5-1982

Numerical Methods (Finite Element) for Time-Dependent Partial Differential Equations

Masaji Watanabe
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/etd

Part of the Mathematics Commons

Recommended Citation
Watanabe, Masaji, "Numerical Methods (Finite Element) for Time-Dependent Partial Differential Equations" (1982). All Graduate Theses and Dissertations. 6978.
https://digitalcommons.usu.edu/etd/6978
NUMERICAL METHODS (FINITE ELEMENT) FOR TIME-DEPENDENT
PARTIAL DIFFERENTIAL EQUATIONS

by

Masaji Watanabe

A thesis submitted in partial fulfillment
of the requirements for the degree
of
MASTER OF SCIENCE
in
Mathematics

UTAH STATE UNIVERSITY
Logan, Utah
1982
ACKNOWLEDGMENT

I deeply appreciate my major professor Dr. Chris Coray who has enthusiastically shared time and knowledge in the entire preparation of this paper and without whom this work would not have been possible. I also would like to thank each member of the committee; Dr. Lawrence Cannon and Dr. Russell Thompson, for their helpful guidance, and especially Dr. Ian Anderson for helping me with calculus of variations.

Masaji Watanabe
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENT</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>iv</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td><strong>Chapter</strong></td>
<td></td>
</tr>
<tr>
<td>I. PRELIMINARIES</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>The Variational Principle and Elliptic Equations</td>
<td>3</td>
</tr>
<tr>
<td>Hyperbolic and Parabolic Equations</td>
<td>6</td>
</tr>
<tr>
<td>Hilbert Spaces</td>
<td>10</td>
</tr>
<tr>
<td>II. METHODS FOR BOUNDARY VALUE PROBLEMS</td>
<td>15</td>
</tr>
<tr>
<td>Ritz Method</td>
<td>15</td>
</tr>
<tr>
<td>Self-Adjoint Positive Definite Operators</td>
<td>23</td>
</tr>
<tr>
<td>Galerkin Method</td>
<td>29</td>
</tr>
<tr>
<td>$H_0^1(\Omega)$-Elliptic Bilinear Forms</td>
<td>35</td>
</tr>
<tr>
<td>Time-Dependent Problems</td>
<td>36</td>
</tr>
<tr>
<td>III. METHODS FOR TIME-DEPENDENT PROBLEMS</td>
<td>44</td>
</tr>
<tr>
<td>Laplace Transform</td>
<td>44</td>
</tr>
<tr>
<td>Variational Methods</td>
<td>50</td>
</tr>
<tr>
<td>Semi-Discrete Galerkin Method</td>
<td>65</td>
</tr>
<tr>
<td>Error Analysis of Semi-Discrete Galerkin Method</td>
<td>74</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>84</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>86</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. Approximate solutions to the heat equation .......... 89
Figure 2. Approximate solutions to the wave equation .......... 89
This paper surveys reasons why the Ritz method and the Galerkin method are not efficient and why these methods can not be applied directly, for time dependent problems. It also introduces methods that are used for those problems. For a linear boundary value problem defined by a positive definite symmetric (self-adjoint) operator, the existence and the convergence of the Ritz approximation are guaranteed. In non-symmetric case, Lax-Milgram lemma assures the existence and the convergence of the Galerkin approximation for $H^1_0(\Omega)$-elliptic operator. Since time-dependent problems are hyperbolic or parabolic, the existence and the convergence of approximations by those methods are not guaranteed. Moreover, those methods were originally developed for boundary value problems. Thus new techniques are introduced in order to extend those methods to initial-boundary value problems.
CHAPTER I
PRELIMINARIES

Introduction

The finite element method was originated by engineers in 1950's for the approximate solutions for boundary value problems and has been developed considerably for the last two decades. A historical account of the finite element method is found in Tong and Rossettos [18]. It is a method to construct the approximate solution to a differential equation of the form

$$ Au = f, \quad u = u(x), \quad f = f(x), \quad x \in \Omega \subset \mathbb{R}^n \quad (1.1), $$

with arbitrary boundary conditions, where $A$ is a differential operator. Suppose $L = L(u)$ is a functional whose necessary condition to minimise over a certain set of functions $\mathcal{U}$ (the set of admissible functions) satisfies the partial differential equation (1.1). Instead of minimising $I$ over the infinite dimensional space $\mathcal{U}$, we minimise $I$ over

$$ u(x) \in \sum_{i=1}^{n} \alpha_i \phi_i (x), $$

where $\alpha_i$'s are unknown coefficients to be determined and $\phi_i$'s are known functions (trial functions or basis functions) that are specifically piecewise polynomials. The idea is to break up the domain $\Omega$ of the partial differential equation into small pieces $T_1, T_2, \ldots, T_m$, called elements, and to choose trial functions in such a way that each $\phi_i$ is non-zero on some $T_i$'s and vanishes on the most of $T_{i'}$'s. A precise description of these ideas can be found in Oden and Reddy [11].

Roughly speaking, we can classify the finite element method by its formulations:

The finite element method based on;

a) the variational principle,

b) the weak formulation.

Suppose that we are given a boundary value problem,
\[
\begin{align*}
Au &= f, \quad u = u(\bar{x}), \quad f = f(\bar{x}), \quad \bar{x} \in \Omega \subset \mathbb{R}^n \quad (1.1),
\end{align*}
\]

with an arbitrary boundary conditions, where \( A \) is a differential operator. Suppose \( I = I(u) \) is a functional whose necessary condition to minimize over a certain set of functions \( H \) (the set of admissible functions) satisfies the partial differential equation (1.1). Instead of minimizing \( I \) over the infinite dimensional space \( H \), we minimize \( I \) over a finite dimensional subspace \( K_n = \langle \phi_1, \ldots, \phi_n \rangle \) of \( H \), by determining unknown coefficients \( a_i \)'s and obtain the approximate solution

\[
U(\bar{x}) = \sum_{i=1}^{n} a_i \phi_i(\bar{x}).
\]

This is the Ritz method based on the variational principle. In the next section, we will discuss the variational principle and in chapter II, we will illustrate the Ritz method with examples.

The Galerkin method is based on the weak formulation of (1.1), that is, the problem is to find a solution of

\[
(Au, v) = (f, v) \quad \text{for all } v \in H,
\]

where \((u, v)\) is the inner product in \( H \). We seek an approximate solution

\[
U = \sum_{i=1}^{n} a_i \phi_i(\bar{x}) \in K_n
\]

which satisfies

\[
(AU, V) = (f, V) \quad \text{for all } V \in K_n.
\]

We will give detailed discussion of the Galerkin method in chapter II.

It was natural to extend the finite element method to initial-boundary value problems. However, there are some problems associated with this formulation. This thesis discusses the problems associated with the formulations of the method for initial-boundary value problems,
introduces the methods that are used for them.

In chapter I, we will introduce the variational principle, and examples of elliptic, hyperbolic, and parabolic partial differential equations. Sobolev spaces and some results in general Hilbert spaces are also introduced. In chapter II, we will illustrate the Ritz method and Galerkin method with examples, and investigate reasons why these methods are efficient for some partial differential equations and not for some others. We will also discuss the problems associated with formulations of these methods for initial-boundary value problems. In chapter III, we will introduce the methods that are used for initial-boundary value problems. The method of Laplace transform, the variational methods, and the semi-discrete Galerkin method are introduced.

The Variational Principle and Elliptic Equations

Many physical problems can be described by the variational principle, that is, a function which minimizes a certain integral (the fundamental integral) is necessarily a solution of the corresponding differential equation (the Euler-Lagrange equation). For example, let \( u = u(x,y) \) be a \( C^2 \)-function defined on a region \( \Omega \subset \mathbb{R}^2 \), and be fixed on \( \partial \Omega \), the boundary of \( \Omega \), that is, \( u = g \) on \( \partial \Omega \) for some \( g = g(x,y) \). Suppose \( u \) is a minimum for the functional

\[
I(u) = \int_{\Omega} F(x,y,u,u_x,u_y) \, dx \, dy
\]

among \( H \), the set of all \( C^2 \)-functions which agree with \( u \) on \( \partial \Omega \), where \( F \) is a \( C^2 \)-function. Let \( \eta = \eta(x,y) \) be a \( C^2 \)-function which vanishes on \( \partial \Omega \) and let

\[
\phi(\varepsilon) = I(u + \varepsilon \eta) = \int_{\Omega} F(x,y,u + \varepsilon \eta,u_x + \varepsilon \eta_x,u_y + \varepsilon \eta_y) \, dx \, dy .
\]
Then \( u + \varepsilon n \in \mathcal{H} \), and since \( u \) minimize \( I, \frac{dI}{d\varepsilon} \bigg|_{\varepsilon=0} = 0 \) or

\[
\frac{dI}{d\varepsilon} \bigg|_{\varepsilon=0} = \iint_{\Omega} \left( \eta F u + \eta_x F u_x + \eta_y F u_y \right) dx dy
\]

\[
= \iint_{\Omega} \left( \eta F + \eta \frac{\partial F}{\partial x} u_x + \eta \frac{\partial F}{\partial y} u_y \right) dx dy + \iint_{\Omega} \left( \eta u - \frac{\partial F}{\partial x} u_x - \frac{\partial F}{\partial y} u_y \right) dx dy
\]

\[
= \iint_{\Omega} \left( \frac{\partial}{\partial x} (\eta F) + \frac{\partial}{\partial y} (\eta F u_y) \right) dx dy + \iint_{\Omega} \left( \eta u - \frac{\partial F}{\partial x} u_x - \frac{\partial F}{\partial y} u_y \right) dx dy = 0
\]

By Green's theorem,

\[
\iint_{\Omega} \left( \frac{\partial}{\partial x} (\eta F_u) + \frac{\partial}{\partial y} (\eta F u_y) \right) dx dy = \int_{\partial \Omega} \eta F u_x - \eta F u_y dy
\]

and since \( \eta = 0 \) on \( \partial \Omega \),

\[
\int_{\partial \Omega} \eta F u_x - \eta F u_y dy = 0
\]

and we have

\[
\iint_{\Omega} \left( \eta u - \frac{\partial F}{\partial x} u_x - \frac{\partial F}{\partial y} u_y \right) dx dy = 0
\]

Since this must hold for an arbitrary \( C^2 \)-function \( \eta \) which vanishes on \( \partial \Omega \), we conclude that \( u \) satisfies

\[
F_u - \frac{\partial F}{\partial x} u_x - \frac{\partial F}{\partial y} u_y = 0
\]

subject to \( u = g \) on \( \partial \Omega \)

which is the second order partial differential equation. The functional

\[
I(u) = \iint_{\Omega} F(x, y, u, u_x, u_y) dx dy
\]

is called the fundamental integral, the partial differential equation

\[
F_u - \frac{\partial F}{\partial x} u_x - \frac{\partial F}{\partial y} u_y = 0
\]
the Euler-Lagrange equation, and \( H \) the set of admissible functions for the problem. A complete discussion of the variational principle is given in Gelfand and Fomin [8].

Now we turn to a physical problem in which the variational principle is used. The following discussion is based on Courant and Hilbert [5]. Consider a flexible membrane stretched over a region \( \Omega \) in \( x,y \)-plane with a constant tension \( \tau \). Let \( u = u(x,y) \) be a perpendicular displacement of the membrane at \( (x,y) \). Suppose that the membrane is fixed at each point on the boundary of \( \Omega \), that is, \( u = g \) on \( \partial \Omega \) for some \( g = g(x,y) \). The potential energy \( V \) of the membrane is given by the product of the tension \( \tau \) and the change in the area of the membrane,

\[
V = \tau \left( \iint_{\Omega} \sqrt{1 + u_x^2 + u_y^2} \, dx \, dy - S \right)
\]

where \( S \) is the area of \( \Omega \). Suppose that the displacement \( u \) is so small that we can ignore higher powers of \( u_x^2 + u_y^2 \). Since

\[
\sqrt{1 + u_x^2 + u_y^2} = 1 + \frac{1}{2}(u_x^2 + u_y^2) - \frac{1}{8}(u_x^2 + u_y^2)^2 + \cdots,
\]

\[
V = \frac{\tau}{2} \iint_{\Omega} (u_x^2 + u_y^2) \, dx \, dy
\]

The equilibrium position \( u \) of the membrane minimizes the potential energy \( V \) among the set \( H \) of \( C^2 \)-functions which agree with \( g \) on \( \partial \Omega \). Hence, by the variational principle, we obtain the Euler-Lagrange equation

\[
U_{xx} + U_{yy} = 0
\]

subject to \( u = g \) on \( \partial \Omega \),

which is two dimensional Laplace's equation.

If an external force \( f(x,y) \) acts on the membrane at \( (x,y) \), the potential energy \( V \) becomes
and again by the variational principle, we obtain the Euler-Lagrange equation

\[ V = \frac{1}{2} \iint_{\Omega} (u_x^2 + u_y^2) \, dx \, dy - \iint_{\Omega} fu \, dx \, dy \]

which is called Poisson's equation.

In general, a partial differential equation is said to be linear if it is linear in the unknown function and all its derivatives, and it is said to be quasilinear if it is linear in the highest derivatives. The order of a partial differential equation is defined to be the order of the highest derivative that occurs. A second order quasilinear partial differential equation

\[ au_{xx} + bu_{xy} + cu_{yy} + F(x,y,u,u_x,u_y) = f \]

where \( u, f, a, b, \) and \( c \) are functions defined on \( \Omega \subseteq \mathbb{R}^2 \), is said to be

a) elliptic if \( b^2 < 4ac \)

b) hyperbolic if \( b^2 > 4ac \)

c) parabolic if \( b^2 = 4ac \)

for each \((x,y) \in \Omega\). Hence our previous example \( u_{xx} + u_{yy} + \frac{1}{\tau} f = 0 \) is an example of a linear elliptic partial differential equation with the boundary condition \( u = g \) on \( \partial \Omega \).

**Hyperbolic and Parabolic Equations**

In this section, we discuss physical situations in which hyperbolic and parabolic partial differential equations occur. More detailed discussions are given in Courant and Hilbert [5], Sagan [12], Tikhonov and
In many physical problems, the laws of motion are described by Hamilton's principle. Let $J$ be the time integral of the difference between the kinetic and potential energies. Hamilton's principle states:

"The actual motion makes the value of the integral $J$ stationary with respect to all neighboring virtual motions which lead from the initial to the final position of the system in the same interval of time." See [5], p.243.

Using Hamilton's principle, we can derive a simple partial differential equation of hyperbolic type.

Suppose that a flexible string, stretched between two points $0$ and $t$ on the $x$-axis and fixed at each end point, is executing a small vibration. Let $u(x,t)$ be the vertical displacement of the point on the string from the $x$-axis at $x$ at time $t$. Since the string is fixed at $x = 0$ and $x = t$, we have $u(0,t) = u(t,t) = 0$. We assume that the string has a constant tension $T$ and a constant density $\rho$ and that $u$ is so small that higher powers of derivative of $u$ can be ignored compared with lower ones. The Kinetic energy $\Delta T$ of a piece of the string with length $\Delta S$ is given by

$$\Delta T = \frac{\rho \Delta S}{2} u_t^2.$$ 

The length $\Delta S$ of a piece of the string on $[x_0,x_1]$ is

$$\Delta S = \int_{x_0}^{x_1} \sqrt{1 + u_x^2} \, dx = \int_{x_0}^{x_1} \left(1 + \frac{1}{2} u_x^2 - \frac{1}{8} u_x^4 \ldots \right) dx.$$ 

By assumption $\Delta S$ is approximated by

$$\Delta S = \int_{x_0}^{x_1} dx = x_1 - x_0 ,$$

and if $\Delta x$ is the corresponding change in $x$ to the length $\Delta S$,

$$\Delta x \approx \Delta S \text{ and } \Delta T \approx \frac{\rho \Delta x}{2} u_t^2.$$
and the total kinetic energy $T$ over the whole string is

$$T = \frac{1}{2} \int_0^1 u_t^2 \, dx.$$ 

The potential energy $V$ of the string is given by the product of tension $T$ and the increase in length of the string;

$$V = T \left( \int_0^1 \sqrt{1 + u_x^2} \, dx - 1 \right) = \tau \left( \int_0^1 (1 + \frac{1}{2}u_x^2 - \frac{1}{8}u_x^4 \cdots)dx - 1 \right).$$

Again by assumption,

$$V = \frac{1}{2} \int_0^1 u_x^2 \, dx.$$ 

According to Hamilton's principle, if $u$ is the actual motion of the string on the time interval $[t_0,t_1]$, $u$ makes the integral

$$J = \int_{t_0}^{t_1} (T - V) \, dt = \frac{1}{2} \int_{t_0}^{t_1} \left( \rho u_t^2 - \tau u_x^2 \right) \, dx \, dt$$

stationary among the set $H$ of admissible functions which agree with $u$ at $t = t_0$ and $t = t_1$, that is, $u(x,t_0) = u_0(x)$ and $u(x,t_1) = u_1(x)$ for fixed $u_0$ and $u_1$. Therefore, we obtain

$$\tau u_{xx} - \rho u_{tt} = 0, \text{ or } u_{tt} = c^2 u_{xx}, \text{ where } c^2 = \frac{T}{\rho},$$

subject to

$$u(0,t) = u(1,t) = 0$$

$$u(x,t_0) = u_0(x),$$

$$u(x,t_1) = u_1(x),$$

which is the one dimensional wave equation. This is an example of a linear hyperbolic partial differential equation.

In the remainder of this section, we will introduce a physical example of a parabolic partial differential equation, namely, the heat equation. The derivation of the heat equation is given in Tikhonov and
Suppose we have a rod of length \( t \) with end points \( x = 0 \) and \( x = t \), and let \( u = u(x,t) \) be the temperature of the rod at \( x \) at time \( t \). We assume that the rod is so thin that any point in its cross section at \( x \) has a constant temperature. Suppose we know the temperature distribution of the rod at the initial time \( t = t_0 \) and the temperature at each end point at arbitrary time \( t \):

\[
\begin{align*}
  u(x,t_0) &= u_0(x), \\
  u(0,t) &= w_0(t), \\
  u(t,t) &= w_1(t).
\end{align*}
\]

Then heat starts flowing from the place of higher temperature to the place of lower temperature, and \( u = u(x,t) \) satisfies

\[
\frac{\partial}{\partial x}(k \frac{\partial u}{\partial x}) = \rho c \frac{\partial u}{\partial t}
\]

where \( k = k(x) \) is the coefficient of thermal conductivity which is positive and depends on the material of the rod. \( \rho = \rho(x) \) is the density of the material and \( c = c(x) \) is its specific heat. If \( k, c, \) and \( \rho \) are constants, we simply have

\[
\begin{align*}
  u_t &= c^2 u_{xx} \quad \text{where } c^2 = \frac{k}{\rho}, \\
  \text{subject to } u(x,t_0) &= u_0(x), \\
  u(0,t) &= w_0(t), \\
  u(t,t) &= w_1(t).
\end{align*}
\]

This is an example of a linear parabolic partial differential equation with initial-boundary conditions.
Hilbert Spaces

In this section we introduce Sobolev spaces and some important results in Hilbert spaces in general. It is not our intention to prove these results, and references are cited for proofs.

Let $\Omega$ be a bounded domain in $\mathbb{R}^n$, we denote by $L^2(\Omega)$ the set of all functions $f$ defined on $\Omega$ such that

$$\int_{\Omega} f^2(x) \, dx < \infty.$$  

Then $L^2(\Omega)$ is an inner product space with the inner product

$$(u,v) = \int_{\Omega} u(x)v(x) \, dx.$$  

The corresponding norm is denoted by $\| \|$ , so that

$$\| u \|^2 = \int_{\Omega} [u(x)]^2 \, dx.$$  

Lemma 1.1

$L^2(\Omega)$ is a Hilbert space.

Proof. See Adams [1]

Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ where each $\alpha_1$ is a non-negative integer. We define $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_n$ and

$$D^\alpha u = \frac{\partial^{\alpha_1} u}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2} u}{\partial x_2^{\alpha_2}} \ldots \frac{\partial^{\alpha_n} u}{\partial x_n^{\alpha_n}}.$$  

The Sobolev space $H^k_2(\Omega)$ is the set of all functions $u$ defined on $\Omega$ such that

$$\| D^\alpha u \|^2 < \infty \text{ for each } \alpha \text{ such that } |\alpha| \leq k,$$  

where $\| D^\alpha u \|^2 = \int_{\Omega} (D^\alpha u)^2 \, dx$ as defined earlier. Then $H^k_2(\Omega)$ is an inner product space with the inner product
\[(u,v)_{k,\Omega} = \sum_{|\alpha| \leq k} (D^\alpha u, D^\alpha v),\]

and the norm

\[\| u \|_{k,\Omega}^2 = \sum_{|\alpha| \leq k} \| D^\alpha u \|_2^2.\]

We define the semi-norm \[|u|_{k,\Omega}^2\] by

\[|u|_{k,\Omega}^2 = \sum_{|\alpha| = k} \| D^\alpha u \|_2^2.\]

For example if \(\Omega \subseteq \mathbb{R}^2\) and \(u \in H^2(\Omega)\), then

\[\| u \|_{1,\Omega}^2 = \iint_\Omega (u_x^2 + u_y^2) \, dx \, dy\]

and

\[|u|_{1,\Omega}^2 = \iint_\Omega (u_x + u_y) \, dx \, dy.\]

**Lemma 1.2**

\(H^2_0(\Omega)\) is a Hilbert space.

**Proof** See Aubin [3].

**Lemma 1.3**

\(H^2_0(\Omega)\) is separable.

**Proof** See Adams [1].

Define \(H^k_0(\Omega)\) to be the set of all functions \(u \in H^k_0(\Omega)\) such that

\[u = \frac{\partial u}{\partial n} = \ldots = \frac{\partial^{k-1} u}{\partial n^{k-1}} = 0\text{ on } \partial \Omega\]

where \(\frac{\partial u}{\partial n}\) is the normal derivative of \(u\). For example, \(H^2_0(\Omega)\) is the set of all functions in \(H^2_0(\Omega)\) which vanish on the boundary of \(\Omega\).

**Lemma 1.4**

For all \(u \in H^2_0(\Omega)\), there is a constant \(c > 0\) such that

\[\| u \| \leq c |u|_{1,\Omega},\]
where \( \| \cdot \| \) is the norm in \( L^2(\Omega) \).

**Proof** See Oden and Reddy [11].

Since \( \| u \|_1^2 = \int_\Omega (u^2 + \sum_{i=1}^k u_i^2) \, dx = \int_\Omega u^2 \, dx + \sum_{i=1}^k \int_\Omega u_i^2 \, dx = \| u \|_1^2 + \| u \|_1^2 \),

if \( u \in H^1_0(\Omega) \), then \( \| u \|_1^2 = \| u \|_1^2 + \| u \|_1^2 \leq (1 + c^2) \| u \|_1^2 \),

and hence there is a constant \( K \) such that \( \| u \|_1^2 \leq K \| u \|_1^2 \).

Let \( N \) be a normed linear space. A subspace \( M \) of \( N \) is finite dimensional (dimension \( n \)) if \( M = \langle u_1, \ldots, u_n \rangle \) and \( u_1, \ldots, u_n \) are linearly independent.

**Lemma 1.5**

If \( M \) is a finite dimensional subspace of a normed linear space \( N \), then \( M \) is closed.

**Proof** See Schechter [13].

**Lemma 1.6**

Let \( M \) be a closed linear subspace of a Hilbert space \( H \), and let \( x \) be a vector not in \( M \). Then there exists a unique vector \( y_0 \) in \( M \) such that \( \| x - y_0 \| \leq \| x - y \| \) for all \( y \in M \).

**Proof** See Simmons [14].

Let \( N \) be a normed linear space with norm \( \| \cdot \| \). We denote by \( N^* \) the set of all bounded linear functionals \( f \) on \( N \), that is, \( f : N \to \mathbb{R} \) is linear and there is \( C > 0 \) such that \( |f(u)| \leq C\| u \| \) for all \( u \in N \). Then \( N^* \) is a Banach space with norm \( \| f \| = \sup_{u \in N \setminus \{0\}} \frac{|f(u)|}{\| u \|} = \sup_{u \in N \setminus \{0\}} \{ |f(u)| : \| u \| \leq 1 \} \).
See Simmons [14]. It is easy to see that $|f(u)| \leq \| f \| \| u \|$. 

A bilinear form $a = a(u,v)$ on $N$ is a linear mapping from $N \times N$ into $R$, that is, a function $a : N \times N \to R$ such that

\[
a(ax + \beta y, z) = a(x, z) + \beta a(y, z),
\]

\[
a(x, ay + \beta z) = a(x, y) + \beta a(x, z),
\]

for all $\alpha, \beta \in R$ and $x, y, z \in N$. A bilinear form $a = a(u, v)$ is said to be bounded if there is a constant $c > 0$ such that

\[
|a(u, v)| < c \| u \| \| v \|,
\]

for all $u, v \in N$. It is said to be $N$-elliptic if there is a constant $c > 0$ such that

\[
a(u, u) > c \| u \|^2,
\]

for all $u \in N$.

Lemma 1.7 (Lax-Milgram)

If $H$ is a separable Hilbert space, $a = a(u, v)$ is a bounded $H$-elliptic bilinear form, and $g \in H^*$, then there exists a unique $u \in H$ such that

\[
a(u, v) = g(v),
\]

for all $v \in H$. Given a finite dimensional subspace $K$ of $H$, there exists a unique $U \in K$ such that

\[
a(U, V) = g(V),
\]

for all $V \in K$, and there is a constant $c > 0$ such that

\[
\| u - U \| \leq c \inf_{V \in K} \| u - V \|.
\]
Lemma 1.8 (Bramble-Hilbert)

If $F \in [H_2^{k+1}({\Omega})]^*$ and $F(p) = 0$ for all $p \in P_k$ where $P_k$ is the space of polynomials of degree less than or equal to $k$, then there exists a constant $C = C(\Omega)$ such that

$$|F(u)| \leq C \| F \|_1 \| u \|_{k+1, \Omega}$$

for all $u \in H^{k+1}_1(\Omega)$.

Proof See Aubin [2].

Lemma 1.9 (Hahn-Banach theorem)

If $M$ is a linear subspace of a normed linear space $N$ and $f \in M^*$, then $f$ can be extended to $\hat{f} \in N^*$ such that $\| f \| = \| \hat{f} \|$.

Proof See Ciarlet [4].
CHAPTER II

METHODS FOR BOUNDARY VALUE PROBLEMS

Ritz Method

Laplace's equation or Poisson's equation introduced in chapter I

\[ u_{xx} + u_{yy} = f, \quad u = u(x,y), \quad (x,y) \in \Omega \subset \mathbb{R}^2, \]

is called Dirichlet's problem if the boundary condition is such that \( u \) takes a prescribed value on the boundary of \( \Omega \), that is,

\[ u = g \text{ on } \partial\Omega, \]

for some \( g = g(x,y) \), and the boundary condition is called the Dirichlet boundary condition. It is called Neumann's problem if the boundary condition is such that the normal derivative of \( u \) takes a prescribed value on the boundary of \( \Omega \), that is,

\[ \frac{\partial u}{\partial n} = h \text{ on } \partial\Omega, \]

for some \( h = h(x,y) \), and the boundary condition is called the Neumann boundary condition. In this section we describe the Ritz method with examples of Dirichlet's problem and Neumann's problem. For more details, see Mitchell and Wait [10], or Davies [6].

Example 2.1

\[-u_{xx} - u_{yy} = f, \quad u = u(x,y), \quad (x,y) \in \Omega = (-c,c) \times (-c,c), \]

subject to
\[ u(-c,y) = u(c,y) = 0, \]
\[ u(x,-c) = u(x,c) = 0. \]  
\[ \cdots (2.1) \]

We partition \([-c,c) \times [-c,c] \) into \((n+1)^2\) square elements \(E_{00}, E_{01}, \cdots, E_{nn}\) where each

\[ E_{ij} = [-c+h_i,-c+h(i+1)] \times [-c+h_j,-c+h(j+1)], \]
for \(i = 0, \cdots, n\), \(j = 0, \cdots, n\), where \(h = \frac{2c}{n+1}\). We define \(n^2\) basis functions \(\phi_{ij}\), \(i = 1, \cdots, n\), \(j = 1, \cdots, n\), by \(\phi_{ij}(x,y) = \phi_i(x)\phi_j(y)\) where for \(t = x\) or \(y\)

\[ \phi_k(t) = \begin{cases} 
  \frac{t - t_{k-1}}{t_k - t_{k-1}}, & \text{for } t_{k-1} \leq t \leq t_k, \\
  \frac{t_k - t_{k-1}}{t_{k+1} - t_k}, & \text{for } t_k \leq t \leq t_{k+1}, \\
  1, & \text{for } t_{k+1} \leq t.
\end{cases} \]

for \(k = 1, \cdots, n\) where \(t_k = -c + hk\). Observe that \(\phi_k(t_i) = \delta_{ki}\), and for this reason, \((x_i,y_j)\)'s are called nodal points. The corresponding functional \(I = I(u)\) is

\[ I(u) = \frac{1}{2} \iiint_{\Omega} (u_x^2 + u_y^2) \, dx \, dy - \iint_{\Omega} fu \, dx \, dy. \]

The Ritz approximation \(U = \sum_{m=1}^{n} \sum_{k=1}^{n} a_{km} \phi_{km} \in K = \{ \phi_{ij} \mid i,j = 1, \cdots, n \}\) of (2.1) is determined by requiring \(\frac{\partial I(U)}{\partial \alpha_{ij}} = 0\) for each \(i\) and \(j\), or

\[ \frac{\partial}{\partial \alpha_{ij}} \left[ \frac{1}{2} \iiint_{\Omega} \sum_{m=1}^{n} \sum_{k=1}^{n} a_{km} \phi_{km} \frac{\partial^2 \phi_{ij}}{\partial x^2} + \sum_{m=1}^{n} \sum_{k=1}^{n} a_{km} \phi_{km} \frac{\partial^2 \phi_{ij}}{\partial y^2} \right] - \iiint_{\Omega} f \sum_{m=1}^{n} \sum_{k=1}^{n} a_{km} \phi_{km} \, dx \, dy \]

\[ = \iiint_{\Omega} \left[ \sum_{m=1}^{n} \sum_{k=1}^{n} a_{km} \phi_{km} \phi_{ij} \frac{\partial^2 \phi_{ij}}{\partial x^2} + \sum_{m=1}^{n} \sum_{k=1}^{n} a_{km} \phi_{km} \phi_{ij} \frac{\partial^2 \phi_{ij}}{\partial y^2} \right] \, dx \, dy - \iiint_{\Omega} f \phi_{ij} \, dx \, dy \]

\[ = \sum_{m=1}^{n} \sum_{k=1}^{n} a_{km} \left[ \frac{\partial \phi_{km}}{\partial x} \phi_{ij} - \frac{\partial \phi_{ij}}{\partial x} \phi_{km} \right] \frac{\partial \phi_{ij}}{\partial x} \, dx \, dy - \iiint_{\Omega} \left[ \phi_{ij} \frac{\partial^2 \phi_{ij}}{\partial x^2} + \phi_{ij} \frac{\partial^2 \phi_{ij}}{\partial y^2} \right] \, dx \, dy \]

\[ + \sum_{m=1}^{n} \sum_{k=1}^{n} a_{km} \left[ \frac{\partial \phi_{km}}{\partial x} \frac{\partial \phi_{ij}}{\partial y} - \frac{\partial \phi_{ij}}{\partial x} \frac{\partial \phi_{km}}{\partial y} \right] \phi_{ij} \phi_{ij} \, dx \, dy = 0. \]
Since
\[
f^{d_{i}}_{x}f_{k}(x)\frac{d}{dx}f^{i}(x)dx = \begin{cases} \frac{2}{h} & \text{for } k = i \\
-\frac{1}{h} & \text{for } |k-i| = 1 \\
0 & \text{otherwise} \end{cases}
\]
\[
f^{c_{m}}_{x}\phi_{m}(y)\phi_{j}(y)dy = \begin{cases} \frac{2h}{3} & \text{for } m = j \\
\frac{h}{6} & \text{for } |m-j| = 1 \\
0 & \text{otherwise} \end{cases}
\]
\[
f^{c_{k}}_{x}\phi_{k}(x)\phi_{i}(x)dx = \begin{cases} \frac{2h}{3} & \text{for } k = i \\
\frac{h}{6} & \text{for } |k-i| = 1 \\
0 & \text{otherwise} \end{cases}
\]
\[
f^{d_{m}}_{x}\phi_{m}(y)\frac{d}{dx}\phi_{j}(y)dy = \begin{cases} \frac{2}{h} & \text{for } m = j \\
-\frac{1}{h} & \text{for } |m-j| = 1 \\
0 & \text{otherwise} \end{cases}
\]

We obtain $n^2$ linear equations:

\[-\frac{1}{3}\alpha_{i-1,j-1} - \frac{1}{3}\alpha_{i-1,j} - \frac{1}{3}\alpha_{i-1,j+1} - \frac{1}{3}\alpha_{i,j-1} + \frac{8}{3}\alpha_{i,j} - \frac{1}{3}\alpha_{i,j+1} - \frac{1}{3}\alpha_{i+1,j-1} \]
\[-\frac{1}{3}\alpha_{i+1,j} - \frac{1}{3}\alpha_{i+1,j+1} = \iint_{\Omega} f_{ij}dx, \quad i = 1 \cdots n, \quad j = 1 \cdots n \]

for unknown $\alpha_{km}$'s.

In particular, if $c = 2$ and $n = 3$, and $f(x,y) = 1$,

\[
\iint_{\Omega} f_{ij}dx = \iint_{\Omega} f_{ij}dy = \int_{-2}^{2}f_{i}(x)dx \int_{-2}^{2}f_{j}(y)dy = 1,
\]

and we obtain the equations

\[8a_{11} - a_{21} - a_{12} - a_{22} = 3.\]
\[ -a_{11} + 8a_{21} - a_{31} - a_{12} - a_{22} - a_{32} = 3 , \]
\[ -a_{21} + 8a_{31} - a_{22} - a_{32} = 3 , \]
\[ -a_{11} - a_{21} + 8a_{12} - a_{22} - a_{13} - a_{23} = 3 , \]
\[ -a_{11} - a_{21} - a_{31} - a_{12} + 8a_{22} - a_{32} - a_{13} - a_{23} - a_{33} = 3 , \]
\[ -a_{21} - a_{31} - a_{22} + 8a_{32} - a_{23} - a_{33} = 3 , \]
\[ -a_{12} - a_{22} + 8a_{13} - a_{23} = 3 , \]
\[ -a_{12} - a_{22} - a_{32} - a_{13} + 8a_{23} - a_{33} = 3 , \]
\[ -a_{22} - a_{32} - a_{23} + 8a_{33} = 3 . \]

If the boundary condition is non-homogeneous Dirichlet condition, we could define basis functions which are non-zero on the boundary of \( \Omega \), and determine the Ritz approximation

\[ U(x,y) = \sum_i \alpha_i \phi_i(x,y) , \]

by setting

\[ \frac{\partial I(U)}{\partial \alpha_i} = 0 , \]

for each coefficient \( \alpha_i \) on \( \phi_i(x,y) \) which vanish on the boundary of \( \Omega \), and by imposing the boundary conditions to determine the coefficients \( \alpha_i \)'s on \( \phi_i(x,y) \) which are non-zero on the boundary of \( \Omega \) in order to get enough equations. If it is possible to find a function \( g = g(x,y) \) which satisfies the boundary condition, what can be done instead is to define the approximate solution

\[ U(x,y) = \sum_i \alpha_i \phi_i(x,y) + g(x,y) , \]

where each \( \phi_i \) vanishes on the boundary of \( \Omega \) so that \( U \) satisfies the boundary condition and to determine \( \alpha_i \)'s by requiring \( \frac{\partial I(U)}{\partial \alpha_i} = 0 \).
for each $i$. We will illustrate this technique with an example in the Galerkin method.

We turn now to an example of Neumann's problem.

Example 2.2

\[-u_{xx} - u_{yy} = f, \quad u = u(x,y), \quad (x,0) \in \Omega = (-c,c) \times (-c,c),\]

subject to

\[
\frac{\partial u}{\partial x}(-c,y) = \frac{\partial u}{\partial x}(c,y) = 0, \\
\frac{\partial u}{\partial y}(x,-c) = \frac{\partial u}{\partial y}(x,c) = 0,
\]

which are called the natural boundary conditions. If $u$ is a minimum for the functional

\[I(u) = \int_{-c}^{c} \int_{-c}^{c} F(x,y,u,u_x,u_y) \, dx \, dy\]

among the set of $H$ of admissible functions defined on $[-c,c] \times [-c,c]$ and if $\eta$ is any function, then

\[
\frac{dI(u + \epsilon \eta)}{d\epsilon} \bigg|_{\epsilon=0} = \int_{-c}^{c} \int_{-c}^{c} \left[ \frac{\partial^2}{\partial x^2} \eta F_{u_x} + \frac{\partial^2}{\partial y^2} \eta F_{u_y} \right] \, dx \, dy \\
+ \int_{-c}^{c} \int_{-c}^{c} \eta F - \frac{\partial}{\partial x} F_{u_x} - \frac{\partial}{\partial y} F_{u_y} \, dx \, dy = 0.
\]

Hence, if

\[I(u) = \frac{1}{2} \int_{-c}^{c} \left( u_x^2 + u_y^2 \right) \, dx \, dy - \int_{-c}^{c} F \, dx \, dy\]

and $u$ satisfies the natural boundary condition given in (2.2),

\[
\int_{-c}^{c} \int_{-c}^{c} \left[ \frac{\partial^2}{\partial x^2} \eta F_{u_x} + \frac{\partial^2}{\partial y^2} \eta F_{u_y} \right] \, dx \, dy \\
= \int_{-c}^{c} \eta F_{u_x} \bigg|_{x=-c}^{x=c} \, dx + \int_{-c}^{c} \eta F_{u_y} \bigg|_{y=-c}^{y=c} \, dx
\]
and we conclude that

\[
F_u - \frac{\partial^2 F}{\partial x^2} u_x - \frac{\partial^2 F}{\partial y^2} u_y = 0 ,
\]

which gives (2.2). We partition \([-c,c] \times [-c,c]\) into \(n^2\) square elements and define \((n + 1)^2\) basis functions \(\phi_{ij}\), \(i = 0 \ldots n\), \(j = 0 \ldots n\), by

\[
\phi_{ij}(x,y) = \phi_i(x) \phi_j(y) ,
\]

such that for \(t = x\) or \(y\)

\[
\phi_0(t) = \begin{cases} 
\frac{t_1 - t}{t_1 - t_0} & \text{for } t_0 \leq t \leq t_1 \\
0 & \text{otherwise}
\end{cases} ,
\]

\[
\phi_k(t) = \begin{cases} 
\frac{t - t_{k-1}}{t_k - t_{k-1}} & \text{for } t_{k-1} \leq t \leq t_k \\
\frac{t_{k+1} - t}{t_{k+1} - t_k} & \text{for } t_k \leq t \leq t_{k+1} \\
0 & \text{otherwise}
\end{cases} ,
\]

\[
\phi_n(t) = \begin{cases} 
\frac{t - t_{n-1}}{t_n - t_{n-1}} & \text{for } t_{n-1} \leq t \leq t_n \\
0 & \text{otherwise}
\end{cases} ,
\]

where \(t_k = -c + kh, k = 0 \ldots n\), and \(h = \frac{2c}{n}\). Then the Ritz approximation

\[
U = \sum_{m=0}^{n} \sum_{k=0}^{n} \alpha_{km} \phi_{km}
\]

of (2.2) is determined by

\[
\frac{\partial I(u)}{\partial a_{ij}} = 0 ,
\]

for each \(i\) and \(j\) or as in example 2.1, we obtain the equations
\[
\frac{\partial^2}{\partial x^2} \phi_i(x) = \int_{c}^{d} \frac{\partial}{\partial x} \phi_i(x) \, dx + \int_{c}^{d} \phi_i(x) \frac{\partial}{\partial x} \phi_j(y) \, dy - \int_{c}^{d} \phi_j(y) \frac{\partial}{\partial x} \phi_i(x) \, dx
\]

- \int_{c}^{d} \phi_i(x) \, dx = 0 \text{ for each } i \text{ and } j.

If the nodal point \((x_i, y_j)\) is not on the boundary, we obtain an equation as in example 2.1

\[\frac{1}{3} a_{i-1,j-1} - \frac{1}{3} a_{i-1,j} - \frac{1}{3} a_{i-1,j+1} - \frac{1}{3} a_{i,j-1} + \frac{2}{3} a_{i,j} - \frac{1}{3} a_{i,j+1} - \frac{1}{3} a_{i+1,j-1}
\]

\[\frac{1}{3} a_{i+1,j} - \frac{1}{3} a_{i+1,j+1} = 0.\]

If the nodal point \((x_i, y_j)\) is on the side, for example \(i \neq 0 \text{ or } n\), and \(j = 0\), then

\[
\int_{c}^{d} \frac{\partial}{\partial x} \phi_i(x) \, dx = \begin{cases} 
\frac{1}{3} h & \text{for } k = i \\
\frac{h}{3} & \text{for } m = j \\
\frac{2h}{3} & \text{for } k = i \\
\frac{2h}{3} & \text{for } k = i \\
\frac{1}{3} h & \text{for } m = j \\
\frac{1}{3} h & \text{for } m = j \\
0 & \text{otherwise}
\end{cases}
\]

and we obtain an equation
\[ -\frac{1}{6}a_{i-1,j} + \frac{4}{3}a_{i,j} - \frac{1}{6}a_{i+1,j} - \frac{1}{3}a_{i-1,j+1} - \frac{1}{3}a_{i,j+1} - \frac{1}{3}a_{i+1,j+1} = \iint_{\Omega} f_{ij} \, dx \, dy \]

If the nodal point \((x_i, y_j)\) is on the corner, for example \(i = 0\) and \(j = 0\), then

\[
\int_{-c}^{c} \frac{d\phi_k(x)}{dx} \frac{d\phi_i(x)}{dx} \, dx = \begin{cases} 
\frac{1}{h} & \text{for } k = i \\
-\frac{1}{h} & \text{for } k - i = 1 \\
0 & \text{otherwise}
\end{cases}
\]

\[
\int_{-c}^{c} \phi_m(y) \phi_j(y) \, dy = \begin{cases} 
\frac{h}{3} & \text{for } m = j \\
\frac{h}{6} & \text{for } m - j = 1 \\
0 & \text{otherwise}
\end{cases}
\]

\[
\int_{-c}^{c} \phi_k(x) \phi_i(x) \, dx = \begin{cases} 
\frac{h}{3} & \text{for } k = i \\
\frac{h}{6} & \text{for } k - i = 1 \\
0 & \text{otherwise}
\end{cases}
\]

\[
\int_{-c}^{c} \frac{d\phi_j(y)}{dy} \frac{d\phi_j(y)}{dy} \, dy = \begin{cases} 
\frac{1}{h} & \text{for } m = j \\
-\frac{1}{h} & \text{for } m - j = 1 \\
0 & \text{otherwise}
\end{cases}
\]

and we obtain

\[
\frac{2}{3}a_{i,j} - \frac{1}{6}a_{i+1,j} - \frac{1}{6}a_{i,j+1} - \frac{1}{3}a_{i+1,j+1} = \iint_{\Omega} f_{ij} \, dx \, dy.
\]

In particular, if \(n = 2, c = 1\), and \(f(x, y) = 1\), we have

\[
\begin{align*}
\frac{2}{3}a_{00} - \frac{1}{6}a_{10} - \frac{1}{6}a_{01} - \frac{1}{3}a_{11} &= \frac{1}{4}, \\
-\frac{1}{6}a_{00} + \frac{4}{3}a_{10} - \frac{1}{6}a_{20} - \frac{1}{3}a_{01} - \frac{1}{3}a_{11} - \frac{1}{3}a_{21} &= \frac{1}{2}, \\
-\frac{1}{6}a_{10} + \frac{2}{3}a_{20} - \frac{1}{3}a_{11} - \frac{1}{6}a_{21} &= \frac{1}{4}, \\
-\frac{1}{6}a_{00} - \frac{1}{3}a_{10} + \frac{4}{3}a_{01} - \frac{1}{3}a_{11} - \frac{1}{6}a_{02} - \frac{1}{3}a_{12} &= \frac{1}{2}, \\
-\frac{1}{6}a_{00} - \frac{1}{3}a_{10} - \frac{1}{3}a_{20} - \frac{1}{3}a_{01} + \frac{8}{3}a_{11} - \frac{1}{3}a_{21} - \frac{1}{3}a_{02} - \frac{1}{3}a_{12} - \frac{1}{3}a_{22} &= 1,
\end{align*}
\]
Self-Adjoint Positive Definite Operators

In the last section, we described how the Ritz approximation was obtained for the boundary value problem

\[-u_{xx} - u_{yy} = f, \quad u = u(x,y), \quad (x,y) \in \Omega = (-c,c) \times (-c,c),\]

using the functional

\[I(u) = \iint_{\Omega} (u_x^2 + u_y^2) \, dx \, dy - \iint_{\partial \Omega} fu \, dx \, dy\]

where \(u \in H\), the set of admissible functions. Let \(A = -\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}\). In example 2.1, the boundary condition was the homogeneous Dirichlet condition. Then for all \(u,v \in H\),

\[(Au,v) = \iint_{\Omega} (-u_{xx}v - u_{yy}v) \, dx \, dy = \iint_{\Omega} \left[ \frac{\partial^2}{\partial x} (u_xv) - \frac{\partial}{\partial y} (u_yv) \right] \, dx \, dy\]

\[+ \iint_{\Omega} (u_x v_x + u_y v_y) \, dx \, dy = \int_{\partial \Omega} u_y v_x \, dx - u_x v_x \, dy + \int_{\partial \Omega} (u_x v_x + u_y v_y) \, dx \, dy\]

\[= \iint_{\Omega} (u_x v_x + u_y v_y) \, dx \, dy\]

since \(v\) vanishes on \(\partial \Omega\). In example 2.2, we had the natural boundary condition

\[u_y(x,-c) = u_y(x,c) = 0 \quad \text{and} \quad u_x(-c,y) = u_x(c,y) = 0.\]

Then

\[(Au,v) = \iint_{\Omega} (-u_{xx}v - u_{yy}v) \, dx \, dy\]
\[
= \iint_{\Omega} \left[ -\frac{\partial}{\partial x}(u_x v) - \frac{\partial}{\partial y}(u_y v) \right] \, dx \, dy + \iint_{\Omega} (u_x v_x + u_y v_y) \, dx \, dy
\]
\[
= \int_{-c}^{c} -u_x v_x \big|_{x=-c}^{x=c} \, dy + \int_{-c}^{c} -u_y v_y \big|_{y=-c}^{y=c} \, dx + \iint_{\Omega} (u_x v_x + u_y v_y) \, dx \, dy
\]
\[
= \iint_{\Omega} (u_x v_x + u_y v_y) \, dx \, dy
\]
because of the natural boundary condition. In either case, we can write the functional \( I \) as
\[
I(u) = \frac{1}{2} (Au,u) - (f,u).
\]
In general, an operator \( T \) on a Hilbert space \( H \) is said to be positive definite if
\[
(Tx,x) \geq 0 \quad \text{for all} \quad x \in H,
\]
\[
(Tx,x) = 0 \quad \text{if and only if} \quad x = 0.
\]
It is said to be self-adjoint if
\[
(Tx,y) = (x,Ty) \quad \text{for all} \quad x,y \in H.
\]
It is easy to see that our example \( A = -\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} \) is positive definite and self-adjoint on \( H \), the set of admissible functions in case of the homogeneous Dirichlet condition, since
\[
(Au,u) = \iint_{\Omega} (u_x^2 + u_y^2) \, dx \, dy \geq 0,
\]
\[
(Au,u) = 0, \quad \text{if and only if} \quad u_x^2 + u_y^2 = 0, \quad \text{if and only if} \quad u \quad \text{is constant, if and only if} \quad u = 0 \quad \text{on} \quad \partial \Omega,
\]
\[
(Au,v) = \iint_{\Omega} (u_x v_x + u_y v_y) \, dx \, dy = (u,Av) \quad \text{since} \quad u \quad \text{and} \quad v \quad \text{vanish on} \quad \partial \Omega.
\]
Consider the linear boundary value problem
\[
Au = f, \quad u = u(\bar{x}), \quad \bar{x} \in \Omega \subset \mathbb{R}^n \quad \cdots (2.3)
\]
where $A$ is linear, positive definite, and self-adjoint. Assume that the corresponding functional $I$ is written in the form

$$I(u) = \frac{1}{2}(Au, u) - (f, u)$$

with the set $H$ of admissible functions. We will show the existence and the uniqueness of the Ritz approximation of (2.3) and show how good it is.

Let $a(u, v) = (Au, v)$.

**Lemma 2.1**

$a(u, v)$ is an inner product on $H$.

**Proof**

$a(u, v) = (Au, v) = (u, Av) = (Av, u) = a(v, u)$ since $A$ is self-adjoint.

$a(u, v+w) = (Au, v+w) = (Au, v) + (Au, w) = a(u, v) + a(u, w)$ and

$a(cu, v) = (cAu, v) = (cAu, v) = c(Au, v) = ca(u, v)$ since $A$ is linear.

$a(u, u) = (Au, u) \geq 0$ for all $u \neq 0$ since $A$ is positive definite.

We denote $\| u \|_A^2 = a(u, u)$.

The energy space $H_A$ is the subspace of $H$ containing the set of all $u(x)$ and $v(x) \in H$ such that $a(u, v)$ is a bounded bilinear form. If we can transform $(Au, v)$ to $a(u, v)$ using Green's theorem to reduce continuity requirements on $u$, we use this form to define $a(u, v)$ and in this case $H_A$ is known to be a Hilbert space. (See Mitchell and Wait [10]). Hence for our example, $A = \frac{-\delta^2}{\delta x^2} - \frac{\delta^2}{\delta y^2}$

$$a(u, v) = \iint_{\Omega} (u_x v_x + u_y v_y) \, dx \, dy$$

rather than

$$(Au, v) = \iint_{\Omega} (-u_{xx} v - u_{yy} v) \, dx \, dy$$

In order to obtain $(Au, v)$, we choose a finite dimensional subspace $K_0 \subset H$. We assume that $K_0$ is a
and since

\[ |a(u,v)| = |\int \int (u_x v_x + u_y v_y) \, dx \, dy| < \int \int |u_x v_x + u_y v_y| \, dx \, dy = \int \int |u_x||v_x| \, dx \, dy + \int \int |u_y||v_y| \, dx \, dy \]

\[ \leq \|u_x\|\|v_x\| + \|u_y\|\|v_y\| \leq \|u\|_{1,\Omega}\|v\|_{1,\Omega} \]

\[ + \|u\|_{1,\Omega}\|v\|_{1,\Omega} = 2\|u\|_{1,\Omega}\|v\|_{1,\Omega} , \]

\( a \) is bounded on \( H_0^1(\Omega) \). Therefore for \( A = -\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}, H_0^1(\Omega) \subset H_A \) in case of homogeneous Dirichlet condition.

An operator \( P \) of a Hilbert space \( H \) onto the subspace \( K \) is an orthogonal projection if for all \( x \in H \), \((x-Px,y) = 0\) for all \( y \in M \).

**Lemma 2.2**

If \( P \) is an orthogonal projection of a Hilbert space \( H \) onto a subspace \( K \), \( P \) is unique and for all \( x \in H, \|x - Px\| < \|x - y\| \) for any \( y \in K \).

**Proof**

Assume that there are two orthogonal projections \( P \) and \( Q \) onto \( K \). Then for each \( x \in H \), \( Px - Qx \in K \) and \((x - Qx, Px - Qx) = 0\) since \( Q \) is an orthogonal projection. On the other hand, \((x - Qx, Px - Qx) = (x - Px, Px - Qx) + (Px - Qx, Px - Qx) = \|Px - Qx\|^2 = 0\). Therefore \( Px = Qx \). Since \( x \) is an arbitrary element in \( H \), \( P = Q \).

Let \( y \in K \) and \( x \in H \). Then \( \|x - y\|^2 = \|(x - Px) + (Px - y)\|^2 = \|x - Px\|^2 + 2(x - Px, Px - y) + \|Px - y\|^2 = \|x - Px\|^2 + \|Px - y\|^2 \) since \( Px - y \in K \). Therefore \( \|x - Px\|^2 \leq \|x - y\|^2 \) or \( \|x - Px\| \leq \|x - y\| \)

In order to obtain a Ritz approximation of (2.3), we choose a finite dimensional subspace \( K_n = \langle \phi_1, \ldots, \phi_n \rangle \) in \( H \). We assume that \( K_n \) is a
subspace of $H_A$. The Ritz approximation $U(x) = \sum_{i=1}^{n} a_i \phi_i(x)$ is obtained by

$$\frac{\partial I(U)}{\partial a_j} = 0 \quad \text{for each } j.$$ 

By our assumption

$$I(U) = \frac{1}{2} (AU, U) - (f, U),$$

and

$$\frac{\partial I(U)}{\partial a_j} = \frac{\partial}{\partial a_j} \left[ \frac{1}{2} \int \sum_{i=1}^{n} a_i \phi_i(x) \phi_i(x) dx - \int f \sum_{i=1}^{n} a_i \phi_i dx \right]$$

$$= \frac{1}{2} \int \sum_{i=1}^{n} a_i \phi_i(x) \phi_j(x) dx - \int f \phi_j(x) dx$$

$$= \frac{1}{2} \int \sum_{i=1}^{n} a_i \phi_i(x) \phi_j(x) dx - \int f \phi_j(x) dx$$

$$= \sum_{i=1}^{n} a_i \phi_i(x) \phi_j(x) dx - \int f \phi_j(x) dx = \sum_{i=1}^{n} a_i a(\phi_i, \phi_j) - (f, \phi_j) = 0.$$ 

Hence $\sum_{i=1}^{n} a_i a(\phi_i, \phi_j) = (f, \phi_j)$ for each $j$. \quad \cdots (2.4)

**Lemma 2.3**

Let $u_0$ be the true solution of the boundary value problem

$$Au = f, \quad u = u(x), \quad x \in \Omega \subset \mathbb{R}^n,$$

where $A$ is linear, self-adjoint, and positive definite, and the corresponding functional $I$ is

$$I(u) = \frac{1}{2} (Au, u) - (f, u).$$

Then the Ritz approximation

$$U(x) = \sum_{i=1}^{n} a_i \phi_i$$

is the orthogonal projection of the true solution $u_0$ onto the subspace $K_n$. 
of $H_A$ with respect to the inner product $a(u,v)$.

Proof

For each $j$,

$$a(u_0 - \sum_{i=1}^{n} a_i \phi_i, \phi_j) = a(u_0, \phi_j) - \sum_{i=1}^{n} a_i a(\phi_i, \phi_j) = (Au_0, \phi_j) - (f, \phi_j)$$

$$= (f, \phi_j) - (f, \phi_j) = 0$$

by (2.4).

Since $K_n$ is a finite dimensional subspace of $H_A$, by lemma 1.5, we know $K_n$ is closed in $H_A$. Since the Ritz approximation $U$ is the orthogonal projection of the true solution onto $K_n$, by lemma 2.2, $U$ is unique and $\| u_0 - U \|_A < \| u_0 - v \|_A$ for all $v \in K_n$ where $\| v \|_A^2 = a(v, v)$.

By lemma 1.6, we know that such $U$ exists in $K_n$. The next theorem immediately follows from these facts.

Theorem 2.1

There exists the unique Ritz approximation for the boundary value problem

$$Au = f, \quad u = u(x), \quad x \in \Omega \subset \mathbb{R}^n,$$

where $A$ is linear, self-adjoint, and positive definite and the corresponding functional

$$I(u) = \frac{1}{2}(Au, u) - (f, u),$$

and it is the best approximation in $K_n$ in terms of the energy space $H_A$.

We have shown that for example 2.1 in the last section

$$-u_{xx} - u_{yy} = f, \quad u = u(x,y), \quad (x,y) \in (-c,c) \times (-c,c),$$

subject to $u = 0$ on $\partial \Omega$, $H^1_0(\Omega) \subset H_A$, and our choice of the finite
dimensional subspace $K$ of $H$ was hat-functions. Hence $K \subset H^1_0(\Omega) \subset H_A$.

By theorem 2.1, the Ritz approximation $U \in K$ of this problem is the unique best approximation in $K$ in terms of $H_A$.

**Galerkin Method**

In this section, we illustrate the Galerkin method with examples.

More detailed discussion on the Galerkin method is found in Mitchell and Wait [10] or Fairweather [7].

**Example 2.3**

Consider the linear boundary value problem

$$-u_{xx} - u_{yy} = f, \quad u = u(x,y), \quad (x,y) \in \Omega = (-c,c) \times (-c,c),$$

subject to $u = 0$ on $\partial \Omega$. \hspace{1cm} \cdots (2.5)

A weak formulation of (2.5) is: Find $u$ such that

$$(Au,v) = (f,v) \quad \text{for all } v \in H^1_0(\Omega) \quad \cdots (2.6)$$

where $A = -\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}$. Since $v = 0$ on $\partial \Omega$,

$$(Au,v) = \int\int_{\Omega} (-u_{xx}v - u_{yy}) \, dx \, dy$$

$$= \int\int_{\Omega} \left[ -\frac{\partial}{\partial x} (u_x v) - \frac{\partial}{\partial y} (u_y v) \right] \, dx \, dy + \int\int_{\Omega} (u_x v_x + u_y v_y) \, dx \, dy$$

$$= \int_{\partial \Omega} u_x v \, dx - u_x v \, dy + \int\int_{\Omega} (u_x v_x + u_y v_y) \, dx \, dy$$

$$= \int\int_{\Omega} (u_x v_x + u_y v_y) \, dx \, dy .$$

Let

$$a(u,v) = \int\int_{\Omega} (u_x v_x + u_y v_y) \, dx \, dy .$$
The Galerkin approximation of (2.5) is based on the following weak form:

Find \( u \in H^1_0(\Omega) \) such that

\[
a(u,v) = (f,v) \quad \text{for all } v \in H^1_0(\Omega). \quad \cdots (2.7)
\]

Observe that the continuity requirement is reduced on \( u \) as we transform \((Au,v)\) in (2.6) to \( a(u,v)\) in (2.7). We define the approximating subspace \( K \) in \( H^1_0(\Omega) \) to be Hat-functions as in section 2.1, example 2.1. The Galerkin approximation

\[
U = \sum_{m=1}^{n} \sum_{k=1}^{n} \alpha_{km} \phi_{km}
\]

is obtained from

\[
a(U, \phi_{ij}) = (f, \phi_{ij}),
\]
or

\[
\sum_{m=1}^{n} \sum_{k=1}^{n} \alpha_{km} a(\phi_{km}, \phi_{ij}) = (f, \phi_{ij}),
\]

which is the same formulation as obtained by using the Ritz method in example 2.1. For this example, the Galerkin approximation equals the Ritz approximation.

In general, if \( Au = f, \quad u = u(x), \quad x \in \Omega \subset \mathbb{R}^n \)

is a linear second order boundary value problem, then the Galerkin approximation is based on the following weak formulation. Find \( u \) such that

\[
(Au,v) = (f,v) \quad \text{for all } v \in H^1_0(\Omega).
\]

If it is possible to transform \((Au,v)\) to \( a(u,v)\), which involves no
derivative of order 2, using Green’s theorem, we obtain the following weak formulation: Find $u \in H^1_2(\Omega)$ such that

$$a(u, v) = (f, v) \text{ for all } v \in H^1_2(\Omega).$$

Then we choose the finite dimensional subspace $K_n = \langle \phi_1, \ldots, \phi_n \rangle$ of $H^1_2(\Omega)$, and the Galerkin approximation $U = \sum_{i=1}^n a_i \phi_i$ is determined by

$$a(U, \phi_j) = (f, \phi_j) \text{ for each } \phi_j \in K_n,$$

where $K_n = \bigcap_{n=1}^\infty H^1_2(\Omega)$.

If the boundary condition is a non-homogeneous Dirichlet condition, what can be done is to find a function $g(x)$ which satisfies the boundary condition and to define the approximate solution $U$ to be

$$U = \sum_{i=1}^n a_i \phi_i + g$$

where each $\phi_i \in K_n$ so that $U$ satisfies the boundary condition. Each coefficient $a_i$ is determined by equations

$$a(U, \phi_j) = (f, \phi_j) \text{ for each } \phi_j \in K_n.$$

Example 2.4

$$-u_{xx} - u_{yy} = f, \quad u = u(x, y), \quad (x, y) \in \Omega = (-c, c) \times (-c, c),$$

subject to

$$u(-c, y) = u(c, y) = \frac{1}{2}(c^2 + y^2),$$

$$u(x, -c) = u(x, c) = \frac{1}{2}(c^2 + x^2).$$

We define the approximating subspace $K \subset H^1_2(\Omega)$ to be the Hat-functions as in section 2.1, example 2.1, and define $g(x, y) = \frac{1}{2}(x^2 + y^2)$ so that $g$ satisfies the boundary condition. The Galerkin approximation
is determined by

\[ a(U, \phi_{ij}) = (f, \phi_{ij}) \quad \text{for each } i \text{ and } j, \]

where

\[ a(u, v) = \int_{\Omega} (u_x v_x + u_y v_y) \, dx \, dy, \]

or

\[
a(U, \phi_{ij}) = \int_{\Omega} \left[ \frac{\partial}{\partial x} \left( \sum_{m=1}^{n} \sum_{k=1}^{n} \alpha_{km} \phi_m \phi_k + \frac{1}{2} (x^2 + y^2) \right) \frac{\partial \phi_{ij}}{\partial x} \right] dx \, dy
\]

\[ + \int_{\Omega} \left[ \frac{\partial}{\partial y} \left( \sum_{m=1}^{n} \sum_{k=1}^{n} \alpha_{km} \phi_m \phi_k + \frac{1}{2} (x^2 + y^2) \right) \frac{\partial \phi_{ij}}{\partial y} \right] dx \, dy
\]

\[ + \int_{\Omega} [x \frac{\partial \phi_{ij}}{\partial x} + y \frac{\partial \phi_{ij}}{\partial y}] dx \, dy = (f, \psi_{ij}). \]

Now

\[
\int_{\Omega} \frac{\partial \phi_{ij}}{\partial x} dx \, dy = \int_{-c}^{c} \int_{x_i}^{x_{i+1}} \phi_j(x) \, dx \, dy - \int_{-c}^{c} \int_{x_i}^{x_{i+1}} \phi_j(y) \, dy
\]

\[ = -h \left( x_{i+1}^2 - x_i^2 \right), \]

where

\[ h = \frac{2c}{n+1} \quad \text{and} \quad x_k = -c + kh. \]

Similarly

\[
\int_{\Omega} \frac{\partial \phi_{ij}}{\partial y} dx \, dy = \int_{-c}^{c} \int_{y_j}^{y_{j+1}} \phi_i(y) \, dy \, dx - \int_{-c}^{c} \int_{y_j}^{y_{j+1}} \phi_i(x) \, dx \, dy
\]

\[ = -h \left( y_{j+1}^2 - y_j^2 \right), \]

where \( y_k = -c + hk. \) Therefore we obtain the system of equations for \( \alpha_{km} \).
\[
- a_{i-1,j-1} - a_{i-1,j} - a_{i-1,j+1} - a_{i,j-1} + 8a_{i,j} - a_{i,j+1} - a_{i+1,j-1} \\
- a_{i+1,j} - a_{i+1,j+1} = \frac{3h}{2}(x_{i-1}^2 - 2x_i^2 + x_{i+1}^2 + y_{j-1}^2 - 2y_j^2 + y_{j+1}^2) \\
+ 3\int_{\Omega} f \phi_{ij} \, dx \, dy
\]
for \( i = 1 \cdots \cdots n \) and \( j = 1 \cdots \cdots n \).

In particular, if \( c = 2, n = 3, \) and \( f(x,y) = 1, \) we obtain the system of equations:

\[
\begin{align*}
8a_{11} - a_{21} - a_{12} - a_{22} &= 9, \\
a_{11} + 8a_{12} - a_{13} - a_{21} - a_{22} - a_{32} &= 9, \\
a_{21} + 8a_{31} - a_{22} - a_{32} &= 9, \\
a_{11} - a_{21} + 8a_{12} - a_{22} - a_{13} - a_{23} &= 9, \\
a_{11} - a_{21} - a_{31} - a_{12} + 8a_{22} - a_{32} - a_{13} - a_{23} - a_{33} &= 9, \\
a_{11} - a_{21} - a_{31} - a_{12} + 8a_{22} - a_{32} - a_{13} - a_{23} - a_{33} &= 9, \\
a_{21} - a_{11} - a_{23} + 8a_{32} - a_{23} - a_{33} &= 9, \\
a_{11} - a_{22} + 8a_{13} - a_{23} &= 9, \\
a_{12} - a_{22} - a_{32} - a_{13} + 8a_{23} - a_{33} &= 9, \\
a_{22} - a_{32} - a_{23} + 8a_{33} &= 9.
\end{align*}
\]

The Galerkin method can be applied in some cases to non-linear problems.

Example 2.5

\[
- \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = f, \quad u = u(x,y), \quad (x,y) \in \Omega = (-c,c) \times (-c,c),
\]
subject to \( u = 0 \) on \( \partial \Omega \).

Let \( Au = - \frac{\partial}{\partial x} (\frac{\partial u}{\partial x}) - \frac{\partial}{\partial y} (\frac{\partial u}{\partial y}), \) and let \( v \in H^1_0(\Omega). \) By Green's theorem,
\[(Au,v) = \iint_{\Omega} [-\frac{\partial}{\partial x}(u\frac{\partial u}{\partial x})v - \frac{\partial}{\partial y}(u\frac{\partial u}{\partial y})v] \, dx \, dy \]

\begin{align*}
&= \iint_{\Omega} [-\frac{\partial}{\partial x}(u\frac{\partial u}{\partial x}) - \frac{\partial}{\partial y}(u\frac{\partial u}{\partial y})] \, dx \, dy + \iint_{\Omega} (u\frac{\partial u}{\partial x} + u\frac{\partial u}{\partial y}) \, dx \, dy \\
&= \int_{\Omega} \frac{\partial u}{\partial y} \, dx - \int_{\Omega} \frac{\partial u}{\partial x} \, dy + \int_{\Omega} (u\frac{\partial u}{\partial x} + u\frac{\partial u}{\partial y}) \, dx \, dy .
\end{align*}

Since \( v = 0 \) on \( \partial \Omega \),

\[ \int_{\partial \Omega} \frac{\partial u}{\partial y} \, dx - \int_{\partial \Omega} \frac{\partial u}{\partial x} \, dy = 0 , \]

and we have

\[ (Au,v) = \iint_{\Omega} (u\frac{\partial u}{\partial x} + u\frac{\partial u}{\partial y}) \, dx \, dy . \]

Let \( a(u,v) = \iint_{\Omega} (u\frac{\partial u}{\partial x} + u\frac{\partial u}{\partial y}) \, dx \, dy \). Thus we obtain the weak formulation:

Find \( u \in H^1_0(\Omega) \) such that

\[ a(u,v) = (f,v) \quad \text{for all } v \in H^1_0(\Omega). \]

We define the approximating subspace \( K \) of \( H^1_0(\Omega) \) to be the bilinear basis functions as in example 2.1. Then the Galerkin approximation

\[ U(x,y) = \sum_{k=1}^{n} \sum_{m=1}^{n} \alpha_{km} \phi_{km}(x,y) \]

satisfies

\[ a(U,\phi_{ij}) = (f,\phi_{ij}) \]

for each \( i \) and \( j \), or since

\begin{align*}
&= \sum_{p=1}^{n} \sum_{q=1}^{n} \alpha_{pq} \phi_{pq} \int_{\Omega} \phi_{km} \frac{\partial \phi_{pq}}{\partial x} \frac{\partial \phi_{ij}}{\partial x} + \frac{\partial \phi_{pq}}{\partial y} \frac{\partial \phi_{ij}}{\partial y} \, dx \, dy \\
&= \sum_{p=1}^{n} \sum_{q=1}^{n} \alpha_{pq} \phi_{pq} \sum_{k=1}^{n} \sum_{m=1}^{n} \alpha_{km} \int_{\Omega} \phi_{km} \frac{\partial \phi_{pq}}{\partial x} \frac{\partial \phi_{ij}}{\partial x} + \frac{\partial \phi_{pq}}{\partial y} \frac{\partial \phi_{ij}}{\partial y} \, dx \, dy ,
\end{align*}
we have

\[
\sum_{p=1}^{n} \sum_{q=1}^{n} \alpha_{pq} \sum_{k=1}^{n} \sum_{m=1}^{n} \alpha_{km} \int_{\Omega} \phi_{km} \left( \frac{\partial \phi_{pq}}{\partial x} \frac{\partial \phi_{ij}}{\partial x} + \frac{\partial \phi_{pq}}{\partial y} \frac{\partial \phi_{ij}}{\partial y} \right) \, dx \, dy
\]

\[= \int_{\Omega} f \phi_{ij} \, dx \, dy
\]

for each i and j. The integrals can be evaluated and thus we obtain a non-linear system of equations.

**\( H^1_2(\Omega) \)-Elliptic Bilinear Forms**

In this section, we will show the existence and the uniqueness, and derive an error bound of the Galerkin approximation obtained in example 2.3 using lemma 1.8 (Lax-Milgram lemma).

In example 2.3, the Galerkin approximation is based on the weak formulation:

Find \( u \in H^1_2(\Omega) \) such that

\[
a(u,v) = (f,v) \quad \text{for all } v \in H^1_2(\Omega)
\]

\( \cdots (2.8) \)

where

\[
a(u,v) = \int_{\Omega} (u_x v_x + u_y v_y) \, dx \, dy
\]

Define \( \iota : H^1_2(\Omega) \rightarrow \mathbb{R} \) by

\[
\iota(v) = (f,v)
\]

We will show that the bilinear form \( a = a(u,v) \) is bounded on \( H^1_2(\Omega) \) and \( H^1_2(\Omega) \)-elliptic, and \( \iota \in [H^1_2(\Omega)]^* \).

Let \( u \) and \( v \in H^1_2(\Omega) \). Then
\[ |a(u,v)| = |\iint \Omega (u_x v_x + u_y v_y) \, dx \, dy| \leq \iint \Omega |u_x| \, v_x \, dx \, dy + \iint \Omega |u_y| \, v_y \, dx \, dy \]
\[ \leq \|u_x\| \|v_x\| + \|u_y\| \|v_y\| \leq 2 \|u\|_{1,\Omega} \|v\|_{1,\Omega}, \]
by Schwarz's inequality, and hence \( a \) is bounded on \( H_0^2(\Omega) \). Since

\[ a(u,u) = \iint (u_x^2 + u_y^2) \, dx \, dy = \|u\|_{1,\Omega}^2, \]
and \( u \in H_0^1(\Omega) \), by lemma 1.4, there is a \( c > 0 \) such that \( \|u\|_{1,\Omega} \leq c \|u\|_{1,\Omega} \),
and hence \( a(u,u) = \|u\|_{1,\Omega}^2 \geq \frac{1}{c^2} \|u\|_{1,\Omega}^2 \) or \( a \) is \( H_0^1(\Omega) \)-elliptic. Now

\[ |t(v)| = |(f,v)| \leq \|f\| \|v\| \leq \|f\| \|v\|_{1,\Omega}. \]

Therefore

\[ t \in \left[ H_0^1(\Omega) \right]^*. \]

If \( K = \langle \phi_1, \cdots, \phi_n \rangle \) is the finite dimensional subspace of \( H_0^1(\Omega) \),
the Galerkin approximation \( U = \sum_{i=1}^n a_i \phi_i \) is determined by \( a(U,\phi_j) = (f,\phi_j) \)
for each \( j \). If \( V \in K \), then \( V = \sum_{j=1}^n \beta_j \phi_j \) for some \( \beta_j \)'s and

\[ a(U,V) = a(U, \sum_{j=1}^n \beta_j \phi_j) = \sum_{j=1}^n \beta_j a(U, \phi_j) = \sum_{j=1}^n \beta_j (f, \phi_j) = (f, \sum_{j=1}^n \beta_j \phi_j) = (f,V). \]

It follows that the Galerkin approximation \( U \) satisfies \( a(U,V) = (f,V) \) or
\( a(U,V) = t(V) \) for any \( V \in K \). Since \( H_0^1(\Omega) \) is separable by lemma 1.3, by
lemma 1.8, there exists the unique \( U \in K \) such that

\[ \|u - U\|_{1,\Omega} \leq C \inf_{V \in K} \|u - V\|_{1,\Omega}, \]
where \( u \) is the solution of (2.8).

**Time-Dependent Problems**

In this section, we discuss applications of the Ritz method and the
Galerkin method to time-dependent problems.
In section 1.3, we derived the equation of a vibrating string with fixed end points

\[ c^2 u_{xx} - u_{tt} = 0 \ , \ u = u(x,t) , \quad (x,t) \in (0,1) \times (0,t) , \]

subject to

\[ u(0,t) = u(1,t) = 0 \ , \]
\[ u(x,t_0) = u_0(x) , \]
\[ u(x,t_1) = u_1(x) , \]

from the functional

\[ I(u) = \frac{1}{2} \int_0^1 \left( u_t^2 - c^2 u_{xx}^2 \right) dx dt , \]

using Hamilton's principle. If the external force \( f(x,t) \) acts on the string, then the equation takes the form

\[ u_{tt} - c^2 u_{xx} = -f \ , \ u = u(x,t) , \quad (x,t) \in (0,1) \times (t_0,t_1) , \]

subject to

\[ u(0,t) = u(1,t) = 0 \ , \]
\[ u(x,t_0) = u_0(x) , \]
\[ u(x,t_1) = u_1(x) . \quad \text{(Courant and Hilbert [5])} \]

Therefore it makes sense to consider the boundary value problem,

\[ u_{tt} - u_{xx} = g \ , \ u = u(x,t) , \quad (x,t) \in \Omega = (0,2) \times (0,1) , \]

subject to

\[ u(0,t) = u(2,t) = 0 \ , \]
\[ u(x,0) = u(x,1) = 0 . \quad \cdots (2.9) \]
Suppose we are looking for a Ritz approximation. The corresponding functional is

\[ I(u) = \frac{1}{2} \int_{\Omega} (u_x^2 - u_t^2) \, dx \, dt - \int_{\Omega} gu \, dx \, dt \]

Therefore \((Au,v)\) does not define an inner product and a Ritz approximation is not guaranteed to be the best approximation in any finite dimensional subspace of the energy space. It may not even exist.

On the other hand, if \( \Lambda = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} \) and \( v \in H \) the set of admissible functions, then \( v \) vanishes on the boundary of \( \Omega \) and by Green's theorem,

\[
(Au,v) = \int_{\Omega} (u_{tt}v - u_{xx}v) \, dx \, dt = \int_{\Omega} \left[ \frac{\partial}{\partial t} (u_t v) - \frac{\partial}{\partial x} (u_x v) \right] \, dx \, dt + \int_{\Omega} (u_x v_x - u_t v_t) \, dx \, dt
\]

Consequently we can write the functional \( I \) as

\[ I(u) = \frac{1}{2} (Au,u) - (g,u) \]

We have shown in section 2.2 that if a linear boundary value problem

\[ Au = f, \quad u = u(\bar{x}), \quad \bar{x} \in \Omega \subset \mathbb{R}^n \]

has the corresponding functional

\[ I(u) = \frac{1}{2} (Au,u) - (f,u) \]

and the operator \( A \) is positive definite and self-adjoint, then there exists the unique Ritz approximation in the finite dimensional subspace \( K_n \) of the energy space \( H_A \), and it is the best approximation in \( K_n \) with respect to the inner product \( a(u,v) = (Au,v) \).

Now we show that the operator \( \Lambda = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} \) in (2.9) is not positive definite. If \( u(x,t) = x(x-2)t(t-1) \),

\[
(Au,u) = \int_{\Omega} (u_x^2 - u_t^2) \, dx \, dt
\]
\[
\int_\Omega [(4x^2 - 8x + 4)(t^4 - 2t^3 + t^2) - (4t^2 - 4t + 1)(x^4 - 4x^3 + 4x^2)] \, dx \, dy
\]
\[
= \frac{-4}{15} < 0 .
\]

Therefore \((Au,v)\) does not define an inner product and a Ritz approximation is not guaranteed to be the best approximation in any finite dimensional subspace in terms of the energy space. It may not even exist.

If we tried the Galerkin method on (2.9), the weak formulation would be: Find \(u \in H^1_0(\Omega)\) such that

\[
a(u,v) = (q,v) \quad \text{for all } v \in H^1_0(\Omega),
\]

where

\[
a(u,v) = \int_\Omega (u_x v_x - u_t v_t) \, dx \, dt .
\]

As we have shown, the bilinear form \(a(u,v)\) can be negative and fails to be \(H^1_0(\Omega)\)-elliptic. For this reason, we cannot use the Lax-Milgram lemma to guarantee the existence, the uniqueness, or the convergence of the Galerkin approximation. Nevertheless, if we can find a Ritz approximation or a Galerkin approximation to (2.9), it defines an approximate solution. However, there is still a serious problem.

In general, a boundary value problem

\[
Au = f , \quad u = u(\bar{x}) , \quad \bar{x} \in \Omega \subset \mathbb{R}^n ,
\]

is said to be well-posed if there exists a unique solution, and the solution depends continuously on the boundary condition, that is, if the boundary condition changes slightly, the solution changes slightly. It can be shown that the hyperbolic problem
subject to

\[ u(0,t) = u(1,t) = 0 , \]
\[ u(x,t_0) = u_0(x) , \]
\[ u(x,t_1) = u_1(x) . \]

(2.10)

is not well-posed (Mitchell and Wait [10]), but (2.10) is well-posed if it is given with initial-boundary condition

\[ u(0,t) = u(1,t) = 0 , \]
\[ u(x,t_0) = u_0(x) , \]
\[ \frac{\partial u}{\partial t}(x,t_0) = v_0(x). \]

(2.11)

Consequently, it is not appropriate to formulate the Ritz method or the Galerkin method for (2.10) as a boundary value problem, when it is given with the initial-boundary condition (2.11) as a requirement. In other words, we cannot directly apply these methods. In chapter III, we will modify the Ritz method and show how the variational principle can be used to define an approximate solution for initial-boundary value problems, and we will show the finite element method based on the weak formulation for the initial-boundary value problems, namely, the semi-discrete Galerkin method.

In chapter I, we introduced the heat equation

\[ u_t = c^2 u_{xx} , \quad u = u(x,t), \quad 0 < x < 1 , \quad t > t_0 , \]

subject to

\[ u(x,t_0) = u_0(x) , \]
If a heat source is present in the system, then the heat equation takes the form

\[ u_t - \alpha^2 u_{xx} = f, \]

subject to

\[ u(x,t_0) = u_0(x), \]
\[ u(0,t) = w_0(t), \]
\[ u(1,t) = w_1(t). \]

(Tikhonov and Samarskii [17])

It can be shown that this equation has no corresponding functional \( I = I(u) \). In other words, there is no functional of \( u \) alone. In order to find the functional for the heat equation, we need the following concept: Suppose \( u = u(x,y) \) and \( v = v(x,y) \) are \( C^2 \)-functions defined on \( \Omega \subset \mathbb{R}^2 \) such that \( u = f \) and \( v = g \) on \( \partial \Omega \) for some functions \( f \) and \( g \). Let \( H_1 \) and \( H_2 \) be the set of \( C^2 \)-functions which agree with \( f \) and \( g \) on \( \partial \Omega \) respectively:

\[ H_1 = \{ u | u \text{ is } C^2 \text{ and } u = f \text{ on } \partial \Omega \}, \]
\[ H_2 = \{ v | v \text{ is } C^2 \text{ and } v = g \text{ on } \partial \Omega \}. \]

Suppose \( u \in H_1 \) and \( v \in H_2 \) are stationary points for the functional \( I : H_1 \times H_2 \to \mathbb{R} \) such that

\[ I(u,v) = \iint_{\Omega} F(x,y,u,v,u_x,v_x,u_y,v_y)dx\,dy \]

where \( F \) is a \( C^2 \)-function. Let \( \eta = \eta(x,y) \) and \( \gamma = \gamma(x,y) \) be arbitrary \( C^2 \) functions which vanish on \( \partial \Omega \). Then \( u + \varepsilon_1 \eta \in H_1 \) and \( v + \varepsilon_2 \gamma \in H_2 \).
Let $\phi(\varepsilon_1, \varepsilon_2) = I(u + \varepsilon_1 \eta, v + \varepsilon_2 \eta)$. Since $u$ and $v$ are stationary points for $I$,

$$\frac{\partial \phi}{\partial \varepsilon_1} |_{\varepsilon_1=0} + \frac{\partial \phi}{\partial \varepsilon_2} |_{\varepsilon_2=0} = 0,$$

or

$$\iint (F_u \eta + F_{ux} \eta_x + F_{uy} \eta_y) \, dx \, dy + \iint (F_v \eta + F_{vx} \eta_x + F_{vy} \eta_y) \, dx \, dy$$

$$= \iint (F_u \eta - \frac{\partial}{\partial x} F_{ux} \eta - \frac{\partial}{\partial y} F_{uy} \eta) \, dx \, dy + \iint (F_v \eta - \frac{\partial}{\partial x} F_{vx} \eta - \frac{\partial}{\partial y} F_{vy} \eta) \, dx \, dy$$

$$+ \iint \left[ \frac{\partial}{\partial x} (F_u \eta) + \frac{\partial}{\partial y} (F_u \eta) \right] \, dx \, dy + \iint \left[ \frac{\partial}{\partial x} (F_v \eta) + \frac{\partial}{\partial y} (F_v \eta) \right] \, dx \, dy = 0.$$

Since $\eta = 0$ and $\gamma = 0$ on $\partial \Omega$,

$$\iint [\frac{\partial}{\partial x} (F_u \eta) + \frac{\partial}{\partial y} (F_u \eta)] \, dx \, dy = \int_{\partial \Omega} -F_u \eta \, dx + F_{ux} \eta \, dy = 0,$$

and

$$\iint [\frac{\partial}{\partial x} (F_v \eta) + \frac{\partial}{\partial y} (F_v \eta)] \, dx \, dy = \int_{\partial \Omega} -F_v \eta \, dx + F_{vx} \eta \, dy = 0.$$

Therefore

$$\iint \eta (F_u - \frac{\partial}{\partial x} F_{ux} - \frac{\partial}{\partial y} F_{uy}) \, dx \, dy + \iint \gamma (F_v - \frac{\partial}{\partial x} F_{vx} - \frac{\partial}{\partial y} F_{vy}) \, dx \, dy = 0.$$

Since $\eta$ and $\gamma$ are arbitrary functions, we obtain the Euler-Lagrange equations

$$F_u - \frac{\partial}{\partial x} F_{ux} - \frac{\partial}{\partial y} F_{uy} = 0,$$

$$F_v - \frac{\partial}{\partial x} F_{vx} - \frac{\partial}{\partial y} F_{vy} = 0.$$

Using this variational principle, we have to find the functional $I = I(u,v)$ such that one of the Euler-Lagrange equations, $F_v - \frac{\partial}{\partial x} F_{vx} - \frac{\partial}{\partial y} F_{vy} = 0$, gives the heat equation. This is called the adjoint formulation. In
chapter I, we derived the equation of the steady-state membrane by mini-
mizing the functional of the potential energy and we derived the wave
equation from the functional of the difference in the kinetic energy
and the potential energy. Unlike these cases, the adjoint formulation
has no physical meaning, so that it requires the additional work of
constructing the functional $I = I(u,v)$. Some techniques are available
for finding the adjoint formulation, see e.g. Telega [15]. Thus we will
make use of the adjoint formulation for the heat equation in chapter III.
CHAPTER III

METHODS FOR TIME DEPENDENT PROBLEMS

Laplace Transform

Let $E$ be the set of all continuous functions $f : [0, \infty) \rightarrow \mathbb{R}$ such that

$$\int_0^\infty e^{-st}f(t)\,dt \in \mathbb{R}$$

for each $t \in [0, \infty)$. The Laplace transform $L$ is a mapping of $E$ to the set of functions defined on $\mathbb{R}$ such that

$$L(f)(s) = \int_0^\infty e^{-st}f(t)\,dt .$$

Since

$$L(af + \beta g)(s) = \int_0^\infty e^{-st}[af(t) + \beta g(t)]\,dt = a\int_0^\infty e^{-st}f(t)\,dt + \beta \int_0^\infty e^{-st}g(t)\,dt = [aL(f) + \beta L(g)](s),$$

$L$ is a linear transformation. Using this Laplace transform, we make initial-boundary value problems into ordinary differential equations (boundary value problems) and transform the approximate solutions of the ordinary differential equations found by the Galerkin method back to obtain the approximate solutions of the original initial value problems. More detailed discussions of the method of Laplace transform are found in Davies [6].

Specifically, for the heat equation

$$u_t = c^2u_{xx}, \quad u = u(x,t), \quad 0 < x < 1, \quad t > 0 ,$$
subject to \[ u(0,t) = v_1(t), \]
\[ u(1,t) = v_2(t), \]
\[ u(x,0) = u_0(x). \] ... (3.1)

Let \( y(x,s) = \int_0^\infty e^{-st}u(x,t)dt \).

Then\[ \frac{\partial^2 y}{\partial x^2} = \frac{\partial^2}{\partial x^2} \int_0^\infty e^{-st}u(x,t)dt = \int_0^\infty e^{-st} \frac{\partial^2 u}{\partial x^2} dt = \frac{1}{c^2} \int_0^\infty e^{-st} \frac{\partial u}{\partial t} dt \]
\[ = \frac{1}{c^2} \int_0^\infty e^{-st} \left[ u(x,t) \right] dt + \frac{s}{c^2} \int_0^\infty e^{-st} u(x,t) dt \]
\[ = \frac{1}{c^2} \left[ -u(x,0) + s \int_0^\infty e^{-st} u(x,t) dt \right] = \frac{1}{c^2} \left[ -u_0(x) + sy(x,s) \right]. \]

The boundary conditions \( u(0,t) = v_1(t) \) and \( u(1,t) = v_2(t) \) are transformed to give boundary conditions
\[ y(0,s) = \int_0^\infty e^{-st}u(0,t)dt = \int_0^\infty e^{-st}v_1(t)dt, \]
\[ y(1,s) = \int_0^\infty e^{-st}u(1,t)dt = \int_0^\infty e^{-st}v_2(t)dt. \]

Hence, without loss of generality we write the ordinary differential equation
\[ y'' = \frac{s}{c^2} y - \frac{1}{c^2} u_0(x), \quad 0 < x < 1, \]
Subject to \( y(0) = \int_0^\infty e^{-st}v_1(t)dt \) and \( y(1) = \int_0^\infty e^{-st}v_2(t)dt. \)

We find the approximate solution of this problem by the Galerkin method. Then by the linearity of Laplace transform, the value of the approximate solution at \( x = x_1 \) for some \( x_1 \) in \((0,1)\) is transformed to obtain the approximate solution of the original problem (3.1) at \( x = x_1 \).

The next example is illustrative of this technique.

Example 3.1
\[ u_t = u_{xx}, \quad u = u(x,t), \quad 0 < x < 1, \quad t > 0. \]
subject to

\[ \begin{align*}
  u(0,t) &= t, \\
  u(1,t) &= 1 + t, \\
  u(x,0) &= x^2.
\end{align*} \]

Laplace transform gives the ordinary differential equation

\[-y'' = -sy + x^2.\]

with boundary conditions

\[ \begin{align*}
  y(0) &= \int_0^\infty e^{-st} dt = \frac{1}{s^2}, \\
  y(1) &= \int_0^\infty e^{-st}(1 + t) dt = \frac{1}{s} + \frac{1}{s^2}.
\end{align*} \]

Let \( g(x) = \frac{x}{s} + \frac{1}{s^2} \) so that \( g \) satisfies the boundary condition. If \( v \) is a function such that \( v(0) = v(1) = 0 \), then by integrating by parts

\[ \int_0^1 y''v \, dx = \int_0^1 y'v' \, dx. \]

Hence we obtain the weak formulation

\[ \int_0^1 y'v' \, dx + s\int_0^1 yv \, dx = \int_0^1 x^2 v \, dx. \]

Define

\[ \phi(x) = \begin{cases} 
  2x & \text{for } 0 \leq x \leq \frac{1}{2} \\
  -2x + 2 & \text{for } \frac{1}{2} \leq x \leq 1.
\end{cases} \]

Define

\[ U(x) = a\phi(x) + g(x). \]

Then \( a \) is determined by

\[ \int_0^\infty U' \phi' \, dx + s\int_0^\infty U \phi \, dx = \int_0^\infty x^2 \phi \, dx, \]

or

\[ a\int_0^1 \phi'^2 \, dx + \frac{1}{s}\int_0^1 \phi' \, dx + sa\int_0^1 \phi^2 \, dx + \int_0^1 x\phi \, dx + \frac{1}{s}\int_0^1 \phi \, dx = \int_0^1 x^2 \phi \, dx, \]
and we obtain
\[ \alpha = \frac{-5s - 24}{16s(s + 12)} , \]
and
\[ U(x) = \frac{-5s - 24}{16s(s + 12)} \phi(x) + \frac{x}{s} + \frac{1}{s^2} . \]

If we look for the approximate solution at \( x = \frac{1}{2} \)
\[ U\left(\frac{1}{2}\right) = \frac{3}{8s} - \frac{3}{16(s + 12)} + \frac{1}{s^2} . \]

Since \( \int_0^\infty e^{-ts} \, dt = \frac{1}{s} \), \( \int_0^\infty e^{-st} e^{-12t} \, dt = \frac{1}{s + 12} \), and \( \int_0^\infty e^{-st} \, dt = \frac{1}{s^2} \),

by the linearity of Laplace transform, we obtain the approximate solution
at \( x = \frac{1}{2} \)
\[ \frac{3}{8} - \frac{3}{16} e^{-12t} + t . \]

This technique can be used for hyperbolic problems. Consider the
wave equation
\[ u_{tt} = c^2 u_{xx} \, , \quad u = u(x,t) \, , \quad 0 < x < 1 \, , \quad t > 0 \]
subject to the initial-boundary condition
\[ u(0,t) = u(1,t) = 0 \, , \]
\[ u(x,0) = u_0(x) \, , \]
\[ \frac{\partial u}{\partial t}(x,0) = v_0(x) . \]

Let
\[ y(x,s) = \int_0^\infty e^{-st} u(x,t) \, dt . \]
Integrating by parts
\[
\frac{\partial^2 y}{\partial x^2} = \int_0^\infty e^{-st} \frac{\partial^2 u}{\partial x^2} \, dt = \frac{1}{c^2} \int_0^\infty e^{-st} \frac{\partial^2 u}{\partial t^2} \, dt
\]
\[
= \frac{1}{c^2} [v_0(x) + s \int_0^\infty e^{-st} \frac{\partial u}{\partial t} \, dt] = \frac{1}{c^2} [v_0(x) - su_0(x) + s^2 \int_0^\infty e^{-st} u \, dt]
\]
\[
= \frac{1}{c^2} [s^2 y(x,s) - su_0(x) - v_0(x)].
\]

And the boundary condition becomes
\[
y(0,s) = y(1,s) = 0.
\]

Hence we obtain the boundary value problem
\[
-c^2 y'' + s^2 y = su_0(x) + v_0(x).
\]

Then the approximate solution of this ordinary differential equation is found by the Galerkin method and transformed back to give the approximate solution of (3.2).

**Example 3.2**
\[
\begin{align*}
\frac{\partial^2 u}{\partial x^2} &= \frac{\partial^2 u}{\partial t^2}, & 0 < x < 3, & t > 0, \\
\end{align*}
\]
subject to
\[
\begin{align*}
u(0,t) &= u(3,t) = 0, \\
u(x,0) &= x^2 - 3x, \\
\frac{\partial u}{\partial t}(x,0) &= x^2.
\end{align*}
\]

Laplace transform gives
\[
\begin{align*}
-y'' + s^2 y &= s(x^2 - 3x) + x^2 = (s + 1)x^2 - 3sx, & 0 < x < 3, \\
y(0) &= y(3) = 0.
\end{align*}
\]

We define
\[ \phi_1(x) = \begin{cases} x & \text{for } 0 \leq x \leq 1 \\ -x + 2 & \text{for } 1 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases} , \]
\[ \phi_2(x) = \begin{cases} x - 1 & \text{for } 1 \leq x \leq 2 \\ -x + 3 & \text{for } 2 \leq x \leq 3 \\ 0 & \text{otherwise} \end{cases} . \]

The Galerkin approximation

\[ U(x) = a_1 \phi_1(x) + a_2 \phi_2(x) \]

is determined by

\[ \int_0^3 U'(x) \phi'_j(x) dx + s^2 \int_0^3 U(x) \phi_j(x) dx = \int_0^3 [(s + 1)x^2 - 3sx] \phi_j(x) dx , \]

or

\[ \frac{2}{3} a_i \int_0^3 \phi'_i(x) \phi'_j(x) dx + s^2 \sum_{i=1}^{2} a_i \int_0^3 \phi_i(x) \phi_j(x) dx = \int_0^3 [(s + 1)x^2 - 3sx] \phi_j(x) dx \]

for \( j = 1, 2 \), and we obtain

\[ 2a_1 - a_2 + s^2 \left[ \frac{2}{3} a_1 + \frac{1}{6} a_2 \right] = \frac{7}{6} - \frac{47}{6} s , \]

\[ -a_1 + 2a_2 + s^2 \left[ \frac{1}{6} a_1 + \frac{2}{3} a_2 \right] = \frac{25}{6} - \frac{11}{6} s . \]

Then

\[ a_1 = \frac{-59s^3 + s^2 - 210s + 78}{(5s^2 + 6)(s^2 + 6)} = \frac{-29}{5} \left( \frac{s}{s^2 + 6} \right) + \frac{16}{5} \left( \frac{\sqrt{6}}{s^2 + 6} \right) - 6 \left( \frac{s}{s^2 + 6} \right) \]

\[ - \frac{3}{\sqrt{6}} \left( \frac{6}{s^2 + 6} \right) , \]

\[ a_2 = \frac{s^3 - 31s^2 - 138s + 114}{(5s^2 + 6)(s^2 + 6)} = \frac{-29}{5} \left( \frac{s}{s^2 + 6} \right) + \frac{63}{2 \sqrt{30}} \left( \frac{\sqrt{6}}{s^2 + 6} \right) + 6 \left( \frac{s}{s^2 + 6} \right) \]

\[ - \frac{25}{2 \sqrt{6}} \left( \frac{\sqrt{6}}{s^2 + 6} \right) . \]
Since \( U(1) = a_1 \) and \( U(2) = a_2 \), and

\[
\int_0^\infty e^{-st}\cos bt \, dt = \frac{s}{s^2 + b^2} \quad \text{and} \quad \int_0^\infty e^{-st}\sin bt \, dt = \frac{b}{s^2 + b^2} ,
\]

the approximate solution is

\[
\begin{align*}
&\frac{-29}{5}\cos\frac{\sqrt{6}}{5} t + \frac{16}{\sqrt{30}}\sin\frac{\sqrt{6}}{5} t - 6\cos\sqrt{6} t - \frac{3}{\sqrt{6}}\sin\sqrt{6} t \quad \text{at } x = 1 , \\
&\frac{-29}{5}\cos\frac{\sqrt{6}}{5} t + \frac{63}{2\sqrt{30}}\sin\frac{\sqrt{6}}{5} t + 6\cos\sqrt{6} t - \frac{25}{2\sqrt{6}}\sin\sqrt{6} t \quad \text{at } x = 2 .
\end{align*}
\]

**Variational Methods**

We have seen that the Ritz approximation for a linear boundary value problem defined by a positive definite self-adjoint operator is the best approximation in terms of the energy space. This is not true in general for time-dependent problems. For example, the operator which defines the wave equation is not positive definite. Moreover, there are difficulties in formulating the Ritz method or the Galerkin method because of the initial-boundary conditions or in finding functionals for parabolic equations like the heat equation using the adjoint formulation. However, in this section we will show that the variational principle can still be used to define approximate solutions for initial-boundary value problems. More detailed discussions are given in Whiteman [19] or Mitchell and Wait [10].

Consider the wave equation

\[
u_{tt} = c^2 u_{xx} , \quad u = u(x,t) , \quad 0 < x < 1 , \quad t > t_0 ,
\]

subject to

\[
u(x,t) = u(1,t) = 0 ,
\]

\[
u(x,t_0) = u_0(x) ,
\]

\[
\frac{\partial u}{\partial t}(x,t_0) = v_0(x) .
\]
We assume that the initial condition \( \frac{\partial u}{\partial t}(x,t_0) = v_0(x) \) is replaced with the boundary condition \( u(x,t_1) = u_1(x) \) for some \( t_1 > t_0 \). Then by Hamilton's principle, the corresponding functional \( I \) is given by

\[
I(u) = \frac{1}{2} \int_0^{t_0} \left( u_t^2 - c^2 u_x^2 \right) dx dt,
\]

as we have seen in Chapter I. We partition \([0,1]\) into \( 0 = x_0 < x_1 < \ldots < x_{m+1} = 1 \) such that \( x_{i+1} - x_i = h_1 \) for each \( i \) where \( h_1 = \frac{1}{m+1} \), and

\([t_0,\infty) \) into \( t_0 = \tau_0 < \tau_1 < \ldots \) such that \( \tau_{n+1} - \tau = h_2 \) for each \( n = 1,2,\ldots \) for some \( h_2 \). Assume that \( u \) is given for \( u(x,\tau_n) \) and \( u(x,\tau_{n+1}) \) so that on \([0,1] \times [\tau_{n-1},\tau_{n+1}] \) the corresponding functional \( I \) is given by

\[
I(u) = \frac{1}{2} \int_{\tau_{n-1}}^{\tau_{n+1}} \left( u_t^2 - c^2 u_x^2 \right) dx dt.
\]

We define basis functions \( \phi_{ij} \) \( i = 1, \ldots, m \), \( j = n-1, n, n+1 \) on \([0,1] \times [\tau_{n-1},\tau_{n+1}] \) as

\[
\phi_{ij}(x,t) = \phi_i(x) \phi_j(t),
\]

where

\[
\phi_i(x) = \begin{cases} 
\frac{x - x_{i-1}}{h_1} & \text{for } x_{i-1} \leq x \leq x_i \\
\frac{x_{i+1} - x}{h_1} & \text{for } x_i \leq x \leq x_{i+1} \\
0 & \text{otherwise}
\end{cases}
\]

\[
\phi_{n-1}(t) = \begin{cases} 
\frac{\tau_{n-1} - t}{h_2} & \text{for } \tau_{n-1} \leq t \leq \tau_n \\
0 & \text{otherwise}
\end{cases}
\]

\[
\phi_n(t) = \begin{cases} 
\frac{t - \tau_{n-1}}{h_2} & \text{for } \tau_{n-1} \leq t \leq \tau_n \\
\frac{\tau_{n+1} - t}{h_2} & \text{for } \tau_n \leq t \leq \tau_{n+1} \\
0 & \text{otherwise}
\end{cases}
\]
\[\phi_{n+1}(t) = \begin{cases} \frac{t - \tau_n}{h_2} & \text{for } \tau_n \leq t \leq \tau_{n+1} \\ 0 & \text{otherwise} \end{cases} \]

The Ritz approximation \( U \) on \([0,1] \times [\tau_{n-1}, \tau_{n+1}] \) is defined by

\[ U(x,t) = \sum_{j=n-1}^{n+1} \sum_{i=1}^{m} a_{ij} \phi_j(x,t). \]

Since we are assuming \( u \) is given for \( t = \tau_{n-1} \) and \( t = \tau_{n+1} \), we can determine \( a_i,n-1 \) and \( a_i,n+1 \) for each \( i \) by imposing the boundary condition, that is,

\[ a_i,n-1 = u(x_i,\tau_{n-1}) \text{ and } a_i,n+1 = u(x_i,\tau_{n+1}) \text{ for each } i. \]

The coefficients \( a_i,n \) for \( i = 1, \ldots, m \) are determined by equations

\[ \frac{\partial}{\partial a_{k,n}} I(U) = 0, \quad k = 1 \ldots m, \]

or

\[ \frac{\partial}{\partial a_{k,n}} I(U) = \frac{\tau_{n+1}}{\tau_{n-1}} \left( \sum_{j=n-1}^{n+1} \sum_{i=1}^{m} \alpha_{ij} \frac{\partial \phi_j}{\partial t} \right) \frac{\partial \phi_k}{\partial t} - c^2 \left( \sum_{j=n-1}^{n+1} \sum_{i=1}^{m} \alpha_{ij} \frac{\partial \phi_j}{\partial x} \right) \frac{\partial \phi_k}{\partial x} \]

\[ = \sum_{j=n-1}^{n+1} \sum_{i=1}^{m} \alpha_{ij} \int_{\tau_{n-1}}^{\tau_{n+1}} \frac{\partial \phi_j}{\partial t} \frac{\partial \phi_k}{\partial t} dt - c^2 \sum_{j=n-1}^{n+1} \sum_{i=1}^{m} \alpha_{ij} \frac{\partial \phi_j}{\partial x} \frac{\partial \phi_k}{\partial x} \int_{\tau_{n-1}}^{\tau_{n+1}} dt. \]

Since

\[ \int_0^1 \phi_i(x) \phi_k(x) dx = \begin{cases} \frac{2}{h_1} & \text{for } i = k \\ \frac{1}{h_1} & \text{for } |i-k| = 1 \end{cases}, \]

\[ \int_0^1 \frac{d}{dx} \phi_i(x) \frac{d}{dx} \phi_k(x) dx = \begin{cases} \frac{2}{h_1} & \text{for } i = k \\ \frac{1}{h_1} & \text{for } |i-k| = 1 \end{cases}, \]

\[ \int_{\tau_{n-1}}^{\tau_{n+1}} \frac{d}{dt} \phi_j(t) \frac{d}{dt} \phi_k(t) dt = \begin{cases} \frac{2}{h_2} & \text{for } j = n \\ \frac{1}{h_2} & \text{for } |j-n| = 1 \end{cases}. \]
\[ \int_{\tau_{n-1}}^{\tau_n} \phi_j(t) \phi_n(t) \, dt = \begin{cases} \frac{2h_2}{3} & \text{for } j = n \\ h_2 & \text{for } |j-n| = 1 \end{cases} \]

\[ \begin{align*}
\int_{\tau_{n-1}}^{\tau_n} \phi_j(t) \phi_n(t) \, dt &= \begin{cases} \frac{2h_2}{3} & \text{for } j = n \\ h_2 & \text{for } |j-n| = 1 \end{cases} \\
&= 0 
\end{align*} \]

We obtain equations for \( a_{k,n} \)'s

\[ \begin{align*}
&\left( \frac{-h_1^2}{6h_2} + \frac{c^2h_2}{6h_1} \right) a_{k-1,n-l} + \left( \frac{-2h_1}{3h_2} - \frac{c^2h_2}{3h_1} \right) a_{k,n-l} + \left( \frac{-h_1}{6h_2} + \frac{c^2h_2}{6h_1} \right) a_{k+1,n-l} \\
&+ \left( \frac{h_1}{3h_2} + \frac{c^2h_2}{3h_1} \right) a_{k-1,n} + \left( \frac{4h_1}{3h_2} - \frac{4c^2h_2}{3h_1} \right) a_{k,n} + \left( \frac{h_1}{3h_2} + \frac{2c^2h_2}{3h_1} \right) a_{k+1,n} \\
&+ \left( \frac{-h_1}{6h_2} + \frac{c^2h_2}{6h_1} \right) a_{k-1,n+1} + \left( \frac{-2h_1}{3h_2} - \frac{c^2h_2}{3h_1} \right) a_{k,n+1} + \left( \frac{-h_1}{6h_2} + \frac{c^2h_2}{6h_1} \right) a_{k+1,n+1}
\end{align*} \]

\[ = 0 \quad k = 1, \ldots, m \quad \ldots (3.3) \]

We made the assumption that \( u \) is given for \( t = \tau_{n-1} \) and \( t = \tau_{n+1} \) to derive these equations. However, we are given an initial-boundary value problem, not a boundary value problem and \( u \) is not given at \( t = \tau_{n+1} \) for any \( n \) (\( n = 0, 1, 2, \ldots \)). Nevertheless we have initial conditions

\[ u(x,t_0) = u_0(x), \]
\[ \frac{\partial u(x,t_0)}{\partial t} = v_0(x). \]

Hence \( a_{i,0} \)'s are determined by

\[ a_{i,0} = u_0(x_i), \quad i = 1 \ldots m, \]

and we can approximate \( a_{i,1} \)'s by

\[ a_{i,1} = a_{i,0} + h_2 v_0(x_i), \quad i = 1 \ldots m. \]

Then the coefficients \( a_{i,2} \)'s are determined by using equations (3.3), and hence we obtain the approximate solution at \( t = \tau_2 \). Once this is done, using information at \( t = \tau_1 \) and \( t = \tau_2 \), and again by equation (3.3), we can find the approximate solution at \( t = \tau_3 \). Repeating this procedure,
we can find the approximate solution at any time step \( t = \tau_n \). This is called the step-by-step method, and we will illustrate this technique in the next example.

Example 3.3

\[
-u_{tt} - u_{xx} = 0, \quad u = u(x,t), \quad 0 < x < 3, \quad t > 0,
\]

subject to

\[
\begin{align*}
  u(0,t) &= u(3,t) = 0, \\
  u(x,0) &= 3x^2 - x^3, \\
  \frac{\partial u}{\partial t}(x,0) &= x.
\end{align*}
\]

We partition \([0,3]\) into \( 0 < 1 < 2 < 3 \) so that \( h_1 = 1 \) and choose \( h_2 = 1 \).

From the initial condition,

\[
\begin{align*}
  \alpha_{10} &= u(1,0) = 2, \quad \alpha_{20} = u(2,0) = 4, \\
  \alpha_{11} &= \alpha_{10} + \frac{\partial u}{\partial t}(1,0) = 3, \quad \alpha_{21} = \alpha_{20} + \frac{\partial u}{\partial t}(2,0) = 8.
\end{align*}
\]

From (3.3), we get

\[
\begin{align*}
  -\alpha_{10} + \alpha_{21} - \alpha_{12} &= 0, \\
  -\alpha_{20} + \alpha_{11} - \alpha_{22} &= 0,
\end{align*}
\]

and we obtain the approximate solution at \( t = 2 \) defined by \( \alpha_{12} = 4 \) and \( \alpha_{22} = -1 \).

Next, we derive a step-by-step approximation for the heat equation

\[
-u_t = c^2 u_{xx}, \quad u = u(x,t), \quad 0 < x < 1, \quad t > t_0,
\]

subject to

\[
u(0,t) = u(1,t) = 0.
\]
In section 2.5, we showed that if $u = u(x,y)$ and $v = v(x,y)$ are given on $\partial \Omega$, and stationary points for the functional
\[
I(u,v) = \iint_{\Omega} F(x,y,u,v,u_x,v_x,u_y,v_y) \, dx \, dy
\]
among the set $H_1 \times H_2$ where $H_1$ and $H_2$ are the sets of those functions which agree with $u$ and $v$ respectively, then they satisfy the Euler-Lagrange equations
\[
F_u - \frac{\partial^2 F}{\partial x \partial y} u_x - \frac{\partial^2 F}{\partial y} u_y = 0,
\]
\[
F_v - \frac{\partial^2 F}{\partial x \partial y} v_x - \frac{\partial^2 F}{\partial y} v_y = 0.
\]

We will show that if $u$ satisfies the given initial-boundary condition in (3.4) and if $v$ satisfies
\[
v(0,t) = v(1,t) = 0,
\]
\[
v(x,t_1) = 0, \quad \text{for some } t_1 > t_0.
\]

Then one of the Euler-Lagrange equations of the functional
\[
I(u,v) = \int_{t_0}^{t_1} \int_0^1 (u_t v + c^2 u_x v_x) \, dx \, dt
\]
gives the heat equation (3.4). Let $u$ and $v$ are stationary points for the functional
\[
I(u,v) = \int_{t_0}^{t_1} \int_0^1 F(x,y,u,v,u_x,v_x,u_t,v_t) \, dx \, dt = \int_{t_0}^{t_1} (u_t v + c u_x v_x) \, dx \, dt
\]
among $H_1 \times H_2$ where
\[
H_1 = \{u | u \text{ is } C^2, \text{ and } u(0,t) = u(1,t) = 0 \text{ and } u(x,t_0) = u_0(x)\},
\]
\[
H_2 = \{v | v \text{ is } C^2, \text{ and } v(0,t) = v(1,t) = 0 \text{ and } v(x,t_1) = 0\}.
\]
Let \((x,t)\) and \((x,t)\) be any \(C^2\)-functions such that
\[
\eta(0,t) = \eta(1,t) = 0, \quad \gamma(0,t) = \gamma(1,t) = 0,
\]
and
\[
\eta(x,t_0) = 0, \quad \gamma(x,t_1) = 0.
\]

Then \(u + \varepsilon_1 \eta \in H_1\) and \(v + \varepsilon_2 \gamma \in H_2\). Let
\[
\phi(\varepsilon_1, \varepsilon_2) = I(u + \varepsilon_1 \eta, v + \varepsilon_2 \gamma).
\]

Since \(u\) and \(v\) are stationary points for \(I\),
\[
\frac{\partial^2 I}{\partial \varepsilon_1 |_{\varepsilon_1} = 0} + \frac{\partial^2 I}{\partial \varepsilon_2 |_{\varepsilon_2} = 0} = 0,
\]
or
\[
\int_0^1 \int_0^1 \eta \left( F_u \frac{\partial}{\partial x} F_u_x - \frac{\partial}{\partial t} F_u_t \right) dx dt + \int_0^1 \int_0^1 \gamma \left( F_v - \frac{\partial}{\partial x} F_v_x - \frac{\partial}{\partial t} F_v_t \right) dx dt
\]
\[
+ \int_0^1 \int_0^1 \left[ \frac{\partial}{\partial x} (F_u \eta) + \frac{\partial}{\partial t} (F_u \eta) \right] dx dt + \int_0^1 \int_0^1 \left[ \frac{\partial}{\partial x} (F_v \gamma) + \frac{\partial}{\partial t} (F_v \gamma) \right] dx dt = 0.
\]

Now
\[
\int_0^1 \int_0^1 \left[ \frac{\partial}{\partial x} (F_u \eta) + \frac{\partial}{\partial t} (F_u \eta) \right] dx dt = \int_0^1 \left. F_u \eta \right|_{x=1} dt + \int_0^1 \left. F_u \eta \right|_{x=0} dt = 0
\]
\[
= \int_0^1 \eta \left. v \right|_{x=0} dx = 0
\]
and similarly we find
\[
\int_0^1 \int_0^1 \left[ \frac{\partial}{\partial x} (F_v \gamma) + \frac{\partial}{\partial t} (F_v \gamma) \right] dx dt = 0.
\]

Hence, we conclude that
\[
F_v - \frac{\partial}{\partial x} F_v_x - \frac{\partial}{\partial t} F_v_t = 0,
\]
or
\[ u_t = c^2 u_{xx}. \]

Therefore, in order to derive a step-by-step approximation for (3.4), we use the functional

\[ I(u,v) = \int_0^1 \int_0^1 (u_t v + c^2 u_{xx} v_x) \, dx \, dt, \]

where

\[ v(0,t) = v(1,t) = 0, \]
\[ v(x,t_1) = 0. \]

We partition \([0,1]\) into \(0 = x_0 < \ldots < x_{m+1} = 1\) such that \(x_{i+1} - x_i = h_1\) for each \(i\), and \([t_0,\infty)\) into \(t_0 = \tau_0 < \tau_1 < \ldots \) such that \(\tau_{n+1} - \tau_n = h_2\) for each \(n\). On \([0,1] \times [\tau_n,\tau_{n+1}]\), we define basis functions \(\phi_{ij}\), \(i = 1, \ldots, m\) and \(j = n, n+1\) such that

\[ \phi_{ij}(x,t) = \phi_i(x) \phi_j(t), \]

where

\[ \phi_i(x) = \begin{cases} \frac{x - x_{i-1}}{h_1} & \text{if } x_{i-1} \leq x < x_i, \\ \frac{x_{i+1} - x}{h_1} & \text{if } x_i \leq x < x_{i+1}, \\ 0 & \text{otherwise}, \end{cases} \]

\[ \phi_n(t) = \frac{\tau_{n+1} - t}{h_2}, \]

\[ \phi_{n+1}(t) = \frac{t - \tau_n}{h_2}. \]

We define the approximate solution \(U\) of (3.4) on \([0,1] \times [\tau_n,\tau_{n+1}]\) by

\[ U(x,t) = \sum_{j=n}^{n+1} \sum_{i=1}^{m} a_{ij} \phi_{ij}(x,t), \]

and define
\[ V(x,t) = \sum_{i=1}^{m} \beta_i \phi_i(x,t) . \]

Observe that \( V(x,n + 1) = 0 \).

If the approximate solution at \( t = T_n \) is known, that is, \( a_{in} \) are known for \( i = 1 \ldots m \), then \( a_{i,n+1} \) are obtained by equations

\[ \frac{\partial I(U,V)}{\partial \phi_k} = 0 \quad k = 1 \ldots m , \]

or

\[ \frac{3}{\tau_{n+1}} \int_{T_n}^{T_{n+1}} \left[ \left( \sum_{j=1}^{m} a_{ij} \right) \phi_k(x) + c^2 \left( \sum_{i=1}^{m} \beta_i \phi_i(x) \right) \right] dx \right] dt + c^2 \tau_{n+1} \int_{T_n}^{T_{n+1}} a_{ij} \phi_i(x) \phi_j(x) dx dt = 0 . \]

Since

\[ \int_{-h_1}^{h_1} \phi_i(x) \phi_j(x) dx = \begin{cases} \frac{2}{3h_1} & \text{for } i = k \\ \frac{1}{6h_1} & \text{for } |i - k| = 1 , \end{cases} \]

\[ \int_{-h_1}^{h_1} \phi_i'(x) \phi_j(x) dx = \begin{cases} \frac{2}{h_1} & \text{for } i = k \\ \frac{1}{h_1} & \text{for } |i - k| = 1 , \end{cases} \]

\[ \int_{T_n}^{T_{n+1}} \phi_j'(t) \phi_j(t) dt = \begin{cases} \frac{1}{2} & \text{for } j = n + l , \end{cases} \]

\[ \int_{T_n}^{T_{n+1}} \phi_j(t) \phi_j(t) dt = \begin{cases} \frac{h_2}{3} & \text{for } j = n \\ \frac{h_2}{6} & \text{for } j = n + 1 , \end{cases} \]

we obtain equations
\[
\begin{align*}
\left( -\frac{h_2}{12} - \frac{c^2 h_2}{3h_1} \right) a_{k-1,n} &+ \left( -\frac{h_1}{12} - \frac{c^2 h_2}{3h_1} \right) a_{k,n} &+ \left( -\frac{h_1}{12} - \frac{c^2 h_2}{3h_1} \right) a_{k+1,n} \\
+ \left( \frac{1}{12} - \frac{c^2 h_2}{6h_1} \right) a_{k-1,n+1} &+ \left( \frac{1}{3} + \frac{c^2 h_2}{3h_1} \right) a_{k,n+1} &+ \left( \frac{1}{12} - \frac{c^2 h_2}{6h_1} \right) a_{k+1,n+1} &= 0
\end{align*}
\]

for \( k = 1 \cdots m \). \hspace{1cm} \cdots (3.5)

Example 3.4

\[
u_t = u_{xx}, \quad 0 < x < 4, \quad t > 0,
\]

subject to

\[
u(0,t) = u(4,t) = 0,
\]

\[
u(x,0) = x^3 - 4x^2.
\]

Partition \([0,4]\) into \(0 < 1 < 2 < 3 < 4\) so that \(h_1 = 1\) and choose \(h_2 = 1\).

From the initial condition we find

\[
\begin{align*}
a_{10} &= u(1,0) = -3, \\
a_{20} &= u(2,0) = -8, \\
a_{30} &= u(3,0) = -9.
\end{align*}
\]

By equations (3.5), we obtain

\[
\begin{align*}
\frac{2}{3} a_{10} - \frac{5}{6} a_{20} + \frac{4}{3} a_{11} - \frac{1}{6} a_{21} &= 0, \\
-\frac{5}{6} a_{10} + \frac{2}{3} a_{20} - \frac{5}{6} a_{30} - \frac{1}{6} a_{11} + \frac{4}{3} a_{21} - \frac{1}{6} a_{31} &= 0, \\
-\frac{5}{6} a_{20} + \frac{2}{3} a_{30} - \frac{1}{6} a_{21} + \frac{4}{3} a_{31} &= 0
\end{align*}
\]

or

\[
\begin{align*}
8a_{11} - a_{21} &= -28, \\
-\alpha_{11} + 8\alpha_{21} - \alpha_{31} &= -28, \\
-\alpha_{21} + 8\alpha_{31} &= -4
\end{align*}
\]
and we find \( a_{11} = \frac{-249}{62}, a_{21} = \frac{-128}{31}, \) and \( a_{31} = \frac{-63}{62}, \) which define an approximate solution at \( t = 1. \)

An alternative method can be derived using the variational principle for the wave equation

\[
\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \quad u = u(x,t), \quad 0 < x < 1, \quad t > 0 \ldots \tag{3.6}
\]

subject to

\[
\begin{align*}
  u(0,t) &= u(t,t) = 0, \\
  u(x,0) &= u_0(x), \\
  \frac{\partial u}{\partial t}(x,0) &= v_0(x).
\end{align*}
\]

Assume that \( u \) is given for an arbitrary time \( t = \tau > 0, \) that is, \( u(x,\tau) = u_1(x). \) Partition \([0,1]\) into \( 0 = x_0 < x_1 < \cdots < x_{m+1} = 1 \)

such that \( x_{i+1} - x_i = h \) where \( h = \frac{1}{m + 1} \) and define Hat-functions \( \phi_i(x), \) \( i = 1 \cdots m \) on \([0,1]\) as in the last example. We define the approximate solution \( U \) on \([0,1] \times [0,\tau]\) to be

\[
U(x,t) = \sum_{i=1}^{m} \phi_i(x) \psi_i(t),
\]

in which \( \psi_i = \psi_i(t), i = 1 \cdots m, \) are determined using the variational principle. By our assumption that \( u \) is given at \( t = \tau, \) the corresponding functional \( I \) is given by

\[
I(u) = \int_0^\tau \int_0^1 (c^2 \frac{\partial^2 u}{\partial x^2} - u_1^2) \, dx \, dt
\]

Now

\[
I(U) = I\left( \sum_{i=1}^{m} \phi_i(x) \psi_i(t) \right)
\]

\[
= \int_0^\tau \int_0^1 [c^2 \left( \sum_{i=1}^{m} \frac{d}{dx} \phi_i(x) \psi_i(t) \right)^2 - \left( \sum_{i=1}^{m} \frac{d}{dt} \phi_i(x) \psi_i(t) \right)^2] \, dx \, dt
\]
\[
\begin{align*}
&= \int_{0}^{T} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{d}{dx} \phi_{i}(x) \frac{d}{dx} \phi_{j}(x) \psi_{i}(t) \psi_{j}(t) dt - \sum_{i=1}^{m} \sum_{j=1}^{m} \psi_{i}(x) \phi_{j}(x) \frac{d}{dt} \psi_{i}(t) \frac{d}{dt} \psi_{j}(t) \sum_{i=1}^{m} \sum_{j=1}^{m} \phi_{i}(x) \phi_{j}(x) \frac{d}{dt} \psi_{i}(t) \frac{d}{dt} \psi_{j}(t) dt \\
&= \int_{0}^{T} \sum_{i=1}^{m} \sum_{j=1}^{m} \left[ c^{2} \phi_{i}(x) \phi_{j}(x) \frac{d}{dx} \psi_{i}(t) \psi_{j}(t) dt - \int_{0}^{T} \phi_{i}(x) \phi_{j}(x) \frac{d}{dt} \psi_{i}(t) \frac{d}{dt} \psi_{j}(t) dt \right] \\
&= \int_{0}^{T} \sum_{i=1}^{m} \sum_{j=1}^{m} \left[ c^{2} \phi_{i}(x) \phi_{j}(x) \frac{d}{dx} \psi_{i}(t) \psi_{j}(t) dt - \int_{0}^{T} \phi_{i}(x) \phi_{j}(x) \frac{d}{dt} \psi_{i}(t) \frac{d}{dt} \psi_{j}(t) dt \right], \\
where \\
\alpha_{ij} = \int_{0}^{T} \phi_{i}(x) \phi_{j}(x) \psi_{i}(t) \psi_{j}(t) dt \\
\beta_{ij} = \int_{0}^{T} \phi_{i}(x) \phi_{j}(x) \frac{d}{dt} \psi_{i}(t) \frac{d}{dt} \psi_{j}(t) dt,
\end{align*}
\]

Hence the problem becomes: Find the stationary points of the functional

\[
I(\psi_1, \ldots, \psi_m) = \sum_{i=1}^{m} \sum_{j=1}^{m} \left[ c^{2} \alpha_{ij} \psi_{i}(t) \psi_{j}(t) - \beta_{ij} \psi_{i}(t) \psi_{j}(t) \right] dt,
\]

given \( \psi_{i}(0) = u_{0}(x_{i}) \) and \( \psi_{i}(T) = u_{1}(x_{i}) \), \( i = 1 \ldots m \). Define

\[
\psi_{i}(t) = (1 - \frac{t^{2}}{\tau^{2}})u_{0}(x_{i}) + \frac{t^{2}}{\tau^{2}}u_{1}(x_{i}) + \gamma_{i}(t - \frac{t^{2}}{\tau})
\]

so that \( \psi_{i} \) satisfies the condition

\[
\psi_{i}(0) = u_{0}(x_{i}), \quad \psi_{i}(T) = u_{1}(x_{i}), \quad i = 1 \ldots m.
\]

Since

\[
\frac{d}{dt} \psi_{i}(t) = \frac{-2t}{\tau^{2}}u_{0}(x_{i}) + \frac{2t}{\tau^{2}}u_{1}(x_{i}) + \gamma_{i}(1 - \frac{2t}{\tau}),
\]

\[
\frac{d}{dt} \psi_{i}(0) = \gamma_{i}.
\]

Now

\[
I(\psi_1, \ldots, \psi_m)
\]

\[
= \sum_{i=1}^{m} \sum_{j=1}^{m} \left[ c^{2} \alpha_{ij} \left(1 - \frac{t^{2}}{\tau^{2}}\right)u_{0}(x_{i}) + \frac{t^{2}}{\tau^{2}}u_{1}(x_{i}) + \gamma_{i}(t - \frac{t^{2}}{\tau}) \right] \left(1 - \frac{t^{2}}{\tau^{2}}\right)u_{0}(x_{j})
\]

\[
+ \frac{t^{2}}{\tau^{2}}u_{1}(x_{j}) + \gamma_{j}(t - \frac{t^{2}}{\tau}) - \beta_{ij} \left[ \frac{-2t}{\tau^{2}}u_{0}(x_{i}) + \frac{2t}{\tau^{2}}u_{1}(x_{i}) + \gamma_{i}(1 - \frac{2t}{\tau}) \right]
\]

\[
\left[ \frac{-2t}{\tau^{2}}u_{0}(x_{j}) + \frac{2t}{\tau^{2}}u_{1}(x_{j}) + \gamma_{j}(1 - \frac{2t}{\tau}) \right] dt.
\]
Since $\psi_1 \cdots \psi_m$ are stationary points for $I$,

\[ \frac{\partial I}{\partial y_i} = 0 \text{ for each } i, i=1 \cdots m, \text{ or} \]

\[ \frac{\partial I}{\partial y_i} = \sum_{k=1}^{m} \tau t^2 (a_{ik} + \alpha_{ki}) (t - \frac{t^2}{\tau}) \left[ (1 - \frac{t^2}{\tau^2}) u_0(x_k) + \frac{t^2}{\tau^2} u_1(x_k) + \gamma_k (t - \frac{t^2}{\tau}) \right] \]

\[- (\beta_{ik} + \beta_{ki}) (1 - \frac{2t}{\tau}) \left[ \frac{2t}{\tau} u_0(x_k) + \frac{2t}{\tau} u_1(x_k) + \gamma_k (1 - \frac{2t}{\tau}) \right] dt = 0, \]

i = 1 \cdots m.

Since

\[ a_{ij} = \begin{cases} 2 & \text{for } i = j \\ \frac{1}{h} & \text{for } |i-j| = 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \beta_{ij} = \begin{cases} 2h & \text{for } i = j \\ \frac{h}{6} & \text{for } |i-j| = 1 \\ 0 & \text{otherwise} \end{cases}, \]

we obtain

\[ c^2 \left( -2 \frac{t^2}{30h} u_0(x_{i-1}) - \frac{t^2}{10h} u_1(x_{i-1}) - \frac{t^3}{15h} \gamma_{i-1} \right) - \frac{h}{9} u_0(x_{i-1}) + \frac{h}{9} u_1(x_{i-1}) - \frac{h}{9} \gamma_{i-1} \]

\[ + c^2 \left( \frac{t^2}{15h} u_0(x_i) + \frac{t^2}{5h} u_1(x_i) + \frac{2t^3}{15h} \gamma_i \right) - \frac{4h}{9} u_0(x_i) + \frac{4h}{9} u_1(x_i) - \frac{4h}{9} \gamma_i \]

\[ + c^2 \left( -2 \frac{t^2}{30h} u_0(x_{i+1}) - \frac{t^2}{10h} u_1(x_{i+1}) - \frac{t^3}{15h} \gamma_{i+1} \right) - \frac{h}{9} u_0(x_{i+1}) + \frac{h}{9} u_1(x_{i+1}) \]

\[- \frac{h}{9} \gamma_{i+1} = 0, \quad i = 1 \cdots m. \]

Since $\psi_i(0) = \gamma_i$, we let $\gamma_i = v_0(x_i)$, then we obtain equations for the unknowns $u_1(x_1), u_1(x_2), \cdots u_1(x_m)$, i.e.,

\[ (1 - \frac{9}{10} \mu) u_1(x_{i-1}) + (4 + \frac{9}{5} \mu) u_1(x_i) + (1 - \frac{9}{10} \mu) u_1(x_{i+1}) \]

\[ = (1 + \frac{21}{10} \mu) u_0(x_{i-1}) + (4 - \frac{21}{5} \mu) u_0(x_i) + (1 + \frac{21}{10} \mu) u_0(x_{i+1}) \]

\[ + \tau \left[ (1 + \frac{3}{5} \mu) v_0(x_{i-1}) + (4 - \frac{6}{5} \mu) v_0(x_i) + (1 + \frac{3}{5} \mu) v_0(x_{i+1}) \right] \]
for $i = 1 \cdots m$, where $\mu = \frac{c^2 T^2}{h^2}$. \hspace{1cm} \ldots \hspace{1cm} (3.7)$

Example 3.5

$u_{tt} = u_{xx}$, \hspace{0.5cm} $u = u(x,t)$, \hspace{0.5cm} $0 < x < 4$, \hspace{0.5cm} $t > 0$,

subject to

$u(0,t) = u(4,t) = 0$, \hspace{1cm} $u(x,0) = 4x - x^2$, \hspace{1cm} $\frac{\partial u(x,0)}{\partial t} = x^3$.

Choose $\tau = 1$, and $m = 3$ so that $h = 1$. From (3.7), we obtain equations

\[
\frac{1}{10} u_1(x_{i-1}) + \frac{22}{5} u_1(x_i) + \frac{1}{10} u_1(x_{i+1}) = \frac{31}{10} u_0(x_{i-1}) - \frac{1}{5} u_0(x_i) + \frac{31}{10} u_0(x_{i+1}) \\
+ \frac{8}{5} v_0(x_{i-1}) + \frac{14}{5} v_0(x_i) + \frac{8}{5} v_0(x_{i+1}), \hspace{1cm} i = 1,2,3,
\]

or

\[
\frac{29}{5} u_1(1) + \frac{1}{10} u_1(2) = \frac{137}{5}, \hspace{1cm} \frac{1}{10} u_1(1) + \frac{29}{5} u_1(2) + \frac{1}{10} u_1(3) = 85, \hspace{1cm} \frac{1}{10} u_1(2) + \frac{29}{5} u_1(3) = \frac{501}{5},
\]

and we find that

$u_1(1) = 4.4779$, \hspace{0.5cm} $u_1(2) = 14.2844$, \hspace{0.5cm} and $u_1(3) = 17.0296$,

which are values of the approximate solution at $(1.1)$, $(2.1)$, and $(3.1)$ respectively.

After having found values of the approximate solution at $t = \tau$, in order to get to the next time step $t = 2\tau$, we can use the step-by-step method introduced earlier in this section since we have the approximate
solution at time 0 and time \( T \). Alternatively, since

\[
\psi_i(t) = (1 - \frac{t^2}{\tau^2})u_0(x_i) + \frac{t^2}{\tau^2} u_1(x_i) + v_0(x_i)(t - \frac{t^2}{\tau}) ,
\]

\[
\psi'_i(\tau) = \frac{-2}{\tau} u_0(x_i) + \frac{2}{\tau} u_1(x_i) - v_0(x_i) .
\]

Let

\[
v_1(x_i) = \psi'_i(\tau) ,
\]

and use these numbers for

\[
(1 - \frac{9}{10^u})u_2(x_{i-1}) + (4 + \frac{9}{5^u})u_2(x_i) + (1 - \frac{9}{10^u})u_2(x_{i+1})
\]

\[
= (1 + \frac{21}{10^u})u_1(x_{i-1}) + (4 - \frac{21}{5^u})u_1(x_i) + (1 + \frac{21}{10^u})u_1(x_{i+1})
\]

\[
+ \tau [(1 + \frac{3}{5^u})v_1(x_{i-1}) + (4 - \frac{6}{5^u})v_1(x_i) + (1 + \frac{3}{5^u})v_1(x_{i+1})] ,
\]

\[i = 1 \cdots m . \cdots (3.8)\]

For the previous example, since

\[
u_1(1) = 4.4779 , \quad u_1(2) = 14.2844 , \quad u_1(3) = 17.0296 ,
\]

\[
v_1(1) = 1.9558 , \quad v_1(2) = 12.5688 , \quad v_1(3) = 1.0592 ,
\]

from (3.8), we obtain

\[
58u_2(1) + u_2(2) = 689.7238 ,
\]

\[
u_2(1) + 58u_2(2) + u_2(3) = 1038.3301 ,
\]

\[
u_2(2) + 58u_2(3) = 639.5156 ,
\]

and we find

\[
u_2(1) = 11.5898 , \quad u_2(2) = 17.5175 , \text{ and } u_2(3) = 10.7241 .
\]
which are values of the approximate solution at (1,2), (2,2), and (3,2) respectively.

**Semi-Discrete Galerkin Method**

Probably the most widely used finite element method for time-dependent problems is the semi-discrete Galerkin method. In this section, we describe the semi-discrete Galerkin method with examples. More detailed discussions about the semi-discrete Galerkin method are given in Mitchell and Wait [10], and Fairweather [7].

Consider the parabolic initial-boundary value problem

\[ \frac{\partial u}{\partial t} - c^2 \frac{\partial^2 u}{\partial x^2} = f, \quad 0 < x < 1, \quad t > t_0 \]

subject to

\[ u(0,t) = u(1,t) = 0, \]

\[ u(x,t_0) = u_0(x). \]  \hspace{1cm} \ldots (3.9)

If

\[ v = v(x) \in H^1_0(\Omega) \text{ where } \Omega = (0,1), \]

integrating by parts

\[ (-c^2u_{xx}, v) = c^2 \int_0^1 u_{xx} v \, dx = c^2 \int_0^1 -\frac{d}{dx}(u_x v) \, dx + c^2 \int_0^1 u_x v_x \, dx \]

\[ = c^2 \int_0^1 u_x v_x \, dx . \]

Let

\[ a(u,v) = \int_0^1 u_x v_x \, dx . \]

Then the semi-discrete approximation is based on the weak formulation:
Find \( u \in H^1_2(\Omega) \) for each fixed but arbitrary time \( t \) such that

\[
(\frac{\partial u}{\partial t}, v) + c^2 a(u, v) = (f, v) \quad \text{for all } v \in H^1_2(\Omega),
\]

subject to

\[
(u, v)|_{t=t_0} = (u_0, v),
\]

We choose a finite dimensional subspace \( K_n = \langle \phi_1, \ldots, \phi_n \rangle \) of \( H^1_2(\Omega) \).

Then the semi-discrete Galerkin approximation \( U \) of (3.9) is defined by

\[
U(x, t) = \sum_{j=1}^{n} a_j(t) \phi_j(x)
\]

and determined by

\[
(\frac{\partial U}{\partial t}, \phi_i) + c^2 a(U, \phi_i) = (f, \phi_i),
\]

or

\[
\sum_{j=1}^{n} a_j(t) (\phi_j, \phi_i) + c^2 \sum_{j=1}^{n} a_j a(\phi_j, \phi_i) = (f, \phi_i), \quad i = 1 \ldots n
\]

Hence we obtain the first order system of ordinary differential equations

\[
A \dot{\alpha} + B \alpha = b,
\]

where

\[
A_{ij} = (\phi_j, \phi_i), \quad B_{ij} = c^2 a(\phi_j, \phi_i),
\]

\[
\alpha^T = [\alpha_1 \ldots \alpha_n] \quad \text{and} \quad b^T = [(f, \phi_1), \ldots, (f, \phi_n)] \ldots (3.10)
\]

with the initial condition

\[
(u, \phi_i)|_{t=t_0} = (u_0, \phi_i),
\]

or

\[
\sum_{j=1}^{n} a_j(t_0) (\phi_j, \phi_i) = (u_0, \phi_i), \quad i = 1 \ldots n.
\]
Examples 3.6

\[ u_t = u_{xx}, \quad u = u(x, t), \quad 0 < x < 3, \quad t > 0, \]

subject to \[ u(0, t) = u(3, t) = 0, \]

\[ u(x, 0) = u_0(x) = \begin{cases} 
  x & \text{for } 0 \leq x \leq 1 \\
  2 - x & \text{for } 1 \leq x \leq 2 \\
  0 & \text{for } 2 \leq x \leq 3 
\end{cases} \]

We define \( K_2 = \langle \phi_1, \phi_2 \rangle \in H^1_0(\Omega) \) where

\[ \phi_1(x) = \begin{cases} 
  x & \text{for } 0 \leq x \leq 1 \\
  2 - x & \text{for } 1 \leq x \leq 2 \\
  0 & \text{otherwise} 
\end{cases} \]

\[ \phi_2(x) = \begin{cases} 
  x - 1 & \text{for } 1 \leq x \leq 2 \\
  3 - x & \text{for } 2 \leq x \leq 3 \\
  0 & \text{otherwise} 
\end{cases} \]

The semi-discrete Galerkin approximation \( U = \sum_{j=1}^{2} \alpha_j \phi_j \) is determined by

\[ \frac{\partial U}{\partial t} + a(U, \phi_i) = 0, \quad i = 1, 2, \]

and by (3.10) we obtain the system of ordinary differential equations

\[
\begin{bmatrix}
(\phi_1, \phi_1) & (\phi_2, \phi_1) \\
(\phi_1, \phi_2) & (\phi_2, \phi_2)
\end{bmatrix}
\begin{bmatrix}
\alpha_1 \\
\alpha_2
\end{bmatrix}
+ \begin{bmatrix}
a(\phi_1, \phi_1) & a(\phi_2, \phi_1) \\
a(\phi_1, \phi_2) & a(\phi_2, \phi_2)
\end{bmatrix}
\begin{bmatrix}
\alpha_1 \\
\alpha_2
\end{bmatrix}
= \begin{bmatrix}
0 \\
0
\end{bmatrix}
\]

with the initial condition

\[
\begin{bmatrix}
(\phi_1, \phi_1) & (\phi_2, \phi_1) \\
(\phi_1, \phi_2) & (\phi_2, \phi_2)
\end{bmatrix}
\begin{bmatrix}
\alpha_1(0) \\
\alpha_2(0)
\end{bmatrix}
= \begin{bmatrix}
(u_0, \phi_1) \\
(u_0, \phi_2)
\end{bmatrix}.
\]

Since

\[ (\phi_i, \phi_j) = \begin{cases} 
  \frac{2}{3} & \text{for } i = j \\
  \frac{1}{6} & \text{for } |i-j| = 1
\end{cases}, \quad a(\phi_i, \phi_j) = \begin{cases} 
  2 & \text{for } i = j \\
  -1 & \text{for } |i-j| = 1
\end{cases}, \]
\[ (u_0, \phi_1) = \int_0^3 u_0(x) \phi_1(x) \, dx \]
\[ = \int_0^1 x^2 \, dx + \int_1^2 (2 - x)^2 \, dx = \frac{2}{3}, \]

and
\[ (u_0, \phi_2) = \int_0^3 u_0(x) \phi_2(x) \, dx \]
\[ = \int_1^2 (2 - x)(x - 1) \, dx = \frac{1}{6}, \]

we have
\[
\begin{bmatrix}
\frac{2}{3} & \frac{1}{6} \\
\frac{1}{6} & \frac{2}{3}
\end{bmatrix}
\begin{bmatrix}
\alpha_1' \\
\alpha_2'
\end{bmatrix}
\begin{bmatrix}
2 \\
-1
\end{bmatrix}
\begin{bmatrix}
\alpha_1 \\
\alpha_2
\end{bmatrix}
\begin{bmatrix}
0 \\
0
\end{bmatrix},
\]

with
\[
\begin{bmatrix}
\frac{2}{3} & \frac{1}{6} \\
\frac{1}{6} & \frac{2}{3}
\end{bmatrix}
\begin{bmatrix}
\alpha_1(0) \\
\alpha_2(0)
\end{bmatrix}
\begin{bmatrix}
\frac{2}{3} \\
\frac{1}{6}
\end{bmatrix},
\]
or
\[
\begin{bmatrix}
\alpha_1' \\
\alpha_2'
\end{bmatrix}
\begin{bmatrix}
-\frac{18}{5} & \frac{12}{5} \\
\frac{12}{5} & -\frac{18}{5}
\end{bmatrix}
\begin{bmatrix}
\alpha_1 \\
\alpha_2
\end{bmatrix},
\]

with
\[
\begin{bmatrix}
\alpha_1(0) \\
\alpha_2(0)
\end{bmatrix}
\begin{bmatrix}
1 \\
0
\end{bmatrix}.
\]

Since the eigenvalues are \(-\frac{6}{5}\) and \(-6\) with corresponding eigenvectors
\[
k_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad k_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}.
respectively, the general solution is given by

\[
\begin{bmatrix}
    a_1(t) \\
    a_2(t)
\end{bmatrix} = k_1 e^{-6t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + k_2 e^{-6t} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.
\]

Using the initial conditions, we find \( k_1 = \frac{1}{2} \) and \( k_2 = \frac{1}{2} \). Therefore,

\[
a_1(t) = \frac{1}{2} e^{-6t} + \frac{1}{2} e^{-6t},
\]

\[
a_2(t) = \frac{1}{2} e^{-6t} - \frac{1}{2} e^{-6t}.
\]

See the appendix (p. 86) to compare this approximate solution with the theoretical solution.

The semi-discrete Galerkin method can be applied to hyperbolic problems. Consider the wave equation

\[
u_{tt} - c^2 u_{xx} = f, \quad u = u(x,t), \quad 0 < x < 1, \quad t > t_0,
\]

subject to

\[
u(0,t) = u(1,t) = 0, \\
u(x,t_0) = u_0(x), \\
\frac{\partial u}{\partial t}(x,t_0) = v_0(x).
\]

The semi-discrete Galerkin method is based on the weak formulation of (3.11):

Find \( u \in H^1_0(\Omega) \) such that for each fixed but arbitrary \( t \)

\[
(u_{tt},v) + c^2 a(u,v) = (f,v) \quad \text{for all } v \in H^1_0(\Omega),
\]

where \( \Omega = (0,1) \). If \( K_n = \langle \phi_1, \ldots, \phi_n \rangle \subset H^1_0(\Omega) \), the semi-discrete Galerkin approximation

\[
U(x,t) = \sum_{j=1}^{n} a_j(t) \phi_j(x)
\]

is determined by

\[
\left( \frac{\partial^2 U}{\partial t^2}, \phi_i \right) + c^2 a(U, \phi_i) = (f, \phi_i), \quad i = 1, \ldots, n.
\]
or
\[ \sum_{j=1}^{n} \frac{\partial^{2}}{\partial t^{2}}(\phi_{j}, \phi_{i}) + c^{2} \sum_{j=1}^{n} a_{j} a(\phi_{j}, \phi_{i}) = (f, \phi_{i}), \quad i = 1 \cdots n. \]

Therefore, we obtain the first order system
\[ A\dot{u} + B_{u} = b \cdots (3.12) \]
where
\[ A_{ij} = (\phi_{j}, \phi_{i}) \quad \text{and} \quad B_{ij} = c^{2} a(\phi_{j}, \phi_{i}), \]
\[ a^{T} = [a_{1} \cdots a_{n}] \quad \text{and} \quad b^{T} = [(f, \phi_{1}) \cdots (f, \phi_{n})], \]
with the initial conditions
\[ (u, \phi_{i})|_{t=t_{0}} = (u_{0}, \phi_{i}) \]
\[ \left( \frac{\partial u}{\partial t}, \phi_{i} \right)|_{t=t_{0}} = (v_{0}, \phi_{i}) \]
or
\[ \sum_{j=1}^{n} a_{j} (t_{0})(\phi_{j}, \phi_{i}) = (u_{0}, \phi_{i}) \]
\[ \sum_{j=1}^{n} a_{j}'(t_{0})(\phi_{j}, \phi_{i}) = (v_{0}, \phi_{i}), \quad \text{for each } i. \]

Example 3.7
\[ u_{tt} = u_{xx}, \quad u = u(x, t), \quad 0 < x < 3, \quad t > 0, \]
subject to
\[ u(0, t) = u(3, t) = 0, \]
\[ u(x, 0) = u_{0}(x) = \begin{cases} x \text{ for } 0 \leq x \leq 1 \\ 2 - x \text{ for } 1 \leq x \leq 2 \\ 0 \text{ for } 2 \leq x \leq 3 \end{cases}, \]
\[ \frac{\partial u}{\partial t}(x, 0) = v_{0}(x) = x. \]

We choose the approximating subspace \( K_{2} = \langle \phi_{1}, \phi_{2} \rangle \) of \( H_{2}^{1}(\Omega) \) to be hat-functions as defined in the last example. The semi-discrete Galerkin approximation
\[ U(x,t) = \sum_{j=1}^{2} \alpha_j(t) \phi_j(t) \]

is determined by

\[ \frac{\partial^2 U}{\partial t^2} + a(U, \phi_i) = 0 , \quad i = 1,2 , \]

or by (3.12), we obtain

\[
\begin{bmatrix}
(\phi_1, \phi_1) & (\phi_2, \phi_1) \\
(\phi_1, \phi_2) & (\phi_2, \phi_2)
\end{bmatrix}
\begin{bmatrix}
a''_1 \\
a''_2
\end{bmatrix}
+ \begin{bmatrix}
a(\phi_1, \phi_1) & a(\phi_2, \phi_1) \\
a(\phi_1, \phi_2) & a(\phi_2, \phi_2)
\end{bmatrix}
\begin{bmatrix}
a_1 \\
a_2
\end{bmatrix}
= \begin{bmatrix} 0 \\ 0 \end{bmatrix},
\]

with the initial condition

\[
\begin{bmatrix}
(\phi_1, \phi_1) & (\phi_2, \phi_1) \\
(\phi_1, \phi_2) & (\phi_2, \phi_2)
\end{bmatrix}
\begin{bmatrix}
a_1(0) \\
a_2(0)
\end{bmatrix}
= \begin{bmatrix} (u_0, \phi_1) \\ (u_0, \phi_2) \end{bmatrix},
\]

\[
\begin{bmatrix}
(\phi_1, \phi_1) & (\phi_2, \phi_1) \\
(\phi_1, \phi_2) & (\phi_2, \phi_2)
\end{bmatrix}
\begin{bmatrix}
a_1'(0) \\
a_2'(0)
\end{bmatrix}
= \begin{bmatrix} (v_0, \phi_1) \\ (v_0, \phi_2) \end{bmatrix}.
\]

Since

\[
(u_0, \phi_1) = \int_0^1 x^2 dx + \int_1^2 (2 - x)^2 dx = \frac{2}{3},
\]

\[
(u_0, \phi_2) = \int_0^2 (2 - x)(x - 1) dx = \frac{1}{6},
\]

\[
(v_0, \phi_1) = \int_0^1 x^2 dx + \int_0^2 x(2 - x) dx = 1,
\]

\[
(v_0, \phi_2) = \int_1^2 x(x - 1) dx + \int_2^3 x(3 - x) dx = 2,
\]

we have

\[
\begin{bmatrix}
\frac{2}{3} & \frac{1}{6} \\
\frac{1}{6} & \frac{2}{3}
\end{bmatrix}
a'' + \begin{bmatrix}
2 & -1 \\
-1 & 2
\end{bmatrix}
\begin{bmatrix}
a_1 \\
a_2
\end{bmatrix}
= \begin{bmatrix} 0 \\ 0 \end{bmatrix}.
\]
with the initial condition

\[ \begin{bmatrix} \frac{2}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} a_1(0) \\ a_2(0) \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{6} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \frac{2}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} a_1'(0) \\ a_2'(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \]

or

\[ \begin{bmatrix} a_1'' \\ a_2'' \end{bmatrix} = \begin{bmatrix} \frac{-18}{5} & \frac{12}{5} \\ \frac{12}{5} & \frac{-18}{5} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \]

with the initial condition

\[ \begin{bmatrix} a_1(0) \\ a_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a_1'(0) \\ a_2'(0) \end{bmatrix} = \begin{bmatrix} \frac{4}{5} \\ \frac{14}{5} \end{bmatrix}. \]

We can write this as the first order system

\[ \dot{X} = AX \]

where

\[ A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{-18}{5} & \frac{12}{5} & 0 & 0 \\ \frac{12}{5} & \frac{-18}{5} & 0 & 0 \end{bmatrix}, \quad X = \begin{bmatrix} a_1 \\ a_2 \\ a_1' \\ a_2' \end{bmatrix}, \]

with the initial condition

\[ X(0) = \begin{bmatrix} 1 \\ 0 \\ \frac{4}{5} \\ \frac{14}{5} \end{bmatrix}. \]
Since the eigenvalues of $A$ are $\pm \sqrt{30}/5i$ and $\pm \sqrt{6}i$ with the corresponding eigenvectors $[\sqrt{6}/6 \quad -\sqrt{6}/6 \quad 1]^T$ and $[-\sqrt{30}/6 \quad -\sqrt{30}/6 \quad 1]^T$ respectively, we have the complex solutions

$$e^{\sqrt{6}it} = \begin{bmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ -1 \end{bmatrix} \begin{bmatrix} -\sqrt{6}/6 \sin \sqrt{6}t \\ \sqrt{6}/6 \sin \sqrt{6}t \\ \cos \sqrt{6}t \end{bmatrix} + i \begin{bmatrix} \sqrt{6}/6 \cos \sqrt{6}t \\ -\sqrt{6}/6 \cos \sqrt{6}t \\ -\sin \sqrt{6}t \end{bmatrix},$$

$$e^{\sqrt{30}/5it} = \begin{bmatrix} -\sqrt{30}/5 \\ \sqrt{30}/5 \\ 1 \end{bmatrix} \begin{bmatrix} \sqrt{30}/5 \sin \sqrt{30}/5t \\ \sqrt{30}/5 \sin \sqrt{30}/5t \\ \cos \sqrt{30}/5t \end{bmatrix} + i \begin{bmatrix} -\sqrt{30}/5 \cos \sqrt{30}/5t \\ \sqrt{30}/5 \cos \sqrt{30}/5t \\ \sin \sqrt{30}/5t \end{bmatrix}.$$

Since the real parts and the imaginary parts of the complex solution

$$X_1(t) = \begin{bmatrix} -\sqrt{6}/6 \sin \sqrt{6}t \\ \sqrt{6}/6 \sin \sqrt{6}t \\ -\cos \sqrt{6}t \end{bmatrix}, \quad X_2(t) = \begin{bmatrix} \sqrt{6}/6 \cos \sqrt{6}t \\ -\sqrt{6}/6 \cos \sqrt{6}t \\ -\sin \sqrt{6}t \end{bmatrix}, \quad X_3(t) = \begin{bmatrix} \sqrt{30}/5 \sin \sqrt{30}/5t \\ \sqrt{30}/5 \sin \sqrt{30}/5t \\ \cos \sqrt{30}/5t \end{bmatrix}, \quad X_4(t) = \begin{bmatrix} -\sqrt{30}/5 \cos \sqrt{30}/5t \\ \sqrt{30}/5 \cos \sqrt{30}/5t \\ \sin \sqrt{30}/5t \end{bmatrix},$$

are also real solutions, and since $X_1(0), X_2(0), X_3(0),$ and $X_4(0)$ are linearly independent, $X_1(t), X_2(t), X_3(t),$ and $X_4(t)$ are linearly independent, and we have the general solution

$$X(t) = k_1X_1(t) + k_2X_2(t) + k_3X_3(t) + k_4X_4(t).$$
Using the initial condition \( X(0) = [1 \ 0 \ \frac{4}{5} \ \frac{14}{5}]^T \), we get \( k_1 = 1 \)
\( k_2 = \frac{\sqrt{6}}{2} \), \( k_3 = \frac{9}{5} \), and \( k_4 = \frac{\sqrt{30}}{10} \). Therefore

\[
\alpha_1(t) = \frac{\sqrt{6}}{6} \sin \sqrt{6}t + \frac{1}{2} \cos \sqrt{6}t + \frac{3\sqrt{30}}{10} \sin \sqrt{\frac{30}{5}}t + \frac{1}{2} \cos \sqrt{\frac{30}{5}}t,
\]

\[
\alpha_2(t) = \frac{\sqrt{6}}{6} \sin \sqrt{6}t - \frac{1}{2} \cos \sqrt{6}t + \frac{3\sqrt{30}}{10} \sin \sqrt{\frac{30}{5}}t + \frac{1}{2} \cos \sqrt{\frac{30}{5}}t.
\]

See the appendix (p.87) for the theoretical solution.

Error Analysis of Semi-Discrete Galerkin Method

In this section, we will derive an error bound of semi-discrete
Galerkin approximation to the parabolic initial boundary value problem

\[
u_t = u_{xx}, \quad u = u(x,t), \quad x \in (0,1), \quad t > 0,
\]

subject to

\[
u(x,0) = u_0(x),
\]

\[
u(0,t) = u(t,0) = 0,
\]

in terms of Sobolev norms. For detailed discussions, see Mitchell and Wait [10], and Thomee and Wahlbin [16].

In general, given a region \( \Omega \subset \mathbb{R}^n \) on which we want to define a
finite element approximation, we partition \( \Omega \) into elements \( T_1 \ldots T_n \)
and define on \( \Omega \) basis functions (piecewise polynomials) which are non-
zero on some \( T_1 \)'s, but vanish on most of them. In order to find an er-
ror bound of the approximation in terms of Sobolev norms, we define the
standard element \( T_0 \) and the transformation \( F \) of \( T_0 \) onto an arbitrary
element \( T_j \). For example, if \( \Omega = [0,1] \) and the partition \( T_1 \ldots T_n \) is
defined by \( T_j = [x_{j-1}, x_j], j = 1 \ldots n \), where \( x_j = jh \) and \( h = \frac{1}{n} \) as in
examples in the last section, we define \( T_0 = [0,1] \) and \( F : T_0 \rightarrow T_j \) by

\[
F(p) = hp + x_{j-1}, \quad p \in [0,1].
\]
We let $K[T_j]$ be the space spanned by basis functions that are non-zero on $T_j$, and let $K[T_0]$ be the space \{ $\phi \circ F$ $|$ $\phi \in K[T_j]$ \}. For our example, $\Omega = [0,1]$ and $T_j = [x_{j-1}, x_j]$, if basis functions are Hat-functions as defined in the examples in the last section, $K[T_j]$ is the set of all linear functions on $T_j$; the space spanned by $\phi_{j-1}$ and $\phi_j$ where
\[
\phi_{j-1}(x) = \frac{1}{h}(x_j - x) \quad \text{and} \quad \phi_j(x) = \frac{1}{h}(x - x_{j-1}),
\]
and $K[T_0]$ is the space of functions on $[0,1]$ spanned by
\[
\phi_{j-1} \circ F(p) = \phi_{j-1}(hp + x_{j-1}) = 1 - p, \quad \text{and} \quad \phi_j \circ F(p) = \phi_j(hp + x_{j-1}) = p.
\]

Assume that $F : T_0 \rightarrow T_j$ is such that if $u \in H^{k+1}_2(T_j)$, then $u \circ F \in H^{k+1}_2(T_0)$. Define the projection $\Pi_{T_j} : H^{k+1}_2(T_j) \rightarrow K[T_j]$ such that $(\Pi_{T_j} u)(x) \in K[T_j]$ interpolates $u$ at each nodal point on $T_j$, and define the corresponding projection $\Pi : H^{k+1}_2(T_0) \rightarrow K[T_0]$ by
\[
\Pi(u \circ F)(x) = (\Pi_{T_j} u) \circ F(x) = u(x_j)(1 - p) + u(x_{j+1})p.
\]

Let $P_k$ be the space of polynomials of degree less than or equal to $k$. In the following, Q.E.D. will indicate the end of the proof.

Lemma 3.1

Let $\Pi$ be a linear transformation of $H^{k+1}_2(T_0)$ into $H^{k}_2(T_0)$ and be a projection onto $K[T_0]$ for some $r \leq k$ such that $P_k \subset K[T_0] \subset H^{k}_2(T_0)$. Then there is a constant $c > 0$ such that
\[
\| v - \Pi v \|_{r,T_0} < c \| I - \Pi \| \| v \|_{k+1,T_0} \quad \text{for all} \quad v \in H^{k+1}_2(T_0).
\]
Proof

Choose $G \in [H^2_0(T)]^*$, and define $F : H^{k+1}_2(T_0) \rightarrow R$ by $F(v) = G([I - \Pi]v)$ for all $v \in H^{k+1}_2(T_0)$. It is easy to see that $F$ is linear and

$$|F(v)| = |G([I - \Pi]v)| \leq \|G\| \|I - \Pi\| \|v\|_{k+1,T_0}.$$ 

Therefore, $F \in [H^2(T_0)]^*$.

Let $p \in p^k$. Then $F(p) = G([I - \Pi]p) = G(p - \Pi p) = G(0) = 0$ since $p \in K(T_0)$. By lemma 1.8 (Bramble-Hilbert), there is a constant $c > 0$ such that

$$|F(v)| \leq C \|F\|_v \|v\|_{k+1,T_0},$$

where

$$\|F\| = \sup_{w \in H^{k+1}_2(T_0)} \frac{|F(w)|}{\|w\|_{k+1,T_0}} = \sup_{w \in H^{k+1}_2(T_0)} \frac{|G([I - \Pi]w)|}{\|w\|_{k+1,T_0}}.$$ 

Define $T_u : [H^2_0(T_0)]^* \rightarrow R$ by $T_u(G) = G(u)$. Then

$$\|T_u\| = \sup\{ |T_u(G)| : \|G\| \leq 1 \} = \sup\{ |G(u)| : \|G\| \leq 1 \} \leq \sup\{ \|G\| \|u\|_{r,T_0} : \|G\| \leq 1 \} \leq \|u\|_{r,T_0}.$$ 

If $u = 0$, then $\|T_u\| = \|u\|_{r,T_0} = 0$. Suppose $u \neq 0$. Let $M = \langle u \rangle$ so that $M$ is a linear subspace of $H^2_0(T_0)$. Define $Q : M \rightarrow R$ by $Q(au) = a \|u\|_{r,T_0}$. It is clear that $Q \in M^*$, $Q(u) = \|u\|_{r,T_0}$, and

$$\|Q\| = \sup\{ |Q(w)| : w \in \langle u \rangle \text{ and } \|w\|_{r,T_0} \leq 1 \} = \sup\{ \|w\|_{r,T_0} : w \in \langle u \rangle \text{ and } \|w\|_{r,T_0} \leq 1 \} = 1.$$
Since \( M \) is a subspace of \( H^r_2(T_0) \), by the Hahn-Banach Theorem, \( Q \) can be extended to \( \hat{Q} \in [H^r_2(T_0)]^* \) such that \( \|Q\| = \|\hat{Q}\| = 1 \) and \( \hat{Q}(u) = \|u\|_{r,T_0} \).

Since there is at least one such functional \( \hat{Q} \in [H^r_2(T_0)]^* \) that satisfies \( \hat{Q}(u) = \|u\|_{r,T_0} \),

\[
\|T_u\| = \|u\|_{r,T_0},
\]

or

\[
\|T_u\| = \sup_{G \in [H^r_2(T_0)]^*} \frac{|G(u)|}{\|G\|} = \|u\|_{r,T_0}.
\]

In particular,

\[
\| (I - \Pi)v \|_{r,T_0} = \sup_{G \in [H^r_2(T_0)]^*} \frac{|G((I - \Pi)v)|}{\|G\|} \leq \sup_{G \in [H^r_2(T_0)]^*} \frac{c \|F\| \|v\|_{k+1,T_0}}{\|G\|}
\]

\[
= c|v|_{k+1,T_0} \sup_{G \in [H^r_2(T_0)]^*} \frac{\|F\|}{\|G\|} \leq c|v|_{k+1,T_0} \sup_{G \in [H^r_2(T_0)]^*} \sup_{w \in H^{k+1}_2(T_0)} \frac{|G((I - \Pi)w)|}{\|w\|_{k+1,T_0} \|G\|} \leq c|v|_{k+1,T_0} \sup_{G \in [H^r_2(T_0)]^*} \sup_{w \in H^{k+1}_2(T_0)} \left( \frac{1}{\|w\|_{k+1,T_0} \|G\|} \right).
\]

Therefore \( \|v - \Pi v\|_{r,T_0} \leq c\|I - \Pi\| \|v\|_{k+1,T_0}. \)

Q.E.D.

For the next theorem, we assume that the transformation \( F \) of the standard element \( T_0 \) onto an arbitrary element \( T_j \) and the projection \( \Pi \) onto \( K[T_0] \) satisfy the following two hypotheses.

Hypothesis 1 (The regularity hypothesis)

If \( v \in H^r_2(T_j) \) and the diameter of the element \( T_j \) is \( h \), then there are constants \( c_0, c_1, \) and \( c_2 \) such that
where \( J_F \) is the Jacobian of the transformation \( x = F(p) \) and \( |J_F| \) is the determinant of the Jacobian. This means that the semi-norm of the pre-image of \( v \) under the transformation \( x = F(p) \) is equivalent to the product of \( h^r \) and the norm of \( v \).

Hypothesis 2 (The completeness hypothesis)

If \( T_j \) is any element of diameter \( h \), then there is \( k > 0 \) such that \( P_k \subset K[T_0] \) and \( \Pi \) is a projection for which \( \| I - \Pi \| \) is uniformly bounded for all \( h \), that is, there is a constant \( c \) such that

\[
\| (I - \Pi)(v \circ F) \|_{r,T_0} \leq c \| v \circ F \|_{k+1,T_0}
\]

for all \( r \leq k \) and \( v \in H^{k+1}_r(T_j) \) for each \( j \).

Theorem 3.1

Let \( u \in H^{k+1}_r(\Omega) \) and let \( K_\Omega \) be the approximating space. Let \( \tilde{u} \in K_\Omega \) interpolate \( u \) at each nonal point. Assume that the completeness hypothesis is valid for some \( k > 0 \) and that the regularity hypothesis is valid for all \( r \leq k \). Then there is a constant \( c > 0 \) such that

\[
\| u - \tilde{u} \|_{r,\Omega} \leq c h^{k+1-r} \| u \|_{k+1,\Omega},
\]

where \( \Omega \) is a region in \( \mathbb{R}^n \) and \( h \) is a bound on the diameters of the elements in the partition of \( \Omega \).

Proof

Suppose \( \Omega \) is partitioned into elements \( T_1 \cdots T_n \). Then
\| u - \tilde{u} \|_{r,R}^2 = \sum_{j=1}^{n} \| u - \Pi_{T_j} u \|_{r,T_j}^2.

By the regularity hypothesis

\| u - \Pi_{T_j} u \|_{r,T_j} \leq c_2^{-1} (\sup |J_F|) \frac{1}{h^{1-r}} \| u \circ F - (\Pi_{T_j} u) \circ F \|_{r,T_0}

\leq c_2^{-1} (\sup |J_F|) \frac{1}{h^{1-r}} \| u \circ F - (\Pi_{T_j} u) \circ F \|_{r,T_0}.

By lemma 3.1

\| u \circ F - (\Pi_{T_j} u) \circ F \|_{r,T_0} \leq c \| I - \Pi \| \| u \circ F \|_{k+1,T_0}.

Therefore, we have

\| u - \Pi_{T_j} u \|_{r,T_j} \leq c_3 (\sup |J_F|) \frac{1}{h^{1-r}} \| I - \Pi \| \| u \|_{k+1,T_j}.

Again by the regularity hypothesis

\| u \circ F \|_{k+1,T_0} \leq c_1 \frac{h^{k+1}}{(\inf |J_F|)^{\frac{1}{2}}} \| u \|_{k+1,T_j}.

Consequently,

\| u - \Pi_{T_j} u \|_{r,T_j} \leq c_4 (\sup |J_F|) \frac{1}{h^{k+1-r}} \| I - \Pi \| \| u \|_{k+1,T_j},

and

\| u - \Pi_{T_j} u \|_{r,T_j}^2 \leq c_5 \sup |J_F| \| u \|_{h^{2(k+1-r)}} \| I - \Pi \| \| u \|_{k+1,T_j}^2

\leq c_6 h^{2(k+1-r)} \| I - \Pi \| \| u \|_{k+1,T_j}^2.

Since \| I - \Pi \| is uniformly bounded, there is a constant \( c \) such that

\| I - \Pi \| < c

and

\sum_{j=1}^{n} \| u - \Pi_{T_j} u \|_{r,T_j}^2 \leq c_6 h^{2(k+1-r)} \| u \|_{k+1,T_j}^2 = c_6 h^{2(k+1-r)} \| u \|_{k+1,\Omega}^2.

Therefore,
Now consider the parabolic initial boundary value problem

\[
\dot{u} = u_{xx}, \quad u = u(x,t), \quad (x,t) \in \Omega = (0,1) \times (0,t_1),
\]

\[
u(0,t) = u(1,t) = 0,
\]

\[u(x,0) = u_0(x).\]

Let \(K_0 = \langle \phi_1(x) \cdots \phi_n(x) \rangle\), where \(\phi_i(0) = \phi_i(1) = 0\) for each \(i\), be our approximating subspace of \(H^{k+1}_2(\Omega)\) on \(\Omega\) for the semi-discrete Galerkin method. We assume that the regularity hypothesis and the completeness hypothesis for \(K_0\) so that theorem 3.1 is valid and there are \(k \geq 1\) and \(c > 0\) such that for all \(v \in H^{k+1}_2(\Omega)\)

\[
\| v - \tilde{v} \|_{r,\Omega} \leq c h^{k+1-r} \| v \|_{k+1,\Omega}
\]

for each \(r \leq k\) where \(\tilde{v} \in K_0\) interpolates \(v\).

Suppose that the solution \(u = u(x,t)\) is such that \(u \in H^{k+1}_2(\Omega)\) for each fixed but arbitrary \(t \in [0,t_1]\) and \(u \in C^1[0,t_1]\) for each fixed but arbitrary \(x \in [0,1]\). Assume that \(\frac{\partial u}{\partial t} \in H^{k+1}_2(\Omega)\) for each fixed but arbitrary \(t \in [0,t_1]\). Suppose that \(W = W(x,t) \in K_0\) defined by

\[a(W,V) = A(u,V) \quad \text{for all } V \in K_0\]

where \(a(f,g) = \int_0^1 f'(x)g'(x)dx\), belongs to \(C^1[0,t_1]\) for each fixed but arbitrary \(x \in [0,1]\).

Lemma 3.2

The bilinear form \(a(f,g) = \int_0^1 f'(x)g'(x)dx\) is \(H^1_2(\Omega)-\text{elliptic}\).

Proof

\[a(f,f) = \int_0^1 |f'(x)|^2dx = |f'|^2_{1,\Omega}.\]

By lemma 1.4, there is a constant
c_1 such that \( \| f \|_{1, \Omega} \leq c_1 \| f \|_{1, \Omega} \), for all \( f \in H^1_2(\Omega) \), and we have
\[
a(f, f) \geq c \| f \|_{1, \Omega}^2.
\]

Q.E.D.

Since the differential operator \( A = \frac{-d^2}{dx^2} \) is positive definite and self-adjoint on \( H^{k+1}_2(\Omega) \) and \( W \) satisfies \( a(W, V) = a(u, V) \) for all \( V \in K_n \), \( W \) is the orthogonal projection of \( u \) onto \( K_n \) with respect to the norm \( \| v \|_A^2 = a(v, v) \) and by lemma 2.2, \( \| u - W \|_A \leq \| u - V \|_A \) for all \( V \) in \( H^{k+1}_2(\Omega) \).

Suppose that \( Y \in K_n \) interpolates \( u \) for each fixed but arbitrary \( t \in [0, t_1] \) so that
\[
\| u - Y \|_{1, \Omega} \leq ch^k \| u \|_{k+1, \Omega}.
\]

Then by lemma 3.2, there are constants \( c_0 \) and \( c_1 \) such that
\[
\| u - W \|_{1, \Omega} \leq c_0 \| u - W \|_A \leq c_0 \| u - Y \|_A \leq c_0 \| u - Y \|_{1, \Omega}
\]
\[
\leq c_1 h^k \| u \|_{k+1, \Omega}.
\]

The semi-discrete Galerkin approximation \( U(x, t) = \sum_{i=1}^n a_i(t) \phi_i(x) \) satisfies
\[
(U_t, V) + a(U, V) = 0 \quad \text{for all } V \in \hat{K}_n.
\]

Assume that \( U(x, 0) = W(x, 0) \). Then
\[
(U_t - W_t, V) + a(U - W, V) = (U_t, V) - (W_t, V) + a(U, V) - a(W, V)
\]
\[
= -(W_t, V) - a(W, V) = -(U_t, V) - a(u, V)
\]
\[
= (u_t, V) - (W_t, V) = (u_t - W_t, V) \quad \text{for all } V \in K_n.
\]

Let \( V = U_t - W_t \). Then we have
\[ ||u_t - w_t||^2 + a(u - w, u_t - w_t) = (u_t - w_t, u_t - w_t). \]

Since
\[
a(u - w, u_t - w_t) = \int_0^1 \frac{\partial}{\partial x} (u - w) \cdot \frac{\partial}{\partial x} (u_t - w_t) \, dx
\]
\[
= \frac{1}{2} \int_0^1 (\frac{\partial}{\partial x} (u - w) \cdot \partial (u - w) \, dx
\]
\[
= \frac{1}{2} \frac{d}{dt} a(u - w, u - w),
\]
\[
||u_t - w_t||^2 + \frac{1}{2} \frac{d}{dt} a(u - w, u - w) = (u_t - w_t, u_t - w_t).
\]

By the Schwartz inequality,
\[
(u_t - w_t, u_t - w_t) \leq ||u_t - w_t|| \cdot ||u_t - w_t|| \leq 2 ||u_t - w_t|| \cdot ||u_t - w_t||
\]
\[
\leq ||u_t - w_t||^2 + ||u_t - w_t||^2,
\]
and we have
\[
||u_t - w_t||^2 + \frac{1}{2} \frac{d}{dt} a(u - w, u - w) \leq ||u_t - w_t||^2 + ||u_t - w_t||^2,
\]
or
\[
\frac{d}{dt} a(u - w, u - w) \leq 2 ||u_t - w_t||^2.
\]

Since \( a(u, V) = A(W, V) \) for all \( V \in K \),
\[
\int_0^1 \frac{\partial a}{\partial u} \, dV = \int_0^1 \frac{\partial a}{\partial w} \, dV.
\]

Differentiating both sides with respect to \( t \),
\[
\int_0^1 \frac{\partial}{\partial x \partial t} (2u) \, dV = \int_0^1 \frac{\partial}{\partial x \partial t} (2w) \, dV,
\]
since \( V \) is a function of \( x \) only. Therefore, \( a(u_t, V) = a(W_t, V) \) for all
\( v \in K \) and there is a constant \( c \) such that

\[
\| u_t - w_t \| \leq \| u_t - w_t \|_{1, \Omega} \leq c h^k \| u_t \|_{k+1, \Omega}.
\]

Therefore

\[
\frac{d}{dt} a(U - W, U - W) \leq c^2 h^{2k} \| u_t \|_{k+1, \Omega}^2.
\]

Let \( c_3 = \max_{t \in [0, t_1]} c_2 \| u_t \|_{k+1, \Omega}^2. \) Then

\[
\frac{d}{dt} a(U - W, U - W) \leq c_3 h^{2k}.
\]

Since \( U(x, 0) = W(x, 0) \), \( a(U - W, U - W) \big|_{t=0} = 0 \) and

\[
a(U - W, U - W) \leq c_3 h^{2k} \leq c_3 t_1 h^{2k}
\]

for all \( t \in [0, t_1] \). Since the bilinear form \( a = a(f, g) \) is \( H^1_2(\Omega) \) -elliptic, there is a constant \( c_4 \) such that

\[
c_4 \| U - W \|_{1, \Omega}^2 \leq a(U - W, U - W).
\]

Therefore

\[
\| U - W \|_{1, \Omega} \leq c_5 h^k,
\]

where \( c_5 = \sqrt{\frac{c_2}{c_4}}. \) Since

\[
\| u - W \|_{1, \Omega} \leq c_1 h^k \| u \|_{k+1, \Omega},
\]

\[
\| u - u \|_{1, \Omega} \leq \| u - W \|_{1, \Omega} + \| u - W \|_{1, \Omega} \leq (c_1 \| u \|_{k+1, \Omega} + c_5) h^k.
\]

The reason for the introduction of the convergence proof in the first part of this section is to show the results obtainable by the assumption made on the approximating subspace via Bramble-Hilbert lemma.
REFERENCES


True Solutions and Approximate Solutions

In example 3.6, we investigated the heat equation

\[ u_t = u_{xx}, \quad u = u(x,t), \quad 0 < x < 3, \quad t > 0, \]

subject to

\[ u(0,t) = u(3,t) = 0, \]

\[ u(x,0) = \begin{cases} x & \text{for } 0 \leq x \leq 1 \\ 2 - x & \text{for } 1 \leq x \leq 2 \\ 0 & \text{for } 2 \leq x \leq 3 \end{cases}, \]

and we obtained the semi-discrete Galerkin approximation

\[ a_1(t) = \frac{1}{2} e^{-5t} + \frac{1}{2} e^{-6t}, \]

\[ a_2(t) = \frac{1}{2} e^{-5t} - \frac{1}{2} e^{-6t}, \]

to \( u(x,t) \) at \( x = 1 \) and \( 2 \) respectively. The theoretical solution obtained by the separation of variables is

\[ u(x,t) = \sum_{n=1}^{\infty} u_n(x,t), \]

where

\[ u_n(x,t) = \frac{6}{(n\pi)^2} \left(2\sin\frac{n\pi}{3} - \sin \frac{2n\pi}{3}\right) e^{-\left(\frac{n\pi}{3}\right)^2 t} \sin \frac{n\pi x}{3}. \]

Let \( \hat{u}(x,t) = \sum_{n=1}^{5} u_n(x,t) \). Figure 1 shows the comparison between \( a_1(t) \) and \( \hat{u}(1,t) \) and illustrate the accuracy of the semi-discrete Galerkin method.

In example 3.7, we considered the wave equation

\[ u_{tt} = u_{xx}, \quad u = u(x,t), \quad 0 < x < 3, \quad t > 0, \]

subject to

\[ u(0,t) = u(3,t) = 0, \]
\[
\begin{align*}
\begin{cases}
x & \text{for } 0 \leq x \leq 1 \\
2 - x & \text{for } 1 \leq x \leq 2 \\
0 & \text{for } 2 \leq x \leq 3
\end{cases},
\end{align*}
\]
\[\frac{\partial u}{\partial t}(x,0) = x.\]

The semi-discrete Galerkin approximation for this problem is

\[
\begin{align*}
\alpha_1(t) &= \frac{-\sqrt{6}}{6}\sin \sqrt{6}t + \frac{1}{2}\cos \sqrt{6}t + \frac{3\sqrt{30}}{10}\sin \frac{\sqrt{30}}{5}t + \frac{1}{2}\cos \frac{\sqrt{30}}{5}t, \\
\alpha_2(t) &= \frac{\sqrt{6}}{6}\sin \sqrt{6}t - \frac{1}{2}\cos \sqrt{6}t + \frac{3\sqrt{30}}{10}\sin \frac{\sqrt{30}}{5}t + \frac{1}{2}\cos \frac{\sqrt{30}}{5}t.
\end{align*}
\]

The theoretical solution is

\[u(x,t) = \sum_{n=1}^{\infty} u_n(x,t),\]

where

\[
\begin{align*}
u_n(x,t) &= \frac{6}{(n\pi)^2} \left( 2\sin \frac{n\pi}{3} - \sin \frac{2n\pi}{3} \right) \cos \frac{n\pi t}{3} \sin \frac{n\pi x}{3} \\
&\quad + \frac{18}{(n\pi)^2} (-1)^{n+1} \sin \frac{n\pi t}{3} \sin \frac{n\pi x}{3}.
\end{align*}
\]

Figure 2 shows the difference between \(u(x,t) = \sum_{n=1}^{5} u_n(x,t)\) and \(\alpha_1(t)\), and establishes the usefulness of the semi-discrete Galerkin method.
Figure 1. Approximate solutions to the heat equation.

Figure 2. Approximate solutions to the wave equation.