leq (1 + \beta) ||\nabla v - y^\star||^2 + \left(1 + \frac{1}{\beta}\right)\left(1 + \frac{1}{\gamma}\right) C_{\Gamma_N}^2 (1 + C_\Omega^2) ||y^\star.n + g||^2_{L_2(\Gamma_N)}$$

$$+ \left(1 + \frac{1}{\beta}\right)(1 + \gamma) C_\Omega^2 ||div \; y^\star + f||^2, \quad \forall \; v \in V, \forall \; y^\star \in \tilde{H}(\Omega, div),$$

where β and γ are arbitrary positive numbers come from Young's inequality.□

Since u is the exact solution of (4.4), v is any function from V, and $y^\star$ is any function from $\tilde{H}(\Omega, div)$, the estimate (4.7) is an a posteriori error estimate valid for any approximation of the problem (4.4).

Chapter 5

The a posteriori error estimator for reaction-diffusion and convection-diffusion problems

5.1 Reaction-diffusion problems

The reaction-diffusion problem arises naturally in systems consisting of many interacting components as chemical reactions, and are widely used to describe pattern-formation phenomena in variety of biological, chemical and physical systems.

Reaction-diffusion equations describe distributions of temperature, concentrations or of some other variables in space and in time. These equations are characterized by the presence of diffusion and production terms. Originally, diffusion was understood as random motion of atoms and molecules and described by the Laplace operator.

The Reaction-Diffusion Problem takes the form

$$-\varepsilon \Delta u + cu = f, \quad \text{in } \Omega \text{ with some boundary conditions on } \partial\Omega,$$

where ε is a small positive parameter, c and f are continuously differentiable functions on Ω.

5.1.1 A priori error estimation

We start this section by considering the model of elliptic problem with mixed (Dirichlet/Neumann) boundary conditions; find a function u such that

$$\begin{aligned} -\Delta u + cu &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \Gamma_D, \\ n.\nabla u &= g \quad \text{on } \Gamma_N, \end{aligned}$$

where Ω is a bounded domain in $\mathbf{R}^n$ with a Lipschitz continuous boundary Γ such that $\Gamma = \Gamma_N \cup \Gamma_D$, n is the outward normal to the boundary, $f \in L_2(\Omega)$, $g \in L_2(\Gamma_N)$, and $c \in L_\infty(\Omega)$.

The variational formulation of this problem is: Find $u \in H^1_{\Gamma_D}(\Omega)$ such that

$$\int_\Omega \nabla u.\nabla v \; dx + \int_\Omega cuv \; dx = \int_\Omega fv \; dx + \int_{\Gamma_N} gv \; ds, \quad \forall \; v \in H^1_{\Gamma_D}(\Omega), \tag{5.1}$$

$$\text{where} \quad H^1_{\Gamma_D}(\Omega) = \{v \in H^1(\Omega) : v = 0 \quad \text{on } \Gamma_D\}.$$

In addition, we assume that almost everywhere $c \geq 0$ in Ω, and introduce the set $\Omega^c = \{x \in \Omega : c(x) > 0\}$. Let us define the bilinear form $a(.,.)$ and the linear form $F(.)$ as follows

$$\begin{aligned} a(u,v) &= \int_\Omega \nabla u.\nabla v \; dx + \int_\Omega cuv \; dx, \quad u,v \in H^1(\Omega), \\ F(v) &= \int_\Omega fv \; dx + \int_{\Gamma_N} gv \; ds, \quad v \in H^1(\Omega). \end{aligned}$$

Then the weak formulation (5.1) can be rewritten as: Find $u \in H^1_{\Gamma_D}(\Omega)$ such that $a(u,v) = F(v)$, $\forall v \in H^1_{\Gamma_D}(\Omega)$.
Here, we will need Friedrichs inequality, [29], that is

$$||u||_{0,\Omega} \leq C_{\Omega,\Gamma_D} ||\nabla u||_{0,\Omega} \quad ,\forall \; u \in H^1_{\Gamma_D}(\Omega), \tag{5.2}$$

where C_{Ω,Γ_D} is positive constant depending only on Ω, Γ_D, and $||.||_{0,\Omega}$ stands for the standard norm in $L_2(\Omega)$.
Also, we will use the trace theorem (4.6). Let u_h be any function from $H^1_{\Gamma_D}(\Omega)$ considered as an approximation of u. The error $e = u - u_h$ will be estimated in the

energy norm,

$$\begin{aligned} ||e||^2 &= \int_\Omega \nabla(u-u_h).\nabla(u-u_h)\ dx + \int_\Omega c(u-u_h)^2\ dx \qquad (5.3)\\ &= ||\nabla(u-u_h)||^2_\Omega + ||\sqrt{c}(u-u_h)||^2_{\Omega^c}, \end{aligned}$$

where $||y||_\Omega = ||y||_{L^2(\Omega)}$.

5.1.2 A posteriori error estimation

For $y \in L_2(\Omega, \mathbf{R}^n)$, we define

$$H_N(\Omega, div) = \{y \in L_2(\Omega, \mathbf{R}^n) : div\ y \in L_2(\Omega), y.n \in L_2(\Gamma_N)\}.$$

We will use the notation χ_s for the characteristic function of the set S, i.e.,

$$\chi_s(x) = \begin{cases} 1 & , x \in S, \\ 0 & , x \notin S. \end{cases}$$

Theorem 11.
For the error in the energy norm (5.3) we have the following upper estimate [29],

$$\begin{aligned} ||e||^2 \leq \Big|\Big|\frac{1}{\sqrt{c}}(f + div\ y - cu_h)\Big|\Big|^2_{0,\Omega^c} + (1+\alpha)\big|\big|y^\star - \nabla u_h\big|\big|^2_\Omega + \big(1+\frac{1}{\alpha}\big)(1+\beta)\frac{C^2_{\Omega,\Gamma_D}}{c_1} * \quad (5.4)\\ *\big|\big|\chi_{\Omega/\bar{\Omega}^c}(f + div\ y - cu_h)\big|\big|^2_{0,\Omega} + \big(1+\frac{1}{\alpha}\big)(1+\frac{1}{\beta})C^2_{\Omega,\Gamma_D}\big|\big|g - n.y^\star\big|\big|^2_{0,\Gamma_N}, \end{aligned}$$

where α and β are arbitrary numbers and $y^\star$ is any function from $H_N(\Omega, div)$, and $C_{\Omega,\Gamma} = \frac{C_\Gamma\sqrt{1+C^2_{\Omega,\Gamma_D}}}{\sqrt{c_1}}$.

proof.
Since $u - u_h \in H^1_{\Gamma_D}(\Omega)$, then the equation (5.3) holds as follows

$$\begin{aligned} ||e||^2 &= \int_\Omega \nabla(u-u_h).\nabla(u-u_h)\ dx + \int_\Omega c(u-u_h)(u-u_h)\ dx \\ &= \int_\Omega \nabla u.\nabla(u-u_h)\ dx + \int_\Omega cu(u-u_h)\ dx - \int_\Omega \nabla u_h.\nabla(u-u_h)\ dx - \int_\Omega cu_h(u-u_h)\ dx. \end{aligned}$$

By the Green's formula we get

$$\int_\Omega \nabla u.\nabla(u-u_h)\ dx = \int_\Omega -\Delta u(u-u_h)\ dx + \int_{\Gamma_N} g(u-u_h)\ ds,$$

so, we have

$$\begin{aligned} ||e||^2 &= \int_\Omega(-\Delta u+cu)(u-u_h)\ dx+\int_{\Gamma_N} g(u-u_h)\ ds-\int_\Omega \nabla u_h.\nabla(u-u_h)\ dx-\int_\Omega cu_h(u-u_h)\ dx \\ &= \int_\Omega(f-cu_h)(u-u_h)\ dx+\int_{\Gamma_N} g(u-u_h)\ ds-\int_\Omega(\nabla u_h-y^\star).\nabla(u-u_h)\ dx-\int_\Omega y^\star.\nabla(u-u_h)\ dx, \end{aligned} \tag{5.5}$$

where $y^\star$ is any function from the space $H_N(\Omega, div)$. Applying the Green's formula to the last term in (5.5) gives

$$\int_\Omega y^\star.\nabla(u-u_h)\ dx = \int_{\Gamma_N}(n.y^\star)(u-u_h)\ ds - \int_\Omega \nabla.y^\star(u-u_h)\ dx.$$

Substitute this in (5.5), we get

$$||e||^2 = \int_\Omega(y^\star-\nabla u_h).\nabla(u-u_h)\ dx+\int_\Omega(f-cu_h+\nabla.y^\star)(u-u_h)\ dx+\int_{\Gamma_N}(g-n.y^\star)(u-u_h)\ ds. \tag{5.6}$$

Now, we proceed by estimating the three terms in the right-hand side of (5.6). Using the notations,

$$\begin{aligned} E_1 &= \int_\Omega(y^\star-\nabla u_h).\nabla(u-u_h)\ dx, \\ E_2 &= \int_\Omega(f-cu_h+\nabla.y^\star)(u-u_h)\ dx, \\ E_3 &= \int_{\Gamma_N}(g-n.y^\star)(u-u_h)\ ds, \end{aligned}$$

then $||e||^2 = E_1+E_2+E_3$. According to E_1, by Cauchy-Schwarz inequality, we have

$$E_1 \le ||y^\star-\nabla u_h||_\Omega\ ||\nabla(u-u_h)||_\Omega. \tag{5.7}$$

The term E_2 can be treated as follows

$$\begin{aligned} E_2 &= \int_\Omega (f - cu_h + \nabla.y^\star)(u-u_h)\ dx, \quad \text{and by the definition of } \Omega^c \\ &= \int_{\Omega^c} \frac{1}{\sqrt{c}}(f + \nabla.y^\star - cu_h)\sqrt{c}(u-u_h)\ dx + \int_\Omega \chi_{\Omega/\bar{\Omega^c}}(f+\nabla.y^\star - cu_h)(u-u_h)\ dx. \end{aligned}$$

Apply Cauchy-Schwarz inequality to the terms to the right of the equation above to get

$$E_2 \le ||\sqrt{c}(u-u_h)||_{0,\Omega^c}\ ||\frac{1}{\sqrt{c}}(f+\nabla.y^\star - cu_h)||_{0,\Omega^c} + ||\chi_{\Omega/\bar{\Omega^c}}(f+\nabla.y^\star-cu_h)||_{0,\Omega}\ ||u-u_h||_{0,\Omega}.$$

The first term of the right-hand side of E_2 can be estimated by the simple inequality $ab \le \frac{1}{2}(a^2+b^2)$, and the second term estimated using Friedrichs inequality as follows

$$E_2 \le \frac{1}{2}||\sqrt{c}(u-u_h)||^2_{0,\Omega^c} + \frac{1}{2}||\frac{1}{\sqrt{c}}(f+\nabla.y^\star-cu_h)||^2_{0,\Omega^c} + \frac{C_{\Omega,\Gamma_D}}{\sqrt{c_1}}||f+\nabla.y^\star-cu_h||_{0,\Omega/\bar{\Omega^c}}||\nabla(u-u_h)||_{0,\Omega}. \tag{5.8}$$

Finally, the term E_3,

$$E_3 = \int_{\Gamma_N}(g-n.y^\star)(u-u_h)\ ds \le ||g-n.y^\star||_{0,\Gamma_N}\ ||u-u_h||_{0,\Gamma_N}.$$

Use the inequalities (5.2) and (4.6) to get

$$\begin{aligned} E_3 &\le C_\Gamma ||g-n.y^\star||_{0,\Gamma_N}\ ||u-u_h||_{1,\Omega} \\ &\le C_{\Omega,\Gamma}||g-n.y^\star||_{0,\Gamma_N}\ ||\nabla(u-u_h)||_\Omega. \end{aligned} \tag{5.9}$$

Estimate $||e||$ by using E_1, E_2, E_3 ,and apply Young inequality after regrouping the terms, thus

$$\begin{aligned} ||e||^2 \le \frac{1}{2}\big(||y^\star - \nabla u_h||_\Omega + \frac{C_{\Omega,\Gamma_D}}{\sqrt{c_1}}||f+\nabla.y^\star - cu_h||_{0,\Omega/\bar{\Omega^c}} + C_{\Omega,\Gamma}||g-n.y^\star||_{0,\Gamma_N}\big)^2 \\ + \frac{1}{2}||\nabla(u-u_h)||^2_\Omega + \frac{1}{2}||\sqrt{c}(u-u_h)||^2_{0,\Omega^c} + \frac{1}{2}||\frac{1}{\sqrt{c}}(f+\nabla.y^\star-cu_h)||^2_{0,\Omega^c}. \end{aligned}$$

Multiply the last inequality by two and use inequality (5.3), we immediately get for the error in the energy norm that

$$2||e||^2 \le \left(||y^\star - \nabla u_h||_\Omega + \frac{C_{\Omega,\Gamma_D}}{\sqrt{c_1}}||f + \nabla.y^\star - cu_h||_{0,\Omega/\bar{\Omega}^c} + C_{\Omega,\Gamma}||g - n.y^\star||_{0,\Gamma_N}\right)^2 +$$

$$+\underbrace{||\nabla(u-u_h)||^2_\Omega + ||\sqrt{c}(u-u_h)||^2_{0,\Omega^c}}_{||e||^2} + ||\frac{1}{\sqrt{c}}(f + \nabla.y^\star - cu_h)||^2_{0,\Omega^c}.$$

Finally, using two times the inequality $(a+b)^2 \le (1+\lambda)a^2 + \left(1+\frac{1}{\lambda}\right)b^2$, which is valid for any $\lambda > 0$, for the terms in the round brackets in the last inequality, we get estimate (5.4). $\square$

5.2 ***Convection-diffusion problems***

The convection-diffusion problems very often happen that the solution have a convective nature on most of the domain of the problem, and the diffusive part of the differential operator is influenced only in certain small subdomains. They usually have a degree of instability. The goal then is to modify these numerical methods in stable form without loss accuracy. The numerical solution of convection-diffusion problems dates back to the 1950s Allen and Southwell [2], but actually it began in 1970s [50] and continued to this day.

A common source of convection-diffusion problems is the Navier-Stokes equations with large Reynolds number. Morton [34] listed ten examples involving convection-diffusion equations that include the drift-diffusion equations of semiconductor device modeling and the Black-sholes equation from financial modeling.

The Convection-Diffusion Problem takes the form

$$-\varepsilon\Delta u + b.\nabla u + cu = f, \text{ in } \Omega,$$

with some boundary conditions on $\partial\Omega$, where ε is a small positive parameter, b, c and f are continuously differentiable functions on Ω.

5.2.1 The convection-diffusion problems

In this section we consider the convection-diffusion boundary value problem

$$-\varepsilon\Delta u + b.\nabla u + cu = f, \quad \text{in } \Omega,$$

$$u = 0, \text{ on } \partial\Omega,$$

where Ω is some domain in $\mathbf{R}^n$ with boundary $\partial\Omega$, the constant ε is living in $(0,1]$, $b \in H^1(\Omega) \cap L^\infty(\Omega)$, $c \in L^\infty(\Omega)$, and $f \in L_2(\Omega)$. Furthermore, we assume the smoothness property div $b \in L_2(\Omega)$.

The term $-\varepsilon\triangle u$ models diffusion and $b.\nabla u$ models convection. The terminology convection-diffusion problem is used since the convection coefficient has much greater magnitude than the diffusion coefficient :

$$\frac{|\text{coefficient } \nabla u|}{|\text{coefficient } \Delta u|} = \frac{|b|}{\varepsilon} >> 1.$$

Derivation of the variational form (weak formulation) is as follows: Multiply the differential equation by a test function v, with $v = 0$ on $\partial\Omega$, and integrate over Ω to get

$$\int_\Omega (-\varepsilon\Delta u + b.\nabla u + cu)v \, dx \quad = \quad \int_\Omega fv \, dx$$

$$\begin{aligned}
\Longleftrightarrow \quad & \int_{\partial\Omega} -\varepsilon(\nabla u.n)v \, ds + \int_\Omega \varepsilon\nabla u.\nabla v + (b.\nabla u + cu)v \, dx \quad = \quad \int_\Omega fv \, dx \\
\Longleftrightarrow \quad & \int_\Omega \big(\varepsilon\nabla u.\nabla v + b.\nabla uv + cuv\big) \, dx \quad = \quad \int_\Omega fv \, dx
\end{aligned}$$

The integral on the boundary vanishes because of the boundary condition of the test function. The highest order derivative of u has been transferred to v. In functional forms, the weak formulation of this convection-diffusion problem is to find $u \in H_0^1(\Omega)$ such that

$$a(u,v) = F(v), \quad \forall v \in H_0^1, \tag{5.10}$$

where the bilinear form $a(.,.)$ and the linear function $F(.)$ are given by

$$a(u,v) = \int_\Omega \varepsilon\nabla u.\nabla v + b.\nabla uv + cuv \, dx,$$

$$F(v) = \int_\Omega fv\ dx, \qquad u, v \in H_0^1(\Omega).$$

Consider

$$a(v, v) = \int_\Omega (\varepsilon(\nabla v)^2 + b.\nabla vv + cv^2\ )dx.$$

Note that

$$\int_\Omega b.\nabla vv\ dx = -\int_\Omega \nabla.(bv)v\ dx$$

$$= -\int_\Omega (\nabla.b)v^2\ dx - \int_\Omega b.\nabla vv\ dx.$$

It follows that

$$\int_\Omega b.\nabla vv\ dx = -\frac{1}{2}\int_\Omega (\nabla.b)v^2\ dx,$$

so,

$$a(v, v) = \varepsilon \int_\Omega |\nabla v|^2 dx + \int_\Omega \big(c - \frac{1}{2}\nabla.b\big)v^2 dx.$$

If $-\frac{1}{2}\nabla.b + c \geq 0$, then using Poincare inequality; $\forall v \in H_0^1(\Omega)$ we have

$$a(v, v) \geq \int_\Omega \varepsilon|\nabla v|^2\ dx = \varepsilon||\nabla v||^2_{L^2(\Omega)} \geq C||v||^2_{H^1(\Omega)},$$

where C is a positive constant. Then the coercivity of the bilinear form $a(.,.)$ implies the unique solvability of problem (5.10) due to the Lax-Milgram lemma.

5.2.2 A posteriori error estimation

Let u_h be a function of $H_0^1(\Omega)$ considered as an approximation of u. The error $e = u - u_h$ will be estimated in the energy norm

$$||e||^2 = a(e, e) = \varepsilon \int_\Omega |\nabla e|^2 dx + \int_\Omega \big(c - \frac{1}{2}\nabla.b\big)e^2 dx. \tag{5.11}$$

Define the estimation of the error in terms of a suitable global weighted energy norm [30] as :

$$||e||^2_{\lambda,\mu} = \lambda \int_\Omega |\nabla e|^2 dx + \mu \int_\Omega \bar{c}e^2 dx, \tag{5.12}$$

where the weights λ and μ are nonnegative real numbers, and $\bar{c} = (c - \frac{1}{2}\nabla.b)$. In particular, in(5.11)we have $\lambda = \epsilon$ and $\mu = 1$. So,

$$a(e,e) = ||e||^2_{\epsilon,1} = \epsilon||\nabla(u-u_h)||^2 + ||\sqrt{\bar{c}}(u-u_h)||^2. \tag{5.13}$$

To construct a posteriori estimates for the error it is noted that, [30, 32, 25],

$$\begin{aligned}
a(e,e) &= a(u-u_h, u-u_h) \\
&= \varepsilon\int_\Omega \nabla(u-u_h).\nabla(u-u_h)dx + \int_\Omega b.\nabla(u-u_h)(u-u_h)dx + \int_\Omega c(u-u_h)(u-u_h)dx \\
&= \int_\Omega \big(\epsilon\nabla u.\nabla(u-u_h) + b.\nabla u(u-u_h) + cu(u-u_h)\big)dx + \\
&+\epsilon\int_\Omega -\nabla u_h.\nabla(u-u_h)dx + \int_\Omega -b.\nabla u_h(u-u_h)dx + \int_\Omega -cu_h(u-u_h)dx.
\end{aligned}$$

Now, by the Green's formula,

$$\epsilon\int_\Omega \nabla u.\nabla(u-u_h)dx = -\int_\Omega \epsilon\Delta u(u-u_h)dx + \underbrace{\int_{\partial\Omega}\epsilon(n.\nabla u)(u-u_h)ds}_{=0},$$

so,

$$a(e,e) = \int_\Omega(-\epsilon\Delta u + b.\nabla u + cu)(u-u_h)dx - \epsilon\int_\Omega \nabla u_h.\nabla(u-u_h)dx +$$
$$-\int_\Omega b.\nabla u_h(u-u_h)dx - c\int_\Omega u_h(u-u_h)dx.$$

Thus,

$$a(e,e) = \int_\Omega f(u-u_h)dx - \epsilon\int_\Omega \nabla u_h.\nabla(u-u_h)dx - \int_\Omega b.\nabla u_h(u-u_h)dx - \int_\Omega cu_h(u-u_h)dx.$$

Further, we regroup some terms and introduce a function $y^\star \in H(\Omega, div)$, where

$$H(\Omega, div) = \{y \in L_2(\Omega, \mathbf{R}^n) : div\ y \in L_2(\Omega)\}.$$

Hence,

$$a(u-u_h, u-u_h) = \int_\Omega(f - b.\nabla u_h - cu_h)(u-u_h)dx - \int_\Omega(\epsilon\nabla u_h - y^\star + y^\star).\nabla(u-u_h)dx$$

$$= \int_\Omega (f - b.\nabla u_h - cu_h)(u - u_h)dx - \int_\Omega (\epsilon \nabla u_h - y^\star).\nabla(u - u_h)dx - \int_\Omega y^\star.\nabla(u - u_h)dx.$$

Since

$$\int_\Omega -y^\star.\nabla(u - u_h)dx = \int_\Omega \nabla.y^\star(u - u_h)dx,$$

we get

$$a(u - u_h, u - u_h) = \int_\Omega (f - b.\nabla u_h - cu_h + \nabla.y^\star)(u - u_h)dx + \int_\Omega (y^\star - \epsilon\nabla u_h).\nabla(u - u_h)dx.$$

Finally, let us introduce another auxiliary function $v \in H_0^1(\Omega)$ and consider the following

$$\begin{aligned} ||e||^2_{\epsilon,1} &= a(e, e) = a(u - u_h, u - u_h) \\ &= \int_\Omega (f - b.\nabla u_h - cu_h + \nabla.y^\star - cv - b.\nabla v)(u - u_h)dx + \end{aligned}$$

$$+ \int_\Omega (y^\star - \epsilon\nabla u_h + \epsilon\nabla v).\nabla(u - u_h)dx + \int_\Omega \big((cv + b.\nabla v)(u - u_h) - \epsilon\nabla v.\nabla(u - u_h)\big)dx$$

Hence,

$$||e||^2_{\epsilon,1} = E_1 + E_2 + E_3,$$

where the terms E_1, E_2 and E_3 are defined as :

$$\begin{aligned} E_1 &= \int_\Omega (f - b.\nabla u_h - cu_h + \nabla.y^\star - cv - b.\nabla v)(u - u_h)dx. \\ E_2 &= \int_\Omega (y^\star - \epsilon\nabla u_h + \epsilon\nabla v).\nabla(u - u_h)dx. \\ E_3 &= \int_\Omega \big((cv + b.\nabla v)(u - u_h) - \epsilon\nabla v.\nabla(u - u_h)\big)dx. \end{aligned}$$

Now, E_3 is estimated as follows

$$E_3 = \int_\Omega (cv(u - u_h) + b.\nabla v(u - u_h) + \epsilon\nabla v.\nabla u_h - \epsilon\nabla v.\nabla u)dx,$$

but

$$\begin{aligned}
-\int_\Omega \epsilon \nabla v.\nabla u dx &= \int_\Omega \epsilon \Delta u v dx - \underbrace{\int_{\partial\Omega} \epsilon(\nabla u.n) v ds}_{=0} \\
&= \int_\Omega (-fv + b.\nabla u v + cuv) dx.
\end{aligned}$$

So,

$$\begin{aligned}
E_3 &= \int_\Omega (cv(u-u_h) + b.\nabla v(u-u_h) + \epsilon\nabla v.\nabla u_h - fv + b.\nabla u v + cuv) dx \\
&= \int_\Omega \big(2cv(u-u_h) - cv(u-u_h) + b.\nabla v(u-u_h) + \\
&\quad +\epsilon\nabla v.\nabla u_h - fv + b.\nabla(u-u_h)v + b.\nabla u_h v + cuv\big) dx.
\end{aligned}$$

Since

$$\int_\Omega \big(b.\nabla v(u-u_h) + b.\nabla(u-u_h)v\big) dx = \int_\Omega b.\nabla(v(u-u_h)) dx,$$

then,

$$E_3 = \int_\Omega \big(b.\nabla(v(u-u_h)) + 2cv(u-u_h)\big) dx + \int_\Omega \big(\epsilon\nabla v.\nabla u_h + b.\nabla u_h v + cvu_h - fv\big) dx.$$

By the divergence theorem

$$\int_\Omega b.\nabla(v(u-u_h)) dx = \underbrace{\int_{\partial\Omega} b.nv(u-u_h) ds}_{=0} - \int_\Omega (\nabla.b) v(u-u_h) dx.$$

Hence,

$$\int_\Omega \big(b.\nabla(v(u-u_h)) + 2cv(u-u_h)\big) dx = \int_\Omega (-\nabla.b + 2c) v(u-u_h) dx.$$

Thus,

$$\begin{aligned}
E_3 &= 2\int_\Omega (c - \frac{1}{2}\nabla.b) v(u-u_h) dx + \int_\Omega (\epsilon\nabla v.\nabla u_h + b.\nabla u_h v + cvu_h - fv) dx. \\
&= \qquad E_{3,1}(u,v,u_h) \qquad + \qquad E_{3,2}(v,u_h).
\end{aligned}$$

Note that the term $E_{3,2}$ is directly computable once we have the approximation u_h computed and fixed v, but since $E_{3,1}$ containing the unknown exact solution u, then

$$\begin{aligned} E_{3,1} &= 2\int_\Omega \bar{c}v(u-u_h)dx, \text{ where } \bar{c} = c - \frac{1}{2}\nabla.b \\ &= 2\int_\Omega \sqrt{\bar{c}}v\sqrt{\bar{c}}(u-u_h)dx. \end{aligned}$$

Then Cauchy-Schwarz inequality leads to

$$E_{3,1} \leq 2||\sqrt{\bar{c}}v||_{L_2}||\sqrt{\bar{c}}(u-u_h)||_{L_2}.$$

Using the inequality $2ab \leq a^2 + b^2$,

$$E_{3,1} \leq ||\sqrt{\bar{c}}v||^2_{L_2} + ||\sqrt{\bar{c}}(u-u_h)||^2_{L_2}.$$

Hence, for any positive number β, we have

$$E_{3,1} \leq \beta||\sqrt{\bar{c}}v||^2_{L_2} + \frac{1}{\beta}||\sqrt{\bar{c}}(u-u_h)||^2_{L_2}.$$

So,

$$E_3 \leq \beta||\sqrt{\bar{c}}v||^2_{L_2} + \frac{1}{\beta}||\sqrt{\bar{c}}(u-u_h)||^2_{L_2} + \int_\Omega(\epsilon\nabla v.\nabla u_h + b.\nabla u_h v + cvu_h - fv)dx. \quad (5.14)$$

Now, E_1 and E_2 will be estimated as

$$E_1 \leq ||f - b.\nabla u_h - cu_h + \nabla.y^\star - cv - b.\nabla v||_{L_2}||u-u_h||_{L_2},$$

but by poincare inequality,

$$E_1 \leq C_\Omega||f - b.\nabla u_h - cu_h + \nabla.y^\star - cv - b.\nabla v||_{L_2}||\nabla(u-u_h)||_{L_2},$$

where C_Ω is poincare constant. Similarly,

$$E_2 \leq ||y^\star - \epsilon\nabla u_h + \epsilon\nabla v||_{L_2}||\nabla(u-u_h)||_{L_2},$$

we immediately get

$$E_1+E_2 \leq \big(C_\Omega||f-b.\nabla u_h-cu_h+\nabla.y^\star-cv-b.\nabla v||_{L_2}+||y^\star-\epsilon\nabla u_h+\epsilon\nabla v||_{L_2}\big)||\nabla(u-u_h)||_{L_2}.$$

Using the inequality,

$$pq \leq \frac{\alpha}{2}p^2+\frac{1}{2\alpha}q^2, \quad p,q\geq 0 \text{ and } \alpha>0,$$

then

$$E_1+E_2 \leq \frac{\alpha}{2}||\nabla(u-u_h)||^2_{L_2}+\frac{1}{2\alpha}\big(||y^\star-\epsilon\nabla u_h+\epsilon\nabla v||_{L_2}+C_\Omega||f-b.\nabla u_h-cu_h+\nabla.y^\star-cv-b.\nabla v||_{L_2}\big)^2.$$

For the second term in the right-hand side of the above inequality, employ the inequality

$$(p+q)^2 \leq (1+\gamma)p^2+(1+\frac{1}{\gamma})q^2, \quad p,q\geq 0 \text{ and } \gamma>0.$$

Thus,

$$E_1+E_2 \leq \frac{\alpha}{2}||\nabla(u-u_h)||^2_{L_2}+\frac{1}{2\alpha}\Bigg((1+\gamma)||y^\star-\epsilon\nabla u_h+\epsilon\nabla v||^2_{L_2}+ \tag{5.15}$$

$$+(1+\frac{1}{\gamma})C^2_\Omega||f-b.\nabla u_h-cu_h+\nabla.y^\star-cv-b.\nabla v||^2_{L_2}\Bigg).$$

So by (5.14) and (5.15) we get,

$$||e||^2_{\epsilon,1}=E_1+E_2+E_3 \leq \frac{\alpha}{2}||\nabla(u-u_h)||^2_{L_2}+\frac{1}{2\alpha}\Bigg((1+\gamma)||y^\star-\epsilon\nabla u_h+\epsilon\nabla v||^2_{L_2}+$$

$$+(1+\frac{1}{\gamma})C^2_\Omega||f-b.\nabla u_h-cu_h+\nabla.y^\star-cv-b.\nabla v||^2_{L_2}\Bigg)+$$

$$+\beta||\sqrt{\bar{c}}v||^2_{L_2}+\frac{1}{\beta}||\sqrt{\bar{c}}(u-u_h)||^2_{L_2}+\int_\Omega(\epsilon\nabla v.\nabla u_h+b.\nabla u_h v+cvu_h-fv)dx.$$

Let

$$Est_{\alpha,\beta}(\gamma,y^\star,v,u_h)=Est_{\alpha,\beta}=$$

$$= \frac{1}{2\alpha}\Bigg((1+\gamma)||y^\star - \epsilon\nabla u_h + \epsilon\nabla v||^2_{L_2} + (1+\frac{1}{\gamma})C^2_\Omega||f - b.\nabla u_h - cu_h + \nabla.y^\star - cv - b.\nabla v||^2_{L_2}\Bigg)+$$

$$+\beta||\sqrt{c}v||^2_{L_2} + \int_\Omega(\epsilon\nabla v.\nabla u_h + b.\nabla u_h v + cvu_h - fv)dx, \tag{5.16}$$

then,

$$||e||^2_{\epsilon,1} \leq \frac{\alpha}{2}||\nabla(u-u_h)||^2 + \frac{1}{\beta}||\sqrt{c}(u-u_h)||^2 + Est_{\alpha,\beta}. \tag{5.17}$$

Theorem 12.
Let α and β be fixed positive numbers such that $2\epsilon \geq \alpha > 0$, $\beta \geq 1$, then the following a posteriori upper error estimates for the error in the weighted energy norm is obtained, [30, 49],

$$||e||^2_{\lambda,\mu} \leq Est_{\alpha,\beta}, \quad where \quad \lambda = \epsilon - \frac{\alpha}{2}, \mu = 1 - \frac{1}{\beta}.$$

proof.
By (5.13),

$$||e||^2_{\epsilon,1} = \epsilon||\nabla(u-u_h)||^2 + ||\sqrt{c}(u-u_h)||^2,$$

using estimate (5.17),

$$(\epsilon - \frac{\alpha}{2})||\nabla(u-u_h)||^2 + (1-\frac{1}{\beta})||\sqrt{c}(u-u_h)||^2 \leq Est_{\alpha,\beta}.$$

Employing (5.12) with $\lambda = \epsilon - \frac{\alpha}{2}$ and $\mu = 1 - \frac{1}{\beta}$ leads to

$$||e||^2_{\lambda,\mu} = (\epsilon - \frac{\alpha}{2})||\nabla(u-u_h)||^2 + (1-\frac{1}{\beta})||\sqrt{c}(u-u_h)||^2,$$

which is less than or equal $Est_{\alpha,\beta}$. Thus,

$$||e||^2_{\lambda,\mu} \leq Est_{\alpha,\beta}.\square \tag{5.18}$$

Let us assume that we have minimised the upper bound $Est_{\alpha,\beta}$, i.e., that we have found the optimal parameters $\gamma_{opt}, y^\star_{opt}, v_{opt}$, and define $\underline{Est_{\alpha,\beta}}=Est_{\alpha,\beta}(\gamma_{opt}, y^\star_{opt}, v_{opt}, u_h)$.
Let $y^\star = \epsilon\nabla u$ and $v = u - u_h$, then $y^\star \in H(\Omega, div)$ and $v \in H^1_0$. Substitute $y^\star$ and v in (5.16) to get

$$\underline{Est_{\alpha,\beta}} = \frac{1}{2\alpha}\Bigg((1+\gamma)||\epsilon\nabla u - \epsilon\nabla u_h + \epsilon\nabla(u-u_h)||^2_{L_2} + (1+\frac{1}{\gamma})C^2_\Omega||f - b.\nabla u_h - cu_h + \nabla.\epsilon\nabla u - c(u-u_h)+$$

$$-b.\nabla(u-u_h)||^2_{L_2}\Big)+\beta||\sqrt{\bar{c}}(u-u_h)||^2_{L_2}+\int_\Omega(\epsilon\nabla(u-u_h).\nabla u_h+b.\nabla u_h(u-u_h)+c(u-u_h)u_h-f(u-u_h))dx.$$

Note that

$$f+\nabla.\epsilon\nabla u-b.\nabla u-cu=f-(-\epsilon\Delta u+b.\nabla u+cu)=f-f=0.$$

Now, rearrangement terms in above equation implies

$$\underline{Est_{\alpha,\beta}}\leq\frac{1}{2\alpha}(1+\gamma)||2\epsilon\nabla(u-u_h)||^2+\int_\Omega(\epsilon\nabla(u-u_h).\nabla u_h+b.\nabla u_h(u-u_h)+c(u-u_h)u_h+$$

$$-f(u-u_h))dx+\beta||\sqrt{\bar{c}}(u-u_h)||^2$$

$$=\frac{2(1+\gamma)\epsilon^2}{\alpha}||\nabla(u-u_h)||^2+\beta||\sqrt{\bar{c}}(u-u_h)||^2-\int_\Omega\big(\epsilon\nabla(u-u_h).\nabla(u-u_h)+b.\nabla(u-u_h)(u-u_h)+$$

$$+c(u-u_h)(u-u_h)+f(u-u_h)\big)dx+\int_\Omega\big(\epsilon\nabla(u-u_h).\nabla u+b.\nabla u(u-u_h)+c(u-u_h)u\big)dx.$$

But

$$a(u-u_h,u-u_h)=\int_\Omega\big(\epsilon\nabla(u-u_h).\nabla(u-u_h)+b.\nabla(u-u_h)(u-u_h)+c(u-u_h)(u-u_h)\big)dx$$

$$=\epsilon||\nabla(u-u_h)||^2+||\sqrt{\bar{c}}(u-u_h)||^2,\text{ where }\bar{c}=c-\tfrac{1}{2}\nabla.b,$$

and

$$\int_\Omega\big(\epsilon\nabla(u-u_h).\nabla u+b.\nabla u(u-u_h)+c(u-u_h)u\big)dx=\int_\Omega f(u-u_h)dx.$$

So,

$$\underline{Est_{\alpha,\beta}}\leq\big(\frac{2(1+\gamma)\epsilon^2}{\alpha}-\epsilon\big)||\nabla(u-u_h)||^2+(\beta-1)||\sqrt{\bar{c}}(u-u_h)||^2,$$

thus,

$$\underline{Est_{\alpha,\beta}}\leq||e||^2_{(\frac{2(1+\gamma)}{\alpha}\epsilon^2-\epsilon,\beta-1)}\ . \tag{5.19}$$

Theorem 13.

$$||e||^2_{(\epsilon-\frac{\alpha}{2}),(1-\frac{1}{\beta})} \leq \underline{Est_{\alpha,\beta}} \leq ||e||^2_{(\frac{2(1+\gamma)}{\alpha}\epsilon^2-\epsilon,\beta-1)} \,. \tag{5.20}$$

proof.
The proof is as above, where the first inequality is clear since $\underline{Est_{\alpha,\beta}}$ is the minimised value of the upper bound $Est_{\alpha,\beta}$, thus

$$Est_{\alpha,\beta} \leq \underline{Est_{\alpha,\beta}}.$$

Now, by (5.18) the first inequality above is obtained, and the second part of the inequality is a consequence of (5.19).□

Conclusion

In this thesis we reviewed some basic and general theory of the finite element method. We also discussed the variational formulation and discretization of one and two dimensional problems. After that, the error estimation in its both types, a posteriori and a priori is explained.

The main goal of this thesis is to find a posteriori error estimations for Poisson, reaction-diffusion, and convection- diffusion problems with homogeneous and mixed, Dirichlet/Neumann, boundary conditions. At the end, we will give some prospective points for the future work of this thesis; we will deepen more on the subject of the a posterior error estimates and study applied real-life problems. Also, it is of importance to implement the final results of the a posteriori estimators to obtain approximation with least errors for the partial differential equations.

Appendix

Matlab code for example 1

```
clear
clc
Interval_Lower_Bound=input('Interval_Lower_Bound=');
Interval_Upper_Bound=input('Interval_Upper_Bound=');
% inputting n is the number of subintervals needed as a partition.
Number_Of_Subintervals=input('Number_Of_Subintervals=  ');
n=Number_Of_Subintervals;
% calculating the width of each subinterval "h", where assumed a uniform mesh.
h=(Interval_Upper_Bound-Interval_Lower_Bound)/n;
x0=Interval_Lower_Bound;
a=Interval_Lower_Bound;
b=Interval_Upper_Bound;
syms x
for i=1:n-2
    diagonal12_s(i)=int(diff(funk2(x,h,x0+h*i))*diff(funk1(x,h,x0+h*(i+1))),x0+h*i,x0+h*(i+1));
end
for i=1:n-1
    central11_s(i)=int(diff(funk1(x,h,x0+h*i))*diff(funk1(x,h,x0+h*i)),x0+h*(i-1),x0+h*i)+...
        int(diff(funk2(x,h,x0+h*i))*diff(funk2(x,h,x0+h*i)),x0+h*i,x0+h*(i+1));
end
for i=1:n-2
    diagonal21_s(i)=int(diff(funk1(x,h,x0+h*(i+1)))*diff(funk2(x,h,x0+h*i)),x0+h*i,x0+h*(i+1));
```

```
end
for i=1:n-1
    s(i,i)=central11_s(i);
end
for i=1:n-2
    s(i,i+1)=diagonal12_s(i);
    s(i+1,i)=diagonal21_s(i);
end
for i=1:n-1
 bb(i)=int(funk1(x,h,x0+h*i)*f(x),x0+h*(i-1),x0+h*i)+int(funk2(x,h,x0+h*i)*f(x),x0+h*i,x0+h*(i+1));
end
bb;
bb=bb';
xx=double(s)\double(bb);
xx=xx';
z=[0 xx 0];
y=linspace(a,b,n+1);
plot(y,z,':r');
hold on
fplot('7*x-x^3-6',[a b]); % a=1, b=2 f=6x

--------------------------------------------------------------------------------------------------
function y=f(x)
y=6*x;

--------------------------------------------------------------------------------------------------
function y=funk1(x,h,x0)
y=(x-(x0-h))/h;

--------------------------------------------------------------------------------------------------
function y=funk2(x,h,x0)
y=(x0+h-x)/h;
```

Matlab code for example 2

```
%EXAMPLE OF THE FORM   -au"+bu'+cu=f(x),  u'(1)=u'(0)=constant=q. a,b,c in R.
%We take a=-1//b=-5//c=6//f(x)=6x-11//q=2.
syms x
scentral1(1)=int(diff(funk2(x,0.1,0))*diff(funk2(x,0.1,0)),0,0.1);
for i=2:10
    scentral1(i)=int(diff(funk1(x,0.1,0.1*i))*diff(funk1(x,0.1,0.1*i)),0.1*(i-1),0.1*i)...
        +int(diff(funk2(x,0.1,0.1*i))*diff(funk2(x,0.1,0.1*i)),0.1*i,0.1*(i+1));
end
scentral1(11)=int(diff(funk1(x,0.1,1))*diff(funk1(x,0.1,1)),0.9,1);
for i=1:10
    syms x
    scentral2(i)=int(diff(funk1(x,0.1,0.1*i))*diff(funk2(x,0.1,0.1*i)),0.1*i,0.1*(i+1));
end
syms s
for i=1:11
    s(1,1)=scentral1(1);
    s(11,11)=scentral1(11);
    s(i,i)=scentral1(i);
end
for i=1:10
    s(i+1,i)=scentral2(i);
    s(i,i+1)=scentral2(i);
end
s=-s;
%---------------------------------------------------------------------------
%---------------------------------------------------------------------------
syms x
ccentral1(1)=int(diff(funk2(x,0.1,0))*(funk2(x,0.1,0)),0,0.1);
for i=2:10
    ccentral1(i)=int(diff(funk1(x,0.1,0.1*i))*(funk1(x,0.1,0.1*i)),0.1*(i-1),0.1*i)+...
```

```
        int(diff(funk2(x,0.1,0.1*i))*(funk2(x,0.1,0.1*i)),0.1*i,0.1*(i+1));
end
ccentral1(11)=int(diff(funk1(x,0.1,1))*(funk1(x,0.1,1)),0.9,1);

for i=1:10
   syms x
    ccentral2(i)=int(diff(funk1(x,0.1,0.1*(i+1)))*(funk2(x,0.1,0.1*i)),0.1*i,0.1*(i+1));
end
for i=1:10
    syms x
    ccentral3(i)=int((funk1(x,0.1,0.1*(i+1)))*diff(funk2(x,0.1,0.1*i)),0.1*i,0.1*(i+1));
end
syms c
for i=1:11
    c(1,1)=ccentral1(1);
    c(11,11)=ccentral1(11);
    c(i,i)=ccentral1(i);
end
for i=1:10
   c(i+1,i)=ccentral3(i);
  c(i,i+1)=ccentral2(i);
end
c=-5*c;
%---------------------------------------------------------------------------
%---------------------------------------------------------------------------
syms x
mcentral1(1)=int((funk2(x,0.1,0))*(funk2(x,0.1,0)),0,0.1);
for i=2:10
    mcentral1(i)=int((funk1(x,0.1,0.1*i))*(funk1(x,0.1,0.1*i)),0.1*(i-1),0.1*i)+...
        int((funk2(x,0.1,0.1*i))*(funk2(x,0.1,0.1*i)),0.1*i,0.1*(i+1));
end
mcentral1(11)=int((funk1(x,0.1,1))*(funk1(x,0.1,1)),0.9,1);
```

```
for i=1:10
    syms x
   mcentral2(i)=int((funk1(x,0.1,0.1*(i+1)))*(funk2(x,0.1,0.1*i)),0.1*i,0.1*(i+1));
end
for i=1:10
  syms x
   mcentral3(i)=int((funk1(x,0.1,0.1*(i+1)))*(funk2(x,0.1,0.1*i)),0.1*i,0.1*(i+1));
end

syms m
for i=1:11
    m(1,1)=mcentral1(1);
    m(11,11)=mcentral1(11);
    m(i,i)=mcentral1(i);
end
for i=1:10
    m(i+1,i)=mcentral3(i);
    m(i,i+1)=mcentral2(i);
end
m=6*m;
a=s+c+m;
syms x
b(1)=int((6*x-11)*funk2(x,0.1,0),0,0.1)+2;
b(11)=int((6*x-11)*funk1(x,0.1,1),0.9,1)-2;
for i=2:10
    syms x
    b(i)=int((6*x-11)*funk1(x,0.1,0.1*(i-1)),0.1*(i-2),0.1*(i-1))+...
        int((6*x-11)*funk2(x,0.1,0.1*(i-1)),0.1*(i-1),0.1*i);
end
d=double(a)\double(b');
d=d';
 plot(0:0.1:1,d,'r*')
```

```
hold on

fplot('((exp(3)-1)/(2*(exp(3)-exp(2))))*exp(2*x)+((1-exp(2))/(3*(exp(3)-exp(2))))*exp(3*x)+x-1',[0,1

----------------------------------------------------------------------------------------------------
function y=funk1(x,h,x0)

y=(x-(x0-h))/h;

----------------------------------------------------------------------------------------------------
function y=funk2(x,h,x0)

y=(x0+h-x)/h;
```

Matlab code for example 3

```
function poi2D( )
%
% -div(grad u) = f, in [0,1]x[0,1],
%             u = 0, on boundary
%
clear all, clc
% triangulation
g = [2 0 1 0 0 1 0;
     2 1 1 0 1 1 0;
     2 1 0 1 1 1 0;
     2 0 0 1 0 1 0]';
[p,e,t] = initmesh(g,'hmax',0.05);
figure(1); clf
pdemesh(p,e,t)
% assemble
[A,b] = assemble(p,e,t,'f');
% solve
U = A\b;
% visualize
figure(2); clf
pdesurf(p,t,U)
xlabel('x'), ylabel('y'), zlabel('U(x,y)')
% subroutines ------------------------------------------------------------

function z = f(x,y)
z = 2*pi^2*sin(pi*x).*sin(pi*y);

function [A,b] = assemble(p,e,t,f)
Nt = size(t,2);
Np = size(p,2);
```

```
Ne = size(e,2);
A = sparse(Np,Np);
b = zeros(Np,1);
for i = 1:Nt
  n = t(1:3,i);
  x = p(1,n);
  y = p(2,n);
  dx = [y(2)-y(3); y(3)-y(1); y(1)-y(2)];
  dy = [x(3)-x(2); x(1)-x(3); x(2)-x(1)];
  area = 0.5*abs(x(2)*y(3)-y(2)*x(3)-x(1)*y(3)+y(1)*x(3)+x(1)*y(2)-y(1)*x(2));
  A(n,n) = A(n,n) + (dx*dx'+dy*dy')/4/area;
  b(n) = b(n) + area/12*[2 1 1; 1 2 1; 1 1 2]*feval('f',x,y)';
end
% BC
for i = 1:Ne
  n = e(1,i);
  A(n,n) = 1e6;
  b(n) = 0;
end
```

Bibliography

[1] M. Ainsworth and J.T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, John Wiley and Sons Interscience, New York, 2000.

[2] V.B. Andreev and N.V. Kopteva, *Investigation of difference schemes with an approximation of the first derivative by a central difference relation* , Zh. Vychisl. Mat.i Mat. Fiz., pp. 101-117, 36(8)(1996).

[3] J.H. Argyris, *Energy Theorems and Structural Analysis*, Part 1, Aircraft Eng., pp. 383-394, 26 (1954).

[4] J.H. Argyris and S. Kelsey, *Energy Theorems and Structural Analysis*, Butterworths, London, (1960).

[5] M. Asadzadeh, *An Introduction to the Finite Element Method (FEM) for Differential Equations*, (2012).

[6] K. Sharath Babu and N. Srinivasacharyulu, *Numerical study of convection diffusion problem in two- dimensional space*, IJRRAS, (2010).

[7] I. Babuska and W.C. Rheinboldt, *A Posteriori Error Estimate for the Finite Element Method*, SIAM J. Numer. Anal., pp. 1597-1615, 15 (1978).

[8] I. Babuska and W.C. Rheinboldt, *Adaptive Approaches and Reliability Estimates in Finite element Analysis*, Comp. Meth. Appl. Mech. Eng., pp. 519-540, 17/18 (1979).

[9] I. Babuska and T. Strouboulis. *The Finite Element Method and Its Reliability*, Oxford University Press Inc., New York, (2001).

[10] E. Barkanov, *INTRODUCTION TO THE FINITE ELEMENT METHOD*, Institute of Materials and Structures Faculty of Civil Engineering, (2001). Riga Technical University

[11] KJ. Bathe, *Finite element procedures, Englewood Cliffs*, NJ: Prentice Hall, (1996).

[12] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Method*, Springer-Verlag, New York, (1994).

[13] H. Brezis and R. E. L. Turner, *On a class of superlinear elliptic problems. Commun. Partial Differential Equations*, pp. 601-614, 2 (1977).

[14] Ph. Ciarlet, *Lectures on The Finite Element Method*, Tata Institute of Fundamental Research Bombay, (1975).

[15] R.W. Clough, *The Finite Element method in Plane Stress Analysis*, In: Proc. 2nd ASCE Conf. on Electronc Comp, Pittsburgh Pa, (1960).

[16] R.W. Clough and E.W. Wilson, *Stress Analysis of a Gravity Dam by the Finite Element Method*, In: Proc. Symp. on the use of Computers in the Civil Eng. Lab., Lisbon, Portugal, Nacional de Engenharia Civil, (1962).

[17] R. Courant, *Variational Method for the Solution of Problems of Equilibrium and Vibration*, Bulletin of the American Mathematical Society, pp.1-23, 49 (1943).

[18] M. Del Pino, M. Musso and F. Pacard, *Boundary singularities for weak solutions of semi-linear elliptic problems*, J. Funct. Anal., pp. 241-272, 253 (2007).

[19] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, *Computational Differential Equations*, (2009).

[20] J. E. Flaherty, *Finite element analysis lecture notes*, spring, CSCI, math 6860, Troy, New York 12180, (2000).

[21] K.O. Friedrichs, *A Finite Difference Scheme for the Neumann and Dirichlet Problem*, Technical report University Rep. TID 4500, NYO 9760, ACE Computing and Applied Mathemaics Center, Courant Institute of Mathematical Sciences, New York, (1962).

[22] B. Gidas and J. Spruck, *A priori bounds for positive solutions of nonlinear elliptic equations*, Comm. Partial Differential Equations 6, pp. 883-901, 8 (1981).

[23] T.Gratsch, K. Bathe, *A posteriori error estimation techniques in practical finite element analysis*, Computers and Structures, pp. 235-265, 83 (2005).

[24] F. M. Guerra, *Finite element analysis for the adaptive method of rezoning*, PhD thesis, The University of Texas at Austin, (1977).

[25] A. Hannukainen and S. Korotov, *Techniques for a posteriori error estimation in terms of linear functionals for elliptic type boundary value problems*, Far East J. Appl. Math., pp. 289-304, 21 (2005).

[26] A. Hrennikov, *Solution of problems in elasticity by the frame work method*, Journal of Applied Mechanics, pp. 169-175, 8(1941).

[27] C. Johnson, *Numerical solution of partial differential equations by the finite element method*, Cambridge University Press, (1987).

[28] M. Izadi, *Streamline Diffusion Finite Element Method for Coupling Equations of Nonlinear Hyperbolic Scalar Conservation Laws*, M.Sc. Thesis, (2005).

[29] S. Korotov, *A posteriori error estimation for linear elliptic problems with mixed boundary conditions*, Helsinki University of Technology, Institute of Mathematics, Research Reports A495 (2006).

[30] S. Korotov, *Global a posteriori error estimates for convection-reaction-diffusion problems*, Appl. Math. Model., pp. 1579-1586, 32 (2008).

[31] I. Kosirova, *Boundedness, a priori estimates and existence of solutions of nonlinear elliptic problems*, Bratislava, Dissertation Thesis, (2011).

[32] D. Kuzmin, A. Hannukainen, S. Korotov, *A new a posteriori error estimate for convection-reaction-diffusion problems*, J. Comput. Appl. Math., in press, pp. 70 - 78, 218 (2008).

[33] Z. Li, Z. Qiao, T. Tang, *Numerical solutions of partial differential equations-An introduction of finite difference and finite element method*, (2011).

[34] K.W.Morton, *Numerical solution of Convection-Diffusion Problems*, Appl. Math. and Mathematical Computation, Chapman and Hall, London, (1996).

[35] J. Necas, *Les methodes directes en theorie des equations elliptiques*, Academia, Praha, and Masson et Cie, Editeurs, Paris, (1967).

[36] R. H. Nochetto, K. G. Siebert and A.Veeser, *Theory of Adaptive Finite Element Methods: An Introduction*, Multiscale, Nonlinear and Adaptive Approximation, (2009).

[37] R. Nussbaum, *Positive solutions of some nonlinear elliptic boundary value problems*, J. Math. Anal. Appl., pp. 461-482, 51 (1975).

[38] J. T. Oden and L. Demkowicz, *A survey of adaptive finite element methods in computational mechanics*, In A. K. Noor and J. T. Oden, editors, State-of-the-Art Surveys on Computational Mechanics. The American Society of Mechanical Engineers, New York, (1989).

[39] J. T Oden, S. Prudhomme, and L. Demkowicz, *A Posteriori Error Estimation for Acoustic Wave Propagation Problems*, Archives of Computational Methods in Engineering, pp. 343-389, (2005).

[40] L.A. Oganesjan, *Numerical Solution of Plates, In: Solution of Engineering Problems on Computers*, Lensovnarchoz CBTI, Sbronik, Russian, 1663.

[41] A. Ostrowski, Recherches sur la methode de Graeffe et les zeros des polynomes et des series des Laurent. Acta Math., pp. 99-257, 72 (1940).

[42] A. Otto, *Methods of Numerical Simulation in Fluids and Plasmas*, Spring, (2013).

[43] P. Quittner and Ph. Souplet, *A priori estimates and existence for elliptic systems via bootstrap in weighted Lebesgue spaces*, Arch. Rational Mech. Anal., pp. 49-81, 174 (2004).

[44] S.I Repin, *A posteriori error estimation for nonlinear variational problems by duality theory*, Zapiski Nauchnih Seminarov, V.A. Stekov Mathematical Institute (POMI), pp. 201-214, 243 (1997).

[45] S.I. Repin, *A posteriori error estimation for variational problems with uniformly convex functionals*, Math. Comp. 69,pp. 481-500, , 230 (2000).

[46] S. Repin, S. Sauter and A. Smolianski, *A posteriori error estimation for the Dirichlet problem with account of the error in the approximation of boundary conditions*, Computing 70, pp. 205-233, (2003).

[47] S. Repin, S. Sauter, A. Smolianski, *A posteriori error estimation for the Poisson equation with mixed Dirichlet/Neumann boundary conditions*, Journal of Computational and Applied Mathematics, pp. 601-612 ,164 (2004).

[48] W. Ritz, *A new method for the solution of certain variational problems of mathematical physics*, J. Reine Angew. Math, pp. 1-61, 135 (1909).

[49] K. Segeth, *A review of some a posteriori error estimates for adaptive finite element methods*, Mathematics and Computers in Simulation, pp. 1589-1600, 80 (2010).

[50] M. Stynes, *Steady-state convection-diffusion problems*, Acta Numerica, pp. 445-508, (2005).

[51] M. Stynes, *Numerical methods for convection-diffusion problems or The 30 years war*, Department of Mathematics, National University of Ireland, Cork, Ireland, (2013).

[52] E. Suli, *Lecture Notes on Finite Element Methods for Partial Differential Equations*, University of Oxford, (2012).

[53] M.J. Turner, R.W. Clough, H.C. Martin and L.T. Topp, *Stiffness and Deflection Analysis of Complex Structures*, J. Aeronaut. Sci., pp. 805-823, 25 (1956).

[54] R. E. L. Turner, *A priori bounds for positive solutions of nonlinear elliptic equations in two variables*, J. Duke Math., pp. 759-774, 41 (1974).

[55] T. Vejchodsky, *On a posteriori error estimation strategies for elliptic problems*, in J. Privratska, J. Prihonska, D. Andrejsova (Eds.), Proceedings of international conference ICPM'05, Liberec, Czech Republic, pp. 373-386, (2005).

[56] R. Verfurth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-refinement Techniques*, Wiley-Teubner, Stuttgart, (1996).

[57] O. C. Zienkiewicz and Y. K. Cheung, *The Finite Element Method in Structural and Continuum Mechanics*, McGraw-Hill London, (1967).

[58] M. Zlamal, *On the Finite Element Method*, Numer. Math., pp. 394-409, 12 (1968).

[59] F. Zuger, *Solution of Non-Homogeneous Dirichlet Problems with FEM*, zurich university, (2013).