

NUMERICAL METHODS FOR SOLVING PARTIAL DIFFERENTIAL EQUATIONS

A COMPREHENSIVE INTRODUCTION FOR
SCIENTISTS AND ENGINEERS



GEORGE F. PINDER

WILEY

Numerical Methods for Solving Partial Differential Equations

**A Comprehensive Introduction
for Scientists and Engineers**

George F. Pinder

WILEY

This edition first published 2018 by John Wiley and Sons, Inc.

© 2018 by John Wiley & Sons, Inc.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The right of George F Pinder to be identified as the author of this work has been asserted in accordance with law.

Registered Office

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

Editorial Office

111 River Street, Hoboken, NJ 07030, USA

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Limit of Liability/Disclaimer of Warranty In view of ongoing research, equipment modifications, changes in governmental regulations, and the constant flow of information relating to the use of experimental reagents, equipment, and devices, the reader is urged to review and evaluate the information provided in the package insert or instructions for each chemical, piece of equipment, reagent, or device for, among other things, any changes in the instructions or indication of usage and for added warnings and precautions. While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data:

Hardback ISBN: 978-1-119-31611-4

Cover image by Charles Bombard

Cover design by Wiley

Robert N. Farvolden and John D. Bredehoeft my mentors

Contents

[Preface](#)

[CHAPTER 1: Interpolation](#)

[1.1 Purpose](#)

[1.2 Definitions](#)

[1.3 Example](#)

[1.4 Weirstraus Approximation Theorem](#)

[1.5 Lagrange Interpolation](#)

[1.6 Compare \$P_2\(\theta\)\$ and \$\hat{f}\(\theta\)\$](#)

[1.7 Error of Approximation](#)

[1.8 Multiple Elements](#)

[1.9 Hermite Polynomials](#)

[1.10 Error in Approximation by Hermites](#)

[1.11 Chapter Summary](#)

[1.12 Problems](#)

[Bibliography](#)

[CHAPTER 2: Numerical Differentiation](#)

[2.1 General Theory](#)

[2.2 Two-point Difference Formulae](#)

[2.3 Two-point Formulae from Taylor Series](#)

[2.4 Three-point Difference Formulae](#)

[2.5 Chapter Summary](#)

[2.6 Problems](#)

[Notes](#)

[Bibliography](#)

[CHAPTER 3: Numerical Integration](#)

[3.1 Newton-Cotes Quadrature Formulas](#)

[3.2 Chapter Summary](#)

[3.3 Problems](#)

[Bibliography](#)

[CHAPTER 4: Initial Value Problems](#)

[4.1 Euler Forward Int egrat ion Method Example](#)

[4.2 Convergence](#)

[4.3 Consistency](#)

[4.4 Stability](#)

[4.5 Lax Equivalence Theorem](#)

[4.6 Runge-Kutta Type Formulas](#)

[4.7 Chapter Summary](#)

[4.8 Problems](#)

[Notes](#)

[Bibliography](#)

[CHAPTER 5: Weighted Residuals Methods](#)

[5.1 Finite Volume or Subdomain Method](#)

[5.2 Galerkin Method for First Order Equations](#)

[5.3 Galerkin Method for Second Order Equations](#)

[5.4 Finite Volume Method for Second-Order Equations](#)

[5.5 Collocation Method](#)

[5.6 Chapter Summary](#)

[5.7 Problems](#)

[Bibliography](#)

[CHAPTER 6: Initial Boundary-Value Problems](#)

[6.1 Introduction](#)

[6.2 Two Dimensional Polynomial Approximations](#)

[6.3 Finite Difference Approximation](#)

[6.4 Stability of Finite Difference Approximations](#)

[6.5 Galerkin Finite Element Approximations in Time](#)

[6.6 Chapter Summary](#)

[6.7 Problems](#)

[Bibliography](#)

[CHAPTER 7: Finite Difference Methods in Two Space](#)

[7.1 Example Problem](#)

[7.2 Chapter Summary](#)

[7.3 Problems](#)

[Bibliography](#)

[CHAPTER 8: Finite Element Methods in Two Space](#)

[8.1 Finite Element Approximations over Rectangles](#)

[8.2 Finite Element Approximations over Triangles](#)

[8.3 Isoparametric Finite Element Approximation](#)

[8.4 Chapter Summary](#)

[8.5 Problems](#)

[Bibliography](#)

[CHAPTER 9: Finite Volume Approximation in Two Space](#)

[9.1 Finite Volume Formulation](#)

[9.2 Finite Volume Example Problem 1](#)

[9.3 Finite Volume Example Problem Two](#)

[9.4 Chapter Summary](#)

[9.5 Problems](#)

[Bibliography](#)

[CHAPTER 10: Initial Boundary-value Problems](#)

[10.1 Mass Lumping](#)

[10.2 Chapter Summary](#)

[10.3 Problems](#)

[Bibliography](#)

[CHAPTER 11: Boundary-Value Problems in Three Space](#)

[11.1 Finite Difference Approximations](#)

[11.2 Finite Element Approximations](#)

[11.3 Chapter Summary](#)

[Bibliography](#)

[CHAPTER 12: Nomenclature](#)

[Index](#)

[EULA](#)

List of Tables

[Table 1.1](#)

[Table 1.2](#)

[Table 1.3](#)

[Table 1.4](#)

[Table 1.5](#)

[Table 1.6](#)

[Table 9.1](#)

List of Illustrations

[Figure 1.1](#)

[Figure 1.2](#)

[Figure 1.3](#)

[Figure 1.4](#)

[Figure 1.5](#)

[Figure 1.6](#)

[Figure 1.7](#)

[Figure 1.8](#)

[Figure 1.9](#)

[Figure 1.10](#)

[Figure 1.11](#)
[Figure 1.12](#)
[Figure 1.13](#)
[Figure 1.14](#)
[Figure 1.15](#)
[Figure 1.16](#)
[Figure 1.17](#)
[Figure 1.18](#)
[Figure 1.19](#)
[Figure 1.20](#)
[Figure 2.1](#)
[Figure 2.2](#)
[Figure 2.3](#)
[Figure 2.4](#)
[Figure 2.5](#)
[Figure 2.6](#)
[Figure 2.7](#)
[Figure 3.1](#)
[Figure 4.1](#)
[Figure 4.2](#)
[Figure 5.1](#)
[Figure 5.2](#)
[Figure 5.3](#)
[Figure 5.4](#)
[Figure 5.5](#)
[Figure 5.6](#)
[Figure 5.7](#)
[Figure 5.8](#)
[Figure 5.9](#)
[Figure 5.10](#)
[Figure 5.11](#)
[Figure 5.12](#)
[Figure 5.13](#)
[Figure 5.14](#)
[Figure 5.15](#)

[Figure 5.16](#)

[Figure 5.17](#)

[Figure 6.1](#)

[Figure 6.2](#)

[Figure 6.3](#)

[Figure 6.4](#)

[Figure 6.5](#)

[Figure 6.6](#)

[Figure 6.7](#)

[Figure 6.8](#)

[Figure 6.9](#)

[Figure 6.10](#)

[Figure 6.11](#)

[Figure 6.12](#)

[Figure 6.13](#)

[Figure 6.14](#)

[Figure 7.1](#)

[Figure 7.2](#)

[Figure 7.3](#)

[Figure 7.4](#)

[Figure 7.5](#)

[Figure 7.6](#)

[Figure 7.7](#)

[Figure 7.7](#)

[Figure 8.1](#)

[Figure 8.2](#)

[Figure 8.3](#)

[Figure 8.4](#)

[Figure 8.5](#)

[Figure 8.6](#)

[Figure 8.7](#)

[Figure 8.8](#)

[Figure 8.9](#)

[Figure 8.10](#)

[Figure 8.11](#)

[Figure 8.12](#)

[Figure 8.13](#)

[Figure 8.14](#)

[Figure 8.15](#)

[Figure 8.17](#)

[Figure 8.18](#)

[Figure 8.19](#)

[Figure 8.20](#)

[Figure 8.21](#)

[Figure 8.22](#)

[Figure 9.1](#)

[Figure 9.2](#)

[Figure 9.3](#)

[Figure 9.4](#)

[Figure 9.5](#)

[Figure 9.6](#)

[Figure 9.7](#)

[Figure 9.8](#)

[Figure 9.9](#)

[Figure 9.10](#)

[Figure 11.1](#)

[Figure 11.3](#)

Preface

While there are many good books on numerical methods suitable for students of mathematics and many others that are accessible to scientists and engineers, but dedicated to a specific discipline, there is a need for a book that is accessible to students of science and engineering that is not discipline specific, yet rigorous and comprehensive in scope. This book is an effort to fill this need.

Herein I provide the logical underpinnings of all of the commonly encountered numerical methods, namely finite difference, finite element, collocation, and finite volume methods, at a level of sophistication consistent with the needs and interests of science and engineering students. Two mathematical concepts, namely polynomial approximation theory and the method of weighted residuals, form the intellectual framework for the introduction and explanation of all of these methods.

The approach is to first introduce polynomial approximation theory in one space dimension followed by the concept of the methods of weighted residuals. Employing only polynomial approximation theory the finite difference method is easily developed and presented. With the addition of the method of weighted residuals, finite element, collocation and finite volume methods are readily accessible. These concepts are introduced first in one space dimension, then the time dimension, then two space dimensions, and finally two space dimensions and time.

The equations considered are first order, second order, and second order in space and first order in time. By design, the book does not focus on any specific area of science or engineering. It is designed to teach numerical methods as a concept rather than as applied to a specific discipline. The intent is to provide the student with the ability to understand numerical methods as encountered in technical readings specific to his/her discipline and to be able to apply them in practice.

The book assumes a knowledge of matrix algebra and differential equations. A programming language is also needed if the reader is interested in applying numerical methods to example problems. No prior knowledge of numerical methods is assumed. While a few theorems are used, no proofs are presented.

This book stems from a course I teach in Numerical Methods for Engineers. The course is taught as a precept and typically populated by an approximately equal number of senior undergraduates and graduate students from different engineering disciplines. A project of practical significance is assigned that requires the creation of a computer program capable of solving a second-order two-space dimensional equation using finite elements.

I am indebted to Xin Kou, my doctoral student in mathematics, for carefully reviewing the manuscript for his book, identifying notational inconsistencies and making important suggestions as to how to improve the presentation.

Chapter 1

Interpolation

In this chapter, we will introduce interpolation theory, the first of two key topics that will form the foundation of everything that follows in this book. We will find that this concept leads quite naturally to finite difference methods and, when combined with the second key topic, the method of weighted residuals, provides the necessary mathematical concepts needed for all other numerical methods we present. So, let us get started.

1.1 Purpose

Interpolation is a method of constructing new data points between known values or for creating a function that fits exactly a known set of discrete data points defined within a specific range. Interpolation has many applications in science and engineering. In this book, it will be used to form the basis for numerical differentiation, numerical quadrature, numerical integration, and as a key part of several numerical methods used to solve differential and partial differential equations.

1.2 Definitions

We begin by introducing some interpolation notation. Consider a region

$$(a =) x_0 < x_1 < x_2 \dots < x_n (= b) \tag{1.1}$$

as illustrated in [Fig. 1.1](#)



Figure 1.1 Discretized line spanning a to b .

Next assume there exists a function $f(x)$ that is a known function of x . We will use this function momentarily.

Now also consider a function $P_n(x)$ that has the following properties:

1. $P_n(x)$ is a polynomial of degree n , that is.

$$P_n(x) = \sum_{i=0}^n a_i x^i \tag{1.2}$$

where the coefficients a_i are known constants and x^i indicates the variable x to the i th power.

2. $P_n(x_i) = f(x_i)$ where x_i are particular values of x as seen in [Fig. 1.1](#). In other words at

the locations x_i the values of $f(x)$ and $P_n(x)$ are identical.

According to our definition of interpolation, this $P_n(x)$ is an interpolating polynomial of degree n . Note that by convention, i has a lower bound value of 0 rather than 1. To make the above clear, consider the following example.

1.3 Example

Consider the sine function shown in [Fig. 1.2](#). and the information presented in [Table 1.1](#)

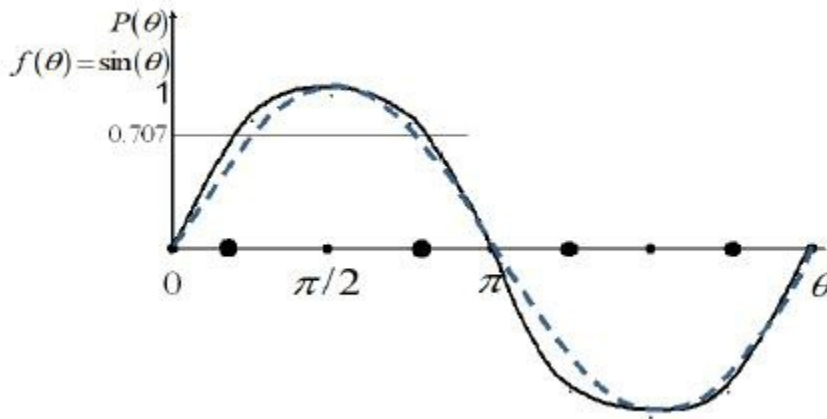


Figure 1.2 Sin (θ) curve with measured points indicated by small black dots (dashed curve) and interpolated values indicated by large black dots (solid curve).

Table 1.1 Comparison of exact and calculated approximations for the sin curve.

θ	$f(\theta)$	$P_2(\theta)$	$E(\theta) \equiv f(\theta) - P_2(\theta)$
0	0.	0.	0.
$\pi/4$	0.707	0.753	- 0.046
$\pi/2$	1.	1.	0.
$3\pi/4$	0.707	0.751	- 0.044
π	0.	0.	0.
$5\pi/4$	- 0.707	- 1.18	0.37
$6\pi/4$	- 1.	- 2.89	1.89

Let the second-degree polynomial $P_2(\theta)$ be given by

$$P_2(\theta) = a_0 + a_1\theta + a_2\theta^2. \quad (1.3)$$

Now create the special version of [Eq. \(1.3\)](#) that satisfies the three known values of $P_2(\theta)$, namely those at $\theta = 0, \pi/2$ and π . Then, substituting values of $P_2(\theta)$ from [Table 1.1](#) we obtain

$$P_2(0) = a_0 + a_1 \cdot 0 + a_2 \cdot 0^2 = 0 \quad (1.4)$$

$$P_2(\pi/2) = a_0 + a_1 (\pi/2) + a_2 (\pi/2)^2 = 1$$

$$P_2(\pi) = a_0 + a_1 (\pi) + a_2 (\pi)^2 = 0$$

from which we can generate the set of equations

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & \pi/2 & (\pi/2)^2 \\ 1 & \pi & (\pi)^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Solving for the coefficients a_0 , a_1 , and a_2 we obtain

$$P_2(\theta) = 0 + 1.27(\theta) - 0.40(\theta^2). \quad (1.5)$$

From this expression, we can obtain values of $P(\theta)$ for any θ . In [Table 1.1](#), we calculate the values of $P(\theta)$ for various values of θ . Notice the difference in the error of the interpolation at $\theta = \pi/4$ and $\theta = (5\pi)/4$. Why did this happen? It is due to the fact that the value $\theta = 5\pi/4$ lies outside of the range of θ used to define $P(\theta)$.

The question now arises as to whether polynomials can be used to represent functions other than the sine. To answer this question we turn to the Weirstraus approximation theorem.

1.4 Weirstraus Approximation Theorem

The Weirstraus approximation theorem basically tells us that it is possible to calculate a polynomial approximation of any desired accuracy, provided you employ a suitably large number of terms in the polynomial. It states:

Theorem 1 *If $f(x)$ is continuous on a finite interval $[a, b]$, then, given any $\epsilon > 0$, there exists an n and a polynomial $P_n(x)$ of degree n such that $|f(x) - P_n(x)| < \epsilon$ for all x in $[a, b]$.*

Although this theorem indicates that a polynomial can be found to represent any function, it does not mean that the coefficients of all polynomials can be calculated. In some cases, especially for large n , the coefficient matrix for the polynomial coefficients can be ill-conditioned (almost singular) and the coefficient values unobtainable. Fortunately, in our work, we will be using polynomials that do not exhibit this pathology.

1.5 Lagrange Interpolation

Let us now focus on one special kind of interpolation that we will use extensively in subsequent material. Consider the approximation of a function $f(x)$ denoted as $\hat{f}(x)$ written as follows:

$$f(x) = \hat{f}(x) + E(x) \quad (1.6)$$

where we will call $E(x)$ the error of the approximation; in other words $E(x)$ is a measure of how well $\hat{f}(x)$ represents $f(x)$.

We now define the form of $\hat{f}(x)$ in a very special way, that is

$$\hat{f}(x) = \sum_{j=0}^n \ell_j^n(x) f(x_j) \quad (1.7)$$

where $\ell_j^n(x)$ is an as yet undefined polynomial of degree n . Next, substitute [Eq. \(1.7\)](#) into [Eq. \(1.6\)](#) to obtain

$$f(x) = \sum_{j=0}^n \ell_j^n(x) f(x_j) + E(x). \quad (1.8)$$

To determine the functional form of the polynomial $\ell_j^n(x)$ we will require that $E(x)$, the error in the approximation, vanishes at selected locations along x , namely at x_i , $i = 0, \dots, n$. We will call these locations nodes and they are indicated in [Fig. 1.1](#) by the black dots. Recall that we required $f(x)$ to equal $f(x_j)$ in our general definition of a polynomial in [Section 1.2](#). We can write this requirement formally, in terms of the errors $E(x_i)$, as

$$E(x_i) = 0 \quad i = 0, \dots, n \quad (1.9)$$

where we have $n + 1$ nodes. Now combine [Eq. \(1.9\)](#) with [Eq. \(1.8\)](#) to give

$$f(x_i) = \sum_{j=0}^n \ell_j^n(x_i) f(x_j) \quad i = 0, \dots, n. \quad (1.10)$$

[Equation \(1.10\)](#) is [Eq. \(1.8\)](#) written for the specific nodal locations x_i where, by definition, the error must vanish. Note that the index i identifies the location, that is $f(x_i)$ and $\ell_j^n(x_i)$ where the polynomial is being evaluated and the index j indicates the term in the polynomial, that is $\ell_j^n(x_i) f(x_j)$. Let us now expand [Eq. \(1.10\)](#) as

$$f(x_1) = \ell_0^n(x_1) f(x_0) + \ell_1^n(x_1) f(x_1) + \dots + \ell_n^n(x_1) f(x_n). \quad (1.11)$$

The form of [Eq. \(1.11\)](#) suggests that the polynomials $\ell_j^n(x_i)$ must have special properties. In order to satisfy the requirement that

$$E(x_i) = 0, \quad i = 0, \dots, n. \quad (1.12)$$

Indeed, at the location x_1 , for example, [Eq. \(1.11\)](#) must yield the following

$$\ell_1^n(x_1) = 1 \quad (1.13)$$

$$\ell_0^n(x_1) = \ell_2^n(x_1) \dots = \ell_n^n(x_1) = 0. \quad (1.14)$$

In fact we can generalize this statement to any nodal location x_i , that is

$$\ell_j(x_i) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad i, j = 0, \dots, n. \quad (1.15)$$

In other words, our polynomial must be unity at the nodal location for which it is defined, that is where the indices i and j are the same, and zero at all other nodes. Writing this in shorthand notation we get that

$$\ell_j^n(x_i) = \delta_{ij} \quad j = 0, \dots, n, \quad i = 0, \dots, n \quad (1.16)$$

where δ_{ij} is the Kronecker delta. The Kronecker delta is defined such that

$$\delta_{i,j} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (1.17)$$

Now let us assume the special case of $n = 1$; that is we are considering a linear polynomial. Let us focus on the j th polynomial, that is.

$$\ell_j^1(x) = a_j + b_j x. \quad (1.18)$$

In light of [Eq. \(1.17\)](#), we can say that the following is true for $\ell_j^1(x)$ evaluated at node x_0

$$\ell_0^1(x_0) = a_0 + b_0 x_0 = 1 \quad (1.19)$$

$$\ell_0^1(x_1) = a_0 + b_0 x_1 = 0. \quad (1.20)$$

Writing this system of equations in matrix form we get

$$\begin{bmatrix} 1 & x_0 \\ 1 & x_1 \end{bmatrix} \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (1.21)$$

which we can solve to obtain

$$\begin{bmatrix} a_0 \\ b_0 \end{bmatrix} = \frac{1}{x_1 - x_0} \begin{bmatrix} x_1 & -x_0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{x_1}{x_1 - x_0} \\ \frac{-1}{x_1 - x_0} \end{bmatrix}. \quad (1.22)$$

Substitution of [Eq. \(1.22\)](#) into [Eq. \(1.18\)](#) yields

$$\ell_0^1(x) = a_0 + b_0 x = \frac{x_1}{x_1 - x_0} - \frac{1}{x_1 - x_0} x = \frac{x_1 - x}{x_1 - x_0} \quad (1.23)$$

One can similarly obtain $\ell_1^1(x) = \frac{x - x_0}{x_1 - x_0}$. The functions $\ell_0^1(x)$ and $\ell_1^1(x)$ are **linear Lagrange Polynomials** and are presented in [Fig. 1.3](#)

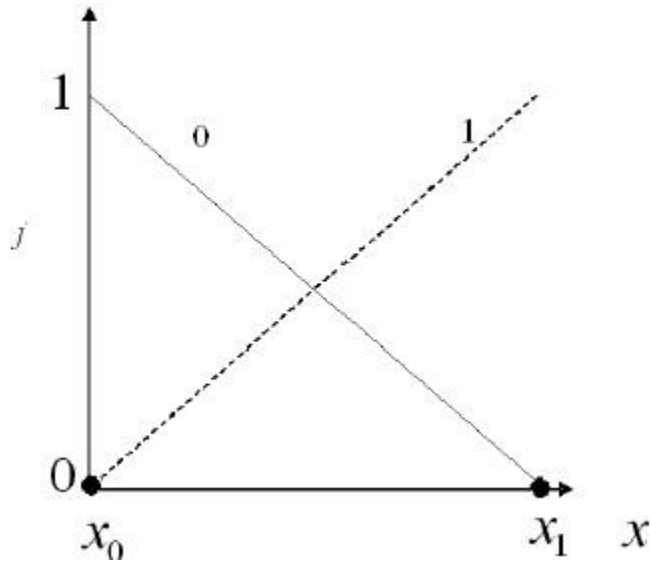


Figure 1.3: Linear Lagrange polynomials, one defined for the node located at x_0 (solid line) and one for the node located at x_1 (dashed line).

Let us check to see if these functions satisfy the requirements stated in [Eqs. \(1.19\)](#) and [\(1.20\)](#):

$$\text{when } x = x_0; \quad \ell_0^1(x_0) = \frac{x_1 - x_0}{x_1 - x_0} = 1 \quad (1.24)$$

$$\text{when } x = x_1; \quad \ell_0^1(x_1) = \frac{x_1 - x_1}{x_1 - x_0} = 0. \quad (1.25)$$

It appears the Lagrange polynomials, as defined above, work for the linear case.

The general form of the n th degree Lagrange polynomial is

$$\ell_j^n(x) = \prod_{\substack{i=0 \\ i \neq j}}^n \frac{x - x_i}{x_j - x_i} \quad (1.26)$$

where the operator $\prod_{\substack{i=0 \\ i \neq j}}^n$ says that for a specific value of j , each term denoted by the subscript i , $i = 0, 1, \dots, n$ will be multiplied together except for the special case of $i = j$, which will be skipped because it would result in a value of zero in the denominator of that term. For example, $\prod_{i=1}^3 x_i = (x_1)(x_2)(x_3)$. Thus, for our linear case we obtain

$$\ell_0^1(x) = \prod_{\substack{i=0 \\ i \neq j}}^1 \frac{x - x_i}{x_0 - x_i} = \frac{x - x_1}{x_0 - x_1} \quad (1.27)$$

which is the same as we obtained in [Eq. \(1.23\)](#) after multiplying the numerator and denominator by (-1) .

The strategy that was used above for the linear polynomial can be extended to define the quadratic. In this case, as we will see below, there are three unknown coefficients and therefore one needs three equations. The equations are obtained by imposing the constraints defined in Eq. (1.16) on the quadratic polynomial. Alternatively, since we have stated the general polynomial form in Eq. (1.16), we can write directly, by selecting $n = 2$ the relationship (where j is now equal to 0):

$$\ell_0^2(x) = \prod_{\substack{i=0 \\ i \neq 0}}^2 \frac{x - x_i}{x_0 - x_i} = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)}. \quad (1.28)$$

The shape of this function is shown in Fig. 1.4. Quadratic Lagrange polynomial identified with node x_0 . By selecting other values of j , that is $j = 1$ or $j = 2$ two additional quadratics will be generated for location x_1 and x_2 .

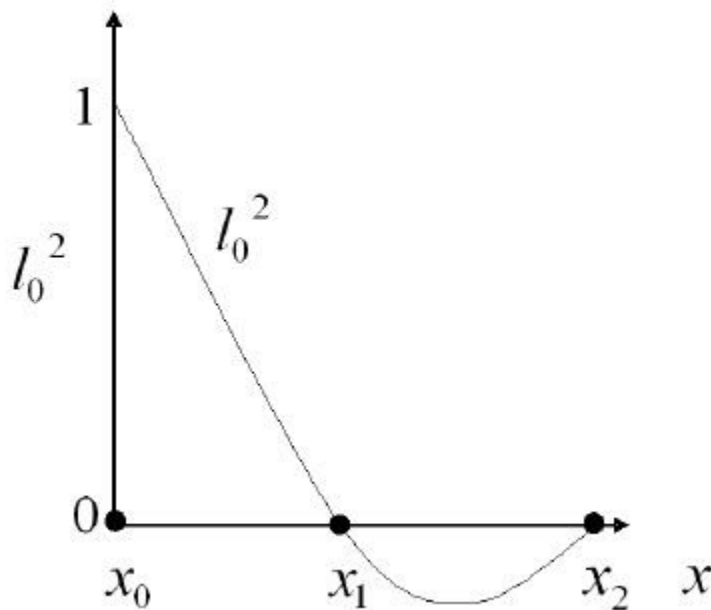


Figure 1.4 Quadratic Lagrange polynomial identified with node x_0 .

1.5.1 Example

Let us use the linear function to represent the function $\ln(x)$. The linear Lagrange polynomials are presented in Fig. 1.3. There will be two terms, one associated with each node, that is $x = x_0$ and $x = x_1$, as seen in Fig. 1.3. Thus we have that $x \in [1, 2]$

$$\hat{f}(x) = \sum_{j=0}^1 \ell_j^1(x) f(x_j) \quad x \in [1, 2] \quad (1.29)$$

From this equation we see that we need $f(x_0)$ and $f(x_1)$. To get this information we need to evaluate $\ln(x)$ at $x = 1.0$ and $x = 2.0$. The following equation shows how this is used:

$$\begin{aligned}
\hat{f}(x) &= \ell_0^1(x)f(x_0) + \ell_1^1(x)f(x_1) & (1.30) \\
&= \ell_0^1(x)f(1.0) + \ell_1^1(x)f(2) \\
&= \ell_0^1(x)(0) + \ell_1^1(x)(0.693) \\
&= \frac{x - x_0}{x_1 - x_0} (0.693) \\
&= \frac{x - 1}{2 - 1} (0.693) \\
&= -0.693 + 0.693x
\end{aligned}$$

A comparison of the function $f(x)$ and the approximation $\hat{f}(x)$ is presented in [Table 1.2](#). Note that at the node points $x = 1.0$ and $x = 2.0$ the solution is exact, as required by our definition of the approximating polynomial.

Table 1.2: Values of the function $f(x) = \ln(x)$, the approximation to $f(x)$, that is $\hat{f}(x)$ and the error $E(x) = f(x) - \hat{f}(x)$.

x	$f(x)$	$\hat{f}(x)$	$E(x)$
1.00	0.00	0.00	0.00
1.25	0.22	0.173	0.047
1.50	0.40	0.347	0.053
1.75	0.56	0.520	0.04.
2.0	.693	0.693	0.00.

It is helpful to examine the information provided in [Fig. 1.5](#). The interpolant \hat{f} is given in the top pane. It is a straight line since it is made up of the weighted sum of two straight lines as can be seen from [Eq. \(1.30\)](#). The approximation is the sum of the linear Lagrange polynomial at $x = 1.0$ multiplied by the value of $\ln(1)$ and the linear Lagrange polynomial defined at $x = 2$ multiplied by $\ln(2)$. The weighted sum of linear polynomials always generates a linear polynomial approximation. The lower pane in this figure shows the comparison between the function $f(x) = \ln(x)$ and its approximation $\hat{f}(x)$.

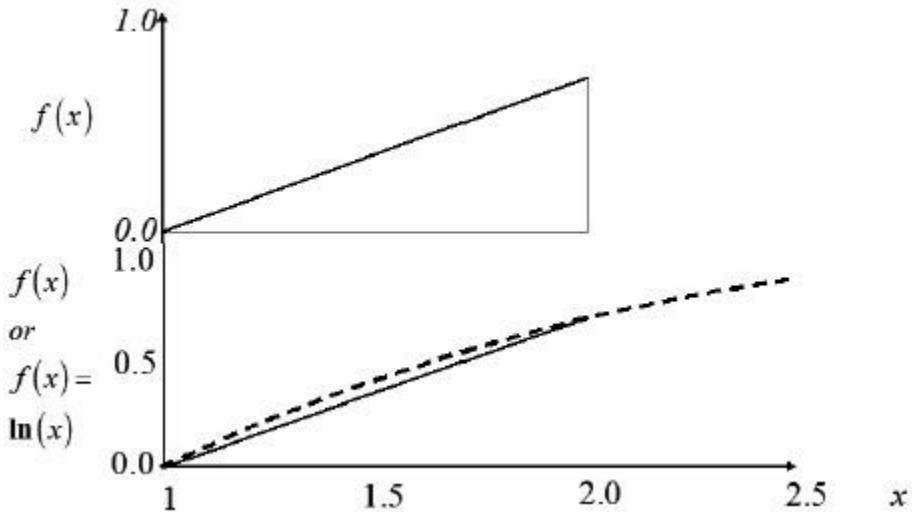


Figure 1.5: Representation of the $\ln(x)$ function using linear Lagrange polynomials.

1.6 Compare $P_2(\theta)$ and $\hat{f}(\theta)$

In this section we want to examine the relationship between the use of a quadratic polynomial, and an approximation based upon quadratic Lagrange polynomials, to interpolate. We start by writing the approximation of the $\sin(\theta)$ function, $\hat{f}(x)$, using the quadratic Lagrange Polynomials, that is

$$\hat{f}(\theta) = \sum_{j=0}^2 \ell_j^2(\theta) f_j(\theta) \quad (1.31)$$

where, using [Eq. \(1.28\)](#) we obtain three polynomials, one for each node in [Fig. 1.4](#)

$$\ell_0^2(\theta) = \frac{(\theta - \theta_1)(\theta - \theta_2)}{(\theta_0 - \theta_1)(\theta_0 - \theta_2)} \quad (1.32)$$

$$\ell_1^2(\theta) = \frac{(\theta - \theta_0)(\theta - \theta_2)}{(\theta_1 - \theta_0)(\theta_1 - \theta_2)} \quad (1.33)$$

$$\ell_2^2(\theta) = \frac{(\theta - \theta_0)(\theta - \theta_1)}{(\theta_2 - \theta_0)(\theta_2 - \theta_1)}. \quad (1.34)$$

Now we multiply each function by the appropriate coefficient value $f(\theta)$ and get

$$\hat{f}(\theta) = \ell_0^2(\theta)(0) + \ell_1^2(\theta)(1) + \ell_2^2(\theta)(0). \quad (1.35)$$

Next we substitute the definitions of $\ell_1^2(\theta)$

$$\ell_1^2(\theta) = \frac{(\theta - 0)(\theta - \pi)}{(\pi/2 - 0)(\pi/2 - \pi)} = \boxed{1.27\theta - 0.406\theta^2}. \quad (1.36)$$

Comparison of this relationship with [Eq. \(1.5\)](#) shows that

$$P_2(\theta) = \hat{f}(\theta). \quad (1.37)$$

In essence, no matter how you manipulate quadratic polynomials, whether or not they are Lagrange polynomials, you will not change the approximation. The reason we use Lagrange polynomials and the $\hat{f}(\theta)$ machinery will become more evident later.

As to the approximation, the value of ℓ_1^2 in [Fig. 1.6](#) is the approximation to the $\sin(\theta)$ function.

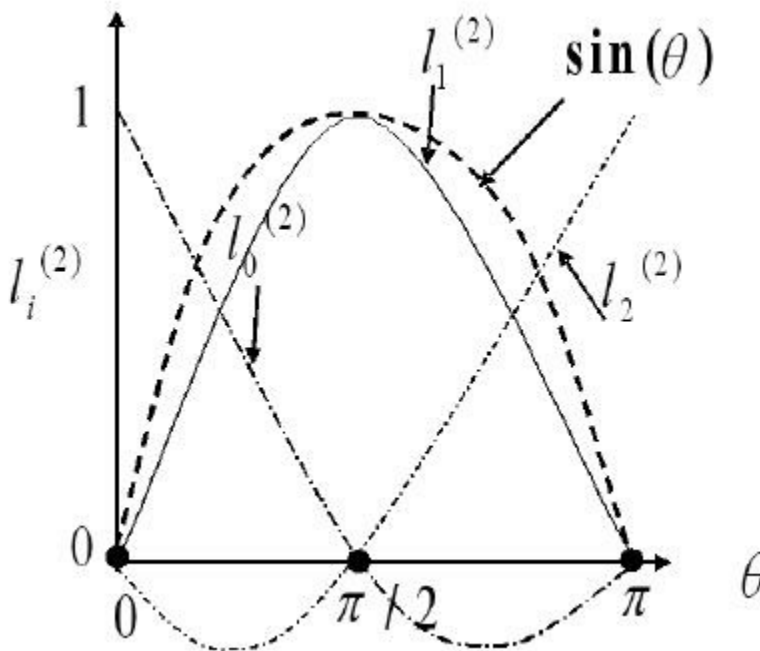


Figure 1.6: Quadratic polynomial approximation of $\sin(\theta)$.

1.7 Error of Approximation

The general idea in this section is to determine, in the absence of the function being approximated being available, how well the polynomial will approximate it. The argument is rather convoluted in that we need to build a set of concepts and then bring

them all together at the end, so please be patient.

Step 1 (Define a function $F(x)$)

We start by defining

$$F(x) = \prod_{i=0}^n (x - x_i). \quad (1.38)$$

For example for the special case of $n = 2$ we have the function

$$F(x) = \prod_{i=0}^2 (x - x_i) = (x - x_0)(x - x_1)(x - x_2) \quad (1.39)$$

which is presented in [Fig. 1.7](#).

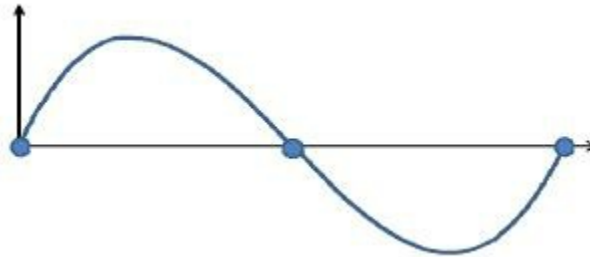


Figure 1.7: Function $F(x)$ with zeros at $x_0, x_1,$ and x_2 . The locations indicated by an x along the axis are where $\frac{dF(x)}{dx} = 0$.

It is important in this development to note that

$$F(x) = \prod_{i=0}^n (x - x_i) = 0 \quad \text{for } x = x_i, \quad i = 0, 1, \dots, n \quad (1.40)$$

because when the term $(x_i - x_i)$ arises, it is zero and that eliminates the series. In other words, the function $F(x)$ vanishes at the nodes.

Because the approximation must equal the function at the nodal points x_i , we have

$$\hat{f}(x_i) = f(x_i) \quad i = 0, 1, \dots, n. \quad (1.41)$$

Alternatively, we can write

$$f(x_i) - \hat{f}(x_i) = 0 \quad i = 0, 1, \dots, n. \quad (1.42)$$

Step 2 (Define a function $g(x)$)

Now we change course and define the following function:

$$g(x) = f(x) - \hat{f}(x) - AF(x) \quad (1.43)$$

where A is a constant. Why we do this has no answer at this point, but its relevance becomes evident shortly. Note from [Eq. \(1.40\)](#) and the [Fig. 1.7](#) that there are $n + 1$ points x_i where $F(x)$ is zero. Therefore, at these points, according to [Eq. \(1.42\)](#) and [\(1.40\)](#)

$$g(x_i) = 0 \quad i = 0, 1, \dots, n. \quad (1.44)$$

We now digress once again in Step 3 which follows.

Step 3 (Introduction of the concept of ξ)

Let us choose A such that $g(x) = 0$ at some **arbitrary** point $x_p \in [x_0, x_n]$. Now $g(x)$ is zero at at least $n + 2$ points, that is at $x_0, x_1, \dots, x_p, x_{n-1}, x_n$ (see [Fig. 1.8](#)). Since $g(x)$ is smooth, it must have a minimum or maximum between each pair of zeroes at which points the derivatives of $g(x)$ vanish. These are indicated by the letter x in [Fig. 1.8](#). Then $\frac{dg(x)}{dx}$ has at least $(n + 2) - 1 = n + 1$ zeros in the interval $[x_0, x_n]$. Similarly at the points of inflection $\frac{d^2g(x)}{dx^2}$ has at least $(n + 2) - 1 = n$ zeros. Using similar logic to look at even higher derivatives, we finally arrive at the observation that $\frac{d^{n+1}g(x)}{dx^{n+1}}$ has $(n + 2) - (n + 1) = 1$ zero.

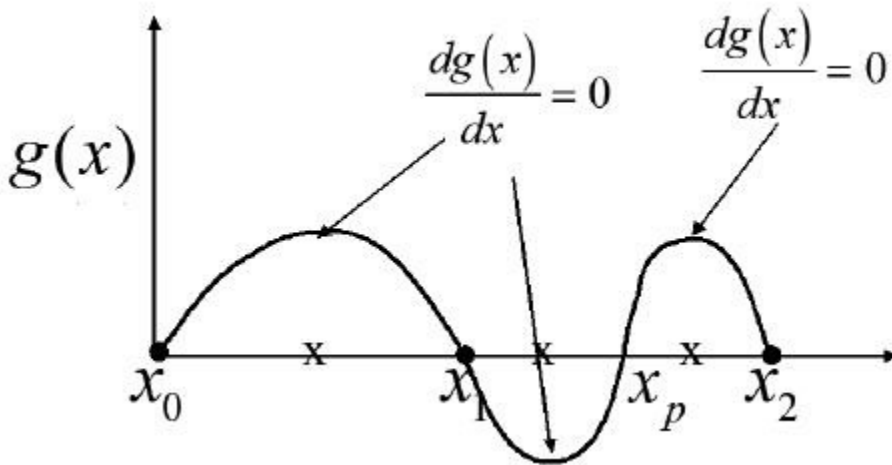


Figure 1.8: Function $g(x)$ as presented in [Eq. 1.43](#).

Let $x = \xi$ represent the location of this zero, that is

$$\boxed{\frac{d^{n+1}g(x)}{dx^{n+1}} \Big|_x = \xi = 0} \quad (1.45)$$

Since $\hat{f}(x)$ is a polynomial of degree n , $\frac{d^{n+1}\hat{f}(x)}{dx^{n+1}} = 0$. This observation will be used a little later.

Step 4 (Calculation of $\frac{d^{n+1}F(x)}{dx^{n+1}}$)

It is not obvious, but true, that by differentiating

$$F(x) = \prod_{i=0}^n (x - x_i) \quad (1.46)$$

we obtain

$$\boxed{\frac{d^{n+1}F}{dx^{n+1}} = (n+1)!} \quad (1.47)$$

To show this to be the case in at least one situation we provide the following example. This is not a proof, but provides some degree of confidence.

Example

Consider the example of $n = 1$

$$F(x) = (x - x_0)(x - x_1) \quad (1.48)$$

$$\frac{dF(x)}{dx} = (x - x_0) + (x - x_1) \quad (1.49)$$

$$\frac{d^2F(x)}{dx^2} = 1 + 1 = 2 = 2! \quad (1.50)$$

Step 5 (Calculation of A)

Let us now differentiate $g(x)$ $n + 1$ times (see [Eq. \(1.43\)](#)) to give

$$\frac{d^{n+1}g(x)}{dx^{n+1}} = \frac{d^{n+1}f(x)}{dx^{n+1}} - \frac{d^{n+1}\hat{f}(x)}{dx^{n+1}} - A \frac{d^{n+1}F(x)}{dx^{n+1}}. \quad (1.51)$$

Consider what is happening at $x = \xi$ in [Eq. \(1.51\)](#) (we now add letters to refer to the terms):

$$\underbrace{\left. \frac{d^{n+1}g(x)}{dx^{n+1}} \right|_x}_D = \xi = \underbrace{\left. \frac{d^{n+1}f(x)}{dx^{n+1}} \right|_x}_B = \xi - \underbrace{\left. \frac{d^{n+1}\hat{f}(x)}{dx^{n+1}} \right|_x}_B = \xi - \underbrace{A \left. \frac{d^{n+1}F(x)}{dx^{n+1}} \right|_x}_C = \xi \quad (1.52)$$

We now address each of the terms in [Eq. \(1.52\)](#) one at a time. Term D is zero because, from [Eq. \(1.45\)](#)

$$\left. \frac{d^{n+1}g(x)}{dx^{n+1}} \right|_{x=\xi} = 0. \quad (1.53)$$

Term B is zero because we are taking the $n + 1$ derivative of an n th degree polynomial.

$$\left. \frac{d^{n+1}\hat{f}(x)}{dx^{n+1}} \right|_{x=\xi} = 0 \quad (1.54)$$

From [Eq. \(1.47\)](#) we know that

$$\left. \frac{d^{n+1}F(x)}{dx^{n+1}} \right|_{x=\xi} = (n+1)! \quad (1.55)$$

We now combine this information with [Eq. \(1.52\)](#) to give

$$\left. \frac{d^{n+1}f}{dx^{n+1}} \right|_{x=\xi} = A \left. \frac{d^{n+1}F}{dx^{n+1}} \right|_{x=\xi} = A(n+1)! \quad (1.56)$$

or, solving for A

$$\boxed{A = \frac{1}{(n+1)!} \left. \frac{d^{n+1}f(x)}{dx^{n+1}} \right|_x = \xi} \quad (1.57)$$

Remember that A is chosen in such a way that $g(x_p) = 0$.

Step 6 (Calculation of $E(x)$)

Now we know by the definition of x_p , that at $x = x_p$, $g(x) = 0$. Thus we have

$$g(x_p) = f(x_p) - \hat{f}(x_p) - AF(x_p) = 0. \quad (1.58)$$

Rearranging [Eq. \(1.58\)](#) and substituting for A we obtain

$$\boxed{f(x_p) - \hat{f}(x_p) = \frac{1}{(n+1)!} \left. \frac{d^{n+1}f(x)}{dx^{n+1}} \right|_x = \xi (F(x_p))} \quad (1.59)$$

Finally, because x_p was selected arbitrarily, we can replace it with x to obtain:

$$f(x) - \hat{f}(x) = \frac{1}{(n+1)!} \left. \frac{d^{n+1}f(x)}{dx^{n+1}} \right|_{x=\xi} (F(x)) \quad (1.60)$$

so we can write, using the definition of $F(x)$ from [Eq. \(1.46\)](#)

$$\boxed{E(x) = f(x) - \hat{f}(x) = \frac{1}{(n+1)!} \left. \frac{d^{n+1}f(x)}{dx^{n+1}} \right|_x = \xi (\prod_i = 0^n (x - x_i)) \quad \xi \in [x_0, x_n].} \quad (1.61)$$

Take a closer look at [Eq. \(1.61\)](#). We see that the error in the interpolation is inversely proportional to a function of the number of nodes, that is the term $\frac{1}{(n+1)!}$. It is directly proportional to the value of the derivative, that is term $\left. \frac{d^{n+1}f(x)}{dx^{n+1}} \right|_{x=\xi}$, as one examines the values along x . Finally it is directly proportional to a function of the size of the distance between the location x and the nodal locations x_i , $i = 0, \dots, n$; such that the smaller the spacing between nodes, the smaller the error. For a given domain length, this is related to the value of n . So a large n yields a small distance between nodes and they work together to yield a smaller error.

Example 1

To see how one might use the above concept of error by revisiting the case presented earlier of

$$f(x) = \ln(x) \quad x \in [1, 2]. \quad (1.62)$$

From [Eq. \(1.30\)](#) on page 7 we have the approximation of $\ln(x)$ using a piecewise linear Lagrange polynomial given as

$$\hat{f}(x) = -0.693 + 0.693x \quad (1.63)$$

and for $x = 1.25$ we have the following computed values and computed error

$$\hat{f}(1.25) = 0.173 \quad (1.64)$$

$$f(1.25) = 0.22 \quad (1.65)$$

$$E(1.25) = 0.047. \quad (1.66)$$

Now let us calculate the *theoretical* error using [Eq. \(1.61\)](#). Substituting the values for $x = 1.25$ we have

$$E(x) = f(x) - \hat{f}(x) = \frac{1}{(n+1)!} \frac{d^{n+1}f(x)}{dx^{n+1}} \Big|_{x=\xi} \left(\prod_{i=0}^n (x - x_i) \right) \quad \xi \in [x_0, x_n] \quad (1.67)$$

which upon substitution of $f(x) = \ln(x)$ and $n = 1$ gives

$$E(x) = f(x) - \hat{f}(x) = \frac{1}{(1+1)!} \frac{d^{1+1}\ln(x)}{dx^{1+1}} \Big|_{x=\xi} \left(\prod_{i=0}^1 (x - x_i) \right) \quad \xi \in [1, 2] \quad (1.68)$$

which yields

$$E(1.25) = \frac{1}{2!} \frac{d^2\ln(x)}{dx^2} \Big|_{x=\xi} (x - x_0)(x - x_1) \quad (1.69)$$

$$\begin{aligned}
E(1.25) &= \frac{1}{2!} \left. \frac{d^2 \ln(x)}{dx^2} \right|_{x=\xi} (1.25 - 1)(1.25 - 2) & (1.70) \\
&= \frac{1}{2!} \left(-\frac{1}{\xi^2} \right) (1.25 - 1)(1.25 - 2) \\
&= \frac{0.094}{\xi^2}.
\end{aligned}$$

So what do we do with this? Well we recognize that

$$1 \leq \xi \leq 2 \quad (1.71)$$

so, using the limiting (upper and lower bound) values of ξ we have

$$\frac{0.094}{(2)^2} \leq E(1.25) \leq 0.094. \quad (1.72)$$

or

$$0.023 \leq E(1.25) \leq 0.094.$$

Note that the actual error according to [Eq. \(1.64\)](#) is $E(1.25) = 0.047$, which is within the bounds indicated by [Eq. \(1.72\)](#).

Example 2

Let us consider the quadratic polynomial approximation to the $\ln(x)$ $x \in [1, 2]$. We have

$$\hat{f}(x) = -1.21 + 1.33x - 0.214x^2 \quad (1.73)$$

which evaluated at $x = 1.25$ gives

$$\hat{f}(1.25) = 0.213 \quad (1.74)$$

$$f(1.25) = 0.22 \quad (1.75)$$

$$E(1.25) = 0.006. \quad (1.76)$$

From the error expression we have

$$E(x) = f(x) - \hat{f}(x) = \frac{1}{(n+1)!} \left. \frac{d^{n+1}f(x)}{dx^{n+1}} \right|_{x=\xi} \left(\prod_{i=0}^n (x - x_i) \right) \quad \xi \in [x_0, x_n] \quad (1.77)$$

which, upon substitution of the function $f(x) = \ln(x)$ and using $n = 2$, yields

(1.78)

$$E(x) = f(x) - \hat{f}(x) = \frac{1}{(2+1)!} \frac{d^{2+1} \ln(x)}{dx^{2+1}} \Big|_{x=\xi} \left(\prod_{i=0}^2 (x - x_i) \right) \quad \xi \in [1, 2]$$

or, simplifying,

$$E(x) = f(x) - \hat{f}(x) = \frac{1}{3!} \frac{d^3 \ln(x)}{dx^3} \Big|_{x=\xi} ((x - x_0)(x - x_1)(x - x_2)). \quad (1.79)$$

Now we select the point of interest, that is $x = 1.25$, and substitute it into [Eq. \(1.79\)](#) to get

$$\begin{aligned} E(1.25) &= \frac{1}{3!} \left(\frac{2}{\xi^3} \right) (1.25 - 1)(1.25 - 1.5)(1.25 - 2.0) \\ &= \frac{2}{(3)(2)\xi^3} (0.0469) = \frac{0.0156}{\xi^3}. \end{aligned} \quad (1.80)$$

Now we need to determine a choice of ξ . If we use the upper and lower bounds of the interval $x \in [1, 2]$ and therefore substitute $\xi = 1$ and $\xi = 2$ in [Eq. \(1.80\)](#) we obtain

$$0.00195 \leq E(1.25) \leq 0.0156. \quad (1.81)$$

Since the true error is $E(1.25) = 0.006$, [Eq. \(1.81\)](#) shows that it lies within the computed interval.

1.8 Multiple Elements

To this point we have been dealing with one interval. We will now introduce some new notation to consider multiple intervals. We define the interval over which one polynomial is defined as an *element*. That element may have any number of nodes depending upon the degree of the polynomial. The higher the degree of Lagrange polynomial, the larger the number of nodes we would use per element. We have thus far considered as high as a quadratic Lagrange polynomial which required three nodes per element.

In numerical methods we find the use of one element rather uninteresting. Rather we want to concatenate several elements, and use a low degree Lagrange polynomial for each element. In this section we extend our earlier work to consider multiple elements.

Firstly, consider a two element approximation using linear Lagrange polynomials in each element. Such an arrangement is found in [Fig. 1.9](#). Notice that there is a node in common at the beginning and end of all internal elements. The elements on the ends of the domain of interest share only one internal node. To introduce and illustrate the multiple-element concept, we will use once again the function $f(x) = \ln(x)$ as our example. The values of $f(x) = \ln(x)$ at each node, that is $\ln(x_j)$ where $j = 1, 2, 3$ are used to define the linear approximations

$$\hat{f}(x) = \hat{\ln}(x) = \sum_{j=1}^2 \ell_j(x) \ln(x_j) \quad x \in [1, 1.5] \quad (1.82)$$

$$\hat{f}(x) = \hat{\ln}(x) = \sum_{j=2}^3 \ell_j(x) \ln(x_j) \quad x \in [1.5, 2.0]. \quad (1.83)$$

Note that the indices in the summation now refer to node numbers, for example x_1 denotes the value of x and node 1 and that to simplify notation from this point forward, **we will drop the superscript on ℓ_j for the linear Lagrange polynomial**.

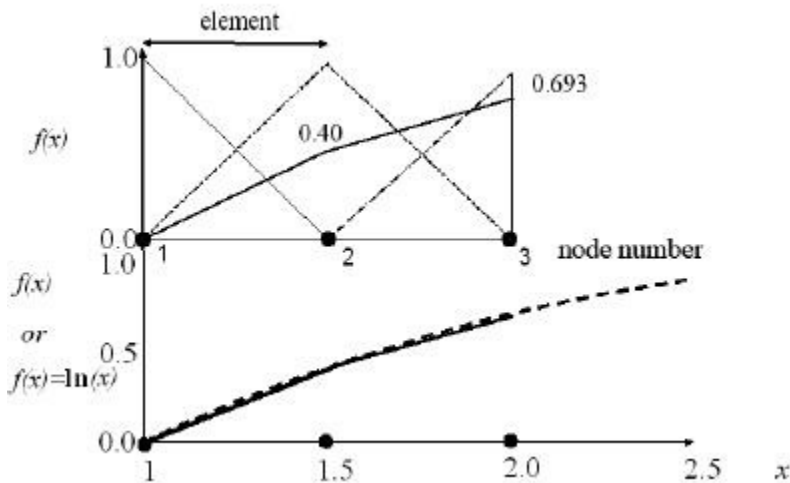


Figure 1.9: Two linear elements and their approximation (solid lines) of the logarithmic function (dashed line).

These two equations describe the piecewise linear approximations between the nodes in the upper panel of [Fig. 1.9](#). In panel two, we see the relationship between this approximation and the function being approximated, that is $\ln(x)$. To proceed we need to find a convenient way to relate information at the local level, for example that is that associated with the element, to that of the global system for example that is associated with the assemblage of elements in which the original problem is defined. To achieve this goal consider the information provided in [Fig. \(1.10\)](#) and [Tables 1.3](#) and [1.4](#). [Figure. 1.10](#) shows the relationship between the global and local coordinate systems. In the local coordinate system, each element sees the world from the same perspective. In other words, an observer at a point in the local system sees only what is happening on the element on which he/she resides. He/she does not see beyond the nodes defining each end of the element. The same observer when associating his/her position with respect to the global system sees the entire domain of interest inclusive of all elements.

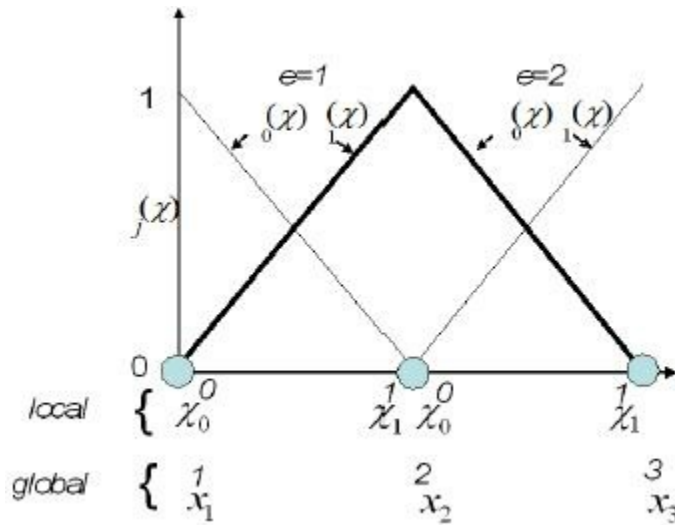


Figure 1.10: Two linear elements in the global (x) and local (χ) coordinate systems.

Table 1.3: Relationship between local and global node numbering.

Element	Local Node	Global Node
1	0	1
1	1	2
2	0	2
2	1	3

It is helpful to distinguish more clearly between the two types of coordinates. In this spirit let us define the local coordinate as χ . On the left is always node 0 and coordinate χ_0 and on the right is always node 1 and coordinate χ_1 . Thus $0 \leq \chi \leq 1$. for each and every element. Similarly the Lagrange polynomials are always represented as a function of χ , that is we have $l_0(\chi)$ and $l_1(\chi)$ no matter which element we are in.

From the global perspective you see in [Fig. 1.10](#) that the node numbers are increasing from left to right as are the coordinate values. [Table 1.3](#) presents the relationship between nodal numbering at the local or element scale and numbering at the global scale. Similarly, [Table 1.4](#) shows the relationship between the global and nodal coordinates.

Table 1.4: Relationship between local and global coordinates

Element	χ_0	χ_1
1	$x = 1$	$x = 1.5$
2	$x = 1.5$	$x = 2$

Let us see how we can derive a relationship between the two coordinate systems. We know that in element 1 when $\chi = \chi_0$, $x = x_1$. We also know that at $\chi = \chi_1$, $x = x_2$. It is clear that x is a linear function of χ , so we will write

$$x = a + b\chi. \quad (1.84)$$

From our earlier observations we have

$$x_1 = a + b\chi_0 \quad (1.85)$$

$$x_2 = a + b\chi_1 \quad (1.86)$$

or

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & \chi_0 \\ 1 & \chi_1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}. \quad (1.87)$$

Solving for a and b we have

$$\begin{aligned} \begin{bmatrix} a \\ b \end{bmatrix} &= \begin{bmatrix} 1 & \chi_0 \\ 1 & \chi_1 \end{bmatrix}^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \frac{1}{\chi_1 - \chi_0} \begin{bmatrix} \chi_1 & -\chi_0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \end{aligned} \quad (1.88)$$

Substitution of [Eq. \(1.88\)](#) into [Eq. 1.84](#) yields

$$\begin{aligned} x &= \frac{\chi_1 x_1 - \chi_0 x_2}{\chi_1 - \chi_0} + \frac{x_2 - x_1}{\chi_1 - \chi_0} \chi \\ &= \frac{\chi_1 - \chi}{\chi_1 - \chi_0} x_1 + \frac{\chi - \chi_0}{\chi_1 - \chi_0} x_2. \end{aligned} \quad (1.89)$$

We have seen this structure before; it is an expansion using Lagrange polynomials, in this case defined in the coordinate system χ . In other words, we can write

$$\begin{aligned} x &= \ell_0(\chi) x_1 + \ell_1(\chi) x_2 \\ &= \sum_{j=0}^1 \ell_j(\chi) x(\chi_j) \end{aligned} \quad (1.90)$$

where

$$\ell_0(\chi) = \frac{\chi_1 - \chi}{\chi_1 - \chi_0} \quad (1.91)$$

and

$$\ell_1(\chi) = \frac{\chi - \chi_0}{\chi_1 - \chi_0}. \quad (1.92)$$

Thus we can see that we can move between the coordinate systems; that is we can determine a value of x given a value of χ if we know the nodal locations $x(\chi_0)$ and $x(\chi_1)$ which are x_1 , and x_2 , respectively in our example. If we were to rewrite [Eq. \(1.84\)](#) as

$$\chi = a + bx \tag{1.93}$$

we could show that for element 2 ($e = 2$) in [Fig. 1.10](#)

$$\chi = \ell_2^2(x) \chi \Big|_{x=x_2} + \ell_3^2(x) \chi \Big|_{x=x_3} \tag{1.94}$$

where the element number is indicated by the superscript on the left hand side of the variable. [Equation \(1.94\)](#) states that, given a value of x , we can determine the value of χ .

1.8.1 Example

Suppose we want to find the value of x at location $\chi = 0.5$ in element 2 in [Fig. 1.10](#). Using [Eq. \(1.90\)](#) we have

$$\begin{aligned} x &= \ell_0(\chi) x_2 + \ell_1(\chi) x_3 & (1.95) \\ x|_{\chi=0.5} &= \ell_0(0.5) x_2 + \ell_1(0.5) x_3 \\ &= (0.5) x_2 + (0.5) x_3 \\ &= \frac{x_2 + x_3}{2}. \end{aligned}$$

If we assume $x_3 = 2$ and $x_2 = 1.5$ as is the case in our approximation of $\ln(x)$ as shown in [Fig. 1.9](#) we obtain

$$\begin{aligned} x|_{\chi=0.5} &= \frac{x_2 + x_3}{2} & (1.96) \\ &= \frac{1.5 + 2}{2} \\ &= \frac{3.5}{2} \end{aligned}$$

For multiple elements we have what is shown in [Fig. 1.11](#).

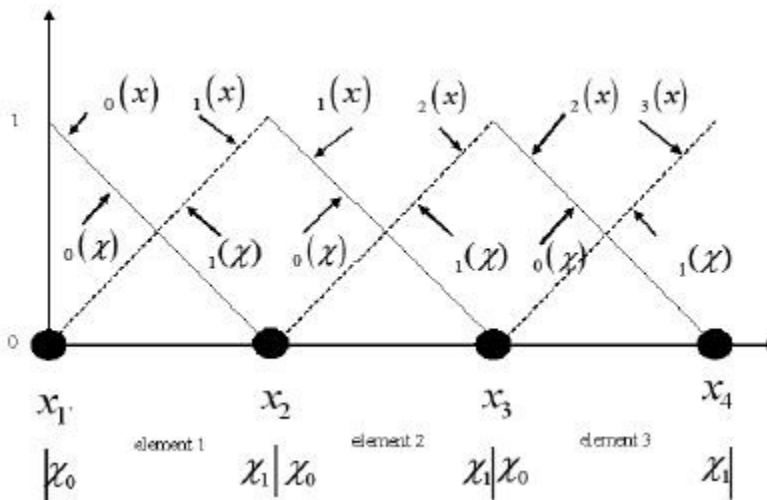


Figure 1.11: Global and local basis functions (Lagrange Polynomials) for four-node problem.

Let us now consider the approximations of our logarithm in each element in [Fig. 1.9](#). In terms of the local coordinate χ it will be of the form (recall that in this context **the element being considered is represented by the superscript 1 or 2**):

$$\hat{f}^1(\chi) = \sum_{j=0}^1 \ell_j^1(\chi) f^1(\chi_j) = \ell_0^1(\chi) f^1(\chi_0) + \ell_1^1(\chi) f^1(\chi_1) \quad x \in [1, 1.5] \quad (1.97)$$

$$\hat{f}^2(\chi) = \sum_{j=0}^1 \ell_j^2(\chi) f^2(\chi_j) = \ell_0^2(\chi) f^2(\chi_0) + \ell_1^2(\chi) f^2(\chi_1) \quad x \in [1.5, 2]. \quad (1.98)$$

To obtain our approximation to the logarithm, we substitute appropriate values for $f(\chi_j)^k$ where k is here used as a generalization of that used above specifically as the element number. For example, in [Fig. 1.9](#)

$f(x_1) = f(x) \Big|_{x=1} = f^1(\chi) \Big|_{\chi=0} = \ln(1) = 0$. Substituting the appropriate values observed at the nodes we get

$$\hat{f}^1(\chi) = \ell_0^1(\chi) (0) + \ell_1^1(\chi) (0.40) \quad x \in [1, 1.5] \quad (1.99)$$

$$\hat{f}^2(\chi) = \ell_0^2(\chi) (0.4) + \ell_1^2(\chi) (0.693) \quad x \in [1.5, 2.0] \quad (1.100)$$

If we now substitute in the definition of the linear Lagrange polynomials we obtain

$$\hat{f}^1(\chi) = \left(\frac{\chi_1 - \chi}{\chi_1 - \chi_0} \right)^1 (0) + \left(\frac{\chi_0 - \chi}{\chi_0 - \chi_1} \right)^1 (0.40) \quad (1.101)$$

$$\hat{f}^1(\chi) = \left(\frac{1.0 - \chi}{1.0} \right)^1 (0) + \left(\frac{0 - \chi}{-1.0} \right)^1 (0.40)$$

$$\hat{f}^1(\chi) = 0.0 + 0.40\chi^1 \quad x \in [1, 1.5]$$

and

$$\hat{f}^2(\chi) = \left(\frac{\chi_1 - \chi}{\chi_1 - \chi_0} \right)^2 (0.4) + \left(\frac{\chi_0 - \chi}{\chi_0 - \chi_1} \right)^2 (0.693) \quad (1.102)$$

$$\hat{f}^2(\chi) = \left(\frac{1.0 - \chi}{1.0} \right)^2 (0.40) + \left(\frac{0 - \chi}{-1.0} \right)^2 (0.693)$$

$$\hat{f}^2(\chi) = 0.29\chi^2 + 0.40 \quad x \in [1.5, 2.0]$$

which gives us an interpolation of the logarithm using the local coordinate system. Note

that the superscripts on the bracketed quantities in [Eqs. \(1.101\)](#) and [\(1.102\)](#) represent the elements and are not exponents. Now let us calculate the value of $\ln(1.25)$. We obtain

$$\hat{f}^1(\chi = 0.5) = 0.20 \quad (1.103)$$

since $\hat{f}^1(\chi)$ has been obtained. For the location 1.75 we need to use the approximation in element 2. Thus we have

$$\hat{f}(\chi = 0.5) = \hat{f}^2(\chi = (0.29)(0.5) + (0.4)) = 0.546 \quad (1.104)$$

Tabulating the results, including the quadratic approximation we obtained in [Eq. \(1.73\)](#), we get

	$f(1.25)$	$\hat{f}(1.25)$	$E(1.25)$	$f(1.75)$	$\hat{f}(1.75)$	$E(1.75)$
one linear element	0.22	0.173	0.047	0.559	0.52	0.04
two linear elements	0.22	0.20	0.02	0.559	0.546	0.013
one quadratic element	0.22	0.22	-0.006	0.559	0.560	0.001

Note for future reference that [Eqs. \(1.97\)](#) and [\(1.98\)](#) could be written in matrix notation for an arbitrary element e as

$$\left[\hat{f}(\chi) \right]^e = \left[l_0(\chi) \quad l_1(\chi) \right]^e \begin{bmatrix} f(\chi_0) \\ f(\chi_1) \end{bmatrix}^e \quad (1.105)$$

1.9 Hermite Polynomials

In this section we investigate another form of polynomial that allows us to directly interpolate not only the function $f(x)$ but also its derivative $\frac{df(x)}{dx}$; this is the Hermite polynomial. The Hermite polynomial is one of a suite of polynomials that when concatenated make up a numeric function that has a prescribed degree of smoothness. The point of departure for establishing the form of these piecewise Hermite polynomial functions is the following expression:

$$\begin{aligned} f(x) &= \sum_{j=0}^1 \left(h_j^0(x) f(x_j) + h_j^1(x) \left. \frac{df}{dx} \right|_{x_j} \right) + E(x) \\ &= \hat{f}(x) + E(x) \end{aligned} \quad (1.106)$$

where the functions $h_j^0(x)$ and $h_j^1(x)$ are the Hermite polynomials. The superscripts here are identified with the two forms of the Hermite *function*, one, h_j^0 , associated with the value of the function at the node x_j , that is, $f(x_j)$, and the other $h_j^1(x)$ associated with the

derivative at the node, that is $\frac{df}{dx}\bigg|_{x_j}$. There are four functions, two associated with each node in the element, and they are shown in [Fig. 1.12](#). To determine the form of the functions $h^0_j(x)$ and $h^1_j(x)$ we require that, as in the case of the Lagrange polynomials, the approximating function exactly represent at the nodes the function being approximated. However, in addition to these constraints we require that the approximation of the derivatives of the function being approximated also be exact at the nodes. These sets of requirement, two at each of two nodes, yields:

$$\hat{f}(x_0) = f(x_0) \tag{1.107}$$

$$\hat{f}(x_1) = f(x_1) \tag{1.108}$$

$$\frac{d\hat{f}}{dx}\bigg|_{x_0} = \frac{df}{dx}\bigg|_{x_0} \tag{1.109}$$

$$\frac{d\hat{f}}{dx}\bigg|_{x_1} = \frac{df}{dx}\bigg|_{x_1} \tag{1.110}$$

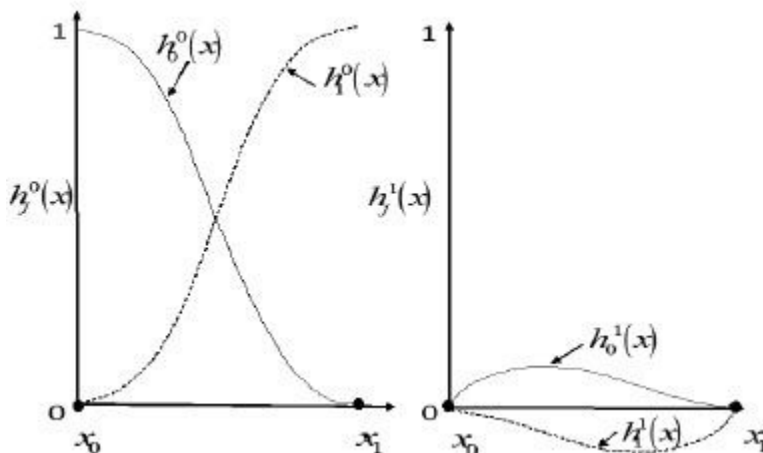


Figure 1.12: Hermite polynomials.

We now expand [Eq. \(1.106\)](#) for each of the above conditions; that is we expand the approximation and then impose the requirements of Eqs. [\(1.107\)](#)-[\(1.110\)](#). The first line of the two associated with each approximate value, for example $\hat{f}(x_0)$, is the expansion and the second is the value the expansion must represent.

$$\begin{aligned} \hat{f}(x_0) &= \boxed{h_0^0(x_0) f(x_0)} + h_1^0(x_0) f(x_1) + h_0^1(x_0) \frac{df}{dx}\bigg|_{x_0} + h_1^1(x_0) \frac{df}{dx}\bigg|_{x_1} \\ &= \boxed{f(x_0)} \end{aligned} \tag{1.111}$$

$$\hat{f}(x_1) = h_0^0(x_1) f(x_0) + \boxed{h_1^0(x_1) f(x_1)} + h_0^1(x_1) \left. \frac{df}{dx} \right|_{x_0} + h_1^1(x_1) \left. \frac{df}{dx} \right|_{x_1} \quad (1.112)$$

$$= \boxed{f(x_1)} \quad (1.113)$$

$$\left. \frac{d\hat{f}}{dx} \right|_{x_0} = \left. \frac{dh_0^0}{dx} \right|_{x_0} f(x_0) + \left. \frac{dh_1^0}{dx} \right|_{x_0} f(x_1) + \boxed{\left. \frac{dh_0^1}{dx} \right|_{x_0} \left. \frac{df}{dx} \right|_{x_0}} + \left. \frac{dh_1^1}{dx} \right|_{x_0} \left. \frac{df}{dx} \right|_{x_1}$$

$$= \boxed{\left. \frac{df}{dx} \right|_{x_0}}$$

$$\left. \frac{d\hat{f}}{dx} \right|_{x_1} = \left. \frac{dh_0^0}{dx} \right|_{x_1} f(x_0) + \left. \frac{dh_1^0}{dx} \right|_{x_1} f(x_1) + \left. \frac{dh_0^1}{dx} \right|_{x_1} \left. \frac{df}{dx} \right|_{x_0} + \boxed{\left. \frac{dh_1^1}{dx} \right|_{x_1} \left. \frac{df}{dx} \right|_{x_1}}$$

$$= \boxed{\left. \frac{df}{dx} \right|_{x_1}}$$

In general, for Eqs. (1.111) through (1.114) to be satisfied, for any one of the four equations, the boxed terms in each expression should be equal. For this to be true, the information appearing in the following tables is required. The first line of Table 1.5, for example, should be read as follows: In order to have $\hat{f}(x_0) = f(x_0)$, we should have $h_0^0 = 1$. In addition, as seen in Table 1.6 $h_1^0(x_0)$ and $h_1^1(x_0)$ must also be zero at both x_0 and x_1 and $h_1^0(x_0) = 0$ where the locations x_0 and x_1 are the two nodal locations in the element.

Table 1.5: Necessary conditions to be imposed on the Hermite polynomials to assure that the approximation and function are the same at the nodes.

	$\hat{f}(x_0) = f(x_0)$	$\hat{f}(x_1) = f(x_1)$
h_0^0	1	0
h_1^0	0	1
h_0^1	0	0
h_1^1	0	0

Table 1.6: Necessary conditions to be imposed on the Hermite polynomials to assure that the approximation of the derivative and the derivative are the same at the nodes.

	$\left. \frac{d\hat{f}}{dx} \right _{x_0} = \left. \frac{df}{dx} \right _{x_0}$	$\left. \frac{d\hat{f}}{dx} \right _{x_1} = \left. \frac{df}{dx} \right _{x_1}$

$\frac{dh_0^0}{dx}$	0	0
$\frac{dh_1^0}{dx}$	0	0
$\frac{dh_0^1}{dx}$	1	0
$\frac{dh_1^1}{dx}$	0	1

From the first row in [Tables 1.5](#) and [1.6](#) we see that there are four conditions to be imposed on $h_0^0(x)$ (read horizontally across in each of these tables). For example, $h_0^0(x)$ must take on a value of 1 at x_0 and 0 at x_1 and $\frac{dh_0^0(x)}{dx}$ must be 0 at both x_0 and x_1 . Since a cubic polynomial is completely defined by four conditions (which is the number we have available for each of the functions $h_0^0(x)$, $h_0^1(x)$, $h_1^0(x)$, and $h_1^1(x)$), let us assume that each is a cubic polynomial of the form, for example:

$$h_0^0(x) = a_0 + b_0x + c_0x^2 + d_0x^3. \quad (1.115)$$

The coefficients a_0 , b_0 , c_0 , and d_0 can be determined using the information in the above tables. For example one equation for the case of $h_0^0(x)$ we can obtain using the first element of the first row in [Table 1.5](#); that is $h_0^0(x)$ must be 1 at x_0 . We get:

$$h_0^0(x_0) = a_0 + b_0x_0 + c_0x_0^2 + d_0x_0^3 = 1. \quad (1.116)$$

If we now take the derivative of [Eq. \(1.115\)](#) and evaluate it at x_0 we know it must equal zero from the first row in [Table 1.6](#). If we now collect this information for each element in row 1 of these tables we obtain the following set of equations for the coefficients a_0 , b_0 , c_0 and d_0

$$\begin{bmatrix} 1 & x_0 & x_0^2 & x_0^3 \\ 0 & 1 & 2x_0 & 3x_0^2 \\ 1 & x_1 & x_1^2 & x_1^3 \\ 0 & 1 & 2x_1 & 3x_1^2 \end{bmatrix} \begin{bmatrix} a_0 \\ b_0 \\ c_0 \\ d_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (1.117)$$

$$\begin{aligned} h_0^0(x_0) &= 1 \\ \frac{dh_0^0}{dx} \Big|_{x_0} &= 0 \\ h_0^0(x_1) &= 0 \\ \frac{dh_0^0(x_1)}{dx} \Big|_{x_1} &= 0. \end{aligned}$$

In reading [Eq. \(1.117\)](#) the information appearing to the right of the matrix equation identifies the conditions giving rise to each row of the equation.

Solving for a_j , b_j , c_j , and d_j and substituting in cubic polynomial we obtain

$$h_j^0(x) = \left(\ell_j^{(1)}(x) \right)^2 \left(1 - 2(x - x_j) \frac{d\ell_j^{(1)}}{dx} \Big|_{x_j} \right) \quad (1.118)$$

where the Lagrange polynomial approximations are used in the definition. One can make similar arguments to obtain

$$h_j^1(x) = \left(\ell_j^1(x)^2 \right) (x - x_j). \quad (1.119)$$

As noted earlier, the resulting functions are provided in [Fig. 1.12](#). Note that the curves in this figure satisfy the constraints provided in [Tables 1.5](#) and [1.6](#).

Example

Consider the function

$$f(x) = \exp(x) \quad x \in [0.5, 1] \quad (1.120)$$

which upon differentiation yields

$$\frac{df(x)}{dx} = \exp(x) \quad x \in [0.5, 1]. \quad (1.121)$$

The approximation $\hat{f}(x)$ is given by introducing the values in [Eq. \(1.120\)](#) and [\(1.121\)](#) into [Eq. \(1.106\)](#); that is

$$\hat{f}(x) = h_0^0(x) \exp(0.5) + h_0^1(x) \exp(0.5) + h_1^0(x) \exp(1.0) + h_1^1(x) \exp(1.0) \quad (1.122)$$

A visual representation of the resulting approximation is provided in [Fig. 1.13](#)

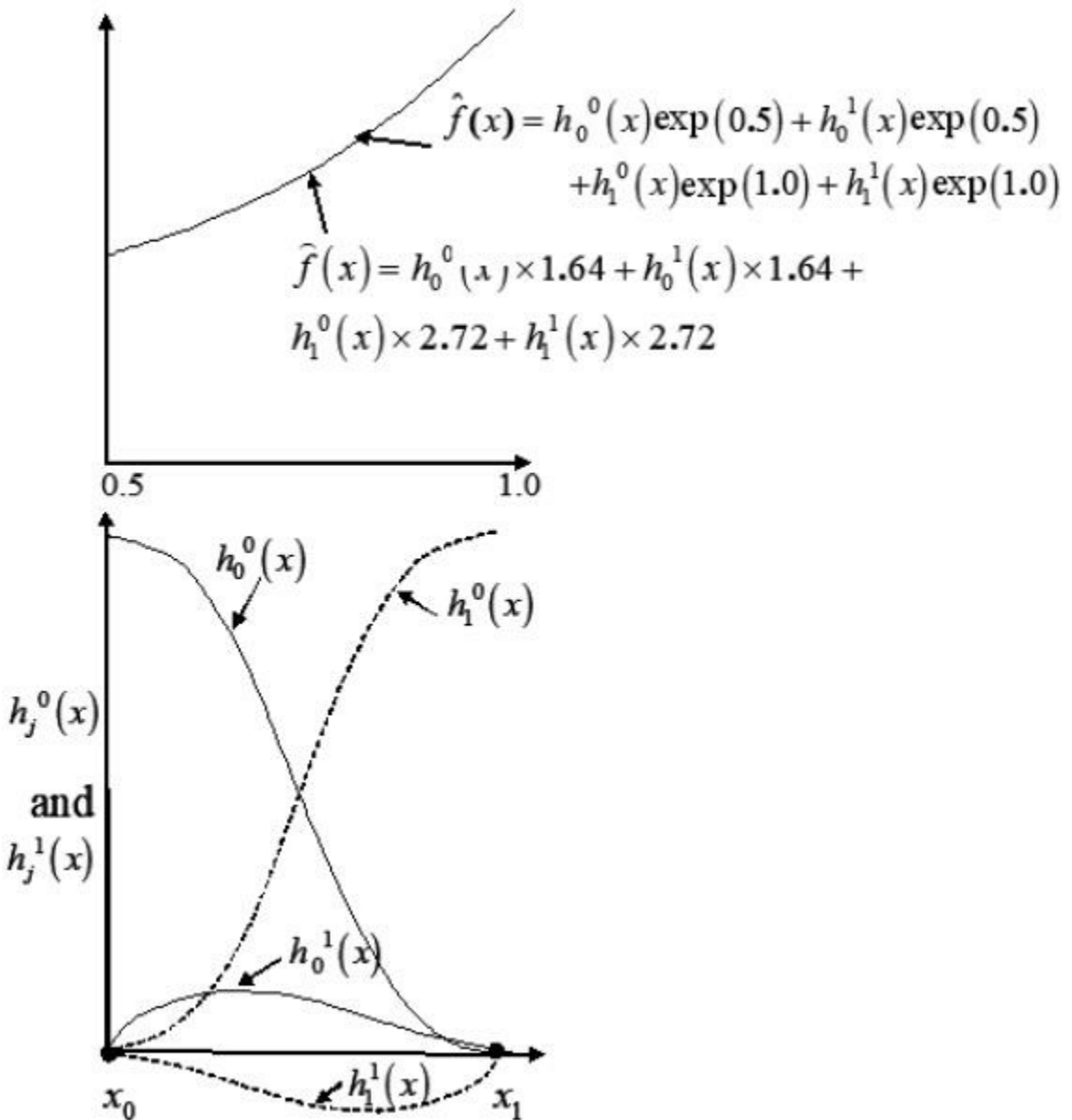


Figure 1.13: Hermite approximation of the exponential function.

hermite approx to exp

1.10 Error in Approximation by Hermites

We are not going to develop the error of the approximation using Hermites, but simply provide it for the our case (cubic with two nodes) below, viz.

$$E(x) = \frac{1}{(4)!} \left[\prod_{i=0}^1 (x - x_i) \right]^2 \frac{d^4 f}{dx^4} \Big|_{x=\xi} \quad (1.123)$$

Example of Hermite Error Approximation

(1.124)

$$\begin{aligned}
 E(0.6) &= \frac{1}{4!} [(x - x_0)(x - x_1)]^2 \exp(x) \Big|_{x=\xi} \quad x \in [0.5, 1.0] \\
 &= \frac{1}{4!} [(0.6 - 0.5)(0.6 - 1.0)]^2 \exp(x) \Big|_{x=\xi} \\
 &= 6.67 \times 10^{-5} \exp(x) \Big|_{x=\xi}.
 \end{aligned}$$

Since

$$0.5 \leq \xi \leq 1.0 \quad (1.125)$$

then using these limits the error bound is

$$1.1 \times 10^{-4} \leq E(0.6) \leq 1.81 \times 10^{-4}. \quad (1.126)$$

The measured error is

$$E(0.6) = 1.82212 - 1.82198 = 1.4 \times 10^{-4} \quad (1.127)$$

which lies within the error bounds calculated.

1.11 Chapter Summary

Polynomial approximation theory, the subject of this chapter, is a fundamental theoretical underpinning for numerical methods. In this chapter we introduce this topic in the context of using polynomial approximation theory to approximate functions. The Lagrangian polynomials and the Hermite polynomials are considered along with the error associated with using them in approximating functions. The concept of discretizing a domain, say a length along x , into a set of concatenated linear segments called elements is introduced.

1.12 Problems

1. Determine (derive) the form of $\ell_1^1(x)$ (see [Fig. 1.14](#)). The result should be an algebraic equation that describes the function $\ell_1^e(x)$ for any interval (x_0, x_1) . The strategy is to write the general form of the linear equation and then impose the conditions required of a linear approximating function.

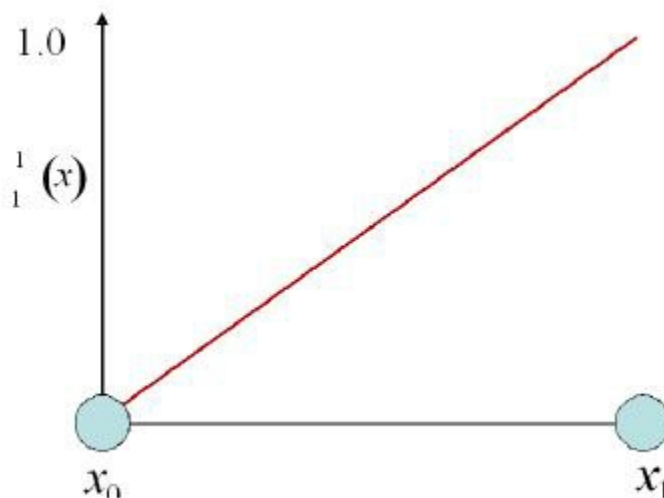


Figure 1.14: Linear basis function.

2. Using [Eq. \(1.128\)](#) below write the form of the second degree or quadratic Lagrange polynomial $\ell_1^2(x)$ shown below? How does it satisfy the requirements of a Lagrange polynomial?

$$\ell_j^{(n)}(x) = \prod_{\substack{i=0 \\ i \neq j}}^n \frac{x - x_i}{x_j - x_i}. \quad (1.128)$$

The shape of this function is shown in [Fig. 1.15](#) below.

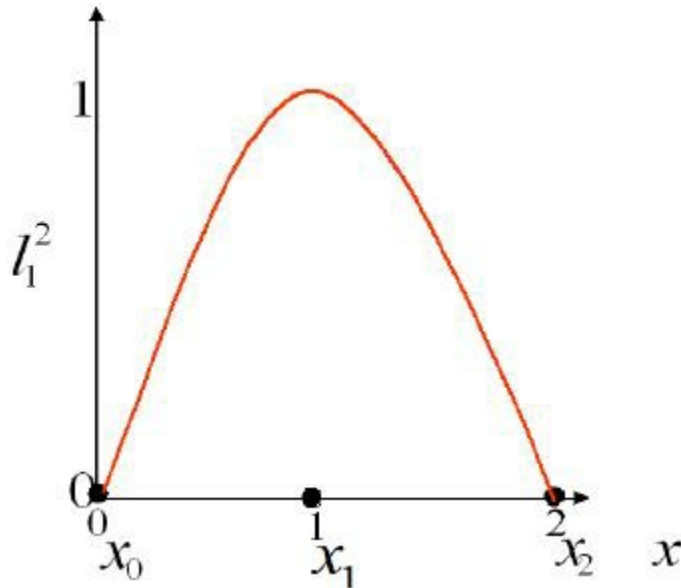


Figure 1.15: Quadratic Lagrange basis function defined over the interval $x_0 - x_2$.

3. $P_n(\theta)$ is a polynomial of degree n , that is,

$$P_n(\theta) = \sum_{i=0}^n a_i \theta^i \quad (1.129)$$

$P_n(\theta_i) = f(\theta_i)$ where θ_i are particular values of θ . Assume $f(x) = \sin(\theta)$ (see [Fig. 1.16](#)) and that $n = 3$ (the polynomial is a cubic). Determine the polynomial represented by [1.129](#) using points $\theta_i = 0, \pi/4, 3\pi/4, \pi$. Compare these errors to those presented in the example in the notes in [Section 1.3](#). $\sin(\theta)$ curve with measured points indicated by small black dots and interpolated values indicated by large black dots.

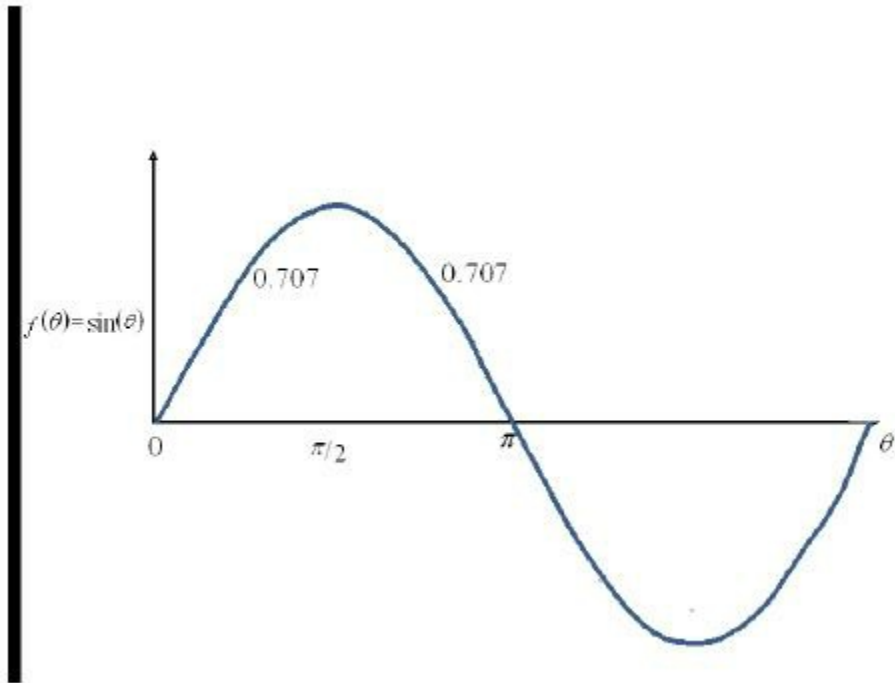


Figure 1.16: Plot of $\sin(\theta)$ over the range $\theta = 0$ to $\theta = 2\pi$.

4. Consider the function

$$f(x) = e^x \quad x \in [0, 1]. \quad (1.130)$$

Use a piecewise linear polynomial to determine $e^{0.5}$. Now calculate the error of your approximation, that is, determine the difference between the exact value of $e^{0.5}$ and your estimate. Using the following equation for the theoretical error

$$E(x) = f(x) - \hat{f}(x) = \frac{1}{(n+1)!} \left. \frac{d^{n+1}f(x)}{dx^{n+1}} \right|_{x=\xi} (\prod_{i=0}^n (x-x_i)) \quad \xi \in [x_0, x_n] \quad (1.131)$$

Show that the error of your estimate is consistent with the theoretical estimate.

5. Consider the situation presented in (Fig. 1.17) You know the value of x at the location at 1.5 in inches and you want to determine the same location in centimeters. Using the transformations given by

$$x = a + b\chi \quad (1.132)$$

where x is the distance in inches and χ is the distance in centimeters determine the location X in centimeters.



Figure 1.17: Representation of a line in terms of inches and centimeters.

6. Please do the following:
 - a. Using the information in Fig 1.18, draw the linear Lagrange functions $\ell_i(x)$ and $\ell_i(\chi)$. In the local coordinate system use a solid line for $\ell_0(\chi)$ and a dashed line for $\ell_1(\chi)$. Label the horizontal and vertical axis with appropriate values. Give each of the symbolic x_i and χ_i actual values, (in the sense of values of x_0 and χ_0 etc.) and actual numerical values (in the sense of 1.0 etc.) in the spaces provided. You will have five values of x_i of your choosing and the same number of values of χ , but there will be duplicate values of χ at most of the nodes.
 - b. What is the value of $h = x_{i+1} - x_i$?
 - c. What is the value of $\chi_1 - \chi_0$?

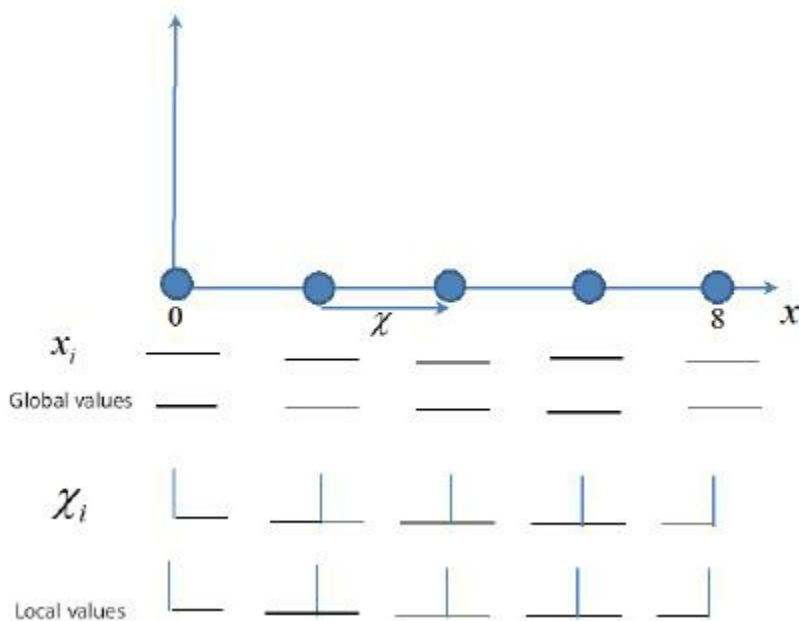


Figure 1.18: Representation of five points in global (x) and local (χ) coordinate systems.

7. Given the formula for the Lagrange polynomial, that is

$$\ell_j^n(x) = \prod_{\substack{i=0 \\ i \neq j}}^n \frac{x - x_i}{x_j - x_i} \tag{1.133}$$

write the form of the cubic Lagrange polynomial. Then use this expression to write the approximation of the $\sin(\theta)$ segment from 0 to π . Use four equally spaced values of π to obtain this approximation (four π values). Now evaluate your polynomial at the points indicated in Table 1.1, to compute the error, and note how the errors compare with those already recorded in this table.

8. From the equation

$$x = \sum_{j=0}^1 \ell_j(\chi) x(\chi_j) \quad \chi \in (0, 1), \quad x \in (1, 2) \quad (1.134)$$

determine the value of χ at $x = 0.75$.

9. State the four requirements (constraints) necessary to obtain a Hermite interpolating polynomial h^0 as used in the expression

$$\hat{u}(x) = u_0 h_0^0 + u_1 h_1^0 + \left. \frac{du}{dx} \right|_0 h_0^1 + \left. \frac{du}{dx} \right|_1 h_1^1 \quad (1.135)$$

assuming you start with the standard cubic polynomial.

- o. Use a linear polynomial to approximate $\sin \theta$, $0 \leq \theta \leq \pi$. shown in [Fig. 1.19](#). Determine whether the error at $\pi/2$ falls within the range predicted by the error formula.

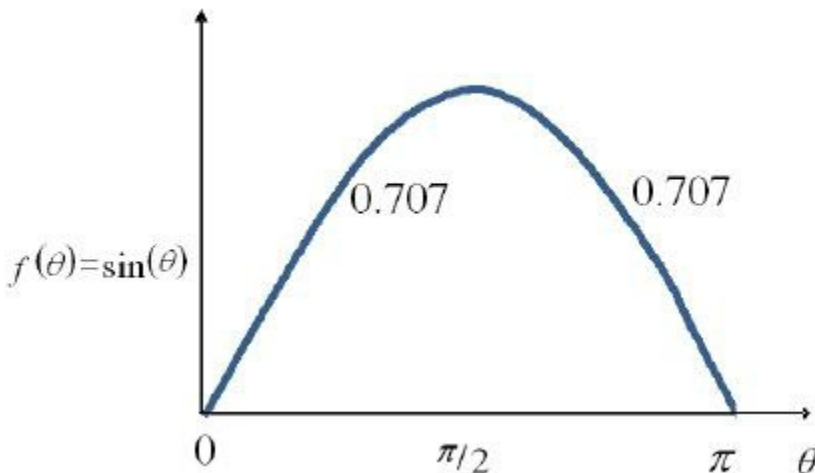


Figure 1.19: The $\sin(\theta)$ function from 0 to π .

11. Use quartic (fourth degree) Lagrange polynomials to approximate $\sin \theta$ using nodes at $0, \pi/2, \pi, (3/2)\pi$ and π . Determine the value of $\sin \theta$ at $(5/4)\pi$. Compare the estimate to the value obtained using the quadratic Lagrange polynomial. Is it better; why (or why not)?
12. Given the $\sin(\theta)$ function shown above and linear polynomials written in terms of the local coordinate system χ , determine the value of the $\sin(\theta)$ at $\chi = 0.5$.
13. The goal here is to determine if one can use any two points along the line from $x = 0$ to $x = 2$ in [Fig. 1.20](#) to estimate the value at $x = 1.0$. Please solve this problem by following these steps:

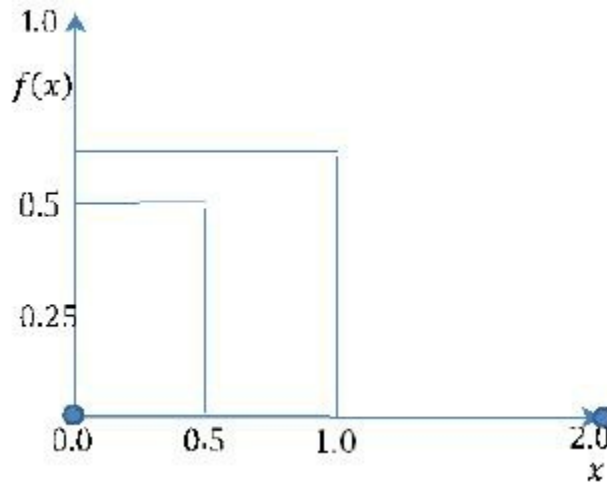


Figure 1.20: Problem definition.

- a. Assume $f(x)$ can be approximated by

$$\hat{f}(x) = \sum_{j=0}^1 \ell_j(x) f_j \tag{1.136}$$

where $\ell_j(x)$ is the linear Lagrange polynomial and f_j is the value of $f(x)$ at $x = x_j$.

- b. Expand this expression to represent that needed for this problem.
- c. Evaluate the resulting expression at the point $x = 0.0$ and $x = 0.5$. Assume $\hat{f}(0.0) = 0.25$ and $\hat{f}(0.5) = 0.5$. You should have two equations in the two unknown values of f_1 and f_2 where the subscripts are the two nodal identifiers, that is node 1 and node 2.
- d. Solve for the two unknown values of f_1 and f_2 .
- e. Use these values to estimate the value of $f(x)$, where $x = 1.0$.

Hint: remember that the linear Lagrange polynomials are 1 at the node for which they are defined and 0 elsewhere, so in this case they are two straight lines crossing at $x = 1$ and $\ell(1) = 0.5$.

Bibliography

Cheney, E.W. (1982) *Introduction to Approximation Theory*, 2nd ed., Chelsea Publishing Company, New York

Davis, P.J. (1975) *Interpolation and Approximation*, 2nd ed.A, Dover, Mineola, New York

Hildebrand, F.B. (1974) *Introduction to Numerical Analysis*, Dover, Mineola, New York

Jeffreys, H. and Jeffreys, B. S.(1988) “Weierstrass’s Theorem on Approximation by Polynomials and Extension of Weierstrass’s approximation theory.,” in *Methods of*

Mathematical Physics, 3rd ed. Cambridge University Press, Cambridge, England pp. 446-448

Mason, J.C. and M.G. Cox (1987) *Algorithms for Approximation*, Clarendon Press, New York.

Stoer, J. (1993) *Introduction to Numerical Analysis*, 2nd ed., Springer-Verlag, New York, NY.

Süli, E. and D. Mayers (2003) *An Introduction to Numerical Analysis*, Cambridge University Press, Cambridge, UK.

Chapter 2

Numerical Differentiation

2.1 General Theory

We are now ready to use some of the theory we developed above to derive the approximation of a continuous derivative that will yield an algebraic approximate equation that can be used to solve a differential equation numerically using elementary algebra. Let us begin with the following interpolation formula where $\ell_j^n(x)$ are the Lagrange polynomials of degree n

$$f(x) = \sum_{j=0}^n \ell_j^n(x) f(x_j) + E(x). \quad (2.1)$$

Now differentiate $f(x)$ to give

$$\frac{df(x)}{dx} = \sum_{j=0}^n \frac{d\ell_j^n(x)}{dx} f(x_j) + \frac{dE(x)}{dx}. \quad (2.2)$$

Notice that in differentiating $f(x)$ the value of the function at location x_j is not affected because it is a number, not a function of x .

From the equation defining the error associated with approximating functions using Lagrange polynomials, we have from Eq. (1.61) on page 12

$$E(x) = \frac{1}{(n+1)!} \prod_{i=0}^n (x - x_i) \left. \frac{d^{(n+1)}f(x)}{dx^{(n+1)}} \right|_{x=\xi}. \quad (2.3)$$

Again, differentiating we have

$$\frac{dE(x)}{dx} = \frac{d}{dx} \left[\frac{1}{(n+1)!} \prod_{i=0}^n (x - x_i) \left. \frac{d^{(n+1)}f(x)}{dx^{(n+1)}} \right|_{x=\xi} \right] \quad (2.4)$$

where we must remember that ξ is an unknown function of x . In order to proceed, we need to introduce the results of a theorem that will help us understand the term on the right hand side of [Eq. \(2.4\)](#). A proof of this theorem is found in Ralston, (1965, pg. 77).

Theorem 2 *Let*

$$E(x) = \frac{1}{(n+1)!} \prod_{i=0}^n (x - x_i) \left. \frac{d^{(n+1)}f(x)}{dx^{(n+1)}} \right|_{x=\xi} \quad (2.5)$$

be the error term in the Lagrangian interpolation formula with ξ in the interval

spanned by x_0, x_1, \dots, x_n and x . Then if $\frac{d^{(n+1)}f(x)}{dx^{(n+1)}}$ is continuous,

$$\frac{1}{(n+1)!} \frac{d}{dx} \left(\frac{d^{(n+1)}f(x)}{dx^{(n+1)}} \Big|_{x=\xi} \right) = \frac{1}{(n+2)!} \frac{d^{(n+2)}f(x)}{dx^{(n+2)}} \Big|_{x=\eta} \quad (2.6)$$

where η is also in the interval spanned by x_0, x_1, \dots, x_n where $x_i = ih$ and

$$h \equiv x_i - x_{i-1}. \quad (2.7)$$

Combination of [Eq. \(2.4\)](#) with the results of the above theorem give

$$\begin{aligned} \frac{dE(x)}{dx} &= \frac{d}{dx} \left[\frac{1}{(n+1)!} \prod_{i=0}^n (x - x_i) \frac{d^{(n+1)}f(x)}{dx^{(n+1)}} \Big|_{x=\xi} \right] \\ &= \prod_{i=0}^n (x - x_i) \frac{1}{(n+2)!} \frac{d^{(n+2)}f(x)}{dx^{(n+2)}} \Big|_{x=\eta} \\ &\quad + \frac{1}{(n+1)!} \left(\frac{d^{(n+1)}f(x)}{dx^{(n+1)}} \Big|_{x=\xi} \right) \frac{d}{dx} \left[\prod_{i=0}^n (x - x_i) \right]. \end{aligned} \quad (2.8)$$

Note that at a nodal point x_j the first term on the right-hand side of [Eq. \(2.8\)](#) vanishes (because we are forming a product and when x is a nodal location, one term in the product is zero, so the whole product is zero) and we have

$$\frac{dE}{dx} \Big|_{x_j} = \frac{1}{(n+1)!} \left(\frac{d^{(n+1)}f(x)}{dx^{(n+1)}} \Big|_{x=\xi} \right) \frac{d}{dx} \left[\prod_{i=0}^n (x - x_i) \right] \Big|_{x_j}. \quad (2.9)$$

Notice that we take the derivative $\frac{d}{dx} \left[\prod_{i=0}^n (x - x_i) \right]$ and then evaluate it at x_j . If we were to do the reverse, that is if we introduced x_j first then differentiated, we would be taking the derivative of a number and the term would vanish. Now we can write $\frac{df}{dx} \Big|_{x_i}$ as

$$\frac{df}{dx} \Big|_{x_k} = \sum_{j=0}^n \frac{d\ell_j^n}{dx} \Big|_{x_k} f(x_j) + \frac{1}{(n+1)!} \left(\frac{d^{(n+1)}f(x)}{dx^{(n+1)}} \Big|_{x=\xi} \right) \frac{d}{dx} \left[\prod_{i=0}^n (x - x_i) \right] \Big|_{x_k} \quad (2.10)$$

which says that the derivative $\frac{df(x)}{dx}$ at x_k can be represented by the right-hand-side of [Eq. \(2.10\)](#).

2.2 Two-Point Difference Formulae

We are now in a position to look at some specific difference approximations arising out of the expression presented in [Eq. \(2.10\)](#). Consider

$$\left. \frac{df}{dx} \right|_{x_i} = \left. \frac{d\hat{f}}{dx} \right|_{x_i} + \left. \frac{dE}{dx} \right|_{x_i}. \quad (2.11)$$

Let us take the particular representation of $\left. \frac{d\hat{f}}{dx} \right|_{x_i}$ presented in the following equation:

$$\left. \frac{d\hat{f}}{dx} \right|_{x_i} = \sum_{j=0}^n \left. \frac{d\ell_j^n}{dx} \right|_{x_i} f(x_j). \quad (2.12)$$

This equation should be read as the derivative of the approximation, that is $\left. \frac{d\hat{f}}{dx} \right|_{x_i}$, evaluated at the location x_i . We will look first at the case of using the linear Lagrange polynomial $\ell_j^1(x)$. Expansion of [Eq. \(2.12\)](#) for the two terms associated with having $n = 1$ gives¹

$$\left. \frac{d\hat{f}}{dx} \right|_{x_i} = \left. \frac{d\ell_0}{dx} \right|_{x_i} f(x_0) + \left. \frac{d\ell_1}{dx} \right|_{x_i} f(x_1) \quad (2.13)$$

where the linear Lagrange polynomials are provided by the definitions

$$\ell_0(x) = \frac{x - x_1}{x_0 - x_1} \quad (2.14)$$

and

$$\ell_1(x) = \frac{x - x_0}{x_1 - x_0}. \quad (2.15)$$

The associated derivatives of these two functions are given by

$$\frac{d\ell_0(x)}{dx} = \frac{1}{x_0 - x_1} \quad (2.16)$$

and

$$\frac{d\ell_1(x)}{dx} = \frac{1}{x_1 - x_0} \quad (2.17)$$

Combining [Eqs. \(2.16\)](#) and [\(2.17\)](#) with [\(2.13\)](#) we have

$$\left. \frac{d\hat{f}}{dx} \right|_{x_i} = \frac{1}{x_0 - x_1} f(x_0) + \frac{1}{x_1 - x_0} f(x_1) \quad (2.18)$$

or, rearranging

$$\left. \frac{d\hat{f}}{dx} \right|_{x_i} = \frac{f(x_1) - f(x_0)}{x_1 - x_0} \quad (2.19)$$

or

$$\left. \frac{d\hat{f}}{dx} \right|_{x_i} = \frac{f(x_1) - f(x_0)}{h} \quad x_i = 0, 1. \quad (2.20)$$

We call this a *finite difference approximation* to $\left. \frac{df}{dx} \right|_{x_i}$. Note that we have not specified the value of i , so this approximation can be envisioned as being identified with (located at) either $i = 0$ or $i = 1$. However, it is sometimes important to know where the approximation is perceived to be located, so the following nomenclature is used.

2.2.1 Forward Difference Formula

$$\left. \frac{d\hat{f}}{dx} \right|_{x_0} = \frac{f(x_1) - f(x_0)}{h} \equiv \left. \frac{\Delta f}{h} \right|_{x_0}. \quad (2.21)$$

2.2.2 Backward Difference Formula

$$\left. \frac{d\hat{f}}{dx} \right|_{x_1} = \frac{f(x_1) - f(x_0)}{h} \equiv \left. \frac{\nabla f(x_1)}{h} \right|_{x_1}. \quad (2.22)$$

Note that the symbol $\nabla(\cdot)$ is not the gradient operator $\nabla(\cdot)$ from the vector calculus which is a vector quantity.

2.2.3 Example

Consider differentiation of $\ln(x)$ evaluated at location x_i , that is

$$\left. \frac{df}{dx} \right|_{x_i} = \left. \frac{d\ln(x)}{dx} \right|_{x_i} = \frac{1}{x_i} \quad x \in [1, 2]. \quad (2.23)$$

We begin with the definition of our approximating derivative $\frac{d\hat{f}}{dx}$

$$\frac{df(x)}{dx} = \frac{d\hat{f}(x)}{dx} + \frac{dE(x)}{dx}. \quad (2.24)$$

Neglecting for the moment the error term $\frac{dE}{dx}$, we have from [Eq. \(2.19\)](#)

$$\left. \frac{d\hat{f}}{dx} \right|_{x_i} = \frac{f(x_1) - f(x_0)}{x_1 - x_0}. \quad (2.25)$$

Since, in this example, we have as the values of $f(x_0)$ and $f(x_1)$

$$f(x_0) = f(1) = \ln(1) = 0 \quad (2.26)$$

and

$$f(x_1) = \ln(x_1) = \ln(2) = 0.693 \quad (2.27)$$

we get, through substitution in [Eq. \(2.25\)](#) the following:

$$\left. \frac{d\hat{f}}{dx} \right|_{x_i} = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{0.693 - 0}{2 - 1} = 0.693 \quad i = 0, 1. \quad (2.28)$$

Let us now investigate the error of this approximation. The exact derivative $\frac{df}{dx}$, our approximation $\frac{d\hat{f}}{dx}$, and the error $\frac{dE}{dx}$ are tabulated below:

	$\frac{df}{dx}$		$\frac{dE}{dx}$
x_0	1	0.693	0.307
x_1	0.5	0.693	-0.193

(2.29)

2.2.4 Error of the Approximation

In the above table we have documented the observed error in the derivative approximation when a forward or backward difference is used. Now we will look at how this relates to the error as described [Eq. \(2.9\)](#) on page 34 which we rewrite here for convenience as:

$$\left. \frac{dE}{dx} \right|_{x_j} = \frac{1}{(n+1)!} \left(\frac{d^{(n+1)}f(x)}{dx^{(n+1)}} \right) \Big|_{x=\xi} \frac{d}{dx} \left[\prod_{i=0}^n (x - x_i) \right] \Big|_{x_j}. \quad (2.30)$$

For linear elements this equation becomes

$$\left. \frac{dE}{dx} \right|_{x_j} = \frac{1}{(1+1)!} \left(\frac{d^{(1+1)}f(x)}{dx^{(1+1)}} \right) \Big|_{x=\xi} \frac{d}{dx} \left[\prod_{i=0}^1 (x - x_i) \right] \Big|_{x_j} \quad (2.31)$$

or

$$\left. \frac{dE}{dx} \right|_{x_j} = \frac{1}{2!} \left. \frac{d^2 f(x)}{dx^2} \right|_{x=\xi} \left[(x_j - x_0) + (x_j - x_1) \right]. \quad (2.32)$$

If we evaluate this at x_0 we have

$$\begin{aligned} \left. \frac{dE}{dx} \right|_{x_0} &= \frac{1}{2!} \left. \frac{d^2 f(x)}{dx^2} \right|_{x=\xi} (x_0 - x_1) \\ &= -\frac{1}{2!} \left. \frac{d^2 f(x)}{dx^2} \right|_{x=\xi} h \\ &= O(h) \end{aligned} \quad (2.33)$$

where $O(h)$ is read as 'order h '. Substitution of $f(x) = \ln(x)$ gives

$$\begin{aligned} \left. \frac{dE}{dx} \right|_{x_0} &= -\frac{1}{2!} \left. \frac{d^2 P(x)}{dx^2} \right|_{x=\xi} h \\ &= -\frac{1}{2!} \frac{1}{(-\xi^2)} h \\ &= \frac{1}{2!} \frac{1}{\xi^2} h \end{aligned} \quad (2.34)$$

If we introduce the limits $\xi \in [1, 2]$ and remember that $h = 1$ because $x \in [1, 2]$ and there are only two nodes, one at x_0 and the other at x_1 , we have

$$\left(\frac{1}{2!} \right) \left(\frac{1}{2^2} \right) (1) \leq \left. \frac{dE}{dx} \right|_{x_0} \leq \left(\frac{1}{2!} \right) \left(\frac{1}{1^2} \right) (1) \quad (2.35)$$

or

$$\frac{1}{8} \leq \left. \frac{dE}{dx} \right|_{x_0} \leq \frac{1}{2} \quad (2.36)$$

or

$$0.125 \leq \left. \frac{dE}{dx} \right|_{x_0} \leq 0.5. \quad (2.37)$$

Thus our observed value of the error of 0.307 as shown in the above table lies within the predicted error bounds.

The bounds for $\left. \frac{dE}{dx} \right|_{x_1}$ can be calculated as

$$-0.5 \leq \left. \frac{dE}{dx} \right|_{x_1} \leq -0.125 \quad (2.38)$$

so the computed error at x_1 , namely -0.193 is also within the expected error bounds.

2.3 Two-Point Formulae from Taylor Series

In this section we will present an alternative way of obtaining Two-Point difference formula. We begin with the definition of a Taylor series, viz.:

If $f(x)$ has continuous derivatives of order $k + 1$ in $[a, b]$ then by Taylor's theorem

$$f(x_j + h) = f(x_j) + h \left. \frac{df(x)}{dx} \right|_{x_j} + \cdots + \frac{h^k}{k!} \left. \frac{d^k f(x)}{dx^k} \right|_{x_j} + \frac{h^{(k+1)}}{(k+1)!} \left. \frac{d^{(k+1)} f(x)}{dx^{(k+1)}} \right|_{x_j + \theta_j h} \quad \theta_j \in [0, 1] \quad (2.39)$$

which we denote as a Taylor series.

Let us consider the interval

$$x \in [x_0, x_1]. \quad (2.40)$$

Then from the Taylor series we have (keeping in mind that $h = x_1 - x_0$ and that we are looking backwards and therefore h is negative)

$$\begin{aligned} f(x_1 - h) & \\ &= f(x_1 - (x_1 - x_0)) \\ &= f(x_0) \\ &= f(x_1) - h \left. \frac{df(x)}{dx} \right|_{x_1} + \frac{h^2}{2!} \left. \frac{d^2 f(x)}{dx^2} \right|_{x_1 + \theta h} \quad 0 < \theta < 1 \end{aligned} \quad (2.41)$$

Rearranging the terms in [Eq. \(2.41\)](#) we have:

$$\left. \frac{df(x)}{dx} \right|_{x_1} = \underbrace{\frac{f(x_1) - f(x_0)}{h}}_{\frac{\nabla f(x_1)}{h}} + \underbrace{\frac{h}{2!} \left. \frac{d^2 f(x)}{dx^2} \right|_{x_1 + \theta h}}_{\frac{dE(x_1)}{dx}} \quad \theta \in [0, 1] \quad (2.42)$$

where the labels under the terms indicate their relationship to our earlier development. We now compare [Eq. \(2.42\)](#) with the following expressions developed previously using

the Lagrange polynomial approach

$$\left. \frac{df}{dx} \right|_{x_i} = \left. \frac{d\hat{f}}{dx} \right|_{x_i} + \left. \frac{dE}{dx} \right|_{x_i}. \quad (2.43)$$

Employing [Eq. \(2.25\)](#) we have

$$\left. \frac{d\hat{f}}{dx} \right|_{x_1} = \frac{f(x_1) - f(x_0)}{h} \quad (2.44)$$

and from [Eq. \(2.32\)](#) we obtain

$$\begin{aligned} \left. \frac{dE}{dx} \right|_{x_1} &= \frac{1}{2!} \left. \frac{d^2 f(x)}{dx^2} \right|_{x=\xi} (x_1 - x_0) \\ &= \frac{1}{2!} \left. \frac{d^2 f(x)}{dx^2} \right|_{x=\xi} h. \end{aligned} \quad (2.45)$$

If we combine the above three equations we obtain

$$\left. \frac{df}{dx} \right|_{x_1} = \frac{f(x_1) - f(x_0)}{h} + \frac{1}{2!} \left. \frac{d^2 f(x)}{dx^2} \right|_{x=\xi} h. \quad (2.46)$$

[Equation \(2.46\)](#) is equivalent to [Eq. \(2.42\)](#).

At this point we have shown how to develop an approximation of a first derivative using two equivalent strategies. In the following subsection we demonstrate how one can use such an approximation to solve a simple problem. This is an elementary application of the method of finite differences first introduced by the famous mathematician Euler in 1768.

Example Solution of a First-Order Equation

Consider the physical system illustrated in [Fig. 2.1](#). Fluid is entering the pipe with cross-sectional area A at a rate Q per unit length $[L]$ and has units $\frac{L^2}{T}$ per unit length $[L]$. The velocity entering the pipe is $v = 3$ with units $\frac{L}{T}$. What is the velocity of the pipe along its length? The governing equation, boundary condition, and known flux Q are given as:

$$A \frac{dv}{dx} = Q, \quad x \in [0, 3] \quad (2.47)$$

$$v(0) = 3, \quad (2.48)$$

$$Q = 1. \quad (2.49)$$

The analytical solution is

$$v = Qx + a \quad (2.50)$$

$$v(0) = Q(0) + a \quad (2.51)$$

$$a = 3 \quad (2.52)$$

$$v = x + 3. \quad (2.53)$$

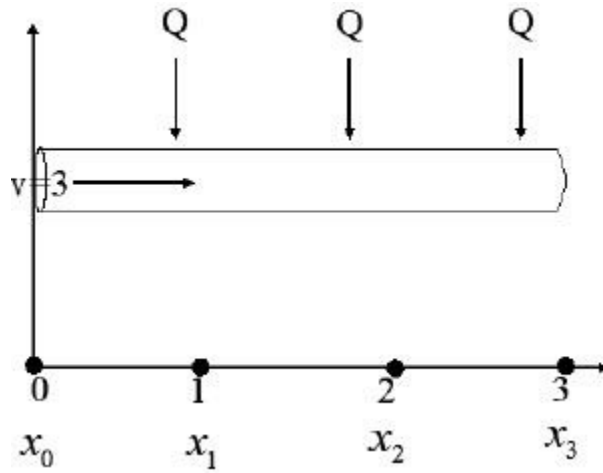


Figure 2.1: Example problem for application of a first derivative finite difference approximation.

Now let us solve this physical problem using our numerical approximation. Employing a first-order finite difference representation for the derivative we get

$$\frac{dv}{dx} = Q \quad (2.54)$$

$$\left. \frac{d\hat{v}}{dx} \right|_{x_0} = \frac{v(x_1) - v(x_0)}{x_1 - x_0} = \frac{v_1 - v_0}{1}. \quad (2.55)$$

Using the same general formula for each nodal location we obtain

$$\left(\frac{d\hat{v}}{dx} - Q \right) \Big|_{x_0} = 0 \Rightarrow \frac{v_1 - v_0}{1} - 1 = 0 \quad (2.56)$$

$$\left(\frac{d\hat{v}}{dx} - Q \right) \Big|_{x_1} = 0 \Rightarrow \frac{v_2 - v_1}{1} - 1 = 0 \quad (2.57)$$

$$\left(\frac{d\hat{v}}{dx} - Q\right)\Big|_{x_2} = 0 \Rightarrow \frac{v_3 - v_2}{1} - 1 = 0 \quad (2.58)$$

In matrix form we have

$$\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \quad (2.59)$$

But we know from the boundary conditions that $v(0) = 3$. Thus we can write

$$\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \quad (2.60)$$

Since we know v_0 to be 3, we can multiply each coefficient in the first column by 3 and, after changing the sign, place the result on the right-hand side of the equation. The first column of the matrix can now be removed. The result follows:

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1+3 \\ 1+0 \\ 1+0 \end{bmatrix}. \quad (2.61)$$

The values of the unknown velocities are now given by

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1+3 \\ 1+0 \\ 1+0 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}. \quad (2.62)$$

The result obtained numerically is the same as we would get if we evaluated [Eq. \(2.53\)](#) for the nodal locations shown in [Fig. 2.1](#). The reason it is exact is that the solution is a straight line and we are representing it with piecewise linear functions. From [Eq. \(2.32\)](#) on page 37 we see that when the solution is linear the error of the approximation vanishes.

2.4 Three-Point Difference Formulae

Now we will extend the above concept to consider a three-node element. We will see that with three nodes we can get a more accurate approximation of the first derivative and an approximation of a second derivative (we could not get an approximation of a second derivative given linear Lagrange polynomials and one element with two nodes).

From the general statement expressing a first derivative in terms of a Lagrangian

expansion, that is [Eq. \(2.2\)](#) on page 33, we have

$$\frac{df(x)}{dx} = \sum_{j=0}^n \frac{d\ell_j^n(x)}{dx} f(x_j) + \frac{dE(x)}{dx}. \quad (2.63)$$

Now consider the second-degree Lagrange polynomial (which involves three nodes and $n = 2$), that is

$$\frac{df(x)}{dx} = \sum_{j=0}^2 \frac{d\ell_j^2(x)}{dx} f(x_j) + \frac{dE(x)}{dx}. \quad (2.64)$$

2.4.1 First-Order Derivative Difference Formulae

From [Eq. \(2.64\)](#) we have for the case of a quadratic Lagrange polynomial

$$\frac{df(x)}{dx} = \frac{d\ell_0(x)}{dx} f(x_0) + \frac{d\ell_1(x)}{dx} f(x_1) + \frac{d\ell_2(x)}{dx} f(x_2) + \frac{dE(x)}{dx} \quad (2.65)$$

which upon substitution of the definitions of the quadratic Lagrange polynomials gives:

$$\begin{aligned} \frac{df(x)}{dx} &= \left(\frac{2x - x_1 - x_2}{2h^2} \right) f(x_0) + \left(\frac{2x - x_0 - x_2}{-h^2} \right) f(x_1) \\ &+ \left(\frac{2x - x_0 - x_1}{2h^2} \right) f(x_2) + \frac{dE(x)}{dx} \end{aligned} \quad (2.66)$$

Let us evaluate this expression at node x_1 . We obtain

$$\begin{aligned} \left. \frac{df}{dx} \right|_{x_1} &= \left(\frac{2x_1 - x_1 - x_2}{2h^2} \right) f(x_0) + \left(\frac{2x_1 - x_0 - x_2}{-h^2} \right) f(x_1) \\ &+ \left(\frac{2x_1 - x_0 - x_1}{2h^2} \right) f(x_2) + \left. \frac{dE}{dx} \right|_{x_1} \\ &= \left(\frac{x_1 - x_2}{2h^2} \right) f(x_0) + \left(\frac{2x_1 - x_0 - x_2}{-h^2} \right) f(x_1) + \left(\frac{x_1 - x_0}{2h^2} \right) f(x_2) + \left. \frac{dE}{dx} \right|_{x_1} \\ &= \left(\frac{-h}{2h^2} \right) f(x_0) + \left(\frac{h - h}{-h^2} \right) f(x_1) + \left(\frac{h}{2h^2} \right) f(x_2) + \left. \frac{dE}{dx} \right|_{x_1} \\ &= \frac{1}{2h} (f(x_2) - f(x_0)) + \left. \frac{dE}{dx} \right|_{x_1}. \end{aligned} \quad (2.67)$$

Notice that this formula is peculiar in that the node at which the derivative is being

defined does not appear in the approximation.

Now let us consider the case for node x_0

$$\begin{aligned}
 \left. \frac{df}{dx} \right|_{x_0} &= \left(\frac{2x_0 - x_1 - x_2}{2h^2} \right) f(x_0) + \left(\frac{2x_0 - x_0 - x_2}{-h^2} \right) f(x_1) \\
 &+ \left(\frac{2x_0 - x_0 - x_1}{2h^2} \right) f(x_2) + \left. \frac{dE}{dx} \right|_{x_0} \\
 &= \left(\frac{x_0 - x_1 + x_0 - x_2}{2h^2} \right) f(x_0) + \left(\frac{x_0 - x_2}{-h^2} \right) f(x_1) + \left(\frac{x_0 - x_1}{2h^2} \right) f(x_2) + \left. \frac{dE}{dx} \right|_{x_0} \\
 &= \left(\frac{-h - 2h}{2h^2} \right) f(x_0) + \left(\frac{-2h}{-h^2} \right) f(x_1) + \frac{-h}{2h^2} f(x_2) + \left. \frac{dE}{dx} \right|_{x_0} \\
 &= \left(\frac{-3}{2h} \right) f(x_0) + \left(\frac{2}{h} \right) f(x_1) + \left(\frac{-1}{2h} \right) f(x_2) + \left. \frac{dE}{dx} \right|_{x_0} \\
 &= \frac{1}{2h} (-3f(x_0) + 4f(x_1) - f(x_2)) + \left. \frac{dE}{dx} \right|_{x_0}
 \end{aligned} \tag{2.68}$$

Note that, based upon what we have provided above, it appears that when using three nodes and quadratic Lagrange polynomials, the derivative approximation has a different form depending on where it is being evaluated. In other words

$$\left. \frac{d\hat{f}}{dx} \right|_{x_0} \neq \left. \frac{d\hat{f}}{dx} \right|_{x_1} \neq \left. \frac{d\hat{f}}{dx} \right|_{x_2}. \tag{2.69}$$

Error of the Approximation

The general definition of the error in representing the first derivative is given by

$$\frac{dE(x)}{dx} = \frac{1}{(n+1)!} \left(\left. \frac{d^{(n+1)} f(x)}{dx^{(n+1)}} \right|_{x=\xi} \right) \frac{d}{dx} [\Pi_{i=0}^n (x - x_i)] \tag{2.70}$$

which, for quadratic Lagrange polynomials, gives for node x_1

(2.71)

$$\frac{dE}{dx} \Big|_{x_1} = \frac{1}{3!} \left(\frac{d^{(3)} f(x)}{dx^{(3)}} \Big|_{x=\xi} \right)$$

$$\left[\underbrace{(x_1 - x_0)(x_1 - x_1)}_0 + \underbrace{(x_1 - x_1)(x_1 - x_2)}_0 + \underbrace{(x_1 - x_0)(x_1 - x_2)}_{-h^2} \right]$$

$$\begin{aligned} \frac{dE}{dx} \Big|_{x_1} &= \frac{-1}{3!} \left(\frac{d^{(3)} f(x)}{dx^{(3)}} \Big|_{x=\xi} \right) h^2 \\ &= O(h^2). \end{aligned} \quad (2.72)$$

Note that this $O(h^2)$ compares with $O(h)$ for the linear polynomial approximation of the first derivative at node x_1 (see [Eq. \(2.33\)](#) on page 37), that is

$$\begin{aligned} \frac{dE}{dx} \Big|_{x_1} &= \frac{1}{(1+1)!} \frac{d^2 f(x)}{dx^2} \Big|_{x=\xi} (x_1 - x_0) \\ &= \frac{1}{2!} \frac{d^2 f(x)}{dx^2} \Big|_{x=\xi} h \\ &= O(h). \end{aligned} \quad (2.73)$$

2.4.2 Second-Order Derivatives

To get an approximation for the second order derivative of $f(x)$, differentiate [Eq. \(2.64\)](#) to give

$$\frac{df^2(x)}{dx^2} = \sum_{j=0}^2 \frac{d^2 \ell_j^2(x)}{dx^2} f(x_j) + \frac{d^2 E(x)}{dx^2}. \quad (2.74)$$

Expanding this equation we obtain

$$\frac{df^2(x)}{dx^2} = \frac{d^2 \ell_0^2(x)}{dx^2} f(x_0) + \frac{d^2 \ell_1^2(x)}{dx^2} f(x_1) + \frac{d^2 \ell_2^2(x)}{dx^2} f(x_2) + \frac{d^2 E(x)}{dx^2}. \quad (2.75)$$

Substitution of the definitions of $\ell_j^2(x)$ yields

(2.76)

$$\begin{aligned} \frac{df^2(x)}{dx^2} &= \frac{d}{dx} \left(\frac{2x - x_1 - x_2}{2h^2} \right) f(x_0) + \frac{d}{dx} \left(\frac{2x - x_0 - x_2}{-h^2} \right) f(x_1) \\ &+ \frac{d}{dx} \left(\frac{2x - x_0 - x_1}{2h^2} \right) f(x_2) + \frac{d^2 E(x)}{dx^2} \end{aligned}$$

Differentiating we obtain

$$\frac{df^2(x)}{dx^2} = \frac{1}{h^2} f(x_0) - \frac{2}{h^2} f(x_1) + \frac{1}{h^2} f(x_2) + \frac{d^2 E(x)}{dx^2}. \quad (2.77)$$

If we neglect the error term we can now write the approximated derivative as

$$\frac{d\hat{f}^2(x)}{dx^2} = \frac{1}{h^2} (f(x_0) - 2f(x_1) + f(x_2)). \quad (2.78)$$

It is interesting to note that, like the Two-Point approximation to the first derivative ([Eq. \(2.19\)](#)), this approximation is the same irrespective of which node you use for the evaluation. In other words, the approximation is the same for each node of the three node series.

Error in approximation

From [Eq. \(2.70\)](#) we have

(2.79)

$$\frac{dE(x)}{dx} = \frac{1}{(n+1)!} \left(\frac{d^{(n+1)} f(x)}{dx^{(n+1)}} \Big|_{x=\xi} \right) \frac{d}{dx} [\Pi_{i=0}^n (x - x_i)]$$

thus, upon differentiation we have

(2.80)

$$\frac{d^2 E(x)}{dx^2} = \frac{d}{dx} \left[\frac{1}{(n+1)!} \left(\frac{d^{(n+1)} f(x)}{dx^{(n+1)}} \Big|_{x=\xi} \right) \frac{d}{dx} [\Pi_{i=0}^n (x - x_i)] \right]$$

or (see page 34)

(2.81)

$$\begin{aligned} \frac{d^2 E(x)}{dx^2} &= \frac{1}{(n+2)!} \left(\frac{d^{(n+2)} f(x)}{dx^{(n+2)}} \Big|_{x=\eta} \right) \frac{d}{dx} [\Pi_{i=0}^n (x-x_i)] \\ &+ \frac{1}{(n+1)!} \left(\frac{d^{(n+1)} f(x)}{dx^{(n+1)}} \Big|_{x=\xi} \right) \frac{d^2}{dx^2} [\Pi_{i=0}^n (x-x_i)]. \end{aligned}$$

If we select $n = 2$ this equation becomes

(2.82)

$$\begin{aligned} \frac{d^2 E(x)}{dx^2} &= \frac{1}{(2+2)!} \left(\frac{d^{(2+2)} f(x)}{dx^{(2+2)}} \Big|_{x=\eta} \right) \frac{d}{dx} [\Pi_{i=0}^2 (x-x_i)] \\ &+ \frac{1}{(2+1)!} \left(\frac{d^{(2+1)} f(x)}{dx^{(2+1)}} \Big|_{x=\xi} \right) \frac{d^2}{dx^2} [\Pi_{i=0}^2 (x-x_i)] \end{aligned}$$

which simplifies to

(2.83)

$$\begin{aligned} \frac{d^2 E(x)}{dx^2} &= \frac{1}{4!} \left(\frac{d^4 f(x)}{dx^4} \Big|_{x=\eta} \right) \frac{d}{dx} (x-x_0)(x-x_1)(x-x_2) \\ &+ \frac{1}{3!} \left(\frac{d^3 f(x)}{dx^3} \Big|_{x=\xi} \right) \frac{d^2}{dx^2} [(x-x_0)(x-x_1)(x-x_2)]. \end{aligned}$$

At the node x_1 this expression becomes

(2.84)

$$\begin{aligned} \frac{d^2 E}{dx^2} \Big|_{x_1} &= \frac{1}{4!} \left(\frac{d^4 f(x)}{dx^4} \Big|_{x=\eta} \right) \\ &\times \left[(x_1-x_0) \underbrace{(x_1-x_1)}_0 + \underbrace{(x_1-x_1)}_0 (x_1-x_2) + (x_1-x_0)(x_1-x_2) \right] \\ &+ \frac{1}{3!} \left(\frac{d^3 f(x)}{dx^3} \Big|_{x=\xi} \right) \left[2(x_1-x_0) + \underbrace{2(x_1-x_1)}_0 + 2(x_1-x_2) \right] \end{aligned}$$

which simplifies to

$$\frac{d^2 E}{dx^2} \Big|_{x_1} = \frac{1}{4!} \left(\frac{d^4 f(x)}{dx^4} \Big|_{x=\eta} \right) [(x_1 - x_0)(x_1 - x_2)] \quad (2.85)$$

$$+ \frac{1}{3!} \left(\frac{d^3 f(x)}{dx^3} \Big|_{x=\xi} \right) [2(x_1 - x_0) + 2(x_1 - x_2)] \quad (2.86)$$

and finally becomes

$$\frac{d^2 E}{dx^2} \Big|_{x_1} = \left\{ \frac{1}{4!} \left(\frac{d^4 f(\eta)}{dx^4} \Big|_{x=\eta} \right) (-h^2) + \frac{2}{3!} \left(\frac{d^3 f(x)}{dx^3} \Big|_{x=\xi} \right) (0) \right\} = O(h^2). \quad (2.87)$$

Example of the Solution of a Second-Order Equation

In this example we will show how to solve a Second-Order equation using a finite difference approach. In this spirit, consider the transport of a contaminant c being transported by a fluid with velocity v and experiencing dispersion with dispersion coefficient D . Let the interval of interest be $x = [0, 1]$. The governing equation is

$$D \frac{d^2 c}{dx^2} - v \frac{dc}{dx} = 0. \quad (2.88)$$

Assume that the concentration at $x = 0$ is 1 and at $x = 1$ is 0. Further assume that we are going to use three nodes located at $x_0 = 0$, $x_1 = 0.5$, and $x_2 = 1.0$. Because the first and last nodes in the three node sequence are known, we only need to solve for the center node. This requires one equation in one unknown, namely the unknown value at $x = 0.5$ which is located at node 1 in the sequence.

Using centered finite difference approximations for each term in [Eq. \(2.88\)](#), we obtain for the equation at node i

$$D \frac{c_{i+1} - 2c_i + c_{i-1}}{h^2} - v \frac{c_{i+1} - c_{i-1}}{2h} = 0. \quad (2.89)$$

For the special case of $n = 1$, this expression becomes

$$D \frac{c_2 - 2c_1 + c_0}{h^2} - v \frac{c_2 - c_0}{2h} = 0. \quad (2.90)$$

We can write the matrix equation as

$$\begin{bmatrix} 1 & 0 & 0 \\ \left(\frac{D}{h^2} + \frac{v}{2h}\right) & \left(\frac{-2D}{h^2}\right) & \left(\frac{D}{h^2} - \frac{v}{2h}\right) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad (2.91)$$

where the first and last rows exist only to incorporate the boundary conditions. More specifically, if you multiply the column of c_i values by the first row of the matrix you obtain $c_0 = 1$, which is the boundary condition at $x = 0$. A similar argument can be made for the representing the boundary condition at $x = 1$ by using the last row of the matrix equation.

Now assume for convenience that $D = v = 1$ and substitute into [Eq. \(2.91\)](#) which now becomes

$$\begin{bmatrix} 1 & 0 & 0 \\ \left(\frac{1}{1/4} + \frac{1}{1}\right) & \left(\frac{(-2)(1)}{1/4}\right) & \left(\frac{1}{1/4} - \frac{1}{1}\right) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \quad (2.92)$$

Focusing now on the equation defined by the second row of [Eq. \(2.92\)](#) we obtain through matrix multiplication

$$\left(-\frac{2}{1/4}\right) c_1 = -\left(\frac{1}{1/4} + \frac{1}{1}\right) \underbrace{c_0}_1 - \left(\frac{1}{1/4} - \frac{1}{1}\right) \underbrace{c_2}_0 \quad (2.93)$$

which can now be solved for the value of c at $x = 0.5$, that is c_1 and the result is

$$c_1 = \left(\frac{1}{8}\right) \left(\frac{5}{1}\right) (1) + \left(\frac{3}{8}\right) (0) = \frac{5}{8}. \quad (2.94)$$

Now, just out of curiosity, let us see what happens as we change the coefficient D . Writing the solution for c_1 symbolically using [Eq. \(2.91\)](#) we have (noting that $c_2 = 0$ via the boundary condition at $x = 1$)

$$c_1 = \left(\frac{1/4}{2D}\right) \left(\frac{D}{1/4} + v\right) c_0 \quad (2.95)$$

or

$$c_1 = \frac{1}{2} + \frac{v}{8D}. \quad (2.96)$$

Notice that as $D \rightarrow \infty$, $c_1 \rightarrow 0.5$ and as $D \rightarrow 0$, $c_1 \rightarrow \infty$. The reason for $c_1 \rightarrow \infty$ is that the

right hand boundary condition is inappropriate and causes a conflict. The solution approaches 1 everywhere when $D = 0$ but this is in conflict with $c_2 = 0$ which is the right-hand boundary condition. As we noted in the previous example, having two boundary conditions specified for a first-order equation (remember $D = 0$ is assumed) is incorrect. The solution for c_1 reflects that fact. In practice [Eq. \(2.88\)](#) is difficult to solve when the ratio of v/D becomes large for a variety of reasons as we will see later.

2.5 Chapter Summary

The polynomials introduced in the preceding chapter are used to create discrete approximations to continuous derivatives. The error in these approximations is introduced and discussed. The resulting discrete approximations to first and second order derivatives are used to represent differential equations and the result is a set of algebraic equations which, when solved, generate an approximate solution to the differential equation. By using different strategies for approximating the continuous derivatives, the discrete equations can have different numbers of nodal locations in their approximations and concomitant differences in the order of their errors. As an illustration of the use of the finite difference methods that arise using this strategy, an example differential equation is solved employing this approach.

2.6 Problems

1. Consider the equation

$$\frac{d^2u}{dx^2} + u = 0 \quad x \in [0, 1] \quad (2.97)$$

$$u(0) = 1 \quad (2.98)$$

$$\left. \frac{du}{dx} \right|_{x=1} = 0. \quad (2.99)$$

Using three nodes located at $x = 0$, $x = 0.5$, and $x = 1.0$, find the value of $u(0.5)$. Note that there is more than one option for the representation of the variable u_1 located at $x = 0.5$.

2. Consider the following equation and associated boundary condition

$$\frac{dv}{dx} + v = 0 \quad (2.100)$$

$$v(0) = 3. \quad (2.101)$$

The analytic solution is

$$v = 3e^{-x} \quad (2.102)$$

so

$$v(1) = \frac{3}{e^1} = \frac{3}{2.718} = 1.1 \quad (2.103)$$

$$v(2) = \frac{3}{e^2} = \frac{3}{7.39} = 0.405. \quad (2.104)$$

Given the mesh shown in [Fig. 2.2](#), calculate $v(1)$ using a backward difference scheme and compare your results to the exact values.

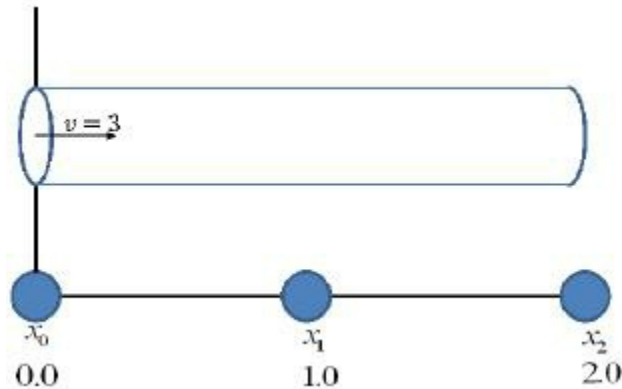


Figure 2.2 Nodal arrangement for mesh for flow in a pipe.

3. Given [Eqs. \(2.105\)-\(2.107\)](#) and a mesh made up of three nodes located at $x_0 = 0$, $x_1 = 0.5$, and $x_2 = 1.0$, use a second order correct finite difference scheme to determine the value of c at $x = 0.5$. Assume $D = 2$ and $v = 1$:

$$D \frac{d^2 u}{dx^2} + u = 0 \quad x \in [0, 1] \quad (2.105)$$

$$u(0) = 1 \quad (2.106)$$

$$u(1) = 1. \quad (2.107)$$

4. Given the error expression

$$\left. \frac{dE}{dx} \right|_{x_j} = \frac{1}{(n+1)!} \left(\frac{d^{(n+1)} f(\xi)}{dx^{(n+1)}} \right) \frac{d}{dx} \left[\prod_{i=0}^n (x - x_i) \right] \Big|_{x_j} \quad (2.108)$$

determine the error bounds for $\frac{d}{dx} (\sin(\theta))$ at $\pi/2$ and π , using the mesh shown in [Fig. 2.3](#) and an approximation using linear Lagrange polynomials ($n = 1$). Are they equivalent and should they be?

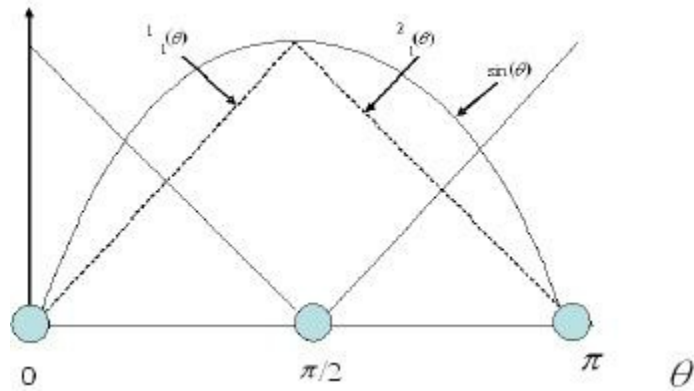


Figure 2.3 Linear Lagrange polynomials at three nodes and the function $\sin(\theta)$

5. Consider the physical system illustrated in [Fig. 2.4](#). Fluid is entering the pipe at a rate $Q = x$ per unit length. The velocity entering the pipe is $v = 3$. What is the velocity of the pipe along its length? The governing equation, boundary condition, and flux are given as:

$$\frac{dv}{dx} = Q \quad (2.109)$$

$$v(0) = 3 \quad (2.110)$$

$$Q = x. \quad (2.111)$$

Compare your solution to the analytical one, namely

$$v = \frac{x^2}{2} - 3. \quad (2.112)$$

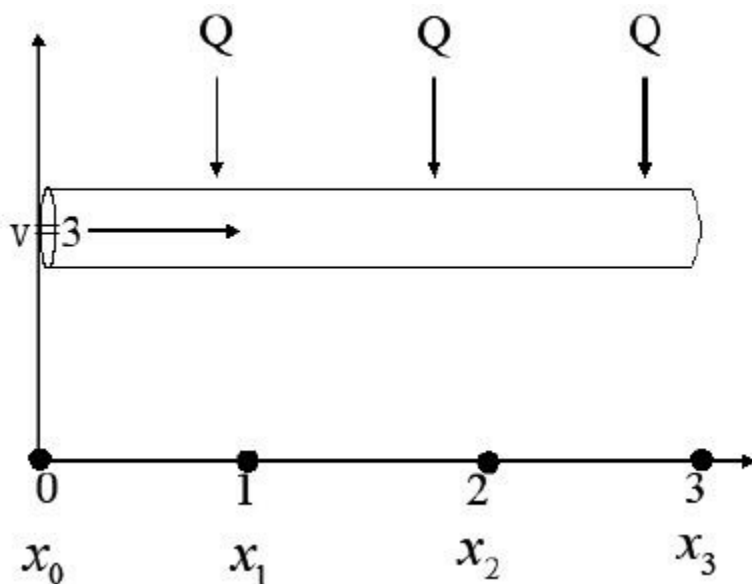


Figure 2.4 Problem formulation where fluid enters a pipe at x_0 and there is a source

$Q = x$ along the length of the pipe. The input velocity is $v = 3$.

6. In [Fig. 2.5](#) is displayed the water-level response to a pumping well drawing from a groundwater reservoir of thickness l . The well produces a drop in water level, h . The water level drops most at the well and less away from the well. It is assumed that at some distance r_e the water level is not disturbed and remains at height H_e . The equation describing the drop in water level is given by

$$Q = K \frac{dh}{dr} 2\pi r l \quad (2.113)$$

where Q is the well discharge. The solution to this equation is

$$H_e - h = \frac{Q}{2\pi K l} \ln \frac{r_e}{r}. \quad (2.114)$$

Assume that the thickness l is 10, that $r_e = 100$, $H_e = 10$, $Q = 1$, and $K = 0.1$. Calculate the value of h at $r = 50$ using a finite difference scheme based on linear Lagrange polynomials and a sequence of element sizes beginning with 50 (3 nodes), then 25 (5 nodes), then 12.5 (9 nodes) and finally 6.25 (17 nodes). Compute the magnitude of the error $h - \hat{h}$, where \hat{h} is the approximate solution at $r = 50$. Plot the logarithm of the error at that point for each scheme on the vertical axis and the logarithm of the element size on the horizontal axis. Calculate the slope of this line. This is an estimate of the order of the global error. Given the method that you have used to approximate the derivative, how does the local error compare to the global error?

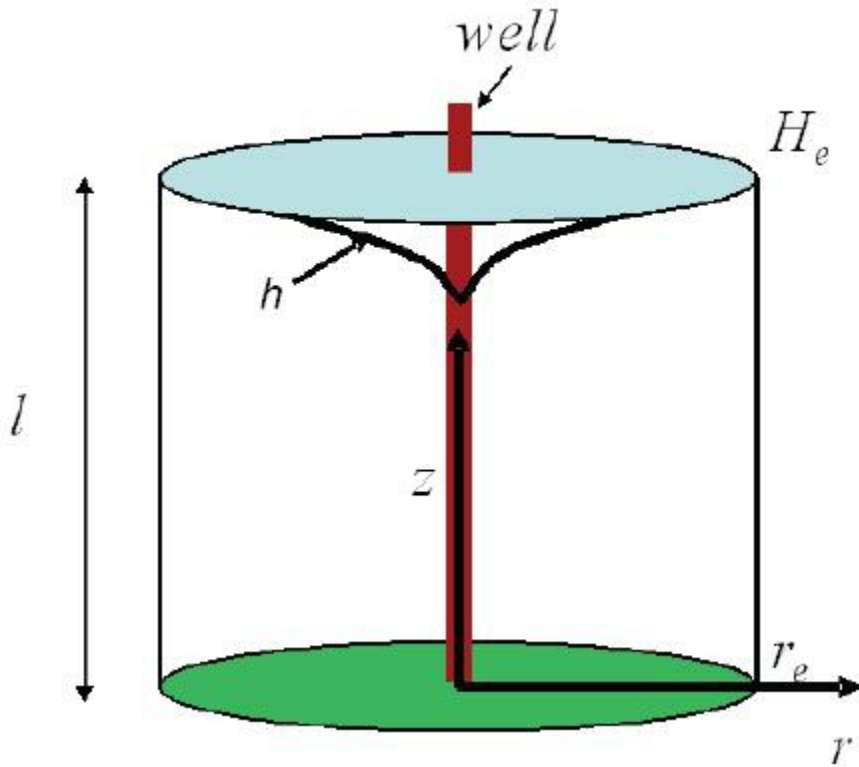


Figure 2.5 Diagrammatic representation for problem involving well discharge.

7. The convection diffusion equation is one of the most often encountered in science and engineering. Solve the following problem that uses this equation [Eq. \(2.116\)](#). The problem is defined for the interval $X \in [0, 1]$ by

$$-\frac{d^2\phi}{dX^2} + P_e \frac{d\phi}{dX} = 0 \quad X \in [0, 1] \quad (2.115)$$

$$\phi(0) = 1 \quad (2.116)$$

$$\phi(1) = 0 \quad (2.117)$$

where $X = \frac{x}{L}$, ϕ is a measure of concentration and the Peclet number

$P_e = \frac{\rho u L}{\Gamma}$, ρ is fluid density, u is fluid velocity, Γ is the diffusivity, and L is a characteristic length. The goal of this exercise is to see under what values of P_e the numerical solution of this equation becomes problematic. Notice that by increasing P_e you are effectively increasing the importance (influence) of the convective term at the expense of the diffusive term.

1. The analytical solution to this problem is

$$\phi = \phi(0) + (\phi(1) - \phi(0)) \frac{\exp(P_e X) - 1}{\exp(P_e) - 1}. \quad (2.118)$$

Proceed as follows:

(a) Replace the value of P_e with the numerical equivalent $P_h = \frac{\rho u h}{\Gamma}$ that uses the spacing h as the value of L . Notice that the h in the P_h term is obtained by multiplying your difference equation through by h , so the h does not

(b) Use two different numerical approximations.

- i. the first is a central difference approximation for both the first and second derivatives.
- ii. the second is a central difference approximation for the second derivative and a backward difference for the first derivative.

(c) Use two different values of P_e .

- i. one case is with $P_h = 1$.
- ii. one case is with $P_h = 3$.

Since you are using only one value of h in all your simulations, by changing P_h , you are really changing either u or Γ .

(d) Using a mesh with 17 nodes, calculate the value of ϕ_i at each node with the boundary conditions indicated. Do this for the two values of P_h and each numerical approximation.

(e) Plot these four solutions along with their analytical equivalent (using $P_e = P_h$). Theory suggests that with a value of P_e greater than 2, you should have problems for the case where the central difference approximation is used for the first derivative and all the other cases should be fine. Also, theoretically, the solution for $P_h = 1$ should be more accurate when the central difference approximation for the first derivative is used (why should this be true?).

8. Given the $\sin(\theta)$ function shown in the [Fig. 2.6](#) and linear polynomials determine the value of the error range of $\frac{dE(\theta)}{d\theta}$ for the derivative of $\sin(\theta)$ at $\theta = 0$. Compare this error to the actual error you get when you attempt to approximate $\left. \frac{d \sin(\theta)}{d\theta} \right|_{\theta=0}$. Is the actual error within the range determined from $\frac{dE(\theta)}{d\theta}$?

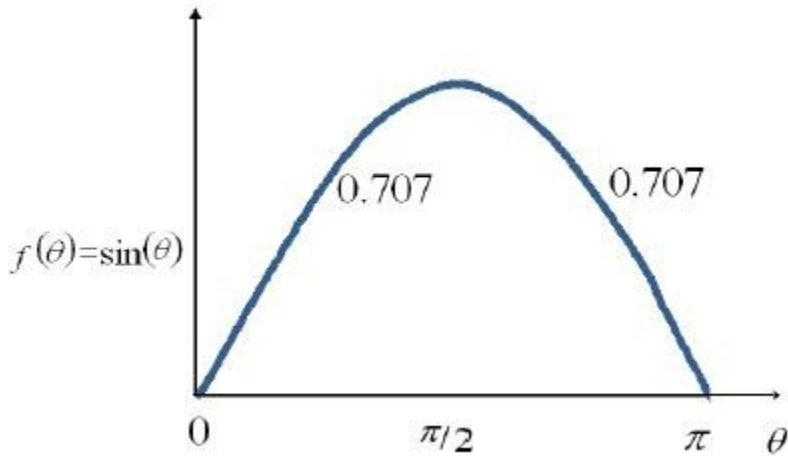


Figure 2.6 Sin function from $0 \leq \theta \leq \pi$.

9. Use quadratic (fourth degree) Lagrange polynomials to approximate $\sin \theta$ using nodes at $0, \pi/2, \pi, (3/2)\pi$ and π . (see Fig. 2.7). Determine the value of $\sin \theta$ at $(5/4)\pi$. Compare the estimate to the value obtained using the quadratic Lagrange polynomial. Is it better; why?

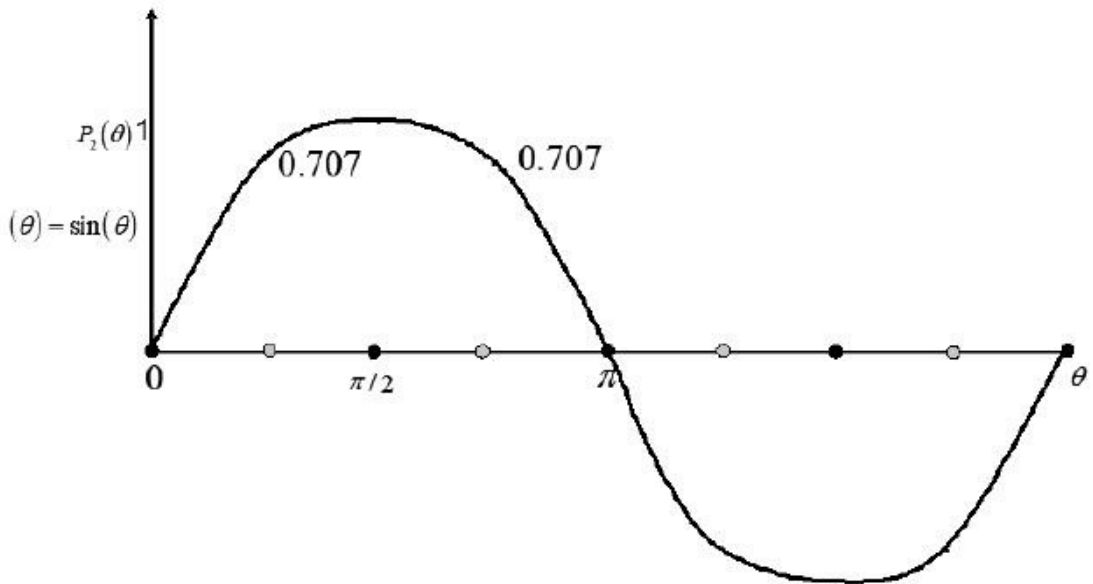


Figure 2.7 Sin (θ) curve with measured points indicated by black dots and interpolated values indicated by grey dots.

- o. Given the equation

$$\frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2} = 0 \quad (2.119)$$

determine whether the expression

$$\frac{u_{i+1} - u_i}{h} + \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} = 0 \quad (2.120)$$

is consistent.

11. Show why you cannot use linear Lagrange polynomial functions and two nodes to approximate a term of the form $\frac{\partial^2 u}{\partial x^2}$.
12. Reconsider the problem presented earlier in this chapter, however with a slight modification. Again we let the velocity entering the pipe be $v = 3$. The governing equation, boundary condition, and known flux Q are given as earlier, but A is now defined as a function of x (see [Fig. 2.8](#)). The problem now is:

$$A \frac{dv}{dx} = Q \quad x \in [0, 3] \quad (2.121)$$

$$v(0) = 3, \quad (2.122)$$

$$Q = 1, \quad (2.123)$$

$$A = 1 - x/6. \quad (2.124)$$

The analytical solution is

$$v = -6 \ln(1 - x/6) + 3. \quad (2.125)$$

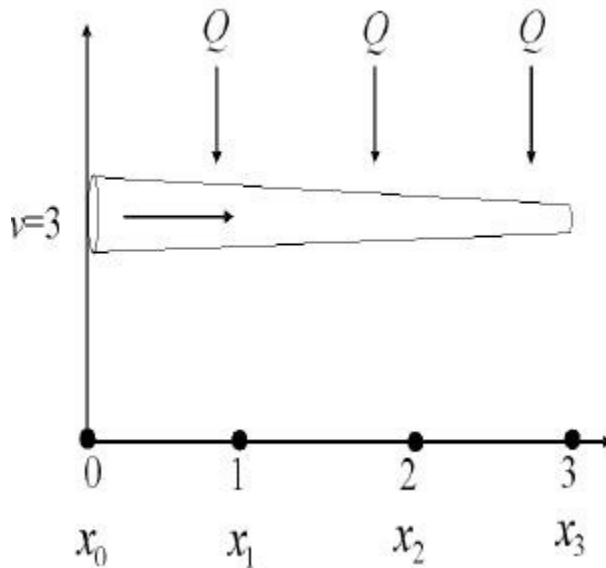


Figure 2.8 Physical system for problem with linearly decreasing tube diameter.

The wrinkle here is how to handle the coefficient A . One approach is to evaluate the coefficient at each node, that is, $A = 1 - x_i/6$.

Write a computer code to calculate the solution at each of the nodes and evaluate the error at each node. Plot the error as a function of the distance along x . Repeat the calculation and analysis using seven equally space nodes. Plot the error at each of the

nodes for which you have values from each calculation. As a measure of error, take the largest error among all the nodes for a particular discretization as the measure of error for that discretization.

Bibliography

Burden, R.L., and J.D. Faires (2000), *Numerical Analysis*, (7th ed.), Brooks/Cole, Pacific Grove CA.

Chang, S. (2000) *Differential Quadrature and Its Application in Engineering: Engineering Applications*, Springer, London.

Lyness, J. N. and C. B. Moler (1967). “Numerical differentiation of analytic functions”. *SIAM J. Numer. Anal.* 4: 202–210.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, (1992) “Numerical derivatives.” Section 5.7 in *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, Cambridge, England, pp. 180-184.

Ralston, A.,(1965) *A First Course in Numerical Analysis*, McGraw-Hill, New York.

Zhang, Y. (2009) *Advanced Differential Quadrature Methods*, Chapman and Hall/CRC Press, Boca Raton.

Note

¹ We will drop the superscript designating the degree of the polynomial when it is clear by the context the degree that is intended.

Chapter 3

Numerical Integration

As we move into a discussion of numerical methods that require the evaluation of integrals, we discover that there is a need to understand some of the elementary aspects of numerical integration. There is a vast literature on this topic and this chapter addresses only those aspects we need later to assist in understanding the finite element and finite volume methods.

3.1 Newton-Cotes Quadrature Formulas

In this subsection we use the techniques we have developed to date to examine the family of integration methods known as Newton-Cotes Quadrature formulas. We begin with reviewing some of the material presented earlier.

3.1.1 Lagrange Interpolation

From the general formula for approximating a function using Lagrange polynomials we have:

$$f(x) = \sum_{j=0}^n \ell_j(x) f(x_j) + \frac{1}{(n+1)!} \prod_{i=0}^n (x - x_i) \frac{d^{n+1} f(\xi)}{dx^{n+1}} \quad \xi \in [a, b]. \quad (3.1)$$

Let us now integrate [Eq. \(3.1\)](#) over an interval from a to b . For n even one can show that (Ralston, 1965) [9]

$$\begin{aligned} \int_a^b f(x) dx &= \sum_{j=0}^n \int_a^b \ell_j^n(x) f(x_j) dx \\ &+ \frac{1}{(n+2)!} \frac{d^{n+2} f(\eta)}{dx^{n+2}} \int_a^b x \prod_{i=0}^n (x - x_i) dx \end{aligned} \quad (3.2)$$

and for n odd

$$\begin{aligned} \int_a^b f(x) dx &= \sum_{j=0}^n \int_a^b \ell_j^n(x) f(x_j) dx \\ &+ \frac{1}{(n+1)!} \frac{d^{n+1} f(\eta)}{dx^{n+1}} \int_a^b \prod_{i=0}^n (x - x_i) dx. \end{aligned} \quad (3.3)$$

We will now see that from these two expressions we can formulate the commonly encountered quadrature formulae for numerical integration.

3.1.2 Trapezoidal Rule

Let us start with the simplest quadrature formula, namely the case when $n = 1$. We consider the interval $x \in [0, h]$ with, as earlier, $h \equiv x_i - x_{i-1}$, $x = ih$ and x_i are nodal locations along the x axis. Since $n = 1$ is odd, we use [Eq. \(3.3\)](#), that is

$$\int_0^h f(x) dx = \int_0^h [\ell_0(x) f(x_0) + \ell_1(x) f(x_1)] dx + \frac{1}{(2)!} \frac{d^2 f(\eta)}{dx^2} \int_0^h (x - x_0)(x - x_1) dx. \quad (3.4)$$

The next step is to substitute our definition of the linear Lagrange polynomial and recognizing that $f(x_i)$ is a number (the value of $f(x)$ at the specific locations x_i) and can be taken out of the integral, we obtain,

$$\int_0^h f(x) dx = \left(\int_0^h \frac{h-x}{h} dx \right) f(x_0) + \left(\int_0^h \frac{x}{h} dx \right) f(x_1) + \frac{1}{(2)!} \frac{d^2 f(\eta)}{dx^2} \int_0^h (x-0)(x-h) dx. \quad (3.5)$$

Now we perform the integrations appearing on the right-hand side of [Eq. \(3.5\)](#) and obtain

$$\left(\int_0^h \frac{h-x}{h} dx \right) dx = \left(x - \frac{x^2}{2h} \right) \Big|_0^h = \left(h - \frac{h^2}{2h} \right) = \frac{h}{2} \quad (3.6)$$

$$\left(\int_0^h \frac{x}{h} dx \right) = \left(\frac{x^2}{2h} \right) \Big|_0^h = \frac{h}{2} \quad (3.7)$$

$$\begin{aligned} & \frac{1}{(2)!} \frac{d^2 f(\eta)}{dx^2} \int_0^h (x-0)(x-h) dx \\ &= \frac{1}{(2)!} \frac{d^2 f(\eta)}{dx^2} \int_0^h (x^2 - hx) dx \\ &= \frac{1}{(2)!} \frac{d^2 f(\eta)}{dx^2} \left(\frac{x^3}{3} - \frac{hx^2}{2} \right) \Big|_0^h \\ &= O(h^3). \end{aligned} \quad (3.8)$$

Substitution of these integrals into [3.5](#) provides us with the final form for the integral $\int_0^h f(x) dx$

$$\begin{aligned} & \int_0^h f(x) dx \\ &= \frac{h}{2} f(0) + \frac{h}{2} f(h) + O(h^3). \end{aligned} \quad (3.9)$$

Equation [3.9](#) tells us that we can obtain an approximation to the integral $\int_0^h f(x) dx$ as the weighted sum of the values at the end points of the interval, namely $f(0)$ and $f(h)$ with weighting coefficients $\frac{h}{2}$. An alternative interpretation is that the integral is approximated by the arithmetic sum of the values at the ends of the interval, multiplied by the interval value h . Now let's extend the concept to the case of three nodes.

3.1.3 Simpson's Rule

In the case of three nodes, $n = 2$. If we assume h to be a constant, the interval over which we are doing our integration becomes $x \in [0, 2h]$. Since n is even, we use [Eq. \(3.2\)](#), that is

$$\begin{aligned}
& \int_0^{2h} f(x) dx && (3.10) \\
&= \int_0^{2h} [\ell_0^2(x) f(x_0) + \ell_1^2(x) f(x_1) + \ell_2^2(x) f(x_2)] dx \\
&+ \frac{1}{(2+2)!} \frac{d^4 f(\eta)}{dx^4} \int_0^{2h} x(x-x_0)(x-x_1)(x-x_2) dx.
\end{aligned}$$

To get a flavor for the process involved in evaluating the terms in [Eq.\(3.10\)](#), consider the first term in the the first integrand appearing on the right-hand side of this equation, viz.

$$\begin{aligned}
& \int_0^{2h} [\ell_0^2(x) f(x_0)] dx && (3.11) \\
&= \left(\int_0^{2h} [\ell_0^2(x)] dx \right) f(x_0) \\
&= \left(\int_0^{2h} \left[\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} \right] dx \right) f(x_0) \\
&= \left(\int_0^{2h} \left[\frac{x^2 - x_1x - x_2x + x_1x_2}{(-h)(-2h)} \right] dx \right) f(x_0) \\
&= \left(\frac{1}{2h^2} \left[\frac{x^3}{3} - x_1 \frac{x^2}{2} - x_2 \frac{x^2}{2} + x_1x_2x \right]_0^{2h} \right) f(x_0) \\
&= \left(\frac{1}{2h^2} \left[\frac{16h^3}{6} - \frac{12h^3}{6} - \frac{24h^3}{6} + \frac{24h^3}{6} \right] \right) f(x_0) \\
&= \frac{1}{2h^2} \left(\frac{2}{3}h^3 \right) = \frac{h}{3} f(x_0).
\end{aligned}$$

The integral in the error term becomes

$$\begin{aligned}
& \int_0^{2h} x^2 (x-h)(x-2h) dx & (3.12) \\
&= \int_0^{2h} x^2 (x^2 - 3hx + 2h^2) dx \\
&= \int_0^{2h} (x^4 - 3hx^3 + 2x^2h^2) dx \\
&= \left(\frac{x^5}{5} - \frac{3}{4}hx^4 + \frac{2h^2}{3}x^3 \right) \Big|_0^{2h} \\
&= \frac{32h^5}{5} - \frac{48h^5}{4} + \frac{16h^5}{3} \\
&= -\frac{4}{15}h^5.
\end{aligned}$$

Combination of these integral terms yields

$$\begin{aligned}
\int_0^{2h} f(x) dx &= \frac{h}{3} [f(x_0) + 4f(x_1) + f(x_2)] + \underbrace{\frac{1}{4!} \left(\frac{-4}{15} \right)}_{-\frac{1}{90}} h^5 \frac{d^4 f(\eta)}{dx^4} & (3.13) \\
&= \frac{h}{3} [f(x_0) + 4f(x_1) + f(x_2)] + O(h^5).
\end{aligned}$$

Equation 3.13 states that we can obtain a value for the integral $\int_0^{2h} f(x) dx$ using three nodal values of $f(x)$, that is $f(x_0)$, $f(x_1)$, and $f(x_2)$ that is $O(h^5)$. Thus we observe that the use of additional nodes increases the accuracy of the numerical approximation; of course it also requires more computational effort.

3.1.4 General Form

The general form of the above is

$$\begin{aligned}
\int_{x_0}^{x_n} f(x) dx &= \sum_{j=0}^n H_j f(x_j) + E & (3.14) \\
H_j &= hAW_j \\
E &= E_1 h^{(k+1)} \frac{d^k f(\eta)}{dx^k}
\end{aligned}$$

where

$$k = n + 1 \quad \text{if } n \text{ is odd} \quad (3.15)$$

$$k = n + 2 \quad \text{if } n \text{ is even.} \quad (3.16)$$

n	A	W_0	W_1	W_2	W_3	E_1
1	1/2	1	1			-1/12
2	1/3	1	4	1		-1/90
3	3/8	1	3	3	1	-3/80

(3.17)

3.1.5 Example using Simpson's Rule

As an example, let us perform the integration as seen in [Fig. 3.1](#). The exact value for this integral is

$$\int_1^3 (x^2 + x + 1) dx = 14.666. \quad (3.18)$$

Let us use a second degree polynomial such that $n = 2$. Then from [Eq. \(3.16\)](#) we have $k = 2 + 2$. We see from [Eq. \(3.14\)](#) that if we use a second degree polynomial, there should be no error since the fourth derivative of our integrand is zero; let's see if this holds for our example. From Simpson's rule

$$\begin{aligned} & \int_1^3 (x^2 + x + 1) dx && (3.19) \\ &= h \left[\frac{1}{3}f(x_0) + \frac{4}{3}f(x_1) + \frac{1}{3}f(x_2) \right] \\ &= 1 \left[\left(\frac{1}{3}\right) 3 + \left(\frac{4}{3}\right) 7 + \left(\frac{1}{3}\right) 13 \right] \\ &= 14.666. \end{aligned}$$

Thus, as hypothesized, the integration is exact.

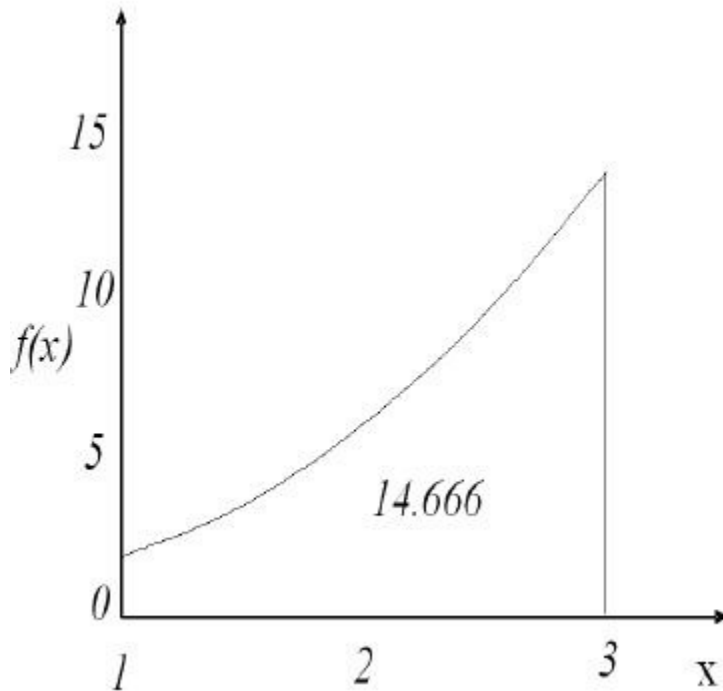


Figure 3.1: Function to be integrated and integral value.

3.1.6 Gauss Legendre Quadrature

In this subsection we look at a quite different method of integration that we will need later in our discussion of isoparametric finite element (see Section 8.3 on page 211) and finite volume methods (see Section 9.2.5 on page 256). This method, called Gauss Legendre Quadrature, can be presented at various levels of sophistication. The following approach is consistent with our needs.

Assume the integral of interest can be approximated as

$$\int_a^b f(x) dx \approx \sum_{i=1}^n c_i f(x_i), \quad (3.20)$$

where both the coefficients c_i and the locations x_i are unknown. We will now determine, using two terms of this series, the values of $c_1, c_2, x_1,$ and x_2 that exactly integrates an equation of the form

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3. \quad (3.21)$$

Substitution of [Eq. \(3.21\)](#) into [Eq. \(3.20\)](#) gives us

$$\int_a^b f(x) dx = \int_a^b (a_0 + a_1x + a_2x^2 + a_3x^3) dx. \quad (3.22)$$

The exact solution to this equation is

$$\int_a^b f(x) dx = a_0(b-a) + a_1 \left(\frac{b^2 - a^2}{2} \right) + a_2 \left(\frac{b^3 - a^3}{3} \right) + a_3 \left(\frac{b^4 - a^4}{4} \right). \quad (3.23)$$

From [Eq. \(3.20\)](#) we have, for $n = 2$

$$\int_a^b f(x) dx \approx \sum_{i=1}^2 c_i f(x_i) = c_1 f(x_1) + c_2 f(x_2). \quad (3.24)$$

Using now the definition of $f(x_i)$ from [Eq. \(3.21\)](#), we obtain for the two values of x_i

$$c_1 f(x_1) + c_2 f(x_2) = c_1 (a_0 + a_1 x_1 + a_2 x_1^2 + a_3 x_1^3) + c_2 (a_0 + a_1 x_2 + a_2 x_2^2 + a_3 x_2^3).$$

Rearrangement of the terms in this expression yields

$$c_1 f(x_1) + c_2 f(x_2) = a_0 (c_1 + c_2) + a_1 (c_1 x_1 + c_2 x_2) + a_2 (c_1 x_1^2 + c_2 x_2^2) + a_3 (c_1 x_1^3 + c_2 x_2^3).$$

We now have two equations that both must be satisfied, these are [Eq. \(3.23\)](#) and [Eq. \(3.24\)](#). For this to be true

$$\begin{aligned} c_1 + c_2 &= b - a & (3.25) \\ c_1 x_1 + c_2 x_2 &= \frac{b^2 - a^2}{2} \\ c_1 x_1^2 + c_2 x_2^2 &= \frac{b^3 - a^3}{3} \\ c_1 x_1^3 + c_2 x_2^3 &= \frac{b^4 - a^4}{4}. \end{aligned}$$

[Equation \(3.25\)](#) represent four non-linear equations in four unknowns. Solution of these equations yields

$$\begin{aligned} c_1 &= \frac{b-a}{2} & (3.26) \\ c_2 &= \frac{b-a}{2} \\ x_1 &= \frac{b-a}{2} \left(-\frac{1}{\sqrt{3}} \right) + \frac{b+a}{2} \\ x_2 &= \frac{b-a}{2} \left(\frac{1}{\sqrt{3}} \right) + \frac{b+a}{2}. \end{aligned}$$

Substitution of these values into [Eq. \(3.24\)](#) yields the integration formula we seek.

In application, the general form of the approximation is modified to accommodate integration limits defined by $a = -1$, $b = +1$. This is the form we will use later in the book.

In this case, the coefficients in [Eq. \(3.26\)](#) become

$$\begin{aligned} c_1 &= \frac{b-a}{2} = 1 \\ c_2 &= \frac{b-a}{2} = 1 \\ x_1 &= -\frac{1}{\sqrt{3}} \\ x_2 &= \frac{1}{\sqrt{3}}. \end{aligned} \tag{3.27}$$

In [Table 3.29](#), are tabulated the values of c_i and x_i for a $n = 1, 2, 3, 4$. In our work we will need to calculate multidimensional integrals. The formula for two dimensions is a straight-forward extension of the one-dimensional case and is for

$$\int_{-1}^1 \int_{-1}^1 f(x, y) dx dy$$

$$\int_{-1}^1 \int_{-1}^1 f(x, y) dx dy = \sum_{i=1}^n \sum_{j=1}^m c_i d_j f(x_i, y_j) \tag{3.28}$$

where n and m are the number of integration points in x and y , c_i and d_j are weighting coefficients and $f(x_i, y_j)$ is the value of the integrand at the specific points $x_j \in [-1, 1]$, $y_i \in [-1, 1]$. In this formula we introduce d_j and y_j to denote values equivalent to c_j and x_j but taken in the y coordinate direction. In our above example we considered a cubic polynomial which we observed was exactly integrated with an $n = 2$. This is consistent with information presented in [Table 3.29](#). In general, **the polynomial that will be exactly integrated by n points is of degree of $2n - 1$** . Notice that there are + and - values for x_i so the number of points being considered for, say, $n = 2$, is two and for $n = 3$ is three.

i	Polynomial degree	$\pm x_i$	c_i
1	Linear	0.000	2.000
2	Cubic	0.577	1.000
3	Quintic	0.775	0.556
		0.000	0.889
4	Septimal	0.861	0.348
		0.340	0.652

(3.29)

Example Gauss Integration

Consider the integral we used as an example in Subsubsection 3.7, that is

$$\int_1^3 (z^2 + z + 1) dz = 14.666. \quad (3.30)$$

Step One: Change Variable Limits,

Change the limits of integration to $[-1, 1]$ using the following equation:

$$x = \frac{1}{2}(2z - 4) \quad (3.31)$$

and verify the transformation is correct.

If $x = 1$,

$$x = \frac{1}{2}(2 - 4) = -1. \quad (3.32)$$

If $x = 3$,

$$x = \frac{1}{2}(6 - 4) = 1. \quad (3.33)$$

Now the integral becomes

$$\int_1^3 (z^2 + z + 1) dz = \int_{-1}^1 (x^2 + 5x + 7) dx. \quad (3.34)$$

This is the interval from -1 to 1 , we stated that we needed to apply Gauss Legendre Quadrature as it is used in this book. For convenience define

$$X = (x^2 + 5x + 7). \quad (3.35)$$

Step Two: Evaluate Series

$$\gamma = \sum_{j=1}^n c_j X(x_j). \quad (3.36)$$

If we choose $n = 2$, we have

$$\begin{aligned} \gamma &= \sum_{j=1}^2 c_j X(x_j) && (3.37) \\ &= 1 \times [(0.577)^2 + 5 \times 0.577 + 7] + 1 [(0.577)^2 - 5 \times 0.577 + 7] \\ &= 14.666. \end{aligned}$$

The integration is exact because $n = 2$ will exactly integrate a cubic and our function is quadratic.

3.2 Chapter Summary

Polynomial approximation theory is used in this brief chapter to integrate functions numerically. The two methods discussed are the Newton-Cotes quadrature and the Gauss-Legendre quadrature. Common formula, such as the trapezoidal rule and Simpson's rule are described and examples of their application provided. An example of the application of the Gauss Legendre method is also provided. The material in this chapter demonstrates that integration of polynomials can be done easily using numerical strategies based upon polynomial approximation theory.

3.3 Problems

1. Use the trapezoidal rule to evaluate the integral of $\sin \theta$ from 0 to π and check to see if the error obtained is consistent with the theoretical limits.
2. Use Simpson's rule to integrate $\sin \theta$ from 0 to π , compute the error in the numerical integration and compare this error to the theoretical limits.
3. Develop the one point integration formula for the linear equation $f(x) = a_0 + a_1x$.
4. What is the value of the truncation error generated when one applies Simpson's rule

$$\int_0^{2h} f(x) dx = h \left(\frac{1}{3}f(x_0) + \frac{4}{3}f(x_1) + \frac{1}{3}f(x_2) \right) - \frac{1}{90} \frac{h^5 d^4 f(\eta)}{dx^4} \quad (3.38)$$

when $f(x)$ is given by $f(x) = a + bx + cx^2 + dx^3$?

5. Use Simpson's rule to evaluate the integral in [Eq. \(3.39\)](#) and check to see if the error obtained is consistent with the theoretical limits.

$$\int_1^3 (x^2 + x + 1) dx = 14.666 \quad (3.39)$$

Simpson's rule is given by [Eq. \(3.38\)](#).

Bibliography

Atkinson, K.E. (1989) *An Introduction to Numerical Analysis*, 2nd ed., John Wiley and Sons.

Burden, R.L. (2000) *Numerical Analysis* 7th Ed., Brooks/Cole, Australia.

Davis, P.J. and P. Rabinowitz (2007) *Methods of Numerical Integration*: 2nd ed., Dover, New York.

Golub, G.H. and J.H. Welsch (1969) "Calculation of Gauss Quadrature Rules," *Math. Comput.*, 23(106), 221-230.

Hazewinkel, M. (2001) "Simpson Formula," *Encyclopedia of Mathematics*, Springer.

Hildebrand, F.B. (1956) *Introduction to Numerical Analysis*, McGraw-Hill, New York.

Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling (1992) "Numerical Derivatives." Section 5.7 in *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, Australia, pp. 180-184.

Quarteroni, A, S. R.Sacco and F. Saleri (2006) *Numerical Mathematics*," Springer, Berlin.

Ralston, A. (1965) *A First Course in Numerical Analysis*" McGraw-Hill, New York.

Stoer, J.B. and R. Bulirsch (2002) *Introduction to Numerical Analysis* 3rd ed., Springer, New York.

Chapter 4

Initial Value Problems

In this section we are interested in expanding the idea of solving differential equations using the finite difference method we first introduced in Chapter 2 on page 2. More specifically we will focus in this chapter on solving ordinary differential equations such as we first encountered in a preliminarily way in Subsection 2.4.2. We will begin this section with equations of the general form

$$\frac{du}{dt} = f(u, t) \quad t \in [a, b]. \quad (4.1)$$

Since [Eq. \(4.1\)](#) is a first-order equation, it requires one auxiliary condition for its solution, in this case an initial condition since the independent variable is time. A suitable condition is of the form

$$u(t_0) = u_0, \quad (4.2)$$

where t_0 is the initial time and u_0 is a specified known value of $u(t)$. It is assumed that $f(u, t)$ is also defined and continuous for all $t \in [a, b]$. Formulating a forward difference approximation of $\frac{du}{dt}$ appearing in [Eq. \(4.1\)](#) yields (see Section 2.3 on page 37)

$$\left. \frac{du}{dt} \right|_{t_i} = \frac{u(t_{i+1}) - u(t_i)}{h} + \frac{1}{2!} \left. \frac{d^2u(x)}{dt^2} \right|_{x=\xi} (t_i - t_{i+1}) \quad (4.3)$$

where

$$\boxed{h = t_{i+1} - t_i} \quad (4.4)$$

Note that we know this is a forward difference because our point of reference for the approximation is at t_i , and we are looking forward to the value at t_{i+1} to create the approximation.

Substitution of [Eq. \(4.3\)](#) into [Eq. \(4.1\)](#) gives

$$\begin{aligned} & \left(\frac{du}{dt} - f(u, t) \right) \Big|_{t_i} \\ &= \frac{u(t_{i+1}) - u(t_i)}{h} - f(u, t_i) - \frac{1}{2!} \left. \frac{d^2u(x)}{dt^2} \right|_{x=\xi} h \\ &= 0, \end{aligned} \quad (4.5)$$

or, rearranging

$$u(t_{i+1}) = u(t_i) + hf(u(t_i), t_i) + O(h^2) \quad (4.6)$$

If we neglect the truncation error, we obtain the *Euler Forward Integration Method*, that is

$$u_{i+1} = u_i + hf(t_i, u_i) \quad (4.7)$$

where u_i is the approximate solution to $u(t_i)$. Examination of [Eq. \(4.7\)](#) reveals that one can step through time explicitly, one step at a time, that is the value of u_{i+1} can be obtained from the known values at u_i . The process gets started using u_0 which is known from the initial condition.

4.1 Euler Forward Integration Method Example

Let us now see how we can apply the above concepts. Consider the equation

$$\begin{aligned} \frac{du}{dt} &= -\frac{0.9}{1+2t}u = f(u, t), \\ u(0) &= u_0 = 1, \end{aligned} \quad (4.8)$$

which has the solution ([4], p. 115).

$$u(t) = (1 + 2t)^{-0.45}. \quad (4.9)$$

From Euler's method [Eq. \(4.7\)](#) we obtain

$$u_{i+1} = u_i + hf(u_i, t_i). \quad (4.10)$$

Substitution of our definition for $f(u, t)$ from [Eq. \(4.8\)](#) gives

$$u_{i+1} = u_i - h \left(\frac{0.9}{1+2t_i} \right) u_i. \quad (4.11)$$

For our example consider

$$h = 0.02 \quad (4.12)$$

such that [Eq. \(4.11\)](#) becomes

$$u_1 = u_0 - 0.02 \left(\frac{0.9}{1+2t_i} \right) u_0. \quad (4.13)$$

After imposing the initial condition, that is

$$u_0 = 1, \quad (4.14)$$

we can write [Eq. \(4.13\)](#) for the first time step as

$$u_1 = 1 - 0.02 \left(\frac{0.9}{1+(2)(0)} \right) 1 = 0.982. \quad (4.15)$$

If we continue this process over several time steps with each newly calculated value

depending only on the preceding time step, we obtain the values in the following table. Here we present, column-wise from left to right, the elapsed time at the beginning of each time step, the computed value of u_i , the exact value of $u(t)$ at time t_i and the error as calculated as the difference $u_i - u(t_i)$. Note that the error reported in the fourth column is cumulative and getting larger as time increases.

t	u_i	$u(t_i)$	$E(t_i)$
0	1.000	1.000	0
0.02	0.9820	0.9825	- 0.0005
0.04	0.9650	0.9659	- 0.0009
0.06	0.9489	0.9502	- 0.0013
0.08	0.9336	0.9354	- 0.0018
0.10	0.9191	0.92123	- 0.0021
⋮	⋮	⋮	⋮
0.20	0.8402	0.8594	- 0.0192
⋮	⋮	⋮	⋮
0.30	0.7862	0.8093	- 0.0231

4.2 Convergence

There are three key properties of numerical approximations like that considered above that need to be addressed. If these three conditions are not met, the numerical approximation is of little practical interest. These properties are convergence, consistency and stability and we will consider each in turn below.

Convergence Definition

A numerical method applied to a given differential equation is termed convergent if, assuming no computer round-off error, the numerical solution approaches the exact solution to the differential equation as $h \rightarrow 0$.

$$\lim_{h \rightarrow 0 \ (i \rightarrow \infty)} u_i = u(t_i) \quad \text{where} \quad (h) = \frac{t_i - t_0}{i} \quad \forall t \in [a, b].^1 \tag{4.16}$$

The above means that as the time step h gets smaller and smaller, the calculated value u_i approaches the exact solution $u(t_i)$.

With this definition in mind, we now examine the convergence properties of the Euler method for this example. From the governing equation we have

$$\left. \frac{du}{dt} \right|_{t_i} = f(u(t_i), t_i) \tag{4.17}$$

and from Taylor's theorem we know

$$u(t_{i+1}) = u(t_i) + h \left. \frac{du}{dt} \right|_{t_i} + \frac{h^2}{2!} \left. \frac{d^2u}{dt^2} \right|_{t=\xi_1} \quad \xi_1 \in [t_i, t_{i+1}]. \quad (4.18)$$

This expression can be rearranged to provide an approximation for $\frac{du(t)}{dt}$. We have used the forward difference approximation in the preceding example, that is

$$u_{i+1} = u_i + h \frac{du_i}{dt} \quad (4.19)$$

$$\frac{du_i}{dt} = \frac{u_{i+1} - u_i}{h}. \quad (4.20)$$

It is important at this stage to realize that $u(t_i)$ and u_i are different as are $\left. \frac{du}{dt} \right|_{t_i}$ and $\frac{du_i}{dt}$.

The value $u(t_i)$ is the exact solution evaluated **at the time** t_i . The value u_i is the value of the function as computed using a first-order difference approximation for [Eq. \(4.17\)](#). Thus u_i inherently contains any errors that have evolved during the calculation of $u_1, u_2, u_3, \dots, u_i$. As a result the accumulated error in the solution can be determined at time t_{i+1} by subtracting $u(t_{i+1})$ from u_{i+1} . Let's do the subtraction, that is subtract [Eq. \(4.18\)](#) from [Eq. \(4.19\)](#). We obtain

$$u_{i+1} - u(t_{i+1}) = u_i - u(t_i) + h \left. \frac{du}{dt} \right|_{t_i} - h \left. \frac{du}{dt} \right|_{t_i} - \frac{h^2}{2!} \left. \frac{d^2u}{dt^2} \right|_{t=\xi_1} \quad (4.21)$$

or, using the information from [Eq. \(4.17\)](#) to replace the first derivative in this expression, we obtain

$$u_{i+1} - u(t_{i+1}) = u_i - u(t_i) + h [f(t_i, u_i) - f(t_i, u(t_i))] - \frac{h^2}{2!} \left. \frac{d^2u}{dt^2} \right|_{t=\xi_1}. \quad (4.22)$$

Now, let us define the error terms as

$$\varepsilon_i = u_i - u(t_i). \quad (4.23)$$

Substitution of these definitions into [Eq. \(4.22\)](#) and defining

$$E = -\frac{h^2}{2!} \left. \frac{d^2u}{dt^2} \right|_{t=\xi_1} \quad (4.24)$$

we obtain

$$\varepsilon_{i+1} = \varepsilon_i + h [f'(t_i, u_i) - f'(t_i - u(t_i))] + E. \quad (4.25)$$

We now require the use of the mean value theorem which states the following:

Given a function $f(t)$ is continuous on the closed interval $[a, b]$, and $a < b$, and that $f(t)$ is differentiable on the open interval (a, b) , then there is a point c in (a, b) such that

$$\left. \frac{df}{dt} \right|_c = \frac{f(b) - f(a)}{b - a}. \quad (4.26)$$

In other words, there exists a point c in the interval (a, b) such that the line joining the function values at the endpoints of the interval $[a, b]$ is parallel to the tangent at c . (see [Fig. 4.1](#)).

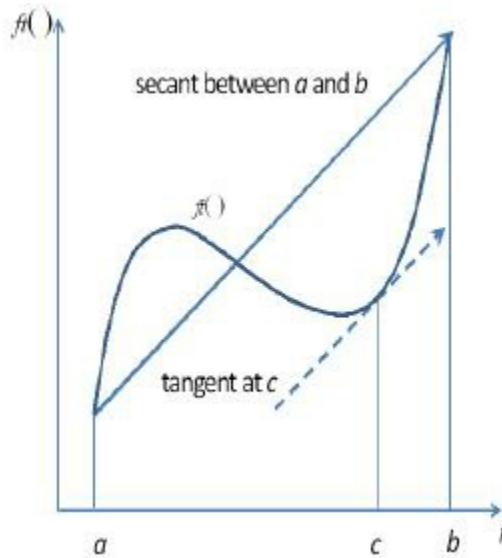


Figure 4.1 Illustration of the mean value theorem. The line connecting the endpoints of the curve is parallel to a tangent of the curve at some point c in the interval $a - b$.

Now we employ the mean value theorem, [Eq. \(4.26\)](#), using $b \equiv u(t_i)$ and $a \equiv u_i$ to write

$$f(t_i, u(t_i)) - f(t_i, u_i) = \left. \frac{df}{du} \right|_{(t_i, \bar{u}_i)} \underbrace{[u(t_i) - u_i]}_{-\varepsilon_i}, \quad \bar{u}_i \in [u_i, u(t_i)], \quad (4.27)$$

or

$$f(t_i, u(t_i)) - f(t_i, u_i) = - \left. \frac{df}{du} \right|_{(t_i, \bar{u}_i)} \varepsilon_i. \quad (4.28)$$

Notice the appearance of the mean value \bar{u}_i in this expression.

Substitution of [Eq. \(4.28\)](#) into [Eq. \(4.25\)](#) gives

$$\varepsilon_{i+1} = \varepsilon_i + h \left. \frac{df}{du} \right|_{(t_i, \bar{u}_i)} \varepsilon_i + E. \quad (4.29)$$

Let

$$p \equiv \left. \frac{df}{du} \right|_{(t_i, \bar{u}_i)} \quad (4.30)$$

such that

$$\varepsilon_{i+1} = \varepsilon_i + hp\varepsilon_i + E. \quad (4.31)$$

To see how the error evolves over time we now make a series of calculations while advancing through time using the following relationship:

$$\varepsilon_{i+1} = \varepsilon_i (1 + hp) + E. \quad (4.32)$$

This will allow us to see what happens to ε_{i+1} as we proceed step-wise from the initial conditions. Here we are considering the special case of p and E constant (which they are not).

We begin by using the observation that when

$$i = 0, \quad (4.33)$$

$$u_{i=0} = u(t_0) = u_0. \quad (4.34)$$

Then, by successive substitution we obtain using [Eq. \(4.32\)](#) the following sequence,

$$\varepsilon_0 = 0 \quad (\text{since } u_{i=0} - u(t_0) = 0) \quad (4.35)$$

$$\varepsilon_1 = E \quad (4.36)$$

$$\begin{aligned} \varepsilon_2 &= E [1 + hp] + E \\ &= E [2 + hp] \end{aligned} \quad (4.37)$$

$$\begin{aligned} \varepsilon_3 &= E [2 + hp] [1 + hp] + E \\ &= 3E + 3Ehp + Eh^2p^2 \\ &= 3E [1 + hp] + Eh^2p^2, \end{aligned} \quad (4.38)$$

$$\begin{aligned}\varepsilon_4 &= [3E [1 + hp] + Eh^2p^2] [1 + hp] + E \\ &= 4E + 6Ehp + 4Eh^2p^2 + Eh^3p^3,\end{aligned}\tag{4.39}$$

$$\varepsilon_{i+1} = (i + 1) E + O(Eh),\tag{4.40}$$

where all the terms of power higher than h are smaller than $O(Eh)$ and are omitted.

Now recall from [Eq. \(4.24\)](#) that

$$E = O(h^2)\tag{4.41}$$

and that

$$\boxed{t_{i+1} = t_0 + (i + 1) h.}\tag{4.42}$$

We can now introduce [Eqs. \(4.41\)](#) and [\(4.42\)](#) into [Eq. \(4.40\)](#) to give an expression for ε_{i+1}

$$\begin{aligned}\varepsilon_{i+1} &\approx \frac{t_{i+1} - t_0}{h} E + O(Eh) \\ &\approx \frac{t_{i+1} - t_0}{h} O(h^2) + O(h^3) \\ &\approx O(h) + \dots\end{aligned}\tag{4.43}$$

The above development shows that the **total error is one order lower than the local error** (that is the error per time step). Notice that as h goes to zero, ε_{i+1} also goes to zero, thereby illustrating **convergence** for the Euler forward integration method.

4.3 Consistency

We will now look at the issue of consistency which, simply stated, is the requirement that the approximating equation approach the original differential equation as the increment h approaches zero. Let's consider an example. Given the equation

$$u(t_{i+1}) = u(t_i) + hf(t_i, u_i) + \underbrace{E \frac{h^2}{2!} \frac{d^2u(t)}{dt^2} \Big|_{t=\xi}}_{\text{truncation error}}\tag{4.44}$$

or, dividing through by h , we get

$$\frac{u(t_{i+1}) - u(t_i)}{h} = f(t_i, u_i) + \underbrace{\frac{h}{2!} \frac{d^2u(t)}{dt^2} \Big|_{t=\xi}}_{E/h}\tag{4.45}$$

then if E/h vanishes as $h \rightarrow 0$, (which it will do because $E/h = O(h^2)/h = O(h)$) so h

appears only in the numerator of E/h), we say the difference equation is **consistent** with the differential equation, viz.

$$\lim_{h \rightarrow 0} \left(\frac{u(t_{i+1}) - u(t_i)}{h} - f(t_i, u_i) - \frac{h}{2!} \frac{d^2 u(x)}{dt^2} \Big|_{x=\xi} \right) = \frac{du(t_i)}{dt} - f(t_i, u). \quad (4.46)$$

Thus, given [Eq. \(4.46\)](#) we have shown that Euler's method is **consistent**, that is the difference equation becomes, in the limit as $h \rightarrow 0$, equal to the original differential equation.

4.4 Stability

The last property to consider is stability. It determines whether, as one advances through time, the numerical solution behaves as expected (stable) or alternatively generates an ever-increasing error due to round off in the computer (unstable). Clearly, an unstable difference equation is of no practical importance.

As an example, consider the approximation to the following equation

$$\frac{du}{dt} = \alpha u \quad (4.47)$$

given by

$$u_{i+1} - u_i = \alpha h u_i \quad t \in [0, b] \quad \boxed{h \equiv t_{i+1} - t_i}. \quad (4.48)$$

Assume an exact solution, free of round-off error, is given by \tilde{u}_i . Then \tilde{u} satisfies

$$\tilde{u}_{i+1} - \tilde{u}_i = \alpha h \tilde{u}_i. \quad (4.49)$$

Now subtract \tilde{u}_i from u_i in [Eq. \(4.48\)](#) to give

$$u_{i+1} - \tilde{u}_{i+1} - (u_i - \tilde{u}_i) = \alpha h (u_i - \tilde{u}_i). \quad (4.50)$$

Let

$$\tilde{e}_i \equiv u_i - \tilde{u}_i \quad (4.51)$$

be the round-off error committed at t_i . If \tilde{e}_i increases as i increases we have a problem.

Let's take a look at what happens to \tilde{e}_i as i increases. To do this we first write [Eq. \(4.50\)](#) as

$$\bar{e}_{i+1} - \bar{e}_i = \alpha h \bar{e}_i \quad (4.52)$$

or

$$\bar{e}_{i+1} = (1 + \alpha h) \bar{e}_i. \quad (4.53)$$

Now consider the cumulative effect of this error as we proceed through time.

$$\tilde{e}_{i+1} = (1 + \alpha h) \tilde{e}_i, \quad (4.54)$$

$$\begin{aligned} \tilde{e}_{i+2} &= (1 + \alpha h) \tilde{e}_{i+1} \\ &= (1 + \alpha h)^2 \tilde{e}_i, \end{aligned} \quad (4.55)$$

$$\begin{aligned} \tilde{e}_{i+3} &= (1 + \alpha h) \tilde{e}_{i+2} \\ &= (1 + \alpha h)^2 \tilde{e}_{i+1} \\ &= (1 + \alpha h)^3 \tilde{e}_{i\dots}, \end{aligned} \quad (4.56)$$

and in general

$$\tilde{e}_{i+n} = (1 + \alpha h)^n \tilde{e}_i. \quad (4.57)$$

Examination of [Eq. \(4.57\)](#) reveals that to **assure no exponential error growth, the following inequality has to hold**,

$$|1 + \alpha h| \leq 1. \quad (4.58)$$

This we consider to be a **stable** approximation, conditional on the requirement expressed in [Eq. \(4.58\)](#).

4.4.1 Example of Stability

To further illustrate the concept of stability we will consider the following example which assumes $\alpha = -1$ in [Eq. \(4.58\)](#). We have

$$|1 + \alpha h| = |1 + (-1)h| = |1 - h|. \quad (4.59)$$

For

$$|1 - h| \leq 1 \quad (4.60)$$

we deduce that

$$-1 \leq (1 - h) \leq 1 \quad (4.61)$$

which requires

$$h \leq 2, \quad h \geq 0. \quad (4.62)$$

Thus the approximation of [Eq. \(4.57\)](#) is **conditionally stable**, that is it is stable provided $h \leq 2$ (since h must always be greater than 0).

4.5 Lax Equivalence Theorem

To determine convergence using the above strategy, one needs to know the solution of

the equations beforehand, but the reason we are trying to find the solution numerically is because the exact solution is not, in fact, generally known. Fortunately, there is a theorem that helps; it is the Lax Equivalence Theorem and it states ([Lax and Richtmyer, 1956]):

Theorem 3 *Given a properly posed initial value problem and a finite-difference approximation to it that satisfies the consistency condition, stability is the necessary and sufficient condition for convergence.*

This is a very important result from a practical perspective because it states that, if we can determine consistency and stability, which do not require knowledge of the exact solution, we can establish convergence.

For the sake of completeness we are now going to provide some very well known, very specific methodology for solving initial value problems that are widely used, these are known as the Runge–Kutta formulas.

4.6 Runge–Kutta Type Formulas

4.6.1 General Form

To illustrate the concepts behind the Runge–Kutta methods, we will consider the equation of interest to be

$$\frac{du}{dt} = f(u, t). \quad (4.63)$$

The general form of the Runge–Kutta type formula for this equation using a two–term expansion is

$$u_{i+1} = u_i + ak_1 + bk_2 \quad (4.64)$$

where

$$k_1 = hf(u_i, t_i) \quad (4.65)$$

$$k_2 = hf(u_i + \beta k_1, t_i + \alpha h). \quad (4.66)$$

4.6.2 Runge–Kutta First Order Form (Euler’s Method)

In [Eq. \(4.64\)](#) let $a = 1$ and $b = 0$. Substitution of these values into [Eq. \(4.64\)](#) we get

$$u_{i+1} = u_i + hf(u_i, t_i) \quad (4.67)$$

or, rewriting this expression

$$\frac{u_{i+1} - u_i}{h} = f(u_i, t_i). \quad (4.68)$$

This is the forward finite difference approximation we derived earlier and denoted as the Euler method.

4.6.3 Runge–Kutta Second Order Form

Now consider the extension of the above method to a two–step scheme. Consider once again the equation

$$\frac{du}{dt} = f(u, t). \quad (4.69)$$

Let the coefficients in [Eq. \(4.64\)](#) take on the values

$$a = b = 1/2 \quad \alpha = \beta = 1. \quad (4.70)$$

Using these values in [Eqs. \(4.64 – 4.66\)](#) we obtain

$$u_{i+1} = u_i + \frac{1}{2}k_1 + \frac{1}{2}k_2 \quad (4.71)$$

where

$$k_1 = hf(u_i, t_i) \quad (4.72)$$

as before and

$$k_2 = hf(u_i + k_1, t_i + h) \quad (4.73)$$

which, when substituted into [Eq. \(4.71\)](#) gives

$$u_{i+1} = u_i + \frac{1}{2} [hf(u_i, t_i) + hf(u_i + hf\{u_i, t_i\}, t_i + h)]. \quad (4.74)$$

Example of Runge–Kutta Second Order Form

Let us now consider an illustrative example in which we consider the equation

$$\frac{du}{dt} = \tan(u) + 1, \quad t \in [1, 1.075], \quad (4.75)$$

with initial condition

$$u_0 = 1, \quad (4.76)$$

and $h = 0.025$. Since the interval is of length 0.075, there will be three time steps. The equation to be used is [Eq. \(4.71\)](#). We will proceed by calculating k_1 and k_2 and then substitute these into [Eq. \(4.71\)](#). With u in radians we obtain for k_1 and k_2

$$\begin{aligned} k_1 &= hf(u_0, t_0) \\ &= 0.025(\tan(1) + 1) \\ &= 0.025(2.56) \\ &= 0.064, \end{aligned} \quad (4.77)$$

$$\begin{aligned}
k_2 &= h[\tan(u_0 + k_1, t_0 + h) + 1] \\
&= h[\tan((1 + 0.064), 1.025) + 1] \\
&= 0.025[\tan(1.064) + 1] \\
&= 0.070.
\end{aligned} \tag{4.78}$$

Substitution of these values into [Eq. \(4.71\)](#) we obtain

$$\begin{aligned}
u_1 &= u_0 + \frac{1}{2}k_1 + \frac{1}{2}k_2 \\
&= 1 + \frac{1}{2}0.064 + \frac{1}{2}0.070 \\
&= 1.067.
\end{aligned} \tag{4.79}$$

Now, for step two, we obtain

$$\begin{aligned}
k_1 &= hf(u_1, t_1) \\
&= 0.025[\tan(1.067) + 1] \\
&= 0.025(1.81 + 1) \\
&= 0.070,
\end{aligned} \tag{4.80}$$

$$\begin{aligned}
k_2 &= h[\tan(u_1 + k_1, t_1 + h) + 1] \\
&= h[\tan((1.067 + 0.070), 1.050) + 1] \\
&= 0.025[(\tan(1.137) + 1)] \\
&= 0.079,
\end{aligned} \tag{4.81}$$

$$\begin{aligned}
u_2 &= u_1 + \frac{1}{2}k_1 + \frac{1}{2}k_2 \\
&= 1.067 + \frac{1}{2}0.070 + \frac{1}{2}0.79 \\
&= 1.14.
\end{aligned} \tag{4.82}$$

And finally, for step three, we have

$$\begin{aligned}
k_1 &= hf(u_2, t_2) \\
&= 0.025[\tan(1.14) + 1] \\
&= 0.025(2.17 + 1) \\
&= 0.0794,
\end{aligned} \tag{4.83}$$

$$\begin{aligned}
 k_2 &= h[\tan(u_2 + k_1, t_2 + h) + 1] & (4.84) \\
 &= h[\tan((1.14 + 0.0794), 1.075) + 1] \\
 &= 0.025[(\tan(1.22) + 1)] \\
 &= .0933,
 \end{aligned}$$

$$\begin{aligned}
 u_3 &= u_2 + \frac{1}{2}k_1 + \frac{1}{2}k_2 & (4.85) \\
 &= 1.14 + \frac{1}{2}0.0794 + \frac{1}{2}0.0933 \\
 &= 1.23.
 \end{aligned}$$

The analytic solution to this problem is

$$0.5(u + \ln |\sin x + \cos x|) + 0.338 - t = 0 \quad (4.86)$$

and for $t = 1.075$ we obtain $u = 1.23$ so the solution is exact to three significant figures.

Geometric Interpretation

We will now use [Fig. 4.2](#) to look at one way to explain the difference between Euler's method and the Runge–Kutta method. The Euler method uses the equation

$$u_{i+1} = u_i + hf(u_i, t_i) \quad (4.87)$$

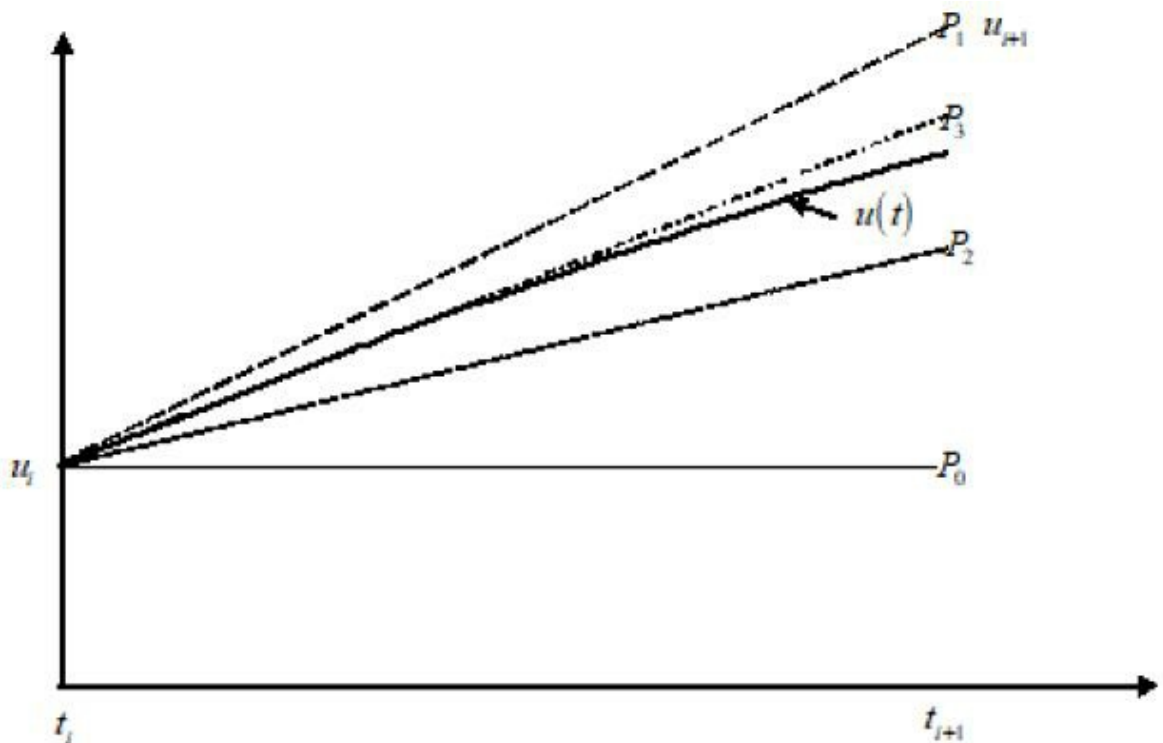


Figure 4.2 Schematic visual representation to assist in the explanation of the difference between Euler's method and the Runge–Kutta method.

which, using the original equation Eq. (4.17) is equivalent to

$$u_{i+1} = u_i + h \left. \frac{du}{dt} \right|_i. \quad (4.88)$$

Thus the Euler method projects from u_i to u_{i+1} as represented by the curve P_1 .

The Runge–Kutta method uses

$$u_{i+1} = u_i + \underbrace{\frac{hf(u_i, t_i)}{2}}_{P_1} + \underbrace{\frac{hf(u_i + hf(u_i, t_i)), t_i + h)}{2}}_{P_2} \quad (4.89)$$

$\underbrace{\hspace{15em}}_{P_3}$

which we can interpret as follows: the P_1 term is the equivalent of the Euler method. The term P_2 is equivalent to

$$P_2 = h \left. \frac{d}{dt} \left(u_i + h \left. \frac{du}{dt} \right|_i \right) \right|_{t_i + h} = h \left. \frac{du}{dt} \right|_{i+1}. \quad (4.90)$$

which is $\left. \frac{du}{dt} \right|_{i+1}$ evaluated at the end of the interval, i.e. at $t = t_0 + (i + 1)h$. The Runge–Kutta method uses the average of P_1 and P_2 to yield

$$P_3 = \frac{P_1 + P_2}{2} \quad (4.91)$$

which are the slopes at t_i and t_{i+1} . Thus the Runge–Kutta method uses the arithmetic mean value of the derivatives evaluated at the end points of the interval.

4.8 Chapter Summary

The solution of initial value problems using numerical approximations is the focus of this chapter. The three fundamental properties of a numerical approximation, namely consistency, convergence and stability are presented and examples are provided to assist in understanding these concepts. Euler's classical method along with Runge–Kutta formulations are also described.

4.9 Problems

1. Consider the equation

$$\frac{du}{dt} + u = 0.$$

Since this is a first-order equation in time we need an initial condition. Assume the initial condition is

$$u(0) = 1.$$

Further assume that you are going to solve this equation using the forward difference scheme

$$\frac{u_{i+1} - u_i}{h} + u_i = 0.$$

and you are using a value of $h = 3$. Calculate the first five time steps and observe whether the approximation is stable or unstable. Now do the same calculation using $h = 3/2$ and make the same observation. How is this result consistent with the theory presented in this chapter?

2. You find it necessary to forecast the rate of heating of a widget. The governing equation is

$$\frac{dT}{dt} - \beta T = 0 \quad t \in (0, \infty), \quad (\beta \text{ is real}) \quad (4.92)$$

where β is an empirical constant and is positive. You decide to use the difference formula:

$$\frac{T_{i+1} - T_i}{h} - \beta T_{i+1} = 0 \quad (4.93)$$

where $h \equiv t_i - t_{i-1}$ and $t_i = ih$. Show how you determine the stability of this scheme and describe what the stability constraints are.

Include the following steps:

1.
 - a. Define your error term and represent it as e_i ;

- b. Substitute the error into equation [4.7](#);
 - c. Show how the error evolves over each time step;
 - d. Determine the requirements on β and Δt for stability.
2. What is the difference between the error of the approximation and the round-off error?
 3. Will the number of significant digits in the representation of a number affect the stability of an otherwise unstable scheme. In other words, can you use a computer which represents numbers with more bits to overcome stability problems?
 4. In the truncation errors involving h , for example in [Eq. \(4.3\)](#), it is assumed that $h \leq 1$ so that as the power of h increases, its magnitude decreases. What happens when $h > 1$? How then do you demonstrate consistency or establish the error of the approximation?

Bibliography

Lambert, J.D. (1991) *Numerical Methods for Ordinary Differential Systems*, John Wiley and Sons, Chichester.

Lapidus, L. and G.F. Pinder (1982) *Numerical Solution of Partial Differential Equations in Science and Engineering*, John Wiley and Sons, New York.

Lax, P.-D. and R.D. Richtmyer (1956) "Survey of the stability of linear finite difference equations," *Comm. Pure Appl. Math.* 9, 267–293.

Ralston, R. and P. Rabinowitz (2012) *A First Course in Numerical Analysis*, 2nd ed., Dover Publications, New York.

Salvadori, M.G. and M.L. Baron (1961) *Numerical Methods in Engineering*, Prentice Hall, Englewood Cliffs.

Smith, G.D. (1985) *Numerical Solution of Differential Equations: Finite Difference Methods*, 3rd ed., Oxford University Press, New York.

Wikipedia contributors (2015) "Runge–Kutta Methods," Wikipedia, The Free Encyclopedia.

Note

¹ The symbol \forall means 'for all'

Chapter 5

Weighted Residuals Methods

We are now going to discuss the concept of the **method of weighted residuals** which, along with the **interpolation theory** presented above, will allow us to create an array of different numerical techniques. To start the discussion we consider the equation

$$\frac{du}{dx} = f(x) \quad x \in [a, b]. \quad (5.1)$$

From Lagrange interpolation (see Section 1.8, page. 14) we can write the approximation to $u(x)$ as

$$u(x) = \sum_{j=0}^n \ell_j(x) u_j + E(x) \quad (5.2)$$

or

$$u(x) = \hat{u}(x) + E(x). \quad (5.3)$$

where

$$\hat{u}(x) = \sum_{j=0}^n \ell_j(x) u_j \quad (5.4)$$

and $\ell_j(x)$ is assumed to be a linear Lagrange polynomial with $n + 1$ nodes on the interval $x \in [a, b]$. Substitution of [Eq. \(5.3\)](#) into [Eq. \(5.1\)](#) gives

$$\frac{d\hat{u}(x)}{dx} - f(x) = -\frac{dE(x)}{dx} \equiv R(x) \quad (5.5)$$

where $R(x) \neq 0$ and is defined as a **residual error** or just a **residual**. The residual is the difference between the value of the differential equation using the exact solution $u(x)$ and the value of the differential equation when we use the approximate solution $\hat{u}(x)$.

Let $w_i(x)$, $i = 0, \dots, n$ be a set of **weighting functions**. At this point, we have not specified what they might be, but we assume they exist. Next, multiply the residual $R(x)$ by each of the weighting functions $w_i(x)$, and integrate each of the resulting products over the domain of interest. Now, for each weighting function $w_i(x)$, $i = 0 \dots n$, set the result to zero. Using mathematical notation we have

$$\int_x R(x) w_i(x) dx = 0 \quad i = 0, 1, 2, \dots, n \quad x \in [x_0 = a, x_n = b]. \quad (5.6)$$

This weighted residual relationship is sometimes written using inner-product notation as

$$\int_x R(x) w_i(x) dx = \langle R(x), w_i(x) \rangle = 0 \quad i = 0, 1, 2, \dots, n \quad x \in [x_0 = a, x_n = b] \quad (5.7)$$

where $\langle \cdot \rangle$ denotes the **inner product**.

Next substitute [Eq. \(5.5\)](#) into [\(5.6\)](#) to get

$$\int_x \left(\frac{d\hat{u}(x)}{dx} - f(x) \right) w_i(x) dx = 0 \quad i = 0, 1, 2, \dots, n. \quad (5.8)$$

If we introduce the definition of $\hat{u}(x)$ found in [Eq. \(5.2\)](#) into [Eq. \(5.8\)](#) we get

$$\int_x \left(\frac{d}{dx} \sum_{j=0}^n \ell_j(x) u_j - f(x) \right) w_i(x) dx = 0 \quad i = 0, 1, 2, \dots, n. \quad (5.9)$$

Since the variables u_j are simply numbers, we can take them out of the integration and obtain

$$\sum_{j=0}^n u_j \int_x \frac{d}{dx} \ell_j(x) w_i(x) dx - \int_x f(x) w_i(x) dx = 0 \quad i = 0, 1, 2, \dots, n. \quad (5.10)$$

Examination of [Eq. \(5.10\)](#) reveals that we have $n + 1$ equations, one for each value of $i = 0, 1, 2, \dots, n$. Each equation has $n + 1$ unknowns, $u(x_j), j = 0, 1, \dots, n$. So, at least in theory, once we select the form of the functions $w_i(x)$, we should be able to solve for the $n + 1$ values of u_j . **Different Lagrange polynomials $\ell_j(x)$ and different weighting functions $w_i(x)$ will give different numerical methods.** Once we have calculated the u_j values we can use [Eq. \(5.4\)](#) to obtain $\hat{u}(x)$ anywhere in the interval $x \in [a, b]$. We will consider a number of different forms for the weighting functions and thereby derive a corresponding number of numerical approaches. We will start with the ‘finite volume’ or ‘subdomain’ method.

5.1 Finite Volume or Subdomain Method

To obtain the finite volume method (also known as the subdomain method) we select as our weighting functions

$$w_i(x) = \begin{cases} 1, & x \in \left[x_i - \frac{h}{2}, x_i + \frac{h}{2} \right] \\ 0, & x \notin \left[x_i - \frac{h}{2}, x_i + \frac{h}{2} \right] \end{cases} \quad (5.11)$$

which are illustrated in [Fig. 5.1](#). In terms of the local coordinate system, that is the one tied to the element, we have

$$w_0(\chi) = \begin{cases} 1, & \chi \in [0, 0.5] \\ 0, & \chi \in [0.5, 1.0] \end{cases} \quad (5.12)$$

and

$$w_1(\chi) = \begin{cases} 1, & \chi \in [0.5, 1.0] \\ 0, & \chi \in [0, 0.5] \end{cases} \quad (5.13)$$

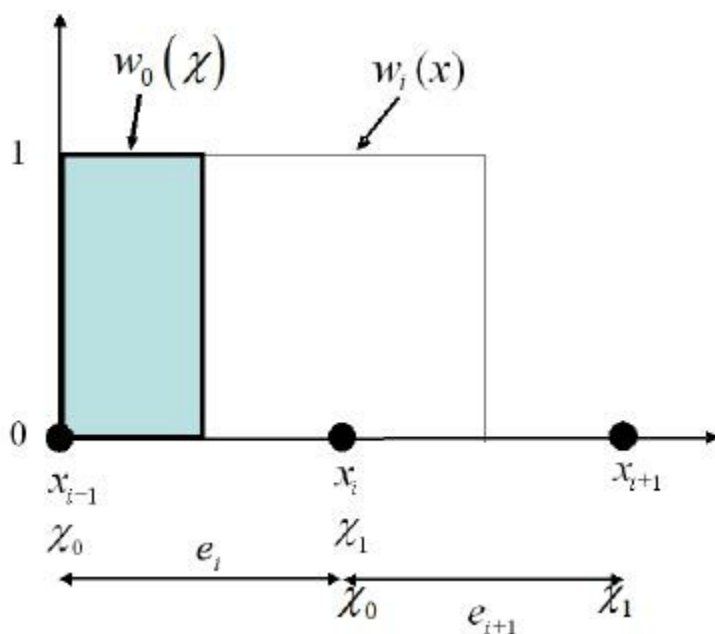


Figure 5.1: Representation of the finite volume weighting function.

Next we decide on the polynomial functions which we wish to use in approximating $u(x)$. We will call these approximating functions **basis functions**. As noted earlier, we will initially use the linear Lagrange polynomials shown in Fig. 1.10 on page 18 as our basis functions. In the following representation of our approximating function $\hat{u}(x)$ we **replace $u(x_j)$ with simply u_j to simplify notation**. With this substitution we obtain for the approximation defined in the local coordinates χ

$$\hat{u}(x) = \sum_{j=0}^1 \ell_j(\chi) u_j \quad \chi \in [0, 1]. \quad (5.14)$$

The linear Lagrange polynomials are given by

$$\ell_0(\chi) = \frac{\chi_1 - \chi}{\chi_1 - \chi_0} \quad (5.15)$$

and

$$\ell_1(\chi) = \frac{\chi - \chi_0}{\chi_1 - \chi_0}. \quad (5.16)$$

The x and χ coordinates are defined by Eq. (1.89) on page 16 and in Fig. 1.10. The combined weighting and basis functions are shown in Fig. 5.2. Keep in mind that we are using two terms to describe the discretization of the x axis. On one hand, we have the finite element, this is the distance between any two nodes in a system characterized by linear basis functions. On the other hand we have the finite volume which is characterized by the segment of the x axis under the finite volume. So, in Fig. 5.2 an element e_i would be characterized by the segment from node x_{i-1} to node x_i and the finite volume associated with the node at x_i would be the segment defined by $\frac{x_i - x_0}{2} + \frac{x_{i+1} - x_i}{2}$.

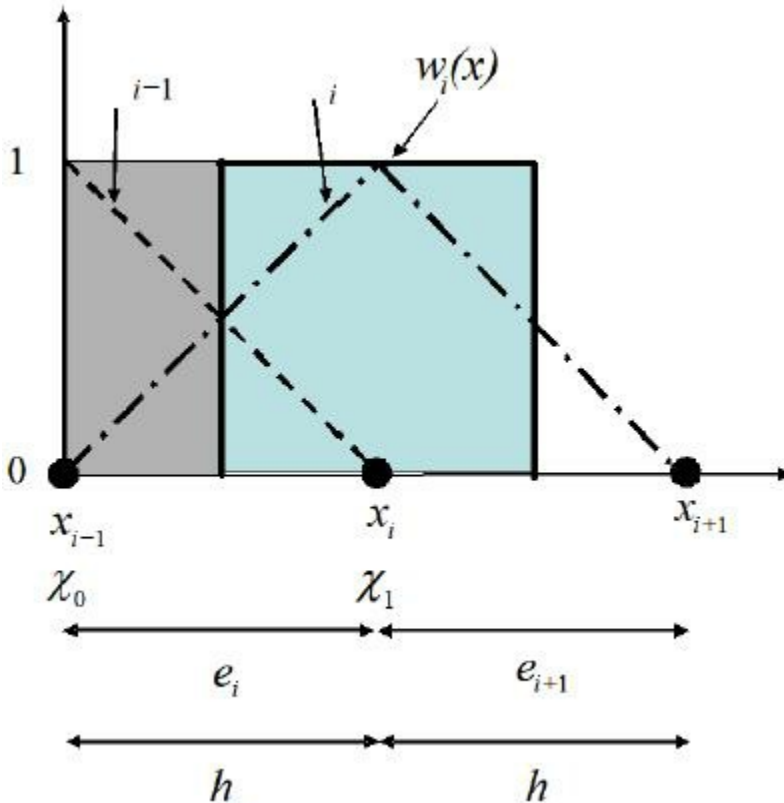


Figure 5.2: Weighting function $w_i(x)$ and basis function $l_i(x)$ in global and local coordinates.

5.1.1 Example

In this subsection we will consider an example of the application of the finite volume approximation to the equation

$$\frac{du(x)}{dx} + u(x) = 1 \quad x \in [0, 1] \quad (5.17)$$

$$u(0) = u_0 = 2. \quad (5.18)$$

The analytic solution for this problem is

$$u(x) = 1 + \exp(-x). \quad (5.19)$$

We first write the residual approximation

$$\frac{d\hat{u}(x)}{dx} + \hat{u}(x) - 1 = R(x). \quad (5.20)$$

Next, we write our linear Lagrange basis functions for two elements (3 nodes). Each element will be of length $e_i = 0.5$, $i = 1, 2$ as will the middle volume. The volumes at each end of the domain are smaller because they are truncated by the existence of the boundaries.

From the method of weighted residuals (Eq. (5.6)) we have (see Fig. 5.3):

$$\int_0^{1.0} R(x) w_i(x) dx = 0 \quad i = 0, 1, 2. \quad (5.21)$$

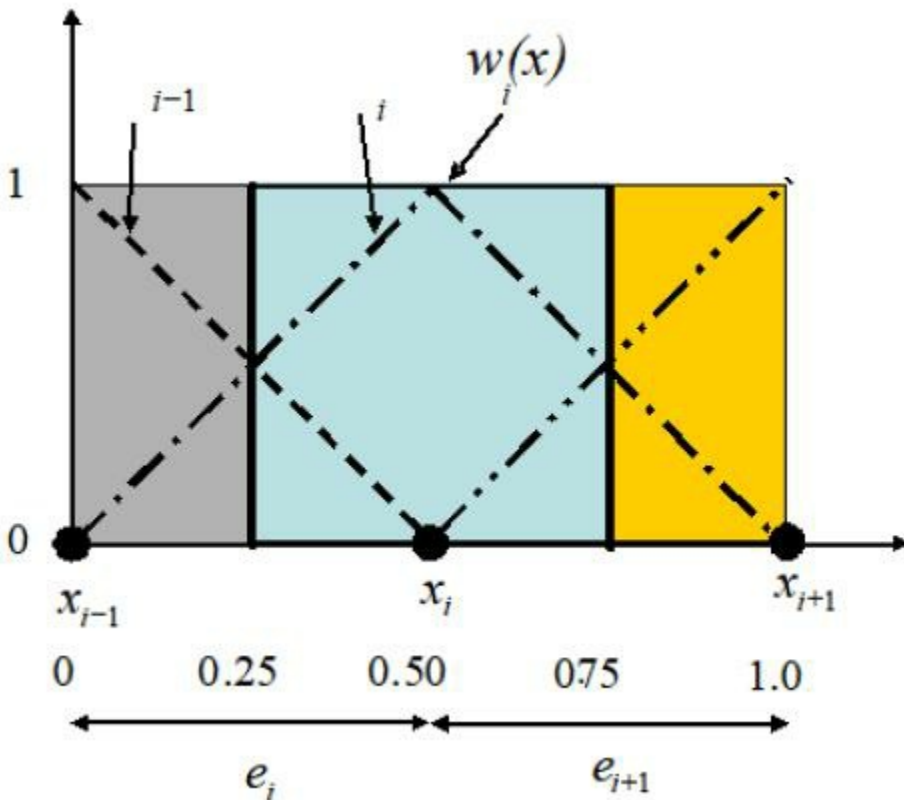


Figure 5.3: Weighting function and basis function for finite volume method example.

In the example discussed in this section the subscripts $i - 1, = 0, i = 1$ and $i + 1 = 2$.

Let us examine the weighting function $w_0(x)$; it has the value of 1 only over one-half of element 1 and 0 elsewhere. Thus, since the integrand is defined only over the first half of the first element, the integration needs to be performed only over the first half of the first element, that is

$$\begin{aligned} & \int_0^{1.0} R(x) w_1(x) dx && (5.22) \\ &= \int_{0.25}^{0.50} R(x) dx + \int_{0.5}^{0.75} R(x) dx \\ &= 0. \end{aligned}$$

For weighting function $w_1(x)$, we have non-zero values of this function over half of the first element and half of the second as shown in [Figs. 5.1](#) and [5.3](#). Thus integration is required over this interval that spans parts of two elements, that is

$$\begin{aligned} & \int_0^{1.0} R(x) w_1(x) dx && (5.23) \\ &= \int_{0.25}^{0.50} R(x) dx + \int_{0.5}^{0.75} R(x) dx \\ &= 0. \end{aligned}$$

Finally, for $w_2(x)$ we have integration over half of element 2 only, that is

$$\begin{aligned} & \int_0^{1.0} R(x) w_2(x) dx && (5.24) \\ &= \int_{0.75}^{1.0} R(x) dx \\ &= 0. \end{aligned}$$

We now perform the integrations, *one element at a time*. Note that each complete weighting function generates one equation. This means that the equation for one weighting function will contain integrals from more than one element. In the case of a one-dimensional problem such as we are considering here, there will be two element integrals for each weighting function. In two dimensions we will see that there are several element integrations per weighting function.

Because the first weighting function $w_0(x)$ occupies only the first element, a complete equation is therefore obtained by integration only over the first element. The appropriate integration therefore is

$$\begin{aligned}
\int_0^{0.25} R(x) dx &= \int_0^{0.25} \left(\frac{d\hat{u}}{dx} + \hat{u} - 1 \right) dx & (5.25) \\
&= \int_0^{0.25} \left[\left(\frac{d\ell_0}{dx} u_0 + \ell_0 u_0 \right) + \left(\frac{d\ell_1}{dx} u_1 + \ell_1 u_1 \right) - 1 \right] dx \\
&= \int_0^{0.25} \left[u_0 \left(\frac{d\ell_0}{dx} + \ell_0 \right) + u_1 \left(\frac{d\ell_1}{dx} + \ell_1 \right) - 1 \right] dx \\
&= \int_0^{0.25} \left[u_0 \left(\frac{-1}{h} + \frac{h-x}{h} \right) + u_1 \left(\frac{1}{h} + \frac{x}{h} \right) - 1 \right] dx \\
&= \frac{u_0}{h} \left(-x + hx - \frac{x^2}{2} \right)_0^{0.25} + \frac{u_1}{h} \left(x + \frac{x^2}{2} \right)_0^{0.25} - x \Big|_0^{0.25} \\
&= u_0 (-5/16) + u_1 (9/16) - 1/4 \\
&= 0.
\end{aligned}$$

or

$$u_0 (-5/16) + u_1 (9/16) - 1/4 = 0.$$

This is our first finite volume equation.

Although it is not really necessary for a one-dimensional problem such as we are considering here, in order to get accustomed to using the local coordinate system χ , we now consider the above integration using χ . The first step is to transform [Eq. \(5.22\)](#) into a form that employs χ . Since χ is a function of x , we can write $\hat{u}(x)$ in terms of χ to give $\hat{u}(\chi)$

$$\int_0^{0.5} R(\chi) \frac{dx}{d\chi} d\chi = \int_0^{0.5} \left(\frac{d\hat{u}(\chi)}{d\chi} \frac{d\chi}{dx} + \hat{u}(\chi) - 1 \right) \frac{dx}{d\chi} d\chi \quad (5.26)$$

where we recall that now the value of h is 1.0. The next challenge is obtaining $\frac{dx}{d\chi}$. We know from [Eq. \(1.89\)](#) on page 16 that is

$$x = \frac{\chi_1 - \chi}{\chi_1 - \chi_0} x_1 + \frac{\chi - \chi_0}{\chi_1 - \chi_0} x_2 \quad (5.27)$$

so we can differentiate x to get, after rearranging,

$$\begin{aligned}
\frac{dx}{d\chi} &= \frac{x_2 - x_1}{\chi_1 - \chi_0} \\
&= \frac{h}{1}.
\end{aligned} \quad (5.28)$$

Similarly, $\frac{d\chi}{dx} = 1/h$. Thus [Eq. \(5.26\)](#) becomes (using the definitions provided by [Eqs.](#)

(5.15) and (5.16))

$$\begin{aligned}
 \int_0^{0.25} R(x) dx &= \int_0^{0.5} \left(\frac{d\hat{u}(\chi)}{d\chi} \frac{d\chi}{dx} + \hat{u}(\chi) - 1 \right) \frac{dx}{d\chi} d\chi & (5.29) \\
 &= \int_0^{0.5} \left(\frac{d\hat{u}(\chi)}{d\chi} \frac{1}{h} + \hat{u}(\chi) - 1 \right) h d\chi \\
 &= \int_0^{0.5} \left[\left(\frac{1}{h} \frac{d\ell_0(\chi)}{d\chi} u_0 + \ell_0(\chi) u_0 \right) + \left(\frac{1}{h} \frac{d\ell_1(\chi)}{d\chi} u_1 + \ell_1(\chi) u_1 \right) - 1 \right] h d\chi \\
 &= \int_0^{0.5} \left[u_0 \left(\frac{1}{h} \frac{d\ell_0(\chi)}{d\chi} + \ell_0(\chi) \right) + u_1 \left(\frac{1}{h} \frac{d\ell_1(\chi)}{d\chi} + \ell_1(\chi) \right) - 1 \right] h d\chi \\
 &= \int_0^{0.5} \left[u_0 \left(\left(\frac{1}{h} \right) \left(\frac{-1}{1} \right) + \frac{1-\chi}{1} \right) + u_1 \left(\left(\frac{1}{h} \right) \left(\frac{1}{1} \right) + \frac{\chi}{1} \right) - 1 \right] h d\chi \\
 &= \left[u_0 \left(\frac{-\chi}{h} + \chi - \frac{\chi^2}{2} \right) \Big|_0^{0.5} + u_1 \left(\frac{\chi}{h} + \frac{\chi^2}{2} \right) \Big|_0^{0.5} - \chi \Big|_0^{0.5} \right] h \\
 &= \left[u_0 \left(\frac{-1/2}{1/2} + 1/2 - \frac{1/4}{2} \right) + u_1 \left(\frac{1/2}{1/2} + \frac{1/4}{2} \right) - 1/2 \right] 1/2 \\
 &= u_0 (-5/16) + u_1 (9/16) - 1/4 \\
 &= 0
 \end{aligned}$$

which is the same result as we obtained in [Eq. \(5.25\)](#). Notice the change in the limits of integration due to the fact that we are now integrating over χ rather than x .

For the second weighting function $w_1(x)$, two elements are involved. For element 1 we have for the global system

$$\begin{aligned}
 \int_{0.25}^{0.5} R(x) dx &= \int_{0.25}^{0.5} \left(\frac{d\hat{u}(x)}{dx} + \hat{u}(x) - 1 \right) dx & (5.30) \\
 &= \int_{0.25}^{0.5} \left[u_0 \left(\frac{d\ell_0(x)}{dx} + \ell_0(x) \right) + u_1 \left(\frac{d\ell_1(x)}{dx} + \ell_1(x) \right) - 1 \right] dx \\
 &= \int_{0.25}^{0.5} \left[u_0 \left(-\frac{1}{h} + \frac{h-x}{h} \right) + u_1 \left(\frac{1}{h} + \frac{x}{h} \right) - 1 \right] dx \\
 &= u_0 \left(-\frac{7}{16} \right) + u_1 \left(\frac{11}{16} \right) - \frac{1}{4}
 \end{aligned}$$

and for the local system (using the definitions of $\ell_0(\chi)$ and $\ell_1(\chi)$ given in Eqs. (1.91) and (1.92) on page 17)

(5.31)

$$\begin{aligned}
& \int_{0.25}^{0.50} R(x) dx \\
&= \int_{0.5}^{1.0} \left(\frac{d\hat{u}(\chi)}{d\chi} \frac{d\chi}{dx} + \hat{u}(\chi) - 1 \right) \frac{dx}{d\chi} d\chi \\
&= \int_{0.5}^{1.0} \left(\frac{d\hat{u}(\chi)}{d\chi} \frac{1}{h} + \hat{u}(\chi) - 1 \right) h d\chi \\
&= \int_{0.5}^{1.0} \left[\left(\frac{1}{h} \frac{d\ell_0(\chi)}{d\chi} u_0 + \ell_0(\chi) u_0 \right) + \left(\frac{1}{h} \frac{d\ell_1(\chi)}{d\chi} u_1 + \ell_1(\chi) u_1 \right) - 1 \right] h d\chi \\
&= \int_{0.5}^{1.0} \left[u_0 \left(\frac{1}{h} \frac{d\ell_0(\chi)}{d\chi} + \ell_0(\chi) \right) + u_1 \left(\frac{1}{h} \frac{d\ell_1(\chi)}{d\chi} + \ell_1(\chi) \right) - 1 \right] h d\chi \\
&= \int_{0.5}^{1.0} \left\{ u_0 \left[\left(\frac{1}{h} \right) \left(\frac{-1}{1} \right) + \frac{1-\chi}{1} \right] + u_1 \left[\left(\frac{1}{h} \right) \left(\frac{1}{1} \right) + \frac{\chi}{1} \right] - 1 \right\} h d\chi \\
&= \left[u_0 \left(\frac{-\chi}{h} + \chi - \frac{\chi^2}{2} \right) \Big|_{0.5}^{1.0} + u_1 \left(\frac{\chi}{h} + \frac{\chi^2}{2} \right) \Big|_{0.5}^{1.0} - \chi \Big|_{0.5}^{1.0} \right] h \\
&= \left\{ u_0 \left[\left(\frac{-1}{h} + 1 - \frac{1}{2} \right) - \left(-\frac{1}{2h} + \frac{1}{2} - \frac{1}{8} \right) \right] + u_1 \left[\left(\frac{1}{h} + \frac{1}{2} \right) - \left(\frac{1}{2h} + \frac{1}{8} \right) \right] - 1/2 \right\} 1/2 \\
&= u_0 \left(-\frac{7}{16} \right) + u_1 \left(\frac{11}{16} \right) - 1/4.
\end{aligned}$$

For element 2, we note that at the local level the integrals are the same as for the case shown in [Eq. \(5.29\)](#), although the values of the coefficients are now u_1 and u_2 .

We now combine the information regarding $w_1(x)$ from both element 1 and element 2. We need both of these integrals to give us the complete integral for $w_1(x)$ and we need the complete integral to form the equation for $w_1(x)$. Thus, combining the integral in [Eq. \(5.29\)](#) and [\(5.31\)](#) we get

(5.32)

$$\begin{aligned}
& \int_{0.25}^{0.50} R(x) dx + \int_{0.50}^{0.75} R(x) dx \\
&= \int_0^{0.5} R(\chi) d\chi + \int_{0.50}^{1.0} R(\chi) d\chi \\
&= \int_{0.5}^{1.0} \left(\frac{d\hat{u}(\chi)}{d\chi} \frac{d\chi}{dx} + \hat{u}(\chi) - 1 \right) \frac{dx}{d\chi} d\chi + \int_0^{0.5} \left(\frac{d\hat{u}(\chi)}{d\chi} \frac{d\chi}{dx} + \hat{u}(\chi) - 1 \right) \frac{dx}{d\chi} d\chi \\
&= [u_0 (-7/16) + u_1 (11/16) - 1/4] \Big|_{e=1} \\
&+ [u_1 (-5/16) + u_2 (9/16) - 1/4] \Big|_{e=2}.
\end{aligned}$$

The two integrals in [Eq. \(5.32\)](#) constitute all the information associated with weighting function $w_1(\chi)$. Thus, we can set their sum to zero and get a complete equation, that is

(5.33)

$$\begin{aligned}
& \int_{0.25}^{0.5} R(x) dx + \int_{0.5}^{0.75} R(x) dx \\
&= [u_0 (-7/16) + u_1 (11/16) - 1/4] \Big|_{\epsilon=1} + [u_1 (-5/16) + u_2 (9/16) - 1/4] \Big|_{\epsilon=2} \\
&= 0.
\end{aligned}$$

Now consider the last weighting function $w_2(x)$. We have only one integral because $w_2(x)$ only appears in the last element. We therefore obtain

(5.34)

$$\begin{aligned}
& \int_{0.75}^{1.0} R(x) dx \\
&= \int_{0.5}^{1.0} \left(\frac{d\hat{u}}{d\chi} \frac{d\chi}{dx} + \hat{u} - 1 \right) \frac{dx}{d\chi} d\chi \\
&= \int_{0.5}^{1.0} \left(\frac{d\hat{u}(\chi)}{d\chi} \frac{1}{h} + \hat{u}(\chi) - 1 \right) h d\chi \\
&= \int_{0.5}^{1.0} \left[\left(\frac{1}{h} \frac{d\ell_0(\chi)}{d\chi} u_0 + \ell_0(\chi) u_0 \right) + \left(\frac{1}{h} \frac{d\ell_1(\chi)}{d\chi} u_1 + \ell_1(\chi) u_1 \right) - 1 \right] h d\chi \\
&= \int_{0.5}^{1.0} \left[u_0 \left(\frac{1}{h} \frac{d\ell_0(\chi)}{d\chi} + \ell_0(\chi) \right) + u_1 \left(\frac{1}{h} \frac{d\ell_1(\chi)}{d\chi} + \ell_1(\chi) \right) - 1 \right] h d\chi \\
&= \int_{0.5}^{1.0} \left\{ u_0 \left[\left(\frac{1}{h} \right) \left(\frac{-1}{1} \right) + \frac{1-\chi}{1} \right] + u_1 \left[\left(\frac{1}{h} \right) \left(\frac{1}{1} \right) + \frac{\chi}{1} \right] - 1 \right\} h d\chi \\
&= \left[u_0 \left(\frac{-\chi}{h} + \chi - \frac{\chi^2}{2} \right) \Big|_{0.5}^{1.0} + u_1 \left(\frac{\chi}{h} + \frac{\chi^2}{2} \right) \Big|_{0.5}^{1.0} - \chi \Big|_{0.5}^{1.0} \right] h \\
&= \left\{ u_0 \left[\left(\frac{-1}{h} + 1 - \frac{1}{2} \right) - \left(-\frac{1}{2h} + \frac{1}{2} - \frac{1}{8} \right) \right] + u_1 \left[\left(\frac{1}{h} + \frac{1}{2} \right) - \left(\frac{1}{2h} + \frac{1}{8} \right) \right] - 1/2 \right\} 1/2 \\
&= u_1 (-7/16) + u_2 (11/16) - 1/4.
\end{aligned}$$

Since this is all the information we will obtain using w_2 , we can write

(5.35)

$$\begin{aligned}
& \int_{0.75}^{1.0} R(x) dx \\
&= u_1 (-7/16) + u_2 (11/16) - 1/4 \\
&= 0.
\end{aligned}$$

We now have three equations, [Eqs. \(5.25\)](#), [\(5.33\)](#) and [\(5.35\)](#) in three unknowns, u_0 , u_1 , and u_2 , which can be written as

$$\begin{aligned}
& [u_0 (-5/16) + u_1 (9/16) - 1/4] |_{e=1} = 0 \quad (5.36) \\
& [u_0 (-7/16) + u_1 (11/16) - 1/4] |_{e=1} + [u_1 (-5/16) + u_2 (9/16) - 1/4] |_{e=2} = 0 \\
& [u_1 (-7/16) + u_2 (11/16) - 1/4] |_{e=2} = 0.
\end{aligned}$$

Let us write this as a matrix equation

$$\begin{bmatrix} -\frac{5}{16} & \frac{9}{16} & 0 \\ -\frac{7}{16} & \frac{11}{16} & \frac{5}{16} \\ 0 & -\frac{7}{16} & \frac{11}{16} \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{bmatrix}. \quad (5.37)$$

However, we have information on u_0 from [Eq. \(5.17\)](#), namely that $u(0) = u_0 = 2$. Thus we can replace u_0 with 2. Now, since we know u_0 we can multiply the first column by 2 and place the result on the right-hand side of the equation. The result is

$$\begin{bmatrix} \frac{9}{16} & 0 \\ \frac{11}{16} & \frac{5}{16} \\ -\frac{7}{16} & \frac{11}{16} \end{bmatrix} \begin{bmatrix} 2 \\ u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} + 2 \left(\frac{5}{16} \right) \\ \frac{1}{2} + 2 \frac{7}{16} \\ \frac{1}{4} \end{bmatrix}. \quad (5.38)$$

Since we know the value of u_0 , we do not need the first line of the matrix and so we remove it along with the first column, which is also no longer needed, to obtain

$$\begin{bmatrix} \frac{3}{8} & \frac{9}{16} \\ -\frac{7}{16} & \frac{11}{16} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \frac{11}{8} \\ \frac{1}{4} \end{bmatrix} \quad (5.39)$$

from which we get

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 1.60 \\ 1.38 \end{bmatrix}. \quad (5.40)$$

With these values of $u_i(x)$ and $u_0(x)$ we can write from the Lagrange expansion

$$\begin{aligned}
\hat{u}(x) &= \ell_0(x) u_0 + \ell_1(x) u_1 + \ell_2(x) u_2 \\
&= \ell_0(x) (2) + \ell_1(x) (1.60) + \ell_2(x) (1.38)
\end{aligned} \quad (5.41)$$

where the x -dependent Lagrange may span two elements. Thus we can obtain a value for the approximation $\hat{u}(x)$ anywhere in the domain of interest. From the analytic solution we have

$$u(0.5) = u_1 = 1 + \exp(-0.5) = 1 + 0.61 = 1.61 \quad (5.42)$$

and

$$u(1.0) = u_2 = 1 + \exp(-1.0) = 1 + 0.37 = 1.37. \quad (5.43)$$

Location	$u(x)$	$\hat{u}(x)$	error
$x = 0.5$	1.61	1.60	0.01
$x = 1.0$	1.37	1.38	0.01

The numerical and analytical solutions are remarkably close given only a three volume approximation was used.

While the above approach to setting up the matrix equation seems like an appropriate way to proceed, this is not the way the problem is solved in practice. Rather, each element is allotted its own 2×2 matrix which is called the **element coefficient matrix**. The elements of the element coefficient matrix are the integrals identified with that element. To see how this works first refer to [Eq. \(5.36\)](#). Note that there is notation indicating from which element each coefficient is obtained. The information in the first row of this equation comes from element 1, the second row from elements 1 and 2 and the third row from element 2. In the following we collect the information from each element.

The element-defined terms in the equations must be (and have been) summed to form complete equations identified with a volume. For example to have all the information associated with volume weighting function w_1 we need to sum the information in the second row of the matrix associated with element 1 and the first row of element 2 in the following [Eq. \(5.44\)](#):

$$\begin{matrix} \begin{bmatrix} -\frac{5}{16} & \frac{9}{16} \\ -\frac{7}{16} & \frac{11}{16} \end{bmatrix} & \begin{bmatrix} u_0 \\ u_1 \end{bmatrix} & \begin{bmatrix} -\frac{5}{16} & \frac{9}{16} \\ -\frac{7}{16} & \frac{11}{16} \end{bmatrix} & \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \end{matrix}, \quad (5.44)$$

Element 1 Element 2

Let us sum all the coefficients in both matrices that are associated with each unknown value of u_i . We get

$$\begin{bmatrix} -\frac{5}{16} & \frac{9}{16} & 0 \\ -\frac{7}{16} & \left(\frac{11}{16} - \frac{5}{16}\right) & \frac{9}{16} \\ 0 & -\frac{7}{16} & \frac{11}{16} \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} \quad (5.45)$$

which is identical to the left hand side of [Eq. \(5.37\)](#). The 3×3 matrix in [Eq. \(5.45\)](#) is called the **global coefficient matrix**. With this information, the equations can be formulated and solved as seen in [Eq. \(5.37\)](#) through [\(5.40\)](#) above. In practice each element coefficient matrix is generated and then the information is transferred from each one to the global matrix.

Recall that we performed the integrations using the local χ coordinate system. To do this, we needed to know the relationship between the location of the nodes in the global x and local χ coordinate systems. This relationship is determined through the use of a table or map such as shown below. For each element the correspondence between the locations of the node numbers in the global and local coordinates are provided.

<i>Element 1</i>		<i>Element 2</i>	
<i>node</i> (x)	<i>node</i> (χ)	<i>node</i> (x)	<i>node</i> (χ)
x_0	χ_0	x_1	χ_0
x_1	χ_1	x_2	χ_1

(5.46)

5.1.2 Finite Difference Interpretation of the Finite Volume Method

Consider an interval $x \in [x_a = 0, x_b = 2h]$. (see [Fig. 5.4](#)). Then the integrals appearing in the preceding example ([Section 5.1.1](#)) for the case of w_1 are

$$\int_0^{2h} R(x) w_1 dx = \int_0^{2h} \left(\frac{d\hat{u}(x)}{dx} + \hat{u}(x) - 1 \right) w_1(x) dx = 0 \quad (5.47)$$

$$\int_0^{2h} \frac{d\hat{u}(x)}{dx} w_1(x) dx = \int_{h/2}^{(3/2)h} \frac{d\hat{u}(x)}{dx} dx = \frac{1}{2} (u_2 - u_0) \quad (5.48)$$

$$\int_0^{2h} \hat{u}(x) w_1(x) dx = \int_{h/2}^{(3/2)h} \hat{u}(x) dx = \left(\frac{1}{8} u_0 + \frac{3}{4} u_1 + \frac{1}{8} u_2 \right) h \quad (5.49)$$

$$\int_0^{2h} (1) w_1(x) dx = \int_{h/2}^{(3/2)h} (1) dx = (1) h. \quad (5.50)$$

Combining the integrals we get

$$\begin{aligned} & \int_0^{2h} \left(\frac{d\hat{u}(x)}{dx} + \hat{u}(x) - 1 \right) w_1(x) dx \\ &= \frac{1}{2} (u_2 - u_0) + \left(\frac{1}{8} u_0 + \frac{3}{4} u_1 + \frac{1}{8} u_2 \right) h - (1) h. \end{aligned} \quad (5.51)$$

Now divide by h , multiply above and below by 2 in the second term, and rearrange the result to get

$$\underbrace{\frac{u_2 - u_0}{2h}}_{\text{second-order correct FD of } \frac{du}{dx}} + \underbrace{\frac{1}{8} (u_0 + 6u_1 + u_2) \frac{2h}{2h}}_{\text{integrated average of } u \text{ over } 2h} - (1). \quad (5.52)$$

Therefore, the finite volume method can be interpreted as being composed of a second-order correct finite difference representation of the spatial derivative and a numerical

estimate of the function integral $\int_0^{2h} u dx$.

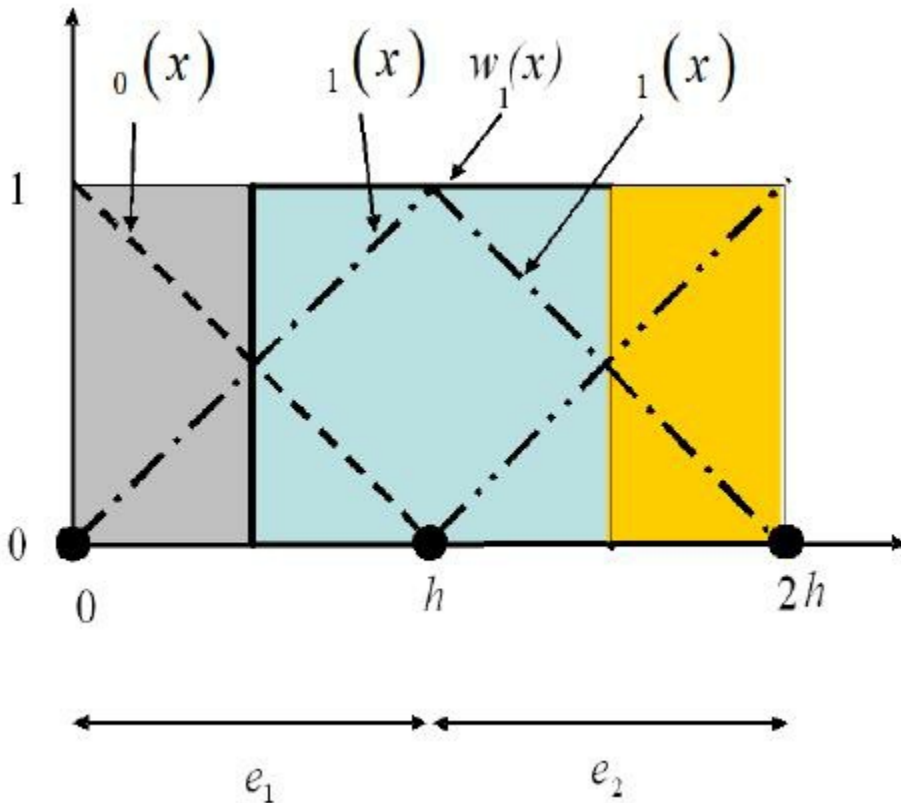


Figure 5.4: Diagrammatic sketch for representing finite volume formula as a finite difference formula.

5.2 Galerkin Method for First Order Equations

In the Galerkin finite element method we simply change the weighting function $w_i(x)$ to be the same as the interpolation or basis function. In other words we let the weighting function be a Lagrange polynomial. To illustrate the concept consider the equation

$$\frac{du}{dx} - f(x) = 0 \quad x \in [x_0 = a, x_n = b] \quad (5.53)$$

and the boundary condition

$$u(0) = u_0. \quad (5.54)$$

From the method of weighted residuals we have

$$\int R(x) w_i(x) dx = 0 \quad i = 0, \dots, n. \quad (5.55)$$

where

$$R(x) = \frac{d\hat{u}(x)}{dx} - f(x). \quad (5.56)$$

Let the weighting function be defined as

$$w_i(x) = \ell_i(x) \quad (5.57)$$

and let us approximate $u(x)$ as

$$u(x) \approx \hat{u}(x) = \sum_{j=1}^n u_j \ell_j(x). \quad (5.58)$$

Then we can substitute [Eq. \(5.58\)](#) into [Eq. \(5.55\)](#) to obtain

$$(5.59)$$

$$\begin{aligned} \int R(x) w_i(x) dx &= \int \left[\frac{d\hat{u}(x)}{dx} - f(x) \right] w_i(x) dx \\ &= \int \left[\frac{d\hat{u}(x)}{dx} - f(x) \right] \ell_i(x) dx \\ &= \int \left[\sum_{j=0}^n \frac{du_j \ell_j(x)}{dx} - f(x) \right] \ell_i(x) dx \quad (\\ &= \sum_{j=0}^n u_j \int \left[\frac{d\ell_j(x)}{dx} \ell_i(x) \right] dx - \int_x f(x) \ell_i(x) dx \\ &= 0 \quad i = 0, \dots, n. \end{aligned}$$

Let us assume that we will use linear Lagrange polynomial basis functions and three elements. Then $n = 3$ (four nodes) and we can write [Eq. \(5.59\)](#) in matrix form as

$$\begin{aligned} & \left[\begin{array}{cccc} \int \frac{d\ell_0(x)}{dx} \ell_0(x) dx & \int \frac{d\ell_1(x)}{dx} \ell_0(x) dx & \int \frac{d\ell_2(x)}{dx} \ell_0(x) dx & \int \frac{d\ell_3(x)}{dx} \ell_0(x) dx \\ \int \frac{d\ell_0(x)}{dx} \ell_1(x) dx & \int \frac{d\ell_1(x)}{dx} \ell_1(x) dx & \int \frac{d\ell_2(x)}{dx} \ell_1(x) dx & \int \frac{d\ell_3(x)}{dx} \ell_1(x) dx \\ \int \frac{d\ell_0(x)}{dx} \ell_2(x) dx & \int \frac{d\ell_1(x)}{dx} \ell_2(x) dx & \int \frac{d\ell_2(x)}{dx} \ell_2(x) dx & \int \frac{d\ell_3(x)}{dx} \ell_2(x) dx \\ \int \frac{d\ell_0(x)}{dx} \ell_3(x) dx & \int \frac{d\ell_1(x)}{dx} \ell_3(x) dx & \int \frac{d\ell_2(x)}{dx} \ell_3(x) dx & \int \frac{d\ell_3(x)}{dx} \ell_3(x) dx \end{array} \right] \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \end{bmatrix} \\ &= \begin{bmatrix} \int f(x) \ell_0(x) dx \\ \int f(x) \ell_1(x) dx \\ \int f(x) \ell_2(x) dx \\ \int f(x) \ell_3(x) dx \end{bmatrix}. \end{aligned} \quad (5.60)$$

Consider next the span of the weighting functions. The span is the number of elements over which a Lagrange polynomial is defined. When linear Lagrange polynomials are used as weighting functions the Lagrange is usually defined over two elements, one on each side of the node. This is not true for the first or last nodes (at x_0 and x_n) which span only one element. The resulting equation is found below as [Eq. \(5.61\)](#). Notice that in the second row, for example, there is a zero in the fourth column. This is because the weighting function $\ell_1(x)$ does not reach the fourth node, that is the one at x_3 . In fact, the maximum number of non-zero coefficients we will see in a row in a problem of this kind is three and they form a band around the diagonal.

$$\begin{aligned}
 & \begin{bmatrix} \int \frac{d\ell_0(x)}{dx} \ell_0(x) dx & \int \frac{d\ell_1(x)}{dx} \ell_0(x) dx & 0 & 0 \\ \int \frac{d\ell_0(x)}{dx} \ell_1(x) dx & \int \frac{d\ell_1(x)}{dx} \ell_1(x) dx & \int \frac{d\ell_2(x)}{dx} \ell_1(x) dx & 0 \\ 0 & \int \frac{d\ell_1(x)}{dx} \ell_2(x) dx & \int \frac{d\ell_2(x)}{dx} \ell_2(x) dx & \int \frac{d\ell_3(x)}{dx} \ell_2(x) dx \\ 0 & 0 & \int \frac{d\ell_2(x)}{dx} \ell_3(x) dx & \int \frac{d\ell_3(x)}{dx} \ell_3(x) dx \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \end{bmatrix} \\
 & = \begin{bmatrix} \int f(x) \ell_0(x) dx \\ \int f(x) \ell_1(x) dx \\ \int f(x) \ell_2(x) dx \\ \int f(x) \ell_3(x) dx \end{bmatrix}.
 \end{aligned} \tag{5.61}$$

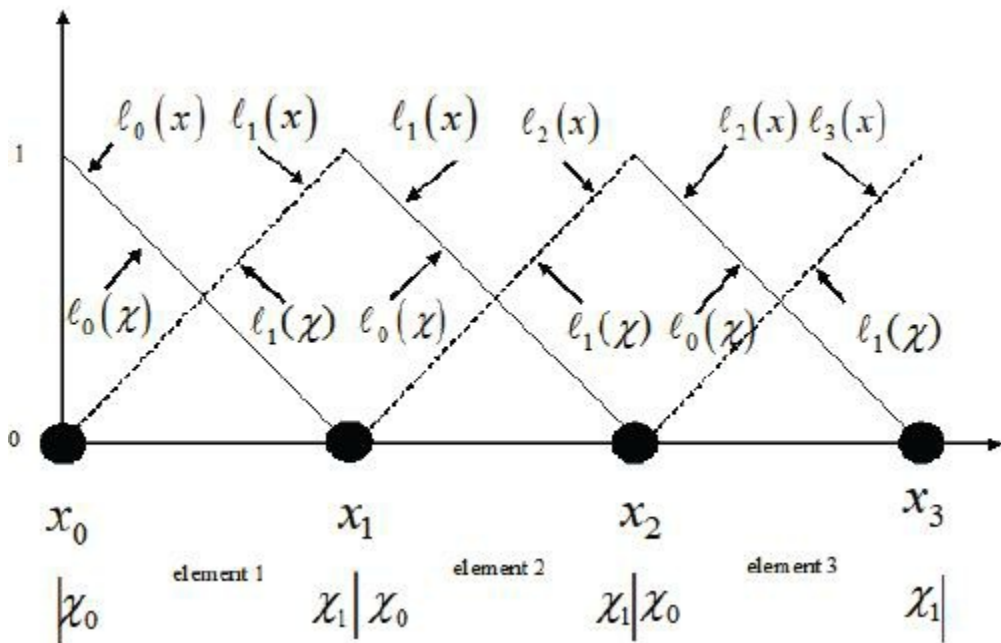


Figure 5.5: Global and local basis functions (Lagrange polynomials) for four-node problem.

Now we can focus on the integrals that remain. For weighting functions that span at most two elements we break the integrations into two one-element integrations. Examination of [Fig. 5.5](#) will facilitate understanding the step that takes us from [Eq.](#)

(5.61) to Eq. (5.62) below. For completeness we also express this matrix equation using Lagrange polynomials defined in terms of χ . We obtain

$$\begin{aligned}
 & \left[\begin{array}{cccc}
 \int_0^1 \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=1} & \int_0^1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=1} & 0 & 0 \\
 \int_0^1 \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \ell_1(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=1} & \int_0^1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \ell_1(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=1} & \int_0^1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=2} & 0 \\
 0 & \int_0^1 \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \ell_1(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=2} & \int_0^1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \ell_1(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=2} & \int_0^1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=3} \\
 0 & 0 & \int_0^1 \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \ell_1(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=3} & \int_0^1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \ell_1(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=3}
 \end{array} \right] \quad (5.62) \\
 & \times \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \int_0^1 f(\chi) \ell_0(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=1} \\ \int_0^1 f(\chi) \ell_1(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=1} + \int_0^1 f(\chi) \ell_0(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=2} \\ \int_0^1 f(\chi) \ell_0(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=2} + \int_0^1 f(\chi) \ell_1(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=3} \\ \int_0^1 f(\chi) \ell_1(\chi) \frac{dx}{d\chi} d\chi \Big|_{e=3} \end{bmatrix}
 \end{aligned}$$

Since integration is only over elements, e.g. spacial increments of $[x_1, x_2]$ or $[\chi_0, \chi_1]$, the Lagrange polynomials are all of the same form for this problem

$$\ell_j(\chi) = \begin{cases} \ell_1(\chi) & \chi \in [0, 1] \\ \ell_0(\chi) & \chi \in [1, 2] \end{cases} \quad (5.63)$$

We now generate the element coefficient matrix for element 1 that contributes to the matrix on the left hand side of 5.62

$$\underbrace{\begin{bmatrix} \int_{x_0}^{x_1} \frac{d\ell_0(x)}{dx} \ell_0(x) dx & \int_{x_0}^{x_1} \frac{d\ell_1(x)}{dx} \ell_0(x) dx \\ \int_{x_0}^{x_1} \frac{d\ell_0(x)}{dx} \ell_1(x) dx & \int_{x_0}^{x_1} \frac{d\ell_1(x)}{dx} \ell_1(x) dx \end{bmatrix}}_{\text{element 1}} \quad (5.64)$$

or using local coordinates

$$\begin{bmatrix} \int_0^1 \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) \frac{dx}{d\chi} d\chi & \int_0^1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) \frac{dx}{d\chi} d\chi \\ \int_0^1 \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \ell_1(\chi) \frac{dx}{d\chi} d\chi & \int_0^1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \ell_1(\chi) \frac{dx}{d\chi} d\chi \end{bmatrix} \quad (5.65)$$

For element 2, we get

$$\underbrace{\begin{bmatrix} \int_{x_1}^{x_2} \frac{d\ell_1(x)}{dx} \ell_1(x) dx & \int_{x_1}^{x_2} \frac{d\ell_2(x)}{dx} \ell_1(x) dx \\ \int_{x_1}^{x_2} \frac{d\ell_1(x)}{dx} \ell_2(x) dx & \int_{x_1}^{x_2} \frac{d\ell_2(x)}{dx} \ell_2(x) dx \end{bmatrix}}_{\text{element 2}} \quad (5.66)$$

or

$$\begin{bmatrix} \int_0^1 \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) \frac{dx}{d\chi} d\chi & \int_0^1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) \frac{dx}{d\chi} d\chi \\ \int_0^1 \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \ell_1(\chi) \frac{dx}{d\chi} d\chi & \int_0^1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \ell_1(\chi) \frac{dx}{d\chi} d\chi \end{bmatrix}. \quad (5.67)$$

As noted the element matrices are the same for elements 1, 2, and 3. This would not be the case if the nodes were not equally spaced since the value of h would change from element to element and this would change the value of the integrals.

As an example of the integration protocol, consider the evaluation of the following two terms that appear in element matrix 1. For the first we have

$$\begin{aligned} & \int_{x_0}^{x_1} \frac{d\ell_0(x)}{dx} \ell_0(x) dx & (5.68) \\ &= \int_0^h \left(-\frac{1}{h}\right) \left(\frac{h-x}{h}\right) dx \\ &= \left(-\frac{x}{h} + \frac{x^2}{2h^2}\right) \Big|_0^h \\ &= -1 + \frac{1}{2} \\ &= -\frac{1}{2} \end{aligned}$$

or in local coordinates

$$\begin{aligned}
& \int_0^1 \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) \frac{dx}{d\chi} d\chi & (5.69) \\
&= \int_0^1 \left(\frac{-1}{1}\right) \left(\frac{1}{h}\right) \left(\frac{1-\chi}{1}\right) \left(\frac{h}{1}\right) d\chi \\
&= \int_0^1 \left(\frac{-1}{1}\right) \left(\frac{1-\chi}{1}\right) d\chi \\
&= \left(-\frac{\chi}{1} + \frac{\chi^2}{2}\right) \Big|_0^1 \\
&= -\frac{1}{2}
\end{aligned}$$

and for the second we obtain

$$\begin{aligned}
& \int_{x_0}^{x_1} \frac{d\ell_1(x)}{dx} \ell_0 dx & (5.70) \\
&= \int_0^h \left(\frac{1}{h}\right) \left(\frac{h-x}{h}\right) dx \\
&= \frac{x}{h} - \frac{x^2}{2h^2} \Big|_0^h \\
&= 1 - \frac{1}{2} \\
&= \frac{1}{2}.
\end{aligned}$$

Expressing the above in local coordinates we can write

$$\begin{aligned}
& \int_0^1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) \frac{dx}{d\chi} d\chi & (5.71) \\
&= \int_0^1 \left(\frac{1}{1}\right) \left(\frac{1}{h}\right) \left(\frac{1-\chi}{1}\right) \left(\frac{h}{1}\right) d\chi \\
&= \int_0^1 \left(\frac{1}{1}\right) \left(\frac{1-\chi}{1}\right) d\chi \\
&= \left(\frac{\chi}{1} - \frac{\chi^2}{2}\right) \Big|_0^1 \\
&= \frac{1}{2}.
\end{aligned}$$

Performing similar integrations for the other elements of the element matrix we obtain

$$\underbrace{\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}}_{\text{element 1}} \quad (5.72)$$

and assuming $f(x) = \alpha$, we get for the term on the right-hand side of [Eq. \(5.62\)](#) (because we assumed α to be constant over an element)

$$\begin{aligned} & \left[\begin{array}{c} \alpha \int_0^1 l_0(x) \frac{dx}{dx} d\chi \\ \int_0^1 f(x) l_1(x) \frac{dx}{dx} d\chi + \int_0^1 f(x) l_0(x) \frac{dx}{dx} d\chi \\ \int_0^1 f(x) l_0(x) \frac{dx}{dx} d\chi + \int_0^1 f(x) l_1(x) \frac{dx}{dx} d\chi \\ \int_0^1 f(x) l_1(x) \frac{dx}{dx} d\chi \end{array} \right] \quad (5.73) \\ &= \left[\begin{array}{c} \alpha \int_0^1 \frac{1-x}{1} h d\chi \\ \alpha \int_0^1 \frac{x}{1} h d\chi + \alpha \int_0^1 \frac{1-x}{1} h d\chi \\ \alpha \int_0^1 \frac{1-x}{1} h d\chi + \alpha \int_0^1 \frac{x}{1} h d\chi \\ \alpha \int_0^1 \frac{x}{1} h d\chi \end{array} \right] \\ &= \left[\begin{array}{c} \alpha \left(x - \frac{x^2}{2} \right) h \Big|_0^1 \\ \alpha \frac{x^2}{2} h \Big|_0^1 + \alpha \left(x - \frac{x^2}{2} \right) h \Big|_0^1 \\ \alpha \left(x - \frac{x^2}{2} \right) h \Big|_0^1 + \alpha \frac{x^2}{2} h \Big|_0^1 \\ \alpha \frac{x^2}{2} h \Big|_0^1 = \alpha \frac{x^2}{2} h \Big|_0^1 \end{array} \right] \\ &= \alpha \begin{bmatrix} \frac{h}{2} \\ h \\ h \\ \frac{h}{2} \end{bmatrix}. \end{aligned}$$

If we now replace the integrals in [Eq. \(5.62\)](#) we obtain

$$\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & \left(\frac{1}{2} - \frac{1}{2}\right) & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \left(\frac{1}{2} - \frac{1}{2}\right) & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \frac{h\alpha}{2} \\ h\alpha \\ h\alpha \\ \frac{h\alpha}{2} \end{bmatrix} \quad (5.74)$$

or

$$\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \frac{h\alpha}{2} \\ h\alpha \\ h\alpha \\ \frac{h\alpha}{2} \end{bmatrix}. \quad (5.75)$$

The manner in which the element matrices are organized to produce the global coefficient matrix in [Eq. \(5.74\)](#) is illustrated in [Fig. 5.6](#). Each element matrix is indicated by the dashed perimeter.

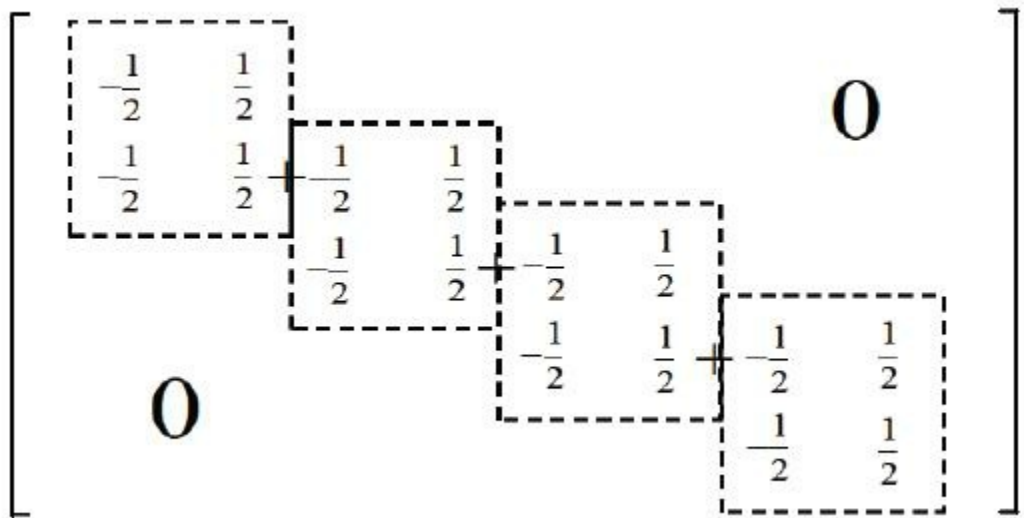


Figure 5.6: Arrangement of element coefficient matrices to form a global matrix. Each box with a dashed perimeter is an element coefficient matrix.

Another way to look at this information is shown in [Fig. 5.7](#). In this rendition one can see the interconnectedness of the information from element to element.

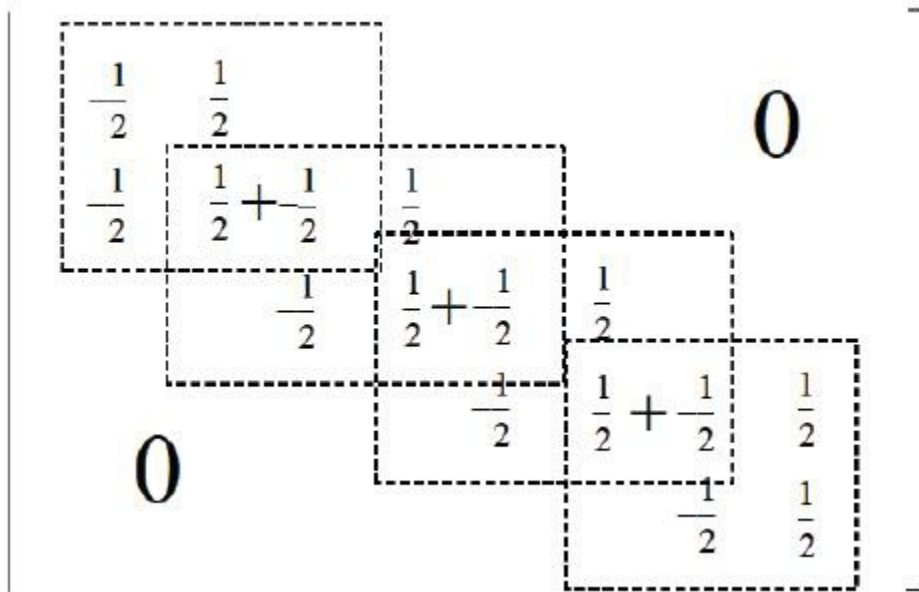


Figure 5.7: Selection of matrix elements to illustrate overlap of element information.

To permit a comparison with the analytic solution let $u_0 = 1$, $\alpha = 2$ and $h = 3$; then we have

$$\begin{bmatrix} 0 & 1/2 & 0 \\ -1/2 & 0 & 1/2 \\ 0 & -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} (3)(2) + \frac{1}{2} \\ (3)(2) \\ \frac{(3)(2)}{2} \end{bmatrix} \quad (5.76)$$

or

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 13 \\ 19 \end{bmatrix}. \quad (5.77)$$

If we now enter these coefficient values into the approximating series for $\hat{u}(x)$ the result is

$$\hat{u}(x) = \ell_0(x) + 7\ell_1(x) + 13\ell_2(x) + 19\ell_3(x). \quad (5.78)$$

The analytic solution is

$$u(x) = 2x + 1. \quad (5.79)$$

If you select an x value at a node such that only one term in [Eq. \(5.78\)](#) is non-zero it is easy to see that the solution is exact. This is to be expected since the solution is a linear function of x , so it can be exactly represented by a sum of linear Lagrange polynomials.

5.2.1 Finite-Difference Interpretation of the Galerkin Approximation

If one expands the matrix product appearing on the left-hand-side of [Eq. \(5.74\)](#) it is seen that the result is a series of finite difference approximations to the first derivative. In viewing the following, keep in mind that each equation is identified with the node that corresponds to the equation. For example, the first equation is identified with an approximation centered at node zero, the second at node 1 etc.

$$\frac{u_1 - u_0}{h} = \alpha \quad \text{a forward difference} \quad (5.80)$$

$$\frac{u_2 - u_0}{2h} = \alpha \quad \text{a centered difference} \quad (5.81)$$

$$\frac{u_3 - u_1}{2h} = \alpha \quad \text{a centered difference} \quad (5.82)$$

$$\frac{u_3 - u_2}{h} = \alpha \quad \text{a backward difference.} \quad (5.83)$$

From the above we observe that, but for the first and last nodes in the sequence, the finite element approximation in this particular case is equivalent to a second-order correct finite difference scheme.

5.3 Galerkin Method for Second-Order Equations

Consider the second-order equation

$$\frac{d^2u(x)}{dx^2} + g_1 \frac{du(x)}{dx} + g_2 u(x) = 0 \quad x \in [x_0, x_n] \quad (5.84)$$

$$u(0) = u_0 \quad (5.85)$$

$$\left. \frac{du(x)}{dx} \right|_{x_n} = u'. \quad (5.86)$$

As earlier, let the approximation for $u(x)$ be defined as

$$u(x) \approx \hat{u}(x) = \sum_{j=0}^n u_j \ell_j(x). \quad (5.87)$$

Let us define the residual associated with [Eq. \(5.84\)](#) as

$$\frac{d^2\hat{u}(x)}{dx^2} + g_1 \frac{d\hat{u}(x)}{dx} + g_2 \hat{u}(x) = R(x). \quad (5.88)$$

As in the case of first-order equations, from the theory underlying the Galerkin finite

element method we use the basis function as the weighting function to define the residual. We obtain

$$\int_{x_0}^{x_n} R(x) \ell_i(x) dx = 0 \quad i = 0, 1, \dots, n. \quad (5.89)$$

Substitute [Eq. \(5.88\)](#) into [\(5.89\)](#) to obtain

$$\int_{x_0}^{x_n} \left(\frac{d^2 \hat{u}(x)}{dx^2} + g_1 \frac{d\hat{u}(x)}{dx} + g_2 \hat{u}(x) \right) \ell_i(x) dx = 0 \quad i = 0, 1, \dots, n. \quad (5.90)$$

Consider the second-order derivative term in this expression. There is a problem here because, if one looks at the behavior of the second derivative of a linear Lagrange polynomial, one observes that at a node the derivative is discontinuous. While such a discontinuity can be accommodated, it is not convenient to do so.

To avoid this problem we proceed as follows. Using **integration by parts** for the interval x_0 to x_n we have

$$\int_{x_0}^{x_n} \frac{d}{dx} \left(\frac{d\hat{u}(x)}{dx} \ell_i(x) \right) dx = \int_{x_0}^{x_n} \left(\frac{d\hat{u}(x)}{dx} \frac{d\ell_i(x)}{dx} + \frac{d^2 \hat{u}(x)}{dx^2} \ell_i(x) \right) dx \quad i = 0, \dots, n \quad (5.91)$$

but

$$\int_{x_0}^{x_n} \frac{d}{dx} \left(\frac{d\hat{u}(x)}{dx} \ell_i(x) \right) dx = \left(\frac{d\hat{u}(x)}{dx} \ell_i(x) \right) \Big|_{x_0}^{x_n}. \quad (5.92)$$

Now let's consider the right-hand-side term in [Eq. \(5.92\)](#). First we expand it to give

$$- \frac{d\hat{u}(x)}{dx} \ell_0(x) \Big|_{x_0} \quad i = 0 \quad (5.93)$$

$$\frac{d\hat{u}(x)}{dx} \ell_i(x) \Big|_{x_0}^{x_n} \quad i = 0, \dots, n \quad (5.94)$$

and

$$\frac{d\hat{u}(x)}{dx} \ell_n(x) \Big|_{x_n} \quad i = n. \quad (5.95)$$

Notice that while there are $n + 1$ possible terms in [Eq. \(5.94\)](#), only two survive [Eq. \(5.95\)](#). This is because the values of ℓ_i for $i \neq 0$ or n vanish at $x = x_0$ and $x = x_n$, the endpoints of the interval. Combination of [Eq. \(5.92\)](#) and [Eq. \(5.91\)](#) gives us

$$\int_{x_0}^{x_n} \frac{d^2 \hat{u}(x)}{dx^2} \ell_i(x) dx = - \int_{x_0}^{x_n} \frac{d\hat{u}(x)}{dx} \frac{d\ell_i(x)}{dx} dx \quad (5.96)$$

$$+ \frac{d\hat{u}(x)}{dx} \ell_i(x) \Big|_{x_0}^{x_n} \quad i = 0, \dots, n.$$

Let us now use [Eq. \(5.96\)](#) to rewrite [Eq. \(5.90\)](#) as

$$\int_{x_0}^{x_n} \left(- \frac{d\hat{u}(x)}{dx} \frac{d\ell_i(x)}{dx} + g_1 \frac{d\hat{u}(x)}{dx} \ell_i(x) + g_2 \hat{u}(x) \ell_i(x) \right) dx \quad (5.97)$$

$$+ \frac{d\hat{u}(x)}{dx} \ell_n(x) \Big|_{x_n} - \frac{d\hat{u}(x)}{dx} \ell_0(x) \Big|_{x_0} = 0 \quad i = 0, \dots, n.$$

Expanding the approximation $\hat{u}(x)$ and combining the last two terms we obtain

$$\sum_{j=0}^n u_j \int_{x_0}^{x_n} \left(- \frac{d\ell_j(x)}{dx} \frac{d\ell_i(x)}{dx} + g_1 \frac{d\ell_j(x)}{dx} \ell_i(x) + g_2 \ell_j(x) \ell_i(x) \right) dx \quad (5.98)$$

$$+ \left(\frac{d\hat{u}}{dx} \ell_i(x) \right) \Big|_{x_0}^{x_n} = 0 \quad i = 0, \dots, n.$$

Now consider element-wise integration of the terms in [Eq. \(5.98\)](#) assuming g_1 and g_2 to be given constants. For the first element we obtain

$$\int_{x_0}^{x_1} \left[-u_0 \frac{d\ell_0(x)}{dx} \frac{d\ell_0(x)}{dx} - u_1 \frac{d\ell_1(x)}{dx} \frac{d\ell_0(x)}{dx} \right. \quad (5.99)$$

$$+ g_1 \left(u_0 \frac{d\ell_0(x)}{dx} \ell_0(x) + u_1 \frac{d\ell_1(x)}{dx} \ell_0(x) \right)$$

$$\left. + g_2 (u_0 \ell_0(x) \ell_0(x) + u_1 \ell_1(x) \ell_0(x)) \right] dx$$

or, using local coordinates

$$\begin{aligned}
& \int_0^1 \left[-u_0 \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} - u_1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \right. \\
& + g_1 \left(u_0 \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) + u_1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) \right) \\
& \left. + g_2 (u_0 \ell_0(\chi) \ell_0(\chi) + u_1 \ell_1(\chi) \ell_0(\chi)) \right] \frac{dx}{d\chi} d\chi.
\end{aligned} \tag{5.100}$$

If we consider one term in [Eq. \(5.100\)](#) at a time we can evaluate the integrals defined in the local χ coordinate system. In the case of the weighting function $\ell_0(\chi)$ we obtain for the second-order terms,

$$\begin{aligned}
& - \int_0^1 \left(\frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \right) \frac{dx}{d\chi} d\chi \\
& = - \int_0^1 \left(\left(-1 \frac{1}{h} \right) \left(-1 \frac{1}{h} \right) \right) \frac{h}{1} d\chi \\
& = - \frac{\chi}{h} \Big|_0^1 = -\frac{1}{h}
\end{aligned} \tag{5.101}$$

and

$$\begin{aligned}
& - \int_0^1 \left(\frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \right) \frac{dx}{d\chi} d\chi \\
& = - \int_0^1 \left(\left(1 \frac{1}{h} \right) \left(-1 \frac{1}{h} \right) \right) \frac{h}{1} d\chi \\
& = \frac{\chi}{h} \Big|_0^1 = \frac{1}{h}.
\end{aligned} \tag{5.102}$$

For the first-order terms and using the weighting function $\ell_0(\chi)$ we get

$$\begin{aligned}
& \int_0^1 \frac{d\ell_0(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) \frac{dx}{d\chi} d\chi \\
& = \int_0^1 \left(-1 \frac{1}{h} \right) \frac{1-\chi}{1} \frac{h}{1} d\chi \\
& = -\chi + \frac{\chi^2}{2} \Big|_0^1 = -\frac{1}{2}
\end{aligned} \tag{5.103}$$

and

$$\begin{aligned}
& \int_0^1 \frac{d\ell_1(\chi)}{d\chi} \frac{d\chi}{dx} \ell_0(\chi) \frac{dx}{d\chi} d\chi & (5.104) \\
&= \int_0^1 \left(\frac{1}{h} \right) \frac{1-\chi}{1} \frac{h}{1} d\chi \\
&= \chi - \frac{\chi^2}{2} \Big|_0^1 = \frac{1}{2}.
\end{aligned}$$

Finally for the zero order terms we obtain

$$\begin{aligned}
& \int_0^1 \ell_0(\chi) \ell_0(\chi) \frac{dx}{d\chi} d\chi & (5.105) \\
&= \int_0^1 \frac{1-\chi}{1} \frac{1-\chi}{1} \frac{h}{1} d\chi \\
&= \left(\chi - \chi^2 + \frac{\chi^3}{3} \right) h \Big|_0^1 = \frac{h}{3}
\end{aligned}$$

and, finally

$$\begin{aligned}
& \int_0^1 \ell_1(\chi) \ell_0(\chi) \frac{dx}{d\chi} d\chi & (5.106) \\
&= \int_0^1 \frac{\chi}{1} \frac{1-\chi}{1} \frac{h}{1} d\chi \\
&= \left(\frac{\chi^2}{2} - \frac{\chi^3}{3} h \right) \Big|_0^1 = \frac{h}{6}.
\end{aligned}$$

At this point we have evaluations for each term in [Eq. \(5.100\)](#). To see how this fits into the scheme of things, we now generate an element coefficient matrix. Since each element has two nodes and each node is identified with both an equation and the approximation of the unknown value of $\hat{u}(\chi)$, each element coefficient matrix will be 2×2 . In looking at the element coefficient matrices remember that the rows are associated with the weighting functions and the columns are associated with the unknown u_i values. Thus the first rows of the following matrices are associated with the weighting function $\ell_0(\chi)$ and the first columns are associated with the term in the expansion of $\hat{u}(\chi)$ associated with u_0 :

$$\begin{aligned}
& - \int_0^1 \left[\begin{array}{cccc} \frac{d\ell_0(x)}{dx} & \frac{d\ell_0(x)}{dx} & \frac{d\ell_0(x)}{dx} & \frac{d\ell_0(x)}{dx} \\ \frac{d\ell_1(x)}{dx} & \frac{d\ell_1(x)}{dx} & \frac{d\ell_1(x)}{dx} & \frac{d\ell_1(x)}{dx} \end{array} \right] \frac{dx}{dx} dx \\
& + \int_0^1 g_1 \left[\begin{array}{cc} \frac{d\ell_0(x)}{dx} \ell_0(x) & \frac{d\ell_1(x)}{dx} \ell_0(x) \\ \frac{d\ell_0(x)}{dx} \ell_1(x) & \frac{d\ell_1(x)}{dx} \ell_1(x) \end{array} \right] \frac{dx}{dx} dx \\
& + \int_0^1 g_2 \left[\begin{array}{cc} \ell_0(x) \ell_0(x) & \ell_1(x) \ell_0(x) \\ \ell_0(x) \ell_1(x) & \ell_1(x) \ell_1(x) \end{array} \right] \frac{dx}{dx} dx.
\end{aligned} \tag{5.107}$$

It is now possible to provide numbers for the coefficients in [Eq. \(5.107\)](#) using [Eqs. \(5.101\)](#) through [\(5.106\)](#). The result is

$$\left[\begin{array}{cc} -\frac{1}{h} & \frac{1}{h} \\ \frac{1}{h} & -\frac{1}{h} \end{array} \right] + g_1 \left[\begin{array}{cc} -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{array} \right] + g_2 \left[\begin{array}{cc} \frac{h}{3} & \frac{h}{6} \\ \frac{h}{6} & \frac{h}{3} \end{array} \right]. \tag{5.108}$$

As long as the elements remain the same length and the functions g_1 and g_2 remain constant, the above element matrices are all that is needed since each element matrix will be the same. One simply uses the same element matrix for each element along the interval $x = x_0$ to $x = x_n$.

Now consider the last term in [Eq. \(5.98\)](#) for the special case of the equation written for the weighting function $\ell_0(x)$. Only information in element 1, the first element, has a contribution to this equation, viz.

$$\left(\frac{d\hat{u}(x)}{dx} \ell_0(x) \right) \Big|_{x_0}. \tag{5.109}$$

But since ℓ_0 is unity at x_0

$$\left(\frac{d\hat{u}(x)}{dx} \ell_0(x) \right) \Big|_{x_0} = \frac{d\hat{u}(x)}{dx} \Big|_{x_0}. \tag{5.110}$$

Since the term $\left(\frac{d\hat{u}}{dx} \right) \Big|_{x_0}$ is associated with the flux evaluated at the boundary, it is a **flux-**

type boundary condition. Thus, a *second type* (also known as a Neumann or flux boundary condition), is imbedded in the equation approximation directly.

Let us now return to our original problem. For interior nodes (not including the boundary nodes of 0 and n) we have, using [Eqs. \(5.90\)](#) and [\(5.91\)](#)

$$\int_{x_0}^{x_n} \left\{ \left(-\frac{d\hat{u}(x)}{dx} \frac{d\ell_i(x)}{dx} \right) + \left[\left(g_1 \frac{d\hat{u}(x)}{dx} + g_2 \hat{u}(x) \right) \ell_i(x) \right] \right\} dx = 0 \quad i = 1, \dots, n-1 \tag{5.111}$$

and after substitution of the approximation in [Eq. \(5.87\)](#) into [Eq. \(5.111\)](#) we get

(5.112)

$$\sum_{j=1}^{n-1} u_j \int_{x_0}^{x_n} \left(-\frac{d\ell_j(x)}{dx} \frac{d\ell_i(x)}{dx} + g_1 \frac{d\ell_j(x)}{dx} \ell_i(x) + g_2 \ell_j(x) \ell_i(x) \right) dx$$

$$= 0 \quad i = 1, \dots, n-1.$$

For node zero we have information only from the first element

$$\sum_{j=0}^1 u_j \int_{x_0}^{x_1} \left(-\frac{d\ell_j(x)}{dx} \frac{d\ell_0(x)}{dx} \right) + \left[\left(g_1 \frac{d\ell_j(x)}{dx} + g_2 \ell_j(x) \right) \ell_0(x) \right] dx - \quad (5.113)$$

$$\left(\frac{d\hat{u}(x)}{dx} \ell_0(x) \right) \Big|_{x_0} = 0. \quad (5.114)$$

Similarly for node n we have

$$\sum_{j=n-1}^n u_j \int_{x_{n-1}}^{x_n} \left(-\frac{d\ell_j(x)}{dx} \frac{d\ell(x)_n}{dx} \right) + \quad (5.115)$$

$$\left[\left(g_1 \frac{d\ell_j(x)}{dx} + g_2 \ell_j(x) \right) \ell(x)_n \right] dx +$$

$$\left(\frac{d\hat{u}}{dx} \ell(x)_n \right) \Big|_{x_n} = 0.$$

This provides us with $n + 1$ equations in $n + 1$ unknowns, which we can assemble into a matrix equation of the form (after accommodation of boundary conditions)

$$\begin{bmatrix} a_{00} & a_{01} & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{10} & a_{11} & a_{12} & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{21} & a_{22} & a_{23} & 0 & 0 & 0 & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & \cdot \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & a_{n(n-1)} & a_{nn} & \cdot \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{bmatrix} = \begin{bmatrix} \frac{d\hat{u}}{dx} \Big|_{x_0} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -\frac{d\hat{u}}{dx} \Big|_{x_n} \end{bmatrix} \quad (5.116)$$

where

$$a_{ij} = \int_{x_0}^{x_n} \left(-\frac{d\ell_j(x)}{dx} \frac{d\ell_i(x)}{dx} + g_1 \frac{d\ell_j(x)}{dx} \ell_i(x) + g_2 \ell_i(x) \right) dx \quad (5.117)$$

$$i = 1, \dots, n-1 \quad (5.118)$$

$$u_i = \hat{u}(x_i).$$

From [Eq. \(5.86\)](#) we have

$$u(0) = u_0 \quad (5.119)$$

$$\left. \frac{du(x)}{dx} \right|_{x_n} = u' \quad (5.120)$$

where u_0 and u' are given. Substitution of this information into [Eq. \(5.116\)](#), then multiplication of the known value u_0 by column 1 and transfer of this information to the right hand side of the equation, and finally elimination of the first row and column, provides

$$\begin{bmatrix} a_{11} & a_{12} & 0 & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} & 0 & 0 & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & a_{n(n-1)} & a_{nn} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{bmatrix} = \begin{bmatrix} -a_{10}u_0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -u' \end{bmatrix}. \quad (5.121)$$

As shown earlier, the necessary integrations can be performed in the local χ coordinate system using linear Lagrange functions $\ell_j(\chi)$ $j = 0, 1$. The integrals for each term, defined in the global coordinate system are presented below for each of three cases depending upon the matrix indices. These are the same integrals we evaluated earlier in [Eqs. \(5.107\)](#) and [\(5.108\)](#) using local coordinates, but now we have combined information from the two elements associated with a common node to get the equivalent global values. This is evident from the fact that the limits of integration in the following equations are from $-h$ to h and therefore represent a span of two elements.

For the second-order term we have:

$$\int_{-h}^h -\frac{d\ell_j(x)}{dx} \frac{d\ell_i(x)}{dx} dx = \frac{1}{h} \quad j < i \quad (5.122)$$

$$\int_{-h}^h -\frac{d\ell_j(x)}{dx} \frac{d\ell_i(x)}{dx} dx = -\frac{2}{h} \quad j = i \quad i, j \neq n, 0 \quad (5.123)$$

$$\int_{-h}^h -\frac{d\ell_j(x)}{dx} \frac{d\ell_i(x)}{dx} dx = \frac{1}{h} \quad j > i \quad (5.124)$$

$$\int_0^h -\frac{d\ell_j(x)}{dx} \frac{d\ell_i(x)}{dx} dx = -\frac{1}{h} \quad j = i = 0 \quad (5.125)$$

$$\int_{-h}^0 -\frac{d\ell_j(x)}{dx} \frac{d\ell_i(x)}{dx} dx = -\frac{1}{h} \quad j = i = n. \quad (5.126)$$

For the first-order term we have

$$\int_{-h}^h g_1 \frac{d\ell_j(x)}{dx} \ell_i(x) dx = -\frac{g_1}{2} \quad j < i \quad (5.127)$$

$$\int_{-h}^h g_1 \frac{d\ell_j(x)}{dx} \ell_i(x) dx = 0 \quad j = i \quad i, j \neq n, 0 \quad (5.128)$$

$$\int_{-h}^h g_1 \frac{d\ell_j(x)}{dx} \ell_i(x) dx = \frac{g_1}{2} \quad j > i \quad (5.129)$$

$$\int_0^h g_1 \frac{d\ell_j(x)}{dx} \ell_i(x) dx = -\frac{g_1}{2} \quad j = i = 0 \quad (5.130)$$

$$\int_{-h}^0 g_1 \frac{d\ell_j(x)}{dx} \ell_i(x) dx = \frac{g_1}{2} \quad j = i = n. \quad (5.131)$$

For the zero order term we have

$$\int_{-h}^h g_2 \ell_j(x) \ell_i(x) dx = \frac{g_2}{6} h \quad j < i \quad (5.132)$$

$$\int_{-h}^h g_2 \ell_j(x) \ell_i(x) dx = \frac{g_2 h}{3} + \frac{g_2 h}{3} = \frac{2g_2 h}{3} \quad j = i \quad i, j \neq n, 0 \quad (5.133)$$

$$\int_{-h}^h g_2 \ell_j(x) \ell_i(x) dx = \frac{g_2 h}{6} \quad j > i \quad (5.134)$$

$$\int_0^h g_2 \ell_j(x) \ell_i(x) dx = \frac{g_2 h}{3} \quad j = i = 0 \quad (5.135)$$

$$\int_{-h}^0 g_2 \ell_j(x) \ell_i(x) dx = \frac{g_2 h}{3} \quad j = i = n. \quad (5.136)$$

As noted earlier, the coefficients of the known state-variable values (first column of the matrix) are multiplied by the known state-variable values and the result is placed on the right-hand side of the equation. The redundant first column and first row in the matrix equation are then removed. Subsequent substitution of the integral values into the resulting element coefficient matrices, and then into the global coefficient matrix, yields the specific form of [Eq. \(5.121\)](#) appropriate for the general form of the problem being considered, that is

$$\begin{bmatrix} -\frac{2}{h} + 0 + \frac{2g_2 h}{3} & \frac{1}{h} + \frac{g_1}{2} + \frac{g_2 h}{6} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{h} - \frac{g_1}{2} + \frac{g_2 h}{6} & -\frac{2}{h} + 0 + \frac{2g_2 h}{3} & \frac{1}{h} + \frac{g_1}{2} + \frac{g_2 h}{6} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{h} - \frac{g_1}{2} + \frac{g_2 h}{6} & -\frac{2}{h} + 0 + \frac{2g_2 h}{3} & \frac{1}{h} + \frac{g_1}{2} + \frac{g_2 h}{6} & 0 & 0 & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\ & & & & & \frac{1}{h} - \frac{g_1}{2} + \frac{g_2 h}{6} & -\frac{1}{h} + \frac{g_1}{2} + \frac{g_2 h}{3} \end{bmatrix} \times \quad (5.137)$$

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{bmatrix} = - \begin{bmatrix} \left(\frac{1}{h} - \frac{g_1}{2} + \frac{g_2 h}{6} \right) u_0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ u \end{bmatrix}.$$

To come up with some concrete numbers *we assume the following*: $h = 1$; $g_1 = 1$; $g_2 = 0.5$; $u_0 = 1$; $\frac{du}{dx} \Big|_{x_n} = 0$; $n = 3$. The resulting set of equations is

$$\begin{bmatrix} -2 + \frac{1}{3} & 1 + \frac{1}{2} + \frac{1}{12} & 0 \\ 1 - \frac{1}{2} + \frac{1}{12} & -2 + \frac{1}{3} & 1 + \frac{1}{2} + \frac{1}{12} \\ 0 & 1 - \frac{1}{2} + \frac{1}{12} & -1 + \frac{1}{2} + \frac{1}{6} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} -1 \left(1 - \frac{1}{2} + \frac{1}{12} \right) \\ 0 \\ 0 \end{bmatrix} \quad (5.138)$$

$$\begin{bmatrix} -1.66 & 1.58 & 0 \\ 0.58 & -1.66 & 1.58 \\ 0 & 0.58 & -0.33 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} -0.58 \\ 0 \\ 0 \end{bmatrix} \quad (5.139)$$

with the solution

$$[u_1, u_2, u_3]^T = [0.234, -0.121, -0.213]^T. \quad (5.140)$$

The analytic solution is

$$u(x) = \exp(-x/2)[\cos(x/2) + \tan(\pi/4 + 3/2)\sin(x/2)] \quad (5.141)$$

and a comparison of numerical and analytical values is presented in the following table

Location	Analytical	Numerical	Difference
$x = 1$	0.200	0.234	0.036
$x = 2$	- 0.158	- 0.121	0.037
$x = 3$	- 0.241	- 0.213	0.028

This is not a bad (but not great) result considering the number of nodes appearing in the approximating equation.

5.3.1 Finite Difference Interpretation of Second-Order Galerkin Method

In this subsection we will show how the finite difference and finite element approximations for the second-order equation considered above are related. Let us consider each term in the second row of [Eq. \(5.137\)](#). After each term is divided by h and the final expression in [Eq. \(5.144\)](#) is multiplied above and below by 2 we obtain

$$\frac{d^2u(x)}{dx^2} \approx \frac{1}{h^2} (u_1 - 2u_2 + u_3) \quad (5.142)$$

$$g_1 \frac{du(x)}{dx} \approx -\frac{g_1 u_1}{2h} + (0) u_2 + \frac{g_1 u_3}{2h} \quad (5.143)$$

$$g_2 u(x) \approx \frac{g_2 h}{6h} (u_1 + 4u_2 + u_3) = \frac{g_2}{6} (u_1 + 4u_2 + u_3) \frac{2h}{2h}. \quad (5.144)$$

Combination of these terms yields

$$\underbrace{\frac{1}{h^2} (u_1 - 2u_2 + u_3)}_{\text{Centered FD}} + \underbrace{\frac{g_1}{2h} (u_3 - u_1)}_{\text{Centered FD}} + \underbrace{\frac{g_2}{6} (u_1 + 4u_2 + u_3) \frac{2h}{2h}}_{\text{Averaged Simpsons rule for } \int_{x_0}^{x_2} g_2 u dx} = 0 \quad (5.145)$$

From this equation we see that each full equation (not the ones associated with the end nodes) consists of a second-order correct finite-difference approximation to both the first and second-order terms in the original differential equation. For the zero-order term the value of $\int_{x_0}^{x_2} g_2 u dx$ is obtained using Simpson's rule divided by the the length of the interval over which the integration is being performed, that is $2h$.

5.4 Finite Volume Method for Second-Order Equations

We have seen how to formulate the finite volume approximations in one dimension for first-order equations. We now extend that work to consider equations with second derivatives. Consider once again the equation

$$\frac{d^2 u(x)}{dx^2} + g_1 \frac{du(x)}{dx} + g_2 u(x) = 0 \quad x \in [x = x_0, x = x_n] \quad (5.146)$$

$$u(0) = u_0 \quad (5.147)$$

$$\left. \frac{du}{dx} \right|_{x_n} = u'. \quad (5.148)$$

From the method of weighted residuals we have

$$\int_{x_0}^{x_n} R(x) w_i(x) dx = 0 \quad i = 0, 1, \dots, n \quad (5.149)$$

where the weighting function $w_i(x)$ is, as earlier, defined as (see [Fig. 5.8](#))

$$w_i(x) = \begin{cases} 1 & x \in \left[x_i - \frac{h}{2}, x_i + \frac{h}{2} \right] \\ 0 & x \notin \left[x_i - \frac{h}{2}, x_i + \frac{h}{2} \right] \end{cases}. \quad (5.150)$$

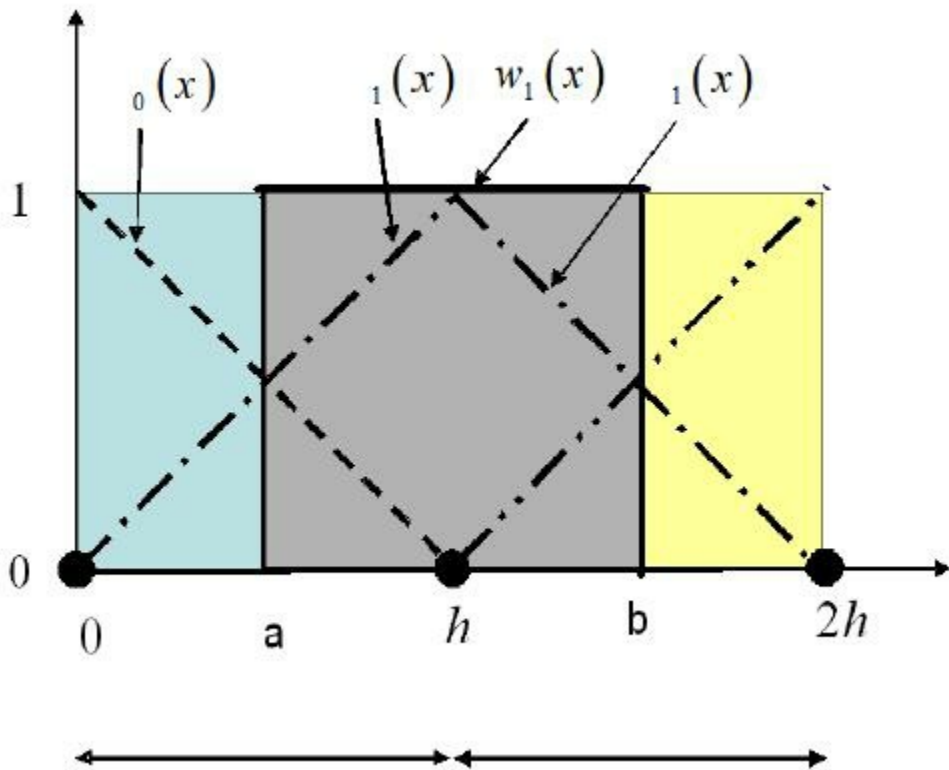


Figure 5.8: Basis functions $\ell_j(x)$ and weighting functions $w_i(x)$ for finite volume method.

The residual is defined as earlier as

$$\frac{d^2 \hat{u}(x)}{dx^2} + g_1 \frac{d\hat{u}(x)}{dx} + g_2 \hat{u}(x) = R(x). \quad (5.151)$$

Substitution of [Eq. \(5.151\)](#) into [Eq. \(5.149\)](#) gives

$$\int_{x_0}^{x_n} \left(\frac{d^2 \hat{u}(x)}{dx^2} + g_1 \frac{d\hat{u}(x)}{dx} + g_2 \hat{u}(x) \right) w_i(x) dx = 0 \quad i = 0, 1, \dots, n. \quad (5.152)$$

Let us now apply integration by parts to the second-order term to obtain

$$\begin{aligned} & \int_{x_0}^{x_n} \left(\frac{d^2 \hat{u}(x)}{dx^2} + g_1 \frac{d\hat{u}(x)}{dx} + g_2 \hat{u}(x) \right) w_i(x) dx \quad (5.153) \\ &= \int_{x_0}^{x_n} \left\{ \left(-\frac{d\hat{u}(x)}{dx} \frac{dw_i(x)}{dx} \right) + \left[\left(g_1 \frac{d\hat{u}(x)}{dx} + g_2 \hat{u}(x) \right) w_i(x) \right] \right\} dx \\ & \quad + \frac{d\hat{u}(x)}{dx} w_i \Big|_{x_n} - \frac{d\hat{u}(x)}{dx} w_i \Big|_{x_0} \\ &= 0 \quad i = 0, 1, \dots, n. \end{aligned}$$

Consider the second-order term integrated over the first element, that is

$$- \int_0^h \frac{d\hat{u}(x)}{dx} \frac{dw_i(x)}{dx} dx \quad i = 0, 1. \quad (5.154)$$

The term of special interest is

$$\frac{dw_i(x)}{dx} \quad i = 0, 1. \quad (5.155)$$

The derivative of a step function such as w_i is defined at the point $x = a$ (as shown in [Fig. 5.8](#)) as

$$\frac{dw_0(x)}{dx} = -\delta(x - a) \quad (5.156)$$

and

$$\frac{dw_1(x)}{dx} = +\delta(x - a) \quad (5.157)$$

where $\delta_i(x - x_i)$ is the **Dirac delta function**. The definition of the Dirac delta function is such that

$$\int_{-\infty}^{\infty} f(x) \delta(x - a) = f(a). \quad (5.158)$$

Thus

$$- \int_0^h \frac{d\hat{u}(x)}{dx} [-\delta(x - a)] dx = \frac{d\hat{u}(x)}{dx} \Big|_{x=a} \quad (5.159)$$

$$- \int_0^h \frac{d\hat{u}(x)}{dx} [\delta(x - a)] dx = - \frac{d\hat{u}(x)}{dx} \Big|_{x=a}. \quad (5.160)$$

If we now introduce our approximation for $u(x)$

$$u(x) \approx \hat{u}(x) = \sum_{j=0}^1 u_j \ell_j(x) \quad (5.161)$$

into the right-hand side of [Eq. \(5.160\)](#), the term $-\frac{d\hat{u}(x)}{dx} \Big|_{x=a}$ can be written

$$- \frac{d\hat{u}(x)}{dx} \Big|_{x=a} = - \left[u_0 \frac{d\ell_0(x)}{dx} + u_1 \frac{d\ell_1(x)}{dx} \right]_{x=a}. \quad (5.162)$$

Let us now consider the terms which, in our Galerkin development, gave rise to the imbedded second type boundary conditions. These are, for one element of length h

$$\frac{d\hat{u}(x)}{dx} w_i(x) \Big|_0^h \quad i = 0, 1 \quad (5.163)$$

which upon expansion gives

$$\frac{d\hat{u}(x)}{dx} w_0(x) \Big|_0^h = \frac{d\hat{u}(x)}{dx} \Big|_h (0) - \frac{d\hat{u}(x)}{dx} \Big|_0 (1) = - \frac{d\hat{u}(x)}{dx} \Big|_0 \quad (5.164)$$

$$\frac{d\hat{u}(x)}{dx} w_1(x) \Big|_0^h = \frac{d\hat{u}(x)}{dx} \Big|_h (1) - \frac{d\hat{u}(x)}{dx} \Big|_0 (0) = \frac{d\hat{u}(x)}{dx} \Big|_h. \quad (5.165)$$

Note that the terms $\frac{d\hat{u}(x)}{dx} \Big|_h (0)$ and $-\frac{d\hat{u}(x)}{dx} \Big|_0 (0)$ vanish because $w_0(x)|_h$ is 0 and $w_1(x)|_0$ is 0.

We now formulate the coefficient matrices for [Eq. \(5.153\)](#). We assume that the values for each element coefficient matrix are

Element 1

(5.166)

u_0	u_1	
$-\int_0^h \frac{dl_0(x)}{dx} \frac{dw_0(x)}{dx} dx$ $+g_1 \int_0^h \frac{dl_0(x)}{dx} dx + g_2 \int_0^a l_0(x) dx$	$-\int_0^h \frac{dl_1(x)}{dx} \frac{dw_0(x)}{dx} dx$ $+g_1 \int_0^h \frac{dl_1(x)}{dx} dx + g_2 \int_0^a l_1(x) dx$	$w_0(x)$
$-\int_0^h \frac{dl_0(x)}{dx} \frac{dw_1(x)}{dx} dx$ $+g_1 \int_a^h \frac{dl_0(x)}{dx} dx + g_2 \int_a^h l_0(x) dx$	$-\int_0^h \frac{dl_1(x)}{dx} \frac{dw_1(x)}{dx} dx$ $+g_1 \int_a^h \frac{dl_1(x)}{dx} dx + g_2 \int_a^h l_1(x) dx$	$w_1(x)$

which can be transformed using [Eqs. \(5.154\)](#) through [\(5.160\)](#) to give

Element 1

(5.167)

u_0	u_1	
$\left. \frac{dl_0(x)}{dx} \right _a + g_1 \int_0^h \frac{dl_0(x)}{dx} dx$ $+g_2 \int_0^a l_0(x) dx$	$\left. \frac{dl_1(x)}{dx} \right _a + g_1 \int_0^h \frac{dl_1(x)}{dx} dx +$ $g_2 \int_0^a l_1(x) dx$	$w_0(x)$
$-\left. \frac{dl_0(x)}{dx} \right _a + g_1 \int_a^h \frac{dl_0(x)}{dx} dx$ $+g_2 \int_a^h l_0(x) dx$	$-\left. \frac{dl_1(x)}{dx} \right _a + g_1 \int_a^h \frac{dl_1(x)}{dx} dx$ $+g_2 \int_a^h l_1(x) dx$	$w_1(x)$

or in local coordinates

Element 1

(5.168)

u_0	u_1	
$\left. \frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} \right _{0.5} + g_1 \int_0^{0.5} \frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} \frac{dx}{d\chi} d\chi$ $+g_2 \int_0^{0.5} l_0(\chi) \frac{d\chi}{dx} d\chi$	$\left. \frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} \right _{0.5} + g_1 \int_0^{0.5} \frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} \frac{dx}{d\chi} d\chi$ $+g_2 \int_0^{0.5} l_1(\chi) \frac{d\chi}{dx} d\chi$	$w_0(\chi)$
$-\left. \frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} \right _{0.5} + g_1 \int_{0.5}^1 \frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} \frac{dx}{d\chi} d\chi$ $+g_2 \int_{0.5}^1 l_0(\chi) \frac{d\chi}{dx} d\chi$	$-\left. \frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} \right _{0.5} + g_1 \int_{0.5}^1 \frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} \frac{dx}{d\chi} d\chi$ $+g_2 \int_{0.5}^1 l_1(\chi) \frac{d\chi}{dx} d\chi$	$w_1(\chi)$

For element 2 we have

Element 2

(5.169)

u_1	u_2	
$-\int_h^{2h} \frac{dl_1(x)}{dx} \frac{dw_1(x)}{dx} dx +$ $g_1 \int_h^b \frac{dl_1(x)}{dx} dx + g_2 \int_h^b l_1(x) dx$	$-\int_h^{2h} \frac{dl_2(x)}{dx} \frac{dw_1(x)}{dx} dx +$ $g_1 \int_h^b \frac{dl_2(x)}{dx} dx + g_2 \int_h^b l_2(x) dx$	$w_1(x)$
$-\int_h^{2h} \frac{dl_1(x)}{dx} \frac{dw_2(x)}{dx} dx +$ $g_1 \int_b^{2h} \frac{dl_1(x)}{dx} dx + g_2 \int_b^{2h} l_1(x) dx$	$-\int_h^{2h} \frac{dl_2(x)}{dx} \frac{dw_2(x)}{dx} dx +$ $g_1 \int_b^{2h} \frac{dl_2(x)}{dx} dx + g_2 \int_b^{2h} l_2(x) dx$	$w_2(x)$

or again using [Eqs. \(5.154\)](#) through [\(5.160\)](#)

Element 2

(5.170)

u_1	u_2	
$\frac{dl_1(x)}{dx} \Big _b + g_1 \int_b^b \frac{dl_1(x)}{dx} dx$ $+g_2 \int_b^b l_1(x) dx$	$\frac{dl_2(x)}{dx} \Big _b + g_1 \int_b^b \frac{dl_2(x)}{dx} dx$ $+g_2 \int_b^b l_2(x) dx$	$w_1(x)$
$-\frac{dl_1(x)}{dx} \Big _b + g_1 \int_b^{2h} \frac{dl_1(x)}{dx} dx$ $+g_2 \int_b^{2h} l_1(x) dx$	$-\frac{dl_2(x)}{dx} \Big _b + g_1 \int_b^{2h} \frac{dl_2(x)}{dx} dx$ $+g_2 \int_b^{2h} l_2(x) dx$	$w_2(x)$

In local coordinates we have

Element 2

(5.171)

u_1	u_2	
$\frac{dl_0(\chi)}{d\chi} \frac{dx}{dx} \Big _{0.5} + g_1 \int_0^{0.5} \frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} \frac{dx}{dx} d\chi$ $+g_2 \int_0^{0.5} l_0(\chi) \frac{dx}{d\chi} d\chi$	$\frac{dl_1(\chi)}{d\chi} \frac{dx}{dx} \Big _{0.5} + g_1 \int_0^{0.5} \frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} \frac{dx}{dx} d\chi$ $+g_2 \int_0^{0.5} l_1(\chi) \frac{dx}{d\chi} d\chi$	$w_0(\chi)$
$-\frac{dl_0(\chi)}{d\chi} \frac{dx}{dx} \Big _{0.5} + g_1 \int_{0.5}^1 \frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} \frac{dx}{dx} d\chi$ $+g_2 \int_{0.5}^1 l_0(\chi) \frac{dx}{d\chi} d\chi$	$-\frac{dl_1(\chi)}{d\chi} \frac{dx}{dx} \Big _{0.5} + g_1 \int_{0.5}^1 \frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} \frac{dx}{dx} d\chi$ $+g_2 \int_{0.5}^1 l_1(\chi) \frac{dx}{d\chi} d\chi$	$w_1(\chi)$

One observes that, as in earlier examples, given h is the same in the two elements, *the element matrices in local (and also in global) coordinates are identical.*

Now we combine the two matrices given by [Eqs. \(5.167\)](#) and [\(5.170\)](#) to create the global matrix. We could also have used the element matrices in local coordinates, that is [Eqs. \(5.168\)](#) and [\(5.171\)](#)

(5.172)

$$\begin{aligned}
 & \left[\begin{array}{ccc}
 \frac{dl_0(x)}{dx} \Big|_a + & \frac{dl_1(x)}{dx} \Big|_a + & \\
 g_1 \int_0^a \frac{dl_0(x)}{dx} dx & g_1 \int_0^a \frac{dl_1(x)}{dx} dx + & 0 \\
 + g_2 \int_0^a l_0(x) dx & g_2 \int_0^a l_1(x) dx & \\
 \\
 -\frac{dl_0(x)}{dx} \Big|_a + & -\frac{dl_1(x)}{dx} \Big|_a + & \frac{dl_2(x)}{dx} \Big|_b + \\
 g_1 \int_a^h \frac{dl_0(x)}{dx} dx + & g_1 \int_a^h \frac{dl_1(x)}{dx} dx + & g_1 \int_h^b \frac{dl_2(x)}{dx} dx + \\
 g_2 \int_a^h l_0(x) dx & g_2 \int_a^h l_1(x) dx + & g_2 \int_h^b l_2(x) dx \\
 \\
 & \frac{dl_1(x)}{dx} \Big|_b + & \\
 & g_1 \int_h^b \frac{dl_1(x)}{dx} dx + & \\
 & g_2 \int_h^b l_1(x) dx - & \\
 \\
 & \frac{dl_1(x)}{dx} \Big|_b + & -\frac{dl_2(x)}{dx} \Big|_b + \\
 & g_1 \int_b^{2h} \frac{dl_1(x)}{dx} dx + & g_1 \int_b^{2h} \frac{dl_2(x)}{dx} dx + \\
 & g_2 \int_b^{2h} l_1(x) dx & g_2 \int_b^{2h} l_2(x) dx
 \end{array} \right] \\
 \\
 \times \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -\frac{d\hat{u}(x)}{dx} \Big|_0 \\ \frac{d\hat{u}(x)}{dx} \Big|_h - \frac{d\hat{u}(x)}{dx} \Big|_h \\ \frac{d\hat{u}(x)}{dx} \Big|_{x=2h} \end{bmatrix}
 \end{aligned}$$

5.4.1 Example of Finite-Volume Solution of a Second-Order Equation

Consider the following equation:

$$\frac{d^2 u(x)}{dx^2} + \frac{du(x)}{dx} + 0.5u(x) = 0 \quad x \in [x=0, x=2] \quad (5.173)$$

$$u(0) = 1 \quad (5.174)$$

$$\left. \frac{du(x)}{dx} \right|_{x=2} = 0. \quad (5.175)$$

Using three nodes and the finite volume method we will determine the solution at $x = 0.5$

$$\int_{x_0}^{x_n} R(x) w_i(x) dx = 0 \quad i = 0, 1, 2. \quad (5.176)$$

The weighting function is, as earlier, defined as (see [Fig. 5.9](#))

$$w_i(x) = \begin{cases} 1 & x \in \left[x_i - \frac{h}{2}, x_i + \frac{h}{2} \right] \\ 0 & x \notin \left[x_i - \frac{h}{2}, x_i + \frac{h}{2} \right] \end{cases}. \quad (5.177)$$

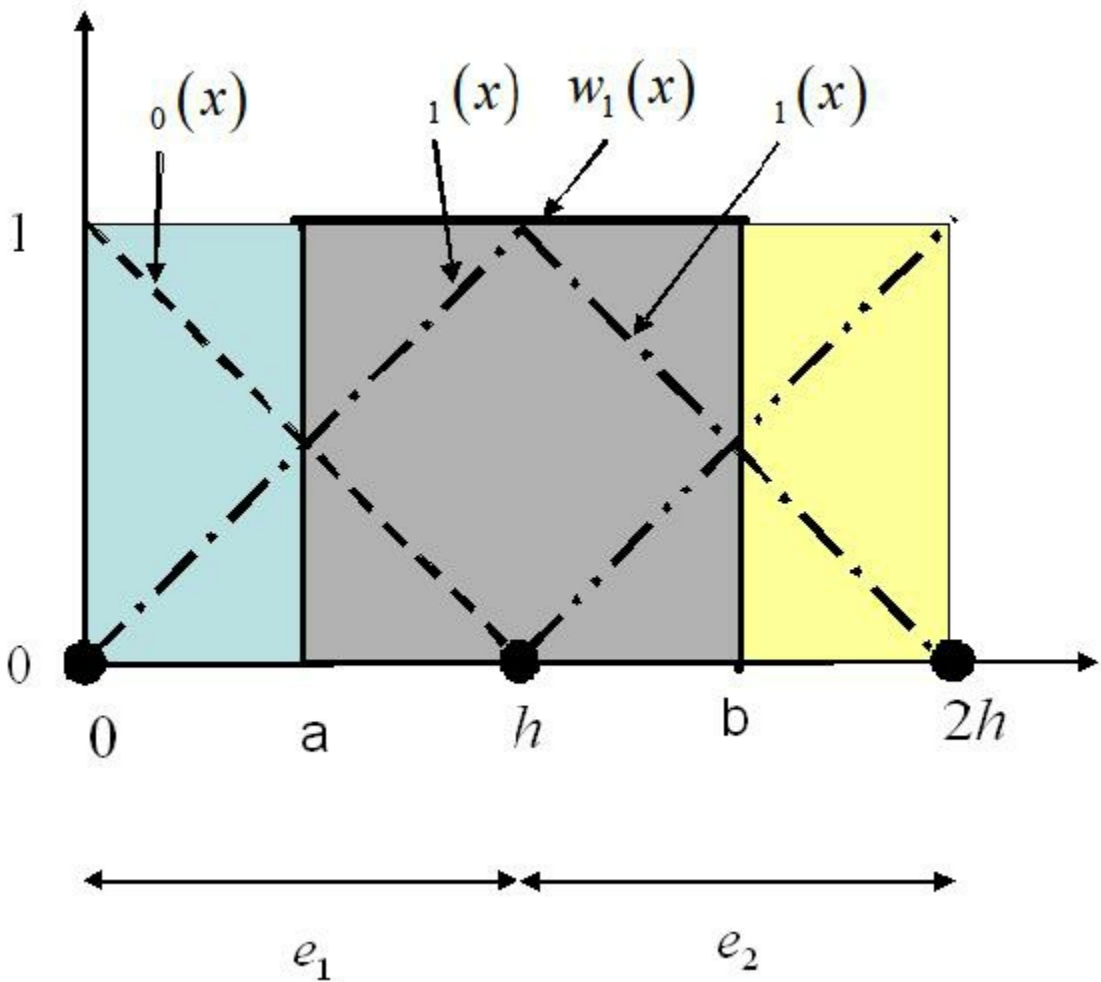


Figure 5.9: As presented earlier, basis functions $\ell_j(x)$ and weighting functions $w_i(x)$ for the finite volume method.

The residual is defined as earlier as

$$\frac{d^2 \hat{u}(x)}{dx^2} + \frac{d\hat{u}(x)}{dx} + 0.5\hat{u}(x) = R(x). \quad (5.178)$$

Substitution of [Eq. \(5.178\)](#) into [\(5.176\)](#) gives

$$\int_0^2 \left(\frac{d^2 \hat{u}(x)}{dx^2} + \frac{d\hat{u}(x)}{dx} + 0.5\hat{u}(x) \right) w_i(x) dx = 0 \quad i = 0, 1, 2. \quad (5.179)$$

Let us now use integration by parts to obtain

$$\begin{aligned} & \int_0^2 \left(\frac{d^2 \hat{u}(x)}{dx^2} + g_1 \frac{d\hat{u}(x)}{dx} + g_2 \hat{u}(x) \right) w_i(x) dx \quad (5.180) \\ &= \int_0^2 \left[\left(-\frac{d\hat{u}(x)}{dx} \frac{dw_i(x)}{dx} \right) + \left(g_1 \frac{d\hat{u}(x)}{dx} + g_2 \hat{u}(x) \right) w_i(x) \right] dx + \\ & \quad \frac{d\hat{u}(x)}{dx} w_i \Big|_2 - \frac{d\hat{u}(x)}{dx} w_i \Big|_0 = 0 \quad i = 0, 1, 2. \end{aligned}$$

Consider the second-order term integrated over the *first element*, that is

$$- \int_0^h \frac{d\hat{u}(x)}{dx} \frac{dw_i(x)}{dx} dx \quad i = 0, 1. \quad (5.181)$$

The term of interest is

$$\frac{dw_i(x)}{dx} \quad i = 0, 1. \quad (5.182)$$

It is defined as

$$\frac{dw_0(x)}{dx} = -\delta(x - h/2) \quad (5.183)$$

and

$$\frac{dw_1(x)}{dx} = +\delta(x - h/2) \quad (5.184)$$

Substitution of [Eqs. \(5.183\)](#) and [\(5.184\)](#) into [Eq. \(5.181\)](#) results in

$$\begin{aligned}
& - \int_0^h \frac{d\hat{u}(x)}{dx} [-\delta_0(x - h/2)] dx = \frac{d\hat{u}(x)}{dx} \Big|_{x=h/2} \quad (5.185) \\
& = \left(u_0 \frac{d\ell_0(x)}{dx} + u_1 \frac{d\ell_1(x)}{dx} \right) \Big|_{x=h/2}
\end{aligned}$$

$$\begin{aligned}
& - \int_0^h \frac{d\hat{u}(x)}{dx} [\delta_1(x - h/2)] dx = - \frac{d\hat{u}(x)}{dx} \Big|_{x=h/2} \quad (5.186) \\
& = - \left(u_0 \frac{d\ell_0(x)}{dx} + u_1 \frac{d\ell_1(x)}{dx} \right) \Big|_{x=h/2}.
\end{aligned}$$

Consider now how to represent the terms

$$\frac{d\hat{u}(x)}{dx} w_i(x) \Big|_0^h \quad i = 0, 1. \quad (5.187)$$

which expanded out, give

$$\begin{aligned}
& \frac{d\hat{u}(x)}{dx} w_0(x) \Big|_0^h \quad (5.188) \\
& = \frac{d\hat{u}(x)}{dx} \Big|_h (0) - \frac{d\hat{u}(x)}{dx} \Big|_0 (1) \\
& = - \frac{d\hat{u}(x)}{dx} \Big|_0
\end{aligned}$$

and

$$\begin{aligned}
& \frac{d\hat{u}(x)}{dx} w_1(x) \Big|_0^h \quad (5.189) \\
& = \frac{d\hat{u}(x)}{dx} \Big|_h (1) - \frac{d\hat{u}(x)}{dx} \Big|_0 (0) \\
& = \frac{d\hat{u}(x)}{dx} \Big|_h.
\end{aligned}$$

If the term represented in the first line of [Eq. \(5.188\)](#) appears on the left hand side of the domain of interest, it represents a second or Neumann type boundary condition. If a term like that found in [Eq. \(5.189\)](#) is located on the right-hand side of the domain, it also represents a second or Neumann type boundary condition.

Finally, consider now the terms

(5.190)

$$\begin{aligned}
\int_0^h (0.5) \hat{u}(x) w_0(x) dx &= (0.5) \int_0^{h/2} \hat{u}(x) dx = (0.5) \int_0^{h/2} u_0(x) \ell_0 + u_1(x) \ell_1(x) dx \\
&= (0.5) u_0 \int_0^{0.5} \frac{h-x}{h} dx + (0.5) u_1 \int_0^{0.5} \frac{x}{h} dx \\
&= (0.5) (u_0) (3/8) + (0.5) (u_1) (1/8).
\end{aligned}$$

We now formulate the coefficient matrices for [Eq. \(5.180\)](#) that is.

Element 1

(5.191)

u_0	u_1	
$\left. \frac{dl_0(x)}{dx} \right _{0.5} +$	$\left. \frac{dl_1(x)}{dx} \right _{0.5} +$	
$\int_0^{0.5} \left(\frac{dl_0(x)}{dx} + \frac{1}{2} \ell_0(x) \right) dx$	$\int_0^{0.5} \left(\frac{dl_1(x)}{dx} + \frac{1}{2} \ell_1(x) \right) dx$	$w_0(x)$
$-\left. \frac{dl_0(x)}{dx} \right _{0.5} +$	$-\left. \frac{dl_1(x)}{dx} \right _{0.5} +$	
$\int_{0.5}^1 \left(\frac{dl_0(x)}{dx} + \frac{1}{2} \ell_0(x) \right) dx$	$\int_{0.5}^1 \left(\frac{dl_1(x)}{dx} + \frac{1}{2} \ell_1(x) \right) dx$	$w_1(x)$

which in local coordinates becomes

Element 1

(5.192)

u_0	u_1	
$\left. \frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} \right _{0.5} +$	$\left. \frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} \right _{0.5} +$	
$\int_0^{0.5} \left(\frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} + \frac{1}{2} \ell_0(\chi) \right) \frac{dx}{d\chi} d\chi$	$\int_0^{0.5} \left(\frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} + \frac{1}{2} \ell_1(\chi) \right) \frac{dx}{d\chi} d\chi$	$w_0(\chi)$
$-\left. \frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} \right _{0.5} +$	$-\left. \frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} \right _{0.5} +$	
$\int_{0.5}^1 \left(\frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} + \frac{1}{2} \ell_0(\chi) \right) \frac{dx}{d\chi} d\chi$	$\int_{0.5}^1 \left(\frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} + \frac{1}{2} \ell_1(\chi) \right) \frac{dx}{d\chi} d\chi$	$w_1(\chi)$

Consider also element 2

Element 2

(5.193)

u_1		u_2		
$\left. \frac{dl_1(x)}{dx} \right _{1.5} +$		$\left. \frac{dl_2(x)}{dx} \right _{1.5} +$		
$\int_1^{1.5} \left(\frac{dl_1(x)}{dx} + \frac{1}{2} l_1(x) \right) dx$		$\int_1^{1.5} \left(\frac{dl_2(x)}{dx} + \frac{1}{2} l_2(x) \right) dx$		$w_1(x)$
$\left. -\frac{dl_1(x)}{dx} \right _{1.5} +$		$\left. -\frac{dl_2(x)}{dx} \right _{1.5} +$		
$\int_{1.5}^2 \left(\frac{dl_1(x)}{dx} + \frac{1}{2} l_1(x) \right) dx$		$\int_{1.5}^2 \left(\frac{dl_2(x)}{dx} + \frac{1}{2} l_2(x) \right) dx$		$w_2(x)$

which using local coordinates yields

Element 2

(5.194)

u_1		u_2		
$\left. \frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} \right _{0.5} +$		$\left. \frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} \right _{0.5} +$		
$\int_0^{0.5} \left(\frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} + \frac{1}{2} l_0(\chi) \right) \frac{dx}{d\chi} d\chi$		$\int_0^{0.5} \left(\frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} + \frac{1}{2} l_1(\chi) \right) \frac{dx}{d\chi} d\chi$		$w_0(\chi)$
$\left. -\frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} \right _{0.5} +$		$\left. -\frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} \right _{0.5} +$		
$\int_{0.5}^1 \left(\frac{dl_0(\chi)}{d\chi} \frac{d\chi}{dx} + \frac{1}{2} l_0(\chi) \right) \frac{dx}{d\chi} d\chi$		$\int_{0.5}^1 \left(\frac{dl_1(\chi)}{d\chi} \frac{d\chi}{dx} + \frac{1}{2} l_1(\chi) \right) \frac{dx}{d\chi} d\chi$		$w_1(\chi)$

Since the element matrices are identical in local coordinates (because the space increment is constant) we will consider only the element 1 matrix hereinafter.

We first introduce the local Lagrange polynomials and the derivative transforms evaluated for this problem into [Eqs. \(5.192\)](#) and [\(5.194\)](#) to obtain

Element 1

(5.195)

u_0		u_1		
$\left(-\frac{1}{1} \right) \left(\frac{1}{1} \right) +$		$\left(\frac{1}{1} \right) \left(\frac{1}{1} \right) +$		
$\int_0^{0.5} \left(\left[-\frac{1}{1} \right] \left[\frac{1}{1} \right] + \frac{1}{2} (1 - \chi) \right) \left(\frac{1}{1} \right) d\chi$		$\int_0^{0.5} \left(\left[\frac{1}{1} \right] \left[\frac{1}{1} \right] + \frac{1}{2} (\chi) \right) \left(\frac{1}{1} \right) d\chi$		$w_0(\chi)$
$\left(\frac{1}{1} \right) \left(\frac{1}{1} \right) +$		$\left(-\frac{1}{1} \right) \left(\frac{1}{1} \right) +$		
$\int_{0.5}^1 \left(\left[-\frac{1}{1} \right] \left[\frac{1}{1} \right] + \frac{1}{2} (1 - \chi) \right) \left(\frac{1}{1} \right) d\chi$		$\int_{0.5}^1 \left(\left[\frac{1}{1} \right] \left[\frac{1}{1} \right] + \frac{1}{2} (\chi) \right) \left(\frac{1}{1} \right) d\chi$		$w_1(\chi)$

Element 1

(5.196)

u_0		u_1		
$-1 + \int_0^{0.5} \left(-1 + \frac{1}{2} (1 - \chi) \right) d\chi$		$1 + \int_0^{0.5} \frac{1}{2} \left(1 + \frac{1}{2} (\chi) \right) d\chi$		$w_0(\chi)$
$1 + \int_{0.5}^1 \left(-1 + \frac{1}{2} (1 - \chi) \right) d\chi$		$-1 + \int_{0.5}^1 \frac{1}{2} \left(1 + \frac{1}{2} (\chi) \right) d\chi$		$w_1(\chi)$

Element 1

u_0	u_1	
$-1 - \frac{1}{2} + \frac{3}{16}$	$1 + \frac{1}{2} + \frac{1}{16}$	w_0
$1 - \frac{1}{2} + \frac{1}{16}$	$-1 + \frac{1}{2} + \frac{3}{16}$	w_1

(5.197)

Now we combine the two element matrices to create the global matrix

$$\begin{aligned}
 & \begin{bmatrix} -1 - \frac{1}{2} + \frac{3}{16} & 1 + \frac{1}{2} + \frac{1}{16} & 0 \\ 1 - \frac{1}{2} + \frac{1}{16} & \left(-1 + \frac{1}{2} + \frac{3}{16}\right) + \left(-1 - \frac{1}{2} + \frac{3}{16}\right) & 1 + \frac{1}{2} + \frac{1}{16} \\ 0 & 1 - \frac{1}{2} + \frac{1}{16} & -1 + \frac{1}{2} + \frac{3}{16} \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} \quad (5.198) \\
 & = \begin{bmatrix} \frac{d\hat{u}}{dx} \Big|_0 \\ \frac{d\hat{u}}{dx} \Big|_1 - \frac{d\hat{u}}{dx} \Big|_1 \\ -\frac{d\hat{u}}{dx} \Big|_2 \end{bmatrix}
 \end{aligned}$$

or, adding up the components of each term we get

$$\frac{1}{16} \begin{bmatrix} -21 & 25 & 0 \\ 9 & -26 & 25 \\ 0 & 9 & -5 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -\frac{d\hat{u}}{dx} \Big|_0 \\ 0 \\ \frac{d\hat{u}}{dx} \Big|_2 \end{bmatrix}. \quad (5.199)$$

Now we impose the boundary conditions and reduce the matrix to get

$$\frac{1}{16} \begin{bmatrix} -26 & 25 \\ 9 & -5 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \frac{1}{16} \begin{bmatrix} 9(1) \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (5.200)$$

Taking the inverse and rearranging the equation we arrive at the solution for u_1 and u_2 , that is

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \frac{1}{(-26)(-5) - (9)(25)} \begin{bmatrix} -5 & -25 \\ -9 & -26 \end{bmatrix} \begin{bmatrix} -9 \\ 0 \end{bmatrix} \quad (5.201)$$

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \frac{1}{-95} \begin{bmatrix} -5 & -25 \\ -9 & -26 \end{bmatrix} \begin{bmatrix} -9 \\ 0 \end{bmatrix} \quad (5.202)$$

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -0.474 \\ -0.853 \end{bmatrix}. \quad (5.203)$$

To obtain a solution at $\hat{u}(0.5)$ you use the approximation for the first element, that is

$$\hat{u}(0.5) = \ell_0(0.5)u_0 + \ell_1(0.5)u_1 \quad (5.204)$$

or

$$\begin{aligned} \hat{u}(0.5) &= 1/2u_0 + 1/2u_1 \\ &= (1/2)1 - (1/2)0.47 \\ &= 0. \end{aligned} \quad (5.205)$$

The analytical solution is

$$u(x) = \exp(-x/2)[\cos(x/2) + \tan(\pi/4 + 1)\sin(x/2)]. \quad (5.206)$$

This gives

$$u(x = 0.5) = -0.129 \quad (5.207)$$

$$u(x = 1) = -0.802 \quad (5.208)$$

$$u(x = 2) = -0.122 \quad (5.209)$$

A comparison between the analytic and numerical results is given in the following table.

x	<i>Analytic</i>	<i>Numerical</i>	<i>Difference</i>
0.5	- 0.129	0.265	0.384
1	- 0.802	- 0.474	0.328
2	- 1.22	- 0.853	0.36

and in [Fig. 5.10](#)

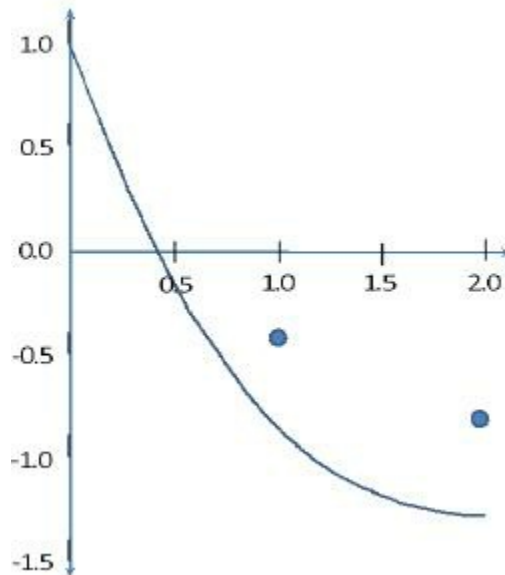


Figure 5.10: Graph of the analytical solution to Eqs. given in [Eqs. \(5.206\)–\(5.147\)](#). The black dots are the computed values using three nodes (two element) and the finite volume method.

This is not a particularly impressive result and this is due to the form of the function $u(x)$ and the way the third term in the differential equations is approximated numerically. To provide confidence that the algorithm provides a solution that does, indeed, converge to the correct solution as h approaches zero, we plot in [Fig. 5.11](#) the relationship between the error between the analytical and numerical solutions and h as h decreases. It is evident from this graph that the method is second-order accurate (the slope of the line is 2.0) and the error is decreasing as the square of h as expected. See Section 4.3 on page 69 for a discussion of convergence.

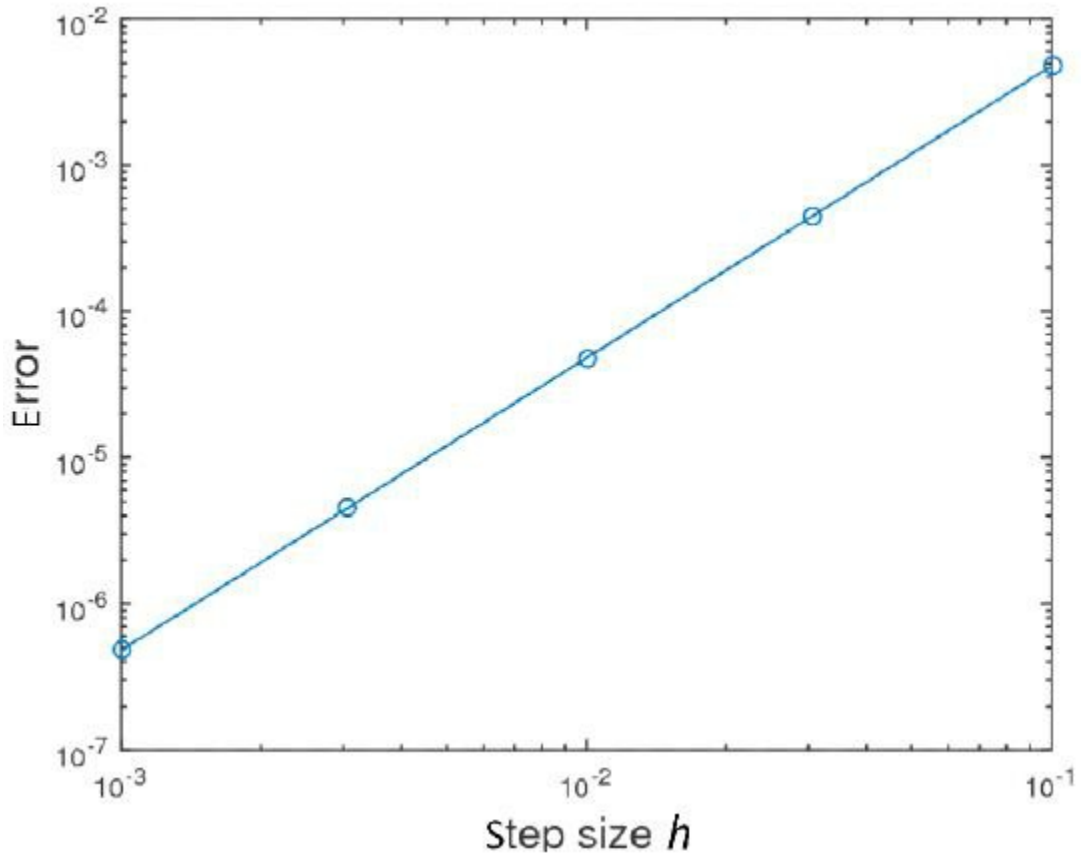


Figure 5.11: Plot of the error between the analytical and numerical solutions as the space increment h is reduced. The slope of the line is 2.0 demonstrating that the method of solution (finite volume) is second-order accurate.

5.4.2 Finite Difference Representation of the Finite-Volume Method for Second-Order Equations

Consider the equation written for the weighting function w_1 . It is formed by the middle row in [Eq. \(5.172\)](#), that is

(5.210)

$$\begin{aligned}
 & \left[\frac{1}{h} + \int_{h/2}^h \left(-\frac{1}{h} + \frac{h-x}{h} \right) dx \right] u_0 \\
 & + \left[-\frac{2}{h} + \int_{h/2}^h \left(\frac{1}{h} + \frac{x}{h} \right) dx + \int_0^{h/2} \left(\frac{-1}{h} + \frac{h-x}{h} \right) dx \right] u_1 \\
 & + \left[\frac{1}{h} + \int_0^{h/2} \left(\frac{1}{h} + \frac{x}{h} \right) dx \right] u_2 = 0.
 \end{aligned}$$

Upon integration we obtain

$$\left(\frac{1}{h} - \frac{h}{2h} + \frac{h^2}{8h}\right) u_0 + \left(\frac{-2}{h} + 0 + \frac{6h}{8}\right) u_1 + \left(\frac{1}{h} + \frac{h}{2h} + \frac{h^2}{8h}\right) u_2 = 0. \quad (5.211)$$

Rearranging this expression we get

$$\frac{1}{h} (u_0 - 2u_1 + u_2) + \frac{1}{2} (u_2 - u_0) + h \left(\frac{1}{8}u_0 + \frac{6}{8}u_1 + \frac{1}{8}u_2\right) = 0. \text{ If we divide this equation by } h \text{ we have}$$

$$\underbrace{\frac{1}{h^2} (u_0 - 2u_1 + u_2)}_{\text{centered FD}} + \underbrace{\frac{1}{2h} (u_2 - u_0)}_{\text{centered FD}} + \underbrace{\left(\frac{1}{8}u_0 + \frac{6}{8}u_1 + \frac{1}{8}u_2\right) \frac{2h}{2h}}_{\text{integrated average of } u} = 0. \quad (5.212)$$

We see from [Eq. \(5.212\)](#) that the finite volume method applied to this one dimensional equation is equivalent to a second-order difference approximation for the second and first derivatives and an integrated average over the two element interval for $u(x)$. One feature of the integration formula is that with the $\frac{1}{8}, \frac{3}{4}, \frac{1}{8}$ weighting there is more emphasis put on the diagonal element of the matrix that with the $\frac{1}{6}, \frac{2}{3}, \frac{1}{6}$ weighting of Simpson's rule encountered in the Galerkin finite element method.

5.5 Collocation Method

5.5.1 Collocation Method for First-Order Equations

In this section we introduce the collocation method. As will be seen, this is a very attractive method inasmuch as it is not necessary to perform any integrations. The downside of the method is that it requires the use of Hermite polynomials, rather than Lagrange polynomials as basis functions for second-order derivatives. In the case of first-order equations as considered in this subsection, we will not require Hermites. So, starting with first-order equations, let's get started. Consider the equation

$$\frac{du(x)}{dx} + gu(x) + f(x) = 0 \quad u(0) = u_0 \quad x \in [0, 1]. \quad (5.213)$$

As usual, let $u(x)$ be represented as

$$u(x) = \hat{u}(x) + E(x) \quad (5.214)$$

and

$$\hat{u}(x) = \sum_{j=0}^n u_j \ell_j(x). \quad (5.215)$$

Further, let the residual $R(x)$ be defined as

$$R(x) \equiv \frac{d\hat{u}(x)}{dx} + g\hat{u}(x) + f(x). \quad (5.216)$$

From the method of weighted residuals

$$\int_0^1 R(x) w_i(x) dx = 0 \quad i = 0, 1, \dots, n. \quad (5.217)$$

The bounds on the integral stem from the definition of the domain as presented in [Eq. \(5.213\)](#).

Let us choose the weighting function $w_i(x)$ to be

$$w_i(x) \equiv \delta(x - \bar{x}_i) \quad (5.218)$$

where $\delta(x - \bar{x}_i)$ is the **Dirac delta function**. **This choice defines the collocation method**.

From the definition of the Dirac delta function, ()we can write the following

$$\int_0^1 R(x) w_i(x) dx = \int_0^1 R(x) \delta(x - \bar{x}_i) dx = R(\bar{x}_i). \quad (5.219)$$

Then our equations for the collocation method are obtained from

$$R(\bar{x}_i) = 0 \quad i = 0, 1, \dots, n \quad (5.220)$$

where \bar{x}_i are called **collocation points**. **Optimal accuracy is achieved by letting the collocation points be the Gauss points**. These are the same points introduced in our discussion of Gaussian –Legendre quadrature integration in Section 3.1.6 on page 59. The required point values of these points are tabulated in reference books.

Next substitute for $R(x)$ using [Eq. \(5.216\)](#) to give

(5.221)

$$\begin{aligned}
& \int_0^1 R(x) \delta(x - \bar{x}_i) dx \\
&= \int_0^1 \frac{d\hat{u}(x)}{dx} + g\hat{u}(x) + f(x) \delta(x - \bar{x}_i) dx \\
&= \left[\frac{d\hat{u}(x)}{dx} + g\hat{u}(x) + f(x) \right]_{x=\bar{x}_i} = 0 \quad i = 0, 1, \dots, n.
\end{aligned}$$

Now substitute the definition of $\hat{u}(x)$ using [Eq. \(5.215\)](#) to give

$$\sum_{j=0}^n u_j \left[\frac{d\ell_j(x)}{dx} + g\ell_j(x) \right]_{x=\bar{x}_i} + f(\bar{x}_i) = 0 \quad i = 0, 1, \dots, n. \quad (5.222)$$

We can now write [Eq. \(5.222\)](#) using matrix notation as

$$\begin{bmatrix} a_{11} & a_{12} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & a_{n(n-1)} & a_{nn} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{bmatrix} = \begin{bmatrix} -e_1 - a_{10}u_0 \\ -e_2 \\ \cdot \\ \cdot \\ \cdot \\ -e_n \end{bmatrix} \quad (5.223)$$

where

$$a_{ij} = \left[\frac{d\ell_j(x)}{dx} + g\ell_j(x) \right]_{x=\bar{x}_i} \quad (5.224)$$

and

$$e_i = f(\bar{x}_i). \quad (5.225)$$

After substitution of the values of $\ell_j(x)$, [Eq. \(5.223\)](#)

$$\hat{u}(x) = \sum_{j=0}^2 u_j \ell_j(x). \quad (5.226)$$

Now evaluate $R(\bar{x}_i)$ $i = 0, 1$ at the Gauss points and set the result to zero for each point. We obtain

$$R(\bar{x}_i) = 0 \quad i = 0, 1 \quad (5.227)$$

or

$$\left[\frac{d\hat{u}(x)}{dx} + \frac{0.9}{1+2x}\hat{u}(x) \right] \Big|_{\bar{x}_i} = 0 \quad i = 0, 1. \quad (5.228)$$

Next substitute for $\hat{u}(x)$ using [Eq. \(5.226\)](#) to get

$$\sum_{j=0}^2 u_j \left[\frac{d\ell_j(x)}{dx} + \frac{0.9}{1+2x}\ell_j(x) \right] \Big|_{\bar{x}_i} = 0 \quad i = 0, 1. \quad (5.229)$$

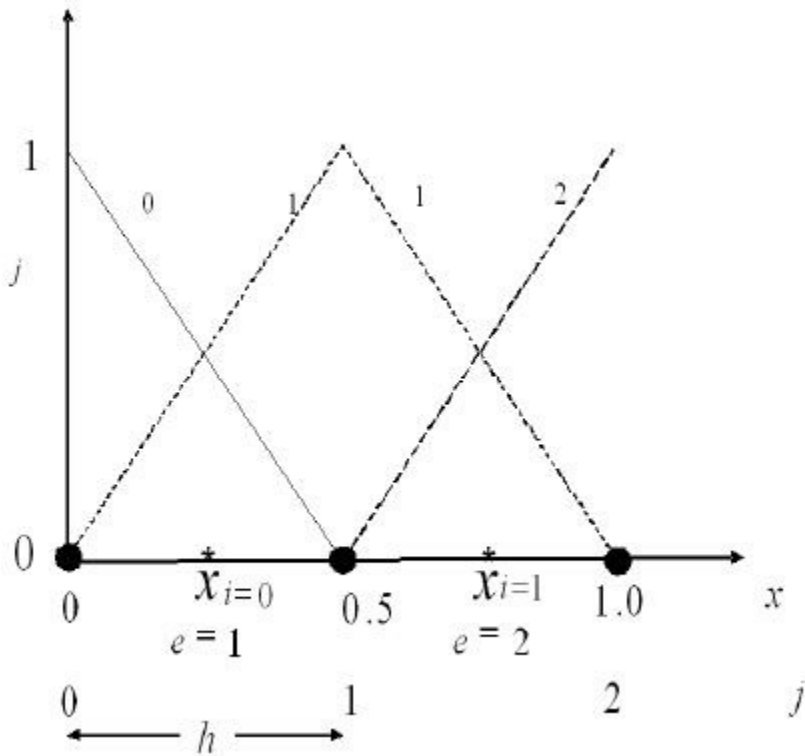


Figure 5.12: Element array for collocation method explanation.

Now, using the Gauss Legendre integration points selected from Table 3.29. Because our basis functions are piecewise linear we require one Gaussian point per element and these points are located at $\bar{x}_0 = 0.25$ and $\bar{x}_1 = 0.75$ (keep in mind that the table assumes a $[-1, 1]$ coordinate system for each element). From [Eq. \(5.229\)](#) we obtain the following matrix equation:

$$\begin{bmatrix} \left[\begin{array}{c} \frac{dl_0(x)}{dx} + \\ \frac{0.9}{1+2x} l_0(x) \end{array} \right]_{\bar{x}=0.25} & \left[\begin{array}{c} \frac{dl_1(x)}{dx} + \\ \frac{0.9}{1+2x} l_1(x) \end{array} \right]_{\bar{x}=0.25} & \left[\begin{array}{c} \frac{dl_2(x)}{dx} + \\ \frac{0.9}{1+2x} l_2(x) \end{array} \right]_{\bar{x}=0.25} \\ \left[\begin{array}{c} \frac{dl_0(x)}{dx} + \\ \frac{0.9}{1+2x} l_0(x) \end{array} \right]_{\bar{x}=0.75} & \left[\begin{array}{c} \frac{dl_1(x)}{dx} + \\ \frac{0.9}{1+2x} l_1(x) \end{array} \right]_{\bar{x}=0.75} & \left[\begin{array}{c} \frac{dl_2(x)}{dx} + \\ \frac{0.9}{1+2x} l_2(x) \end{array} \right]_{\bar{x}=0.75} \end{bmatrix} \times \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (5.230)$$

Notice that this is a rectangular matrix, a phenomenon identified with collocation. No integrations appear in [Eq. \(5.230\)](#), so the coefficients are easily evaluated by inspection of the function values via [Fig. 5.12](#). We obtain through substitution of the collocation point locations

$$\begin{bmatrix} \left(\frac{-\frac{1}{0.9} +}{1+2(0.25)} 0.5 \right) & \left(\frac{\frac{1}{0.9} +}{1+2(0.25)} 0.5 \right) & 0 \\ 0 & \left(\frac{-\frac{1}{0.9} +}{1+2(0.75)} 0.5 \right) & \left(\frac{\frac{1}{0.9} +}{1+2(0.75)} 0.5 \right) \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (5.231)$$

or

$$\begin{bmatrix} -1.7 & 2.3 & 0 \\ 0 & -1.82 & 2.18 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (5.232)$$

The rectangular form of the coefficient matrix would normally be problematic, but we have a solution to the problem. We now impose the boundary condition which states that $u_0 = 1$. After multiplying the first column by this value and transferring the information to the right-hand side we have

$$\begin{bmatrix} 2.3 & 0 \\ -1.82 & 2.18 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -(-1.7)(1) \\ 0 \end{bmatrix} \quad (5.233)$$

which is a solvable system of equations. From this equation we get

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0.740 \\ 0.618 \end{bmatrix} \quad (5.234)$$

while the analytical solution is

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0.732 \\ 0.609 \end{bmatrix} \quad (5.235)$$

which is quite accurate considering we are using only three nodes.

5.5.2 Collocation Method for Second-Order Equations

In this subsection we will extend our discussion of the collocation method to consider second derivatives. It is at this point that we will need to introduce the Hermite polynomials as basis functions. The Hermite polynomial was introduced earlier in Subsection 1.9 on page 20.

Consider now the second-order equation

$$\frac{d^2 u(x)}{dx^2} + \frac{du(x)}{dx} + 0.5u(x) = 0. \quad (5.236)$$

Following the strategy we used in the preceding subsection, we define our weighted residual operator as

$$\int_0^1 R(x) w_i(x) dx = \int_0^1 R(x) \delta(x - \bar{x}_i) dx = R(\bar{x}_i) = 0 \quad i = 0, \dots, n \quad (5.237)$$

where the residual is defined as

$$R(x) = \frac{d^2 \hat{u}(x)}{dx^2} + \frac{d\hat{u}(x)}{dx} + 0.5\hat{u}(x) \quad (5.238)$$

from which we get

$$\left[\frac{d^2 \hat{u}(x)}{dx^2} + \frac{d\hat{u}(x)}{dx} + 0.5\hat{u}(x) \right] \Big|_{\bar{x}_i} = 0 \quad i = 0, 1, \dots, n. \quad (5.239)$$

We see immediately that if we use linear functions to represent $u(x)$ we will have difficulty because the second derivative is zero at the collocation points. Your first inclination probably is to use quadratic Lagrange polynomials, since their second derivatives will be defined everywhere except at the end-of-the-element nodes. However, one cannot ignore these end points since the collocation method does not have restrictions on where the collocation points should be. As noted earlier, the simplest answer is to use *Hermite polynomials* which we introduced in Section 1.9 on page 20, that is use

$$\hat{u}(x) = \sum_{j=0}^n h_j^0(x) u_j + h_j^1(x) \frac{du_j}{dx} \quad (5.240)$$

(see [Fig. 5.13](#)) which provides us with $2(n + 1)$ unknowns. The Hermite basis functions are defined as

$$h_j^0(x) = (\ell_j(x))^2 \left(1 - 2(x - x_j) \frac{d\ell_j}{dx} \right) \quad (5.241)$$

and

$$h_j^1(x) = (\ell_j(x))^2 (x - x_j) \quad (5.242)$$

where, for convenience, the Lagrange polynomial approximations are used in the definition.

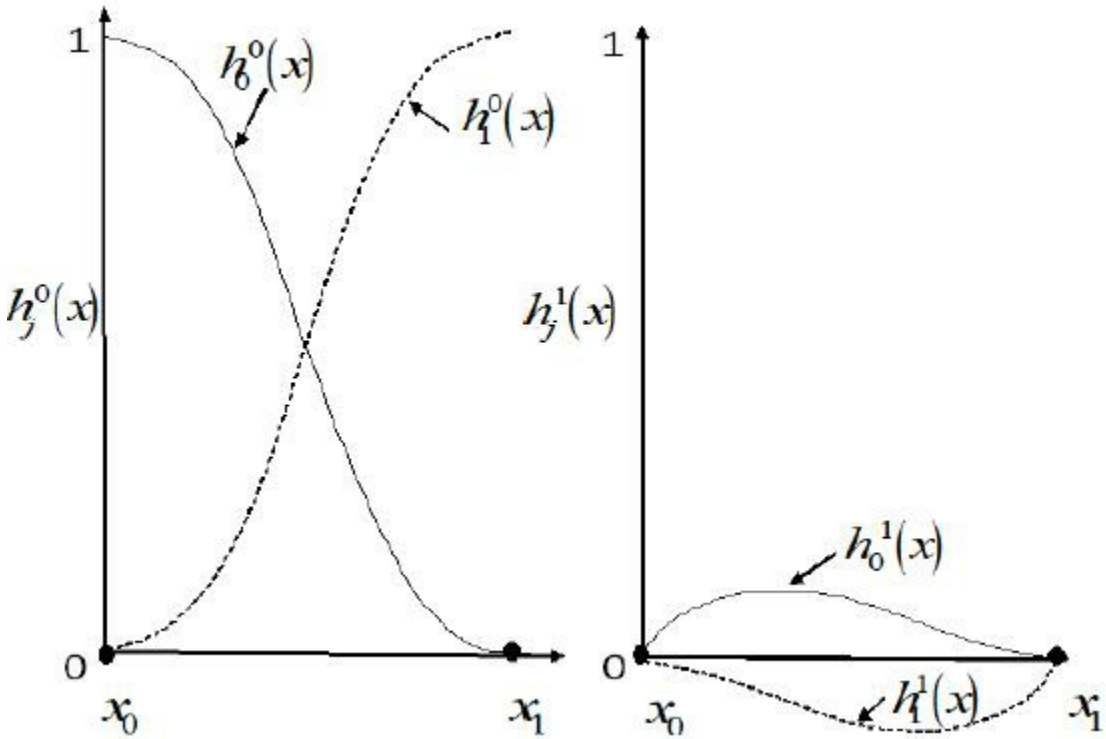


Figure 5.13: Cubic Hermite polynomials.

To make the notation tolerable, let us define the following operator:

$$L(\cdot) \equiv \frac{d^2(\cdot)}{dx^2} + \frac{d(\cdot)}{dx} + 0.5(\cdot). \quad (5.243)$$

where you can replace the dot with any suitable function, for example $u(x)$. Using this operator, and assuming two collocation points per element, our collocation equations become

$$L(\hat{u})_{\bar{x}_i} = 0 \quad i = 0, 1, \dots, (2n - 1) \quad (5.244)$$

or

(5.245)

$$L \left(\sum_{j=0}^n h_j^0(x) u_j + h_j^1(x) \frac{du_j}{dx} \right) \Big|_{\bar{x}_i} = 0 \quad i = 0, 1, \dots, (2n - 1)$$

which gives us $2n$ equations, which as we will see, is the number we need after implementation of boundary conditions.

If we now write [Eq. \(5.245\)](#) for the specific system of functions and collocation points shown in [Figs 5.14](#) and [5.15](#) as a matrix equation. we obtain

$$\begin{bmatrix} Lh_0^0(x)|_{x_1} & Lh_0^1(x)|_{\bar{x}_1} & Lh_1^0(x)|_{\bar{x}_1} & Lh_1^1(x)|_{\bar{x}_1} & 0 & 0 \\ Lh_0^0(x)|_{\bar{x}_2} & Lh_0^1(x)|_{\bar{x}_2} & Lh_1^0(x)|_{\bar{x}_2} & Lh_1^1(x)|_{\bar{x}_2} & 0 & 0 \\ 0 & 0 & Lh_0^0(x)|_{\bar{x}_3} & Lh_0^1(x)|_{\bar{x}_3} & Lh_1^0(x)|_{\bar{x}_3} & Lh_1^1(x)|_{\bar{x}_3} \\ 0 & 0 & Lh_0^0(x)|_{\bar{x}_4} & Lh_0^1(x)|_{\bar{x}_4} & Lh_1^0(x)|_{\bar{x}_4} & Lh_1^1(x)|_{\bar{x}_4} \end{bmatrix} \times \begin{bmatrix} u_0 \\ \frac{du_0}{dx} \\ u_1 \\ \frac{du_1}{dx} \\ u_2 \\ \frac{du_2}{dx} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (5.246)$$

where

$$L [h_0^0(x)]|_{\bar{x}_1} \equiv \left[\frac{d^2 h_0^0(x)}{dx^2} + \frac{dh_0^0(x)}{dx} + h_0^0(x) \right] |_{\bar{x}_1}. \quad (5.247)$$

Note that the Hermite polynomials are defined element-wise and that the collocation points \bar{x}_1 and \bar{x}_2 are in element 1 and \bar{x}_3 and \bar{x}_4 are in element 2.

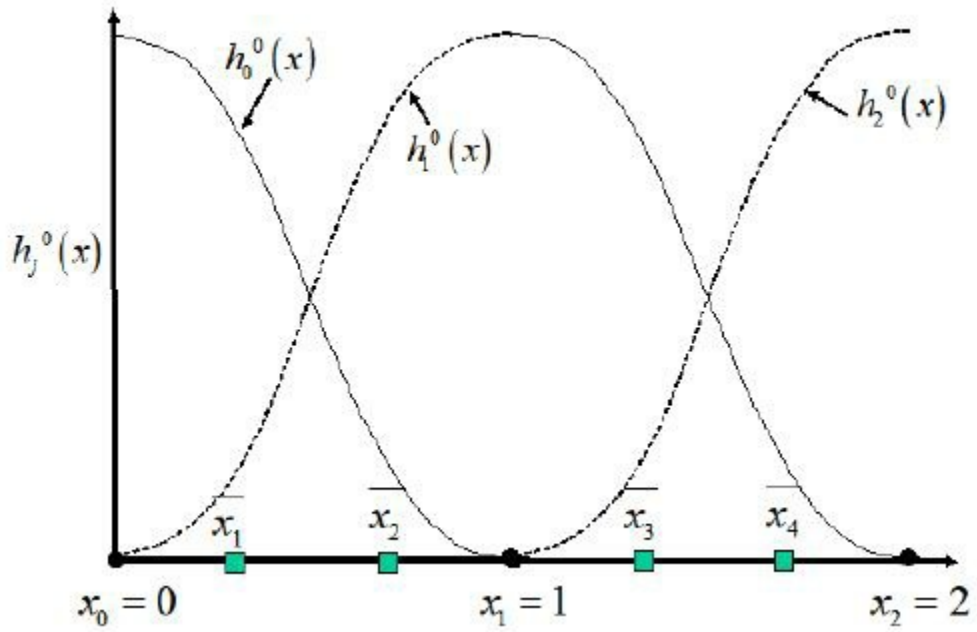


Figure 5.14: Hermite function $h_i^0(x)$.

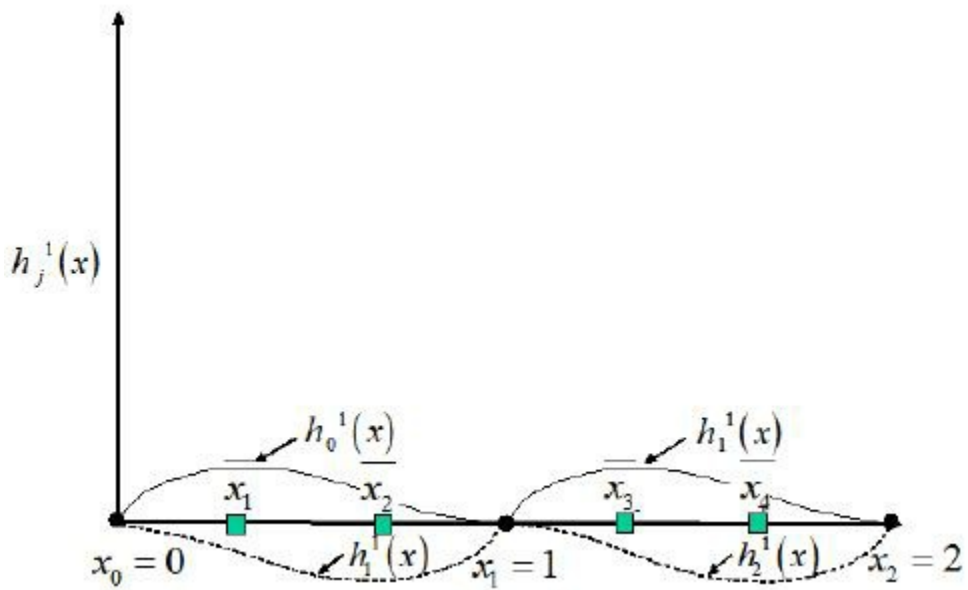


Figure 5.15: Hermite function $h_i^1(x)$.

Assume that the boundary conditions on this problem are

$$u(0) = \bar{u} \tag{5.248}$$

$$\frac{du(x)}{dx} \Big|_{x=2} = \frac{d\bar{u}}{dx}. \tag{5.249}$$

Then [Eq. \(5.246\)](#) becomes, after reducing the set of equations to accommodate these boundary conditions

$$\begin{bmatrix} L(h_0^1(x))|_{\bar{x}_1} & L(h_1^0(x))|_{\bar{x}_1} & L(h_1^1(x))|_{\bar{x}_1} & 0 \\ L(h_0^1(x))|_{\bar{x}_2} & L(h_1^0(x))|_{\bar{x}_2} & L(h_1^1(x))|_{\bar{x}_2} & 0 \\ 0 & L(h_0^0(x))|_{\bar{x}_3} & L(h_1^0(x))|_{\bar{x}_3} & L(h_1^0(x))|_{\bar{x}_3} \\ 0 & L(h_1^0(x))|_{\bar{x}_4} & L(h_1^1(x))|_{\bar{x}_4} & L(h_1^0(x))|_{\bar{x}_4} \end{bmatrix} \begin{bmatrix} \frac{du_0}{dx} \\ u_1 \\ \frac{du_1}{dx} \\ u_2 \end{bmatrix} \quad (5.250)$$

$$= - \begin{bmatrix} L(h_0^0(x))|_{\bar{x}_1} \bar{u} \\ L(h_0^0(x))|_{\bar{x}_2} \bar{u} \\ L(h_1^1(x))|_{\bar{x}_3} \frac{d\bar{u}}{dx} \\ L(h_1^1(x))|_{\bar{x}_4} \frac{d\bar{u}}{dx} \end{bmatrix}$$

or

$$[\mathbf{A}] \{\mathbf{u}\} = \{\mathbf{f}\} \quad (5.251)$$

from which we get formally

$$\{\mathbf{u}\} = [\mathbf{A}]^{-1} \{\mathbf{f}\} . \quad (5.252)$$

Substitution of $\{\mathbf{u}\}$ into [Eq. \(5.240\)](#) gives the required approximate solution across the entire domain:

Let us now apply the collocation method to the equation we used earlier, namely

$$\frac{d^2u(x)}{dx^2} + \frac{du(x)}{dx} + \frac{u(x)}{2} = 0 \quad x \in [x=0, x=2] \quad (5.253)$$

$$u(0) = 1 \quad (5.254)$$

$$\left. \frac{du(x)}{dx} \right|_{x=2} = 0. \quad (5.255)$$

To solve this problem we will **introduce the coordinate system** $-1 \leq \xi \leq 1$ which is the same coordinate system we used in our discussion of Gauss-Legendre quadrature integration (Section 3.1.6 on page 59, but there we used the x notation. In this coordinate system, the collocation equations become, for a problem involving only three nodes (two elements)

$$\begin{aligned}
& \begin{bmatrix} L(h_0^1(\xi))|_{\bar{x}_1(\xi)} & L(h_1^0(\xi))|_{\bar{x}_1(\xi)} & L(h_1^1(\xi))|_{\bar{x}_1(\xi)} & 0 \\ L(h_0^1(\xi))|_{\bar{x}_2(\xi)} & L(h_1^0(\xi))|_{\bar{x}_2(\xi)} & L(h_1^1(\xi))|_{\bar{x}_2(\xi)} & 0 \\ 0 & L(h_0^0(\xi))|_{\bar{x}_3(\xi)} & L(h_1^0(\xi))|_{\bar{x}_3(\xi)} & L(h_0^1(\xi))|_{\bar{x}_3(\xi)} \\ 0 & L(h_0^0(\xi))|_{\bar{x}_4(\xi)} & L(h_1^0(\xi))|_{\bar{x}_4(\xi)} & L(h_0^1(\xi))|_{\bar{x}_4(\xi)} \end{bmatrix} \begin{bmatrix} \frac{du_0}{dx} \\ u_1 \\ \frac{du_1}{dx} \\ u_2 \end{bmatrix} \quad (5.256) \\
& = - \begin{bmatrix} L(h_0^0(\xi))|_{\bar{x}_1} \bar{u} \\ L(h_0^0(\xi))|_{\bar{x}_2} \bar{u} \\ L(h_1^1(\xi))|_{\bar{x}_3} \frac{d\bar{u}}{dx} \\ L(h_1^1(\xi))|_{\bar{x}_4} \frac{d\bar{u}}{dx} \end{bmatrix}
\end{aligned}$$

where

$$L(\cdot) = \frac{d^2(\cdot)}{dx^2} + \frac{d(\cdot)}{dx} + \frac{(\cdot)}{2}.$$

Note that the Hermite polynomials have a very simple form in this ξ coordinate system, that is

$$h_0^0(\xi) = \frac{1}{4}(\xi - 1)^2(\xi + 2) \quad (5.257)$$

$$h_1^0(\xi) = -\frac{1}{4}(\xi + 1)^2(\xi - 2) \quad (5.258)$$

$$h_0^1(\xi) = \frac{h}{8}(\xi + 1)(\xi - 1)^2 \quad (5.259)$$

$$h_1^1(\xi) = \frac{h}{8}(\xi + 1)^2(\xi - 1). \quad (5.260)$$

To approximate [Eq. \(5.253\)](#) we will need the first and second derivatives of [Eq. \(5.257\)](#) through [Eq. \(5.260\)](#), that is

$$\frac{dh_0^0(\xi)}{dx} = \frac{dh_0^0(\xi)}{d\xi} \frac{d\xi}{dx} = 2 \frac{dh_0^0(\xi)}{d\xi} \quad (5.261)$$

$$\frac{dh_1^0(\xi)}{dx} = \frac{dh_1^0(\xi)}{d\xi} \frac{d\xi}{dx} = 2 \frac{dh_1^0(\xi)}{d\xi} \quad (5.262)$$

$$\frac{dh_0^1(\xi)}{dx} = \frac{dh_0^1(\xi)}{d\xi} \frac{d\xi}{dx} = 2 \frac{dh_0^1(\xi)}{d\xi} \quad (5.263)$$

$$\frac{dh_1^1(\xi)}{dx} = \frac{dh_1^1(\xi)}{d\xi} \frac{d\xi}{dx} = 2 \frac{dh_1^1(\xi)}{d\xi} \quad (5.264)$$

and

$$\frac{d^2h_0^0(\xi)}{d^2x} = \frac{d^2h_0^0(\xi)}{d^2\xi} \frac{d\xi}{dx} \frac{d\xi}{dx} = 4 \frac{d^2h_0^0(\xi)}{d^2\xi} \quad (5.265)$$

$$\frac{d^2h_1^0(\xi)}{d^2x} = \frac{d^2h_1^0(\xi)}{d^2\xi} \frac{d\xi}{dx} \frac{d\xi}{dx} = 4 \frac{d^2h_1^0(\xi)}{d^2\xi} \quad (5.266)$$

$$\frac{d^2h_0^1(\xi)}{d^2x} = \frac{d^2h_0^1(\xi)}{d^2\xi} \frac{d\xi}{dx} \frac{d\xi}{dx} = 4 \frac{d^2h_0^1(\xi)}{d^2\xi} \quad (5.267)$$

$$\frac{d^2h_1^1(\xi)}{d^2x} = \frac{d^2h_1^1(\xi)}{d^2\xi} \frac{d\xi}{dx} \frac{d\xi}{dx} = 4 \frac{d^2h_1^1(\xi)}{d^2\xi} \quad (5.268)$$

where the first derivatives are, from [Eq. \(5.257\)](#) through [5.260](#)

$$\frac{dh_0^0(\xi)}{d\xi} = \frac{1}{4} \left[(\xi - 1)^2 + 2(\xi + 2)(\xi - 1) \right] \quad (5.269)$$

$$\frac{dh_1^0(\xi)}{d\xi} = -\frac{1}{4} \left[(\xi + 1)^2 + 2(\xi - 2)(\xi + 1) \right] \quad (5.270)$$

$$\frac{dh_0^1(\xi)}{d\xi} = \frac{h}{8} \left[(\xi - 1)^2 + 2(\xi + 1)(\xi - 1) \right] \quad (5.271)$$

$$\frac{dh_1^1(\xi)}{d\xi} = \frac{h}{8} \left[(\xi + 1)^2 + 2(\xi - 1)(\xi + 1) \right] \quad (5.272)$$

and the second derivatives are

$$\frac{d^2h_0^0(\xi)}{d\xi^2} = \frac{1}{4} \left[2(\xi - 1) + 2\{(\xi + 2) + (\xi - 1)\} \right] \quad (5.273)$$

$$\frac{d^2h_1^0(\xi)}{d\xi^2} = -\frac{1}{4} \left[2(\xi + 1) + 2\{(\xi - 2) + (\xi + 1)\} \right] \quad (5.274)$$

$$\frac{d^2 h_0^1(\xi)}{d\xi^2} = \frac{h}{8} [2(\xi - 1) + 2\{(\xi + 1) + (\xi - 1)\}] \quad (5.275)$$

$$\frac{d^2 h_1^1(\xi)}{d\xi^2} = \frac{h}{8} [2(\xi + 1) + 2\{(\xi - 1) + (\xi + 1)\}]. \quad (5.276)$$

For terms such as $L(h_0^1(\xi))$ we have, using the definitions presented above,

$$\begin{aligned} L(h(\xi)_0^0) &= \frac{1}{4} [2(\xi - 1) + 2\{(\xi + 2) + (\xi - 1)\}] (4) \\ &+ \frac{1}{4} [(\xi - 1)^2 + 2(\xi + 2)(\xi - 1)] (2) \\ &+ \frac{1}{2} \left[\frac{1}{4} (\xi - 1)^2 (\xi + 2) \right] \end{aligned} \quad (5.277)$$

$$\begin{aligned} L(h(\xi)_1^0) &= -\frac{1}{4} [2(\xi + 1) + 2\{(\xi - 2) + (\xi + 1)\}] (4) \\ &- \frac{1}{4} [(\xi + 1)^2 + 2(\xi - 2)(\xi + 1)] (2) \\ &- \frac{1}{2} \left[\frac{1}{4} (\xi + 1)^2 (\xi - 2) \right] \end{aligned} \quad (5.278)$$

$$\begin{aligned} L(h(\xi)_0^1) &= \frac{h}{8} [2(\xi - 1) + 2\{(\xi + 1) + (\xi - 1)\}] (4) \\ &+ \frac{h}{8} [(\xi - 1)^2 + 2(\xi + 1)(\xi - 1)] (2) \\ &+ \frac{1}{2} \left[\frac{h}{8} (\xi + 1)(\xi - 1)^2 \right] \end{aligned} \quad (5.279)$$

$$\begin{aligned}
 L \left(h(\xi)_1^1 \right) &= \frac{h}{8} [2(\xi + 1) + 2\{(\xi - 1) + (\xi + 1)\}] \quad (4) \\
 &+ \frac{h}{8} [(\xi + 1)^2 + 2(\xi - 1)(\xi + 1)] \quad (2) \\
 &+ \frac{1}{2} \left[\frac{h}{8} (\xi - 1)(\xi + 1)^2 \right]
 \end{aligned}
 \tag{5.280}$$

The matrix equation that one obtains after evaluating [Eqs. \(5.277\)](#) through [\(5.280\)](#) at the two Gauss Legendre points, that is at $\xi = \pm 0.577$ (see table on pg. 8.1) and at the end points $x = 0$ (where $\xi = -1$ in the first element) and $x = 2$ (where $\xi = 1$ in the second element) is

$$\begin{bmatrix} 5.5 & 0.5 & 1.5 & 0 \\ -4.52 & -0.628 & -0.345 & 0 \\ 0 & -4.02 & -4.52 & -0.628 \\ 0 & 2.01 & 0.761 & 0.0448 \end{bmatrix} \begin{bmatrix} \frac{du_0}{dx} \\ u_1 \\ \frac{du_1}{dx} \\ u_2 \end{bmatrix} = \begin{bmatrix} 5.5 \\ -2.52 \\ 0 \\ 0 \end{bmatrix}
 \tag{5.281}$$

The solution to this problem is

$$\begin{bmatrix} \frac{du_0}{dx} \\ u_1 \\ \frac{du_1}{dx} \\ u_2 \end{bmatrix} = \begin{bmatrix} -0. - 2.8 \\ -0.804 \\ -0.976 \\ -1.22 \end{bmatrix}
 \tag{5.282}$$

and it is compared with the analytical solution in the following table:

Function	Location	Analytical	Numerical	Difference
$\frac{du_0}{dx}$	$x = 0$	- 2.80	- 2.80	- .000
u_1	$x = 1$	- 0.801	- 0.804	- 003
$\frac{du_1}{dx}$	$x = 1$	- 0.965	- 0.976	- 0.011
u_2	$x = 2$	- 1.22	- 1.22	0

This is again a remarkable result considering that only two elements were employed.

5.6 Chapter Summary

In this chapter we introduced the key elements underlying the theory that we will use for all of the numerical methods presented in this book, except finite differences. The key concepts are the weighted residual formalism and the use of polynomial approximation theory to represent the unknown and sought-after solutions to differential equations and their associated boundary conditions. By selecting different definitions of the weighting functions in the weighted residual formulations, the Galerkin finite element method, the finite volume method and the collocation method are defined. The programming strategy that is generally used to implement these methods is based upon the concept of first evaluating the integrals that are required by the formulation at the element level using a local $0 \leq \chi \leq 1$ coordinate system. Thus there is generated an element coefficient matrix for each element. The information from each element is then collected to form the global coefficient matrix that is actually used to solve the problem. To tie the various methods together, the Galerkin finite element and finite volume methods are interpreted in terms of finite difference formulae. Examples for each numerical approach are presented and their results compared to those generated analytically.

5.7 Problems

1. Consider

$$\frac{du(x)}{dx} + Q = 0 \quad x \in [0, 1] \quad (5.283)$$

$$u(0) = 1 \quad (5.284)$$

$$Q = 2. \quad (5.285)$$

Calculate the value at $u(1)$ using the finite volume method and one element.

2. Consider the equation

$$\frac{d^2u(x)}{dx^2} = 0 \quad x \in [0, 1] \quad (5.286)$$

$$u(0) = 1 \quad (5.287)$$

$$\left. \frac{du(x)}{dx} \right|_{x=1} = 0. \quad (5.288)$$

Use the finite element method with three nodes to solve for $x = 0.5$.

3. Assume we have an integral of the form

$$\int_0^3 \frac{d\phi(L)}{dx} dx. \quad (5.289)$$

Transform this integral into local $0 \leq \chi \leq 1$ coordinates and evaluate the integral.

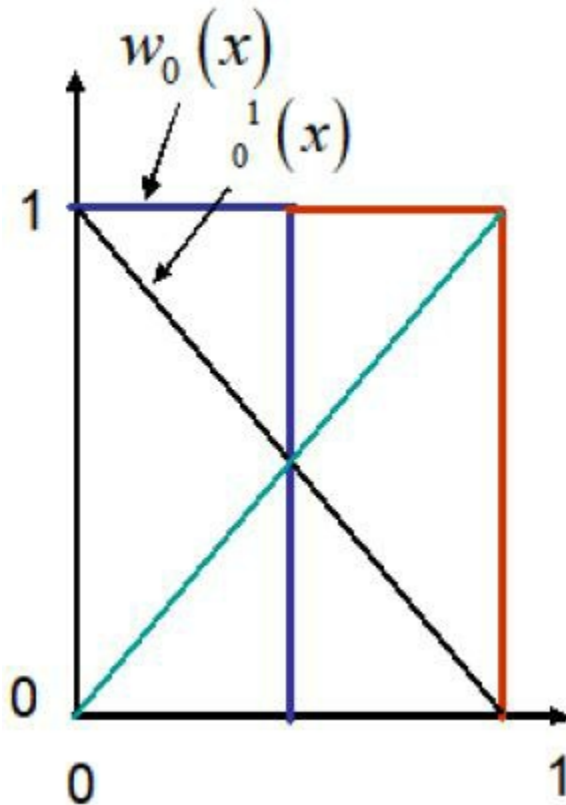


Figure 5.16: Sketch of finite volume configuration for solving problem 1.

4. Consider the equation

$$\frac{du(x)}{dx} + u(x) - 1 = 0 \quad x \in [0, 1] \tag{5.290}$$

$$u(0) = 2.$$

Solve this equation using both the finite difference method and the finite volume or subdomain method and compare the results for the analytical solution

$$u(x) = 1 + \exp(-x). \tag{5.291}$$

You will use 17 nodes. In the case of the finite difference method represent the $u(x)$ term as u_i , and a backward difference approximation. Alternatively, you could imagine you are standing in the middle between two nodes and define $u_{i+\frac{1}{2}} = \frac{(u_i + u_{i-1})}{2}$, as the approximation. You can also use a second-order finite difference approximation for the first derivative.

To do the finite volume model:

1.
 - a. create the element coefficient matrix; it will be 2×2 and the same for every element.

- b. create the global matrix by using a strategy that relates the global node numbers to the local node numbers
- c. impose the boundary conditions
- d. solve the problem

Now, plot the solutions for the two finite-difference methods, subdomain method and the analytic solution. Compare the relative errors of the three methods and discuss why they are different.

5. Consider the equation

$$\frac{d^2 u(x)}{dx^2} = 1. \tag{5.292}$$

Develop the Galerkin finite element representation of this equation at the node i in [Fig. 5.17](#). Do the following:

- a. write the weighted residual approximation using the linear Lagrange polynomial as the weighting function
- b. substitute the above equation replacing $u(x)$ by $\hat{u}(x)$ into the weighted residual formulation
- c. apply integration by parts to the second-order term yielding an integral of the product of two first-order terms and two boundary terms (ignore the boundary terms)
- d. introduce the definitions for $\hat{u}(x)$, that is introduce the Lagrange polynomial expansion.
- e. write the element matrix equations with the coefficients in the matrix presented in symbolic form, that is show them as integrals with appropriate integration limits.
- f. assemble the global matrix equation using the information generated in the preceding step.

Note: Do not do anything in local coordinates and do not do the integrations



Figure 5.17: Nodal arrangement used to solve problem 5.7.

6. In [Section 5.3](#) on page 102 we developed the solution to the problem

$$\frac{d^2u(x)}{dx^2} + \frac{du(x)}{dx} + \frac{u(x)}{2} = 0 \quad x \in [0, 3]$$

$$u(0) = 1$$

$$\left. \frac{du(x)}{dx} \right|_3 = 0$$

with analytic solution

$$u(x) = \exp(-x/2) [\cos(x/2) + \tan(\pi/4 + 3/2) \sin(x/2)].$$

The problem is to determine how your numerical solution to this problem changes as you change the value of h . Please do the following:

- Create a Matlab code that solves the above problem using the Galerkin finite-element formulation shown in [Section 5.3](#);
- Solve the same problem using $h = h/2, h/3, h/4, h/5$; you should have a common node at $x = 1.0$ and 2.0 ;
- Determine the error at $x = 1$ and $x = 2$ for each spacing;
- Plot the log of the error at $x = 1.0$ and 2.0 versus the log of the spacing and calculate the slope. The slope of this line relates to the order n of the error of the approximation, that is $O(h^n)$;
- What is the order of the error, that is, what is n .
- Explain why a cubic Lagrange polynomial cannot be used for the collocation solution to a second-order equation and a cubic Hermite can.
- What are the advantages and disadvantages of using two finite elements with three linear basis functions versus one element using three quadratic basis functions.
- The finite volume method is said to exactly preserve local mass conservation when solving the time independent convective-diffusion equation, that is

$$\frac{d^2u}{dx^2} - v \frac{du}{dx} = 0 \tag{5.293}$$

The reason for this lies in the fact that the finite volume boundaries have the same value of the derivative approximation $\frac{d\hat{u}}{dx}$ on each side. Show why this is the case.

Bibliography

Ascher, U., L. Mattheij, and R. Russel (1988) *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs.

Atkinson, K., W. Han, and D. Stewart (2008) *Numerical Solution of Ordinary Differential Equations*, Wiley Interscience, Hoboken.

Aziz, A. (1975) *Numerical Solutions of Boundary Value Problems for Ordinary Differential Equations*, Academic Press, New York.

Bradie, B. (2006) *A Friendly Introduction to Numerical Analysis*, Pearson Prentice-Hall.

Butcher, J.C. (2008) *Numerical Methods for Ordinary Differential Equations*, 2nd ed. John Wiley, Chichester.

Isaacson, E. and H. Keller (1966) *Analysis of Numerical Methods*, Dover, New York.

Iserles, A. (1996) *A First Course in the Numerical Analysis of Differential Equations*, Cambridge University Press, Cambridge.

Lambert, J. D. (1991) *Numerical Methods for Ordinary Differential Systems*, John Wiley, New York.

Shampine, L.F. (1994) *Numerical Solution of Ordinary Differential Equations*, Taylor & Francis, New York.

Chapter 6

Initial Boundary-Value Problems

6.1 Introduction

In this chapter we are interested in solving problems that have both a spatial and temporal component. As a point of departure, let us consider the equation

$$\frac{\partial}{\partial t} u(x, t) + \frac{\partial}{\partial x} u(x, t) = 0 \quad t \in [t_0 = 0, t_N = T) \quad x \in [A, B]. \quad (6.1)$$

A typical initial condition for this equation would be to specify the state of the system at some time designated as zero such as

$$u(x, 0) = u_0(x) \quad (6.2)$$

and a typical boundary condition that could be a function of time as

$$u(0, t) = u_1(t) \quad t \geq 0. \quad (6.3)$$

6.2 Two Dimensional Polynomial Approximations

To achieve this goal we need to extend our early Lagrange polynomial approximation strategy to consider two dimensions, namely space and time. To this end we define the following quantities:

$$x \in [x_0 = a, x_n = b] \quad (6.4)$$

$$t \in [t_0 = c, t_m = d] \quad (6.5)$$

$$n \equiv \text{degree of the Lagrange polynomial in space, that is } \ell_i(x)^n \quad (6.6)$$

$$m \equiv \text{degree of the Lagrange polynomial in time, that is } \ell_j(t)^m \quad (6.7)$$

$$x = i\Delta x \quad (6.8)$$

$$t = j\Delta t \quad (6.9)$$

Note that this means that to define the Lagrange polynomial $\ell_i^n(x)$ there are $n + 1$ nodes in the x direction between a and b . Similarly, to define the Lagrange polynomial in time $\ell_j^m(t)$ there are $m + 1$ nodes in the time dimension between c and d . The appropriate approximation for $f(x, t)$ is obtained by taking the product of these two functions, that is

$$f(x, t) = \sum_{i=0}^n \sum_{j=0}^m \ell_i^n(x) \ell_j^m(t) f(x_i, t_j) + E_{nm} \quad (6.10)$$

where E_{nm} is the error term. Let

$$\hat{f}(x, t) \equiv \sum_{i=0}^n \sum_{j=0}^m \ell_i^n(x) \ell_j^m(t) f(x_i, t_j). \quad (6.11)$$

A pictorial representation of $\ell_i^n(x) \ell_j^m(t)$ is found in [Fig. 6.1](#) for the case of $n = i$ and $m = j$.

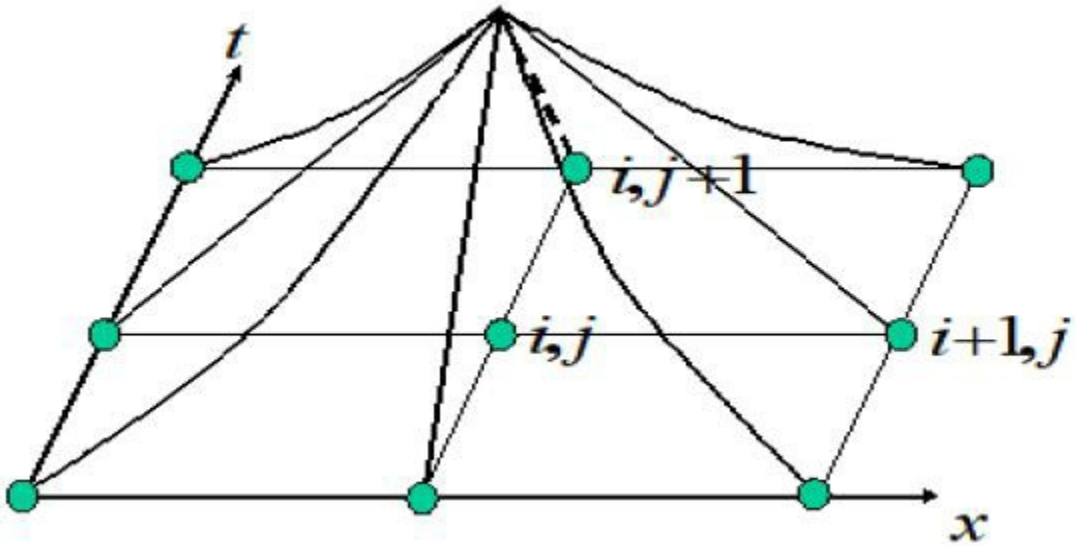


Figure 6.1: A conceptual representation of the function $\ell_i^n(x)\ell_j^m(t)$. The dots represent nodes and the maximum function value is unity.

Recall the definition

$$p_n(x) = \prod_{i=0}^{i=n} (x - x_i) \quad (6.12)$$

where \prod_i is the product operator. Using this notation we can write the form of the error E_{nm} in [Eq. \(6.10\)](#) as (see [1])

$$E_{nm}(x, t) = \frac{p_n(x)}{(n+1)!} \frac{\partial^{(n+1)} f(\xi, t)}{\partial x^{(n+1)}} + \frac{p_m(t)}{(m+1)!} \frac{\partial^{(m+1)} f(x, \eta)}{\partial t^{(m+1)}} \quad (6.13)$$

$$- \frac{p_n(x) p_m(t)}{(m+n+2)!} \frac{\partial^{(n+1)}}{\partial x^{(n+1)}} \frac{\partial f^{(m+1)}}{\partial t^{(m+1)}} f(\xi', \eta')$$

$(a < \xi, \xi' < b; c < \eta, \eta' < d)$

6.2.1 Example of a Two Dimensional Polynomial Approximation

Now we consider an example that illustrates the approximation of a function in two space dimensions using [Eq. \(6.11\)](#). We will consider the representation of $\exp(x, t)$ over $[-1, 1] \times [-1, 1]$ with $m = n = 2$ and $\Delta x = \Delta t = 1$ (see [Fig. 6.2](#)). From [Eq. \(6.11\)](#) we have for the representation of $\exp(x, t)$

$$\exp(x, t) = \sum_{i=0}^2 \sum_{j=0}^2 \exp(x_i, t_j) \ell_i^2(x) \ell_j^2(t) + E_{2,2} \quad (6.14)$$

where

$$\ell_0^2(x) = \frac{1}{2}x(x-1) \quad (6.15)$$

$$\ell_1^2(x) = -(x+1)(x-1) \quad (6.16)$$

$$\ell_2^2(x) = \frac{1}{2}x(x+1). \quad (6.17)$$

The functions $\ell_j^2(t)$ have an analogous form in the time domain.

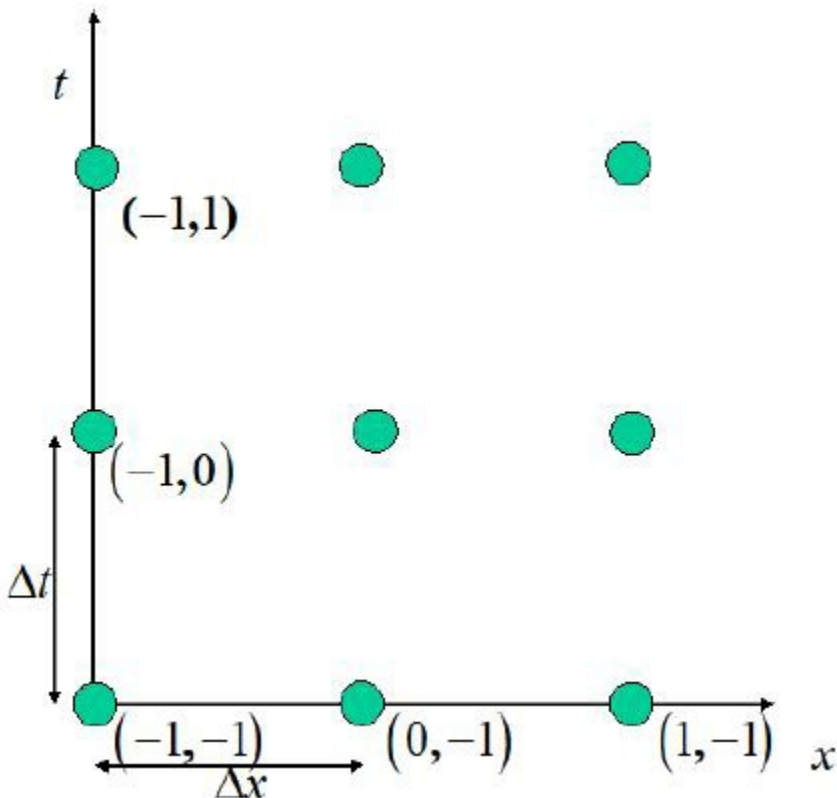


Figure 6.2: Nodal array template for determination of $\widehat{\exp}(x, t)$. A value of $\exp(x, t)$

is required at each nodal location and these values are used in [Eq. \(6.19\)](#).

If we neglect the error term and substitute the definitions of the Lagrange polynomials into [Eq. \(6.14\)](#) we obtain

$$\begin{aligned} \widehat{\text{exp}}(x, t) = & \quad (6.18) \\ & \frac{1}{2}x(x-1) \left[\frac{1}{2}t(t-1) \exp(-1, -1) - (t+1)(t-1) \exp(-1, 0) + \frac{1}{2}t(t+1) \exp(-1, 1) \right] - \\ & \quad (6.19) \\ & (x+1)(x-1) \left[\frac{1}{2}t(t-1) \exp(0, -1) - (t+1)(t-1) \exp(0, 0) + \frac{1}{2}t(t+1) \exp(0, 1) \right] + \\ & \frac{1}{2}x(x+1) \left[\frac{1}{2}t(t-1) \exp(1, -1) - (t+1)(t-1) \exp(1, 0) + \frac{1}{2}t(t+1) \exp(1, 1) \right] \end{aligned}$$

where the $\text{exp}(x, t)$ values in the expansion are those found at the nodes. Thus [Eq. \(6.19\)](#) allows the determination of the value of $\text{exp}(x, t)$ anywhere in the area $x \in [-1, 1]$, $y \in [-1, 1]$ based upon the point values at the nodes as indicated in [Fig. 6.2](#).

6.3 Finite Difference Approximation

Following the logic we used earlier for one-dimensional problems, we now formulate the approximation of a derivative of $f(x, t)$ defined in two dimensions, for example, $\frac{\partial}{\partial x} f(x, t)$. Consider the segment of the grid shown in grey in [Fig. 6.3](#). This area can be expressed as

$$[a, b] = [x_i, x_{i+1}] \quad (6.20)$$

$$[c, d] = [t_j, t_{j+1}]. \quad (6.21)$$

The approximation of $\frac{\partial}{\partial x} f(x_i, t_j)$ over this rectangular area can be obtained directly by differentiation of [Eq. \(6.10\)](#). We obtain

$$\frac{\partial}{\partial x} f(x, t) = \sum_{i=0}^n \sum_{j=0}^m \frac{\partial \ell_i^n(x)}{\partial x} \ell_j^m(t) f(x_i, t_j) + \frac{\partial}{\partial x} E_{nm} \quad (6.22)$$

which, when customized for our problem, becomes

(6.23)

$$\begin{aligned}
\frac{\partial f(x, t)}{\partial x} \Big|_{x_i, t_j} &= f(x_i, t_j) \frac{\partial \ell_i(x)}{\partial x} \Big|_{x_i} \ell_j(t_j) + f(x_{i+1}, t_j) \frac{\partial \ell_{i+1}(x)}{\partial x} \Big|_{x_i} \ell_j(t_j) \\
&+ f(x_i, t_{j+1}) \frac{\partial \ell_i(x)}{\partial x} \Big|_{x_i} \ell_{j+1}(t_j) \\
&+ f(x_{i+1}, t_{j+1}) \frac{\partial \ell_{i+1}(x)}{\partial x} \Big|_{x_i} \ell_{j+1}(t_j) + \frac{d}{dx}(E_{i,j}).
\end{aligned}$$

Substitution of the definitions of the Lagrange polynomials yields

(6.24)

$$\begin{aligned}
\frac{\partial f(x, y)}{\partial x} \Big|_{x_i, t_j} &= f(x_i, t_j) \frac{1}{x_i - x_{i+1}} \underbrace{\frac{t_j - t_{j+1}}{t_j - t_{j+1}}}_{=1} + f(x_{i+1}, t_j) \frac{1}{x_{i+1} - x_i} \underbrace{\frac{t_j - t_{j+1}}{t_j - t_{j+1}}}_{=1} \\
&+ f(x_i, t_{j+1}) \frac{1}{x_i - x_{i+1}} \underbrace{\frac{t_j - t_j}{t_{j+1} - t_j}}_{=0} \\
&+ f(x_{i+1}, t_{j+1}) \frac{1}{x_{i+1} - x_i} \underbrace{\frac{t_j - t_j}{t_{j+1} - t_j}}_{=0} + \frac{d}{dx}(E_{i,j}).
\end{aligned}$$

Upon substitution of nodal coordinates we get (note that $p_1 = (x - x_0)(x - x_1)$)

(6.25)

$$\begin{aligned}
\frac{\partial}{\partial x} f(x_i, t_j) &= \\
\frac{1}{\Delta x} [f(x_{i+1}, t_j) - f(x_i, t_j)] &+ \\
\frac{p_1(x_i)}{2!} \frac{\partial^3}{\partial x^3} f(\xi, t_j) + \boxed{\left(\underbrace{(x_i - x_i)}_{=0} + (x_i - x_{i+1}) \right) \frac{\partial^2}{\partial x^2} f(\xi, t_j)} &+ \\
\frac{p_1(t_j)}{2!} \frac{\partial}{\partial x} \frac{\partial^2}{\partial t^2} f(x_i, \eta) + \underbrace{\frac{\partial^2}{\partial x^2} \frac{\partial^2}{\partial t^2} \left(p_1(t_j) \frac{1}{4!} \left[\underbrace{(x_i - x_i)}_{=0} + (x_i - x_{i+1}) \right] \right)}_{=0 \text{ since } \frac{\partial^2}{\partial t^2} p_1(t_j) = 0} & f(\xi', \eta') - \\
\frac{1}{4!} \underbrace{[p_1(x_i) p_1(t_j)]}_{=0} \frac{\partial^3}{\partial x^3} \frac{\partial^2}{\partial t^2} f(\xi', \eta'). &
\end{aligned}$$

After defining $\Delta x \equiv x_{i+1} - x_i$ and $\Delta t = t_{j+1} - t_j$ we can write

$$\frac{\partial}{\partial x} f(x_i, t_j) = \frac{1}{\Delta x} [f(x_{i+1}, t_j) - f(x_i, t_j)] + O(\Delta x). \quad (6.26)$$

Similarly one can show that

$$\frac{\partial}{\partial t} f(x_i, t_j) = \frac{1}{\Delta t} [f(x_i, t_{j+1}) - f(x_i, t_j)] + O(\Delta t). \quad (6.27)$$

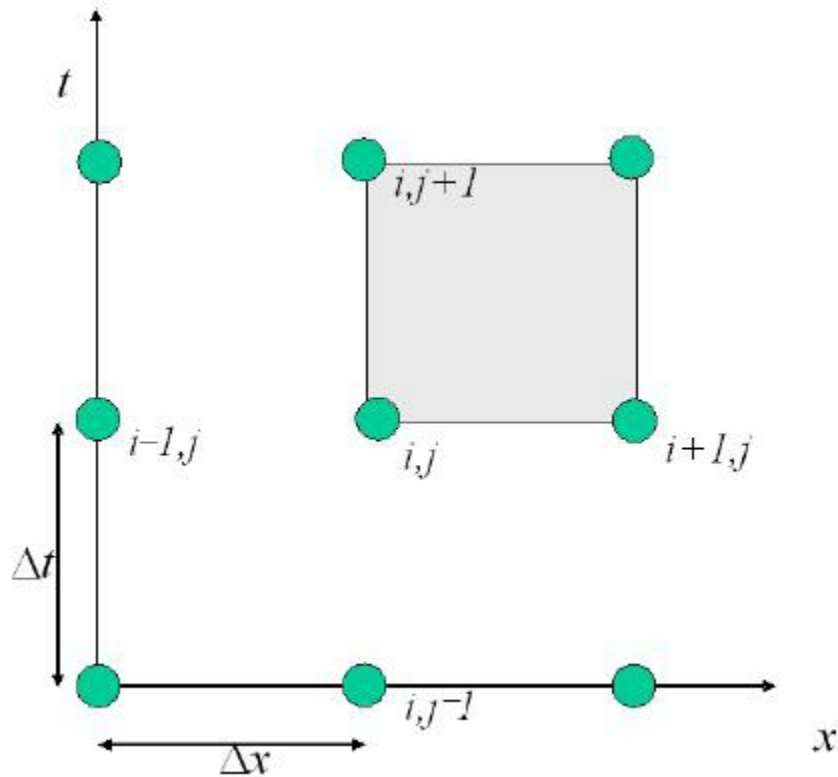


Figure 6.3: Computational molecule for [Eqs. \(6.26\)](#) and [\(6.27\)](#) shown in grey.

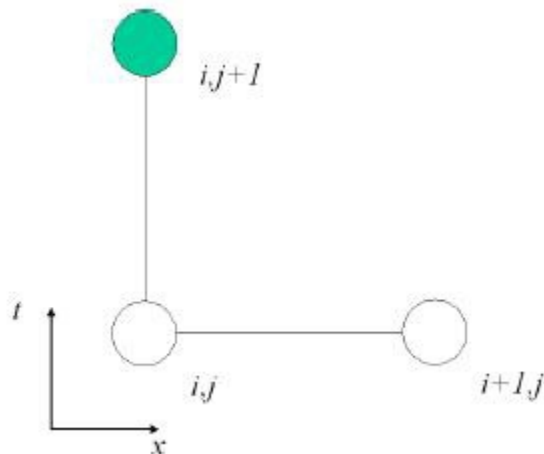


Figure 6.4: Computational molecule for node location i, j associated with [Fig. 6.3](#).

A computational molecule (or template) of the approximation shown in [Eqs. \(6.26\)](#) and [\(6.27\)](#) is found in [Fig. 6.3](#). The gray area is the area represented by this molecule.

6.3.1 Example of Implicit First-Order Accurate Finite Difference Calculation

Consider the type equation presented above, namely

$$\frac{\partial u(x, t)}{\partial x} + \frac{\partial u(x, t)}{\partial t} = 0 \quad (6.28)$$

and impose the following initial and boundary conditions

$$u(x, 0) = 0 \quad (6.29)$$

$$u(0, t) = 1. \quad (6.30)$$

Now evaluate [Eq. \(6.24\)](#) at location $i + 1, j + 1$, that is evaluate $\frac{\partial}{\partial x} u(x_{i+1}, t_{j+1})$. Also obtain $\frac{\partial}{\partial t} u(x_{i+1}, t_{j+1})$. Let $i = 0, 1, 2$ and $j = 0, 1, 2$. Also assume $\Delta x = \Delta t = 1$. For convenience, define $\rho \equiv \frac{\Delta t}{\Delta x} = 1$. The finite-difference expression for [Eq. \(6.28\)](#) using the node x_{i+1}, t_{j+1} as the point of reference is

$$\frac{u_{i+1,j+1} - u_{i+1,j}}{\Delta t} + \frac{u_{i+1,j+1} - u_{i,j+1}}{\Delta x} = 0. \quad (6.31)$$

The computational molecule for this expression is found in [Fig. 6.5](#). Rearranging [Eq. \(6.31\)](#) we obtain

$$\begin{aligned} u_{i+1,j+1} &= \frac{1}{1 + \rho} u_{i+1,j} + \frac{\rho}{1 + \rho} u_{i,j+1} \\ &= \frac{1}{2} u_{i+1,j} + \frac{1}{2} u_{i,j+1}. \end{aligned} \quad (6.32)$$

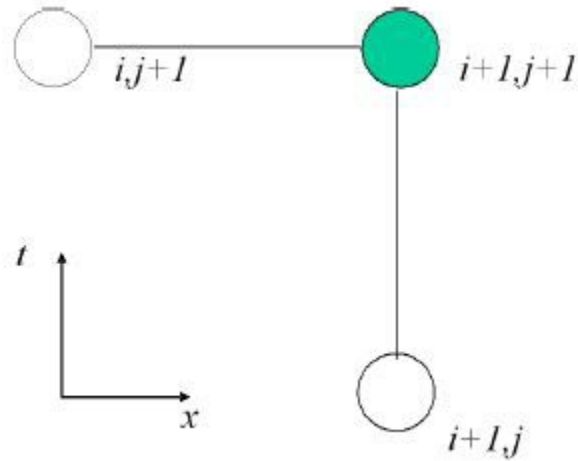


Figure 6.5: Computational molecule for implicit backward in time and backward in space first derivative approximations.

Time Step One

Consider the nodal arrangement in [Fig. 6.6](#). Along the x axis the values are known for the 0 level time step, $t = j\Delta t = 0$ from the initial conditions. Thus there is a value of 0 associated with each of these nodes. For the left-hand boundary, $x = 0$, the values are known for all time by the boundary condition; therefore the value at $x = 0$ for all time is 1. So, the first value we actually need to calculate is at node 1, 1. The equation is

$$\begin{aligned}
 u_{1,1} &= \frac{1}{2}u_{1,0} + \frac{1}{2}u_{0,1} & (6.33) \\
 &= \frac{1}{2}(0) + \frac{1}{2}(1) \\
 &= \frac{1}{2}.
 \end{aligned}$$

Now that we know the value $u_{1,1} = \frac{1}{2}$ we can proceed to node 2, 1. The equation at node 2, 1 is

$$\begin{aligned}
 u_{2,1} &= \frac{1}{2}u_{2,0} + \frac{1}{2}u_{1,1} & (6.34) \\
 &= \frac{1}{2}(0) + \frac{1}{2}\left(\frac{1}{2}\right) \\
 &= \frac{1}{4}.
 \end{aligned}$$

This completes the calculations for time step one.

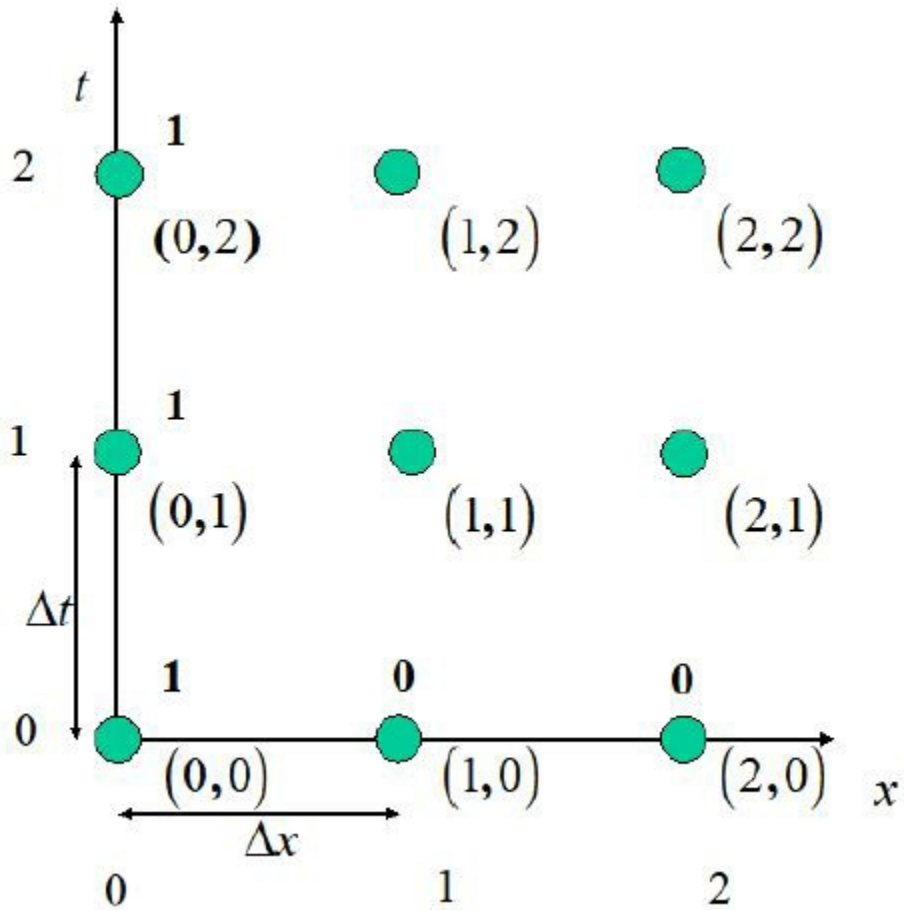


Figure 6.6: Values assigned at beginning of calculation sequence for example problem.

Time Step Two

We are now on the third row of the computational molecule at location $t = 2$. The first unknown value moving left to right along the line at $t = 2$ is

$$\begin{aligned}
 u_{1,2} &= \frac{1}{2}u_{1,1} + \frac{1}{2}u_{0,2} & (6.35) \\
 &= \frac{1}{2} \left(\frac{1}{2} \right) + \frac{1}{2}(1) \\
 &= \frac{3}{4}
 \end{aligned}$$

and once the value at $u_{1,2}$ is known we can proceed to $u_{2,2}$ where the equation is

$$\begin{aligned}
 u_{2,2} &= \frac{1}{2}u_{2,1} + \frac{1}{2}u_{1,2} & (6.36) \\
 &= \frac{1}{2} \left(\frac{1}{4} \right) + \frac{1}{2} \left(\frac{3}{4} \right) \\
 &= \frac{1}{2}.
 \end{aligned}$$

We have now completed two time steps. The solution for each time step is shown graphically in [Fig. 6.7](#)

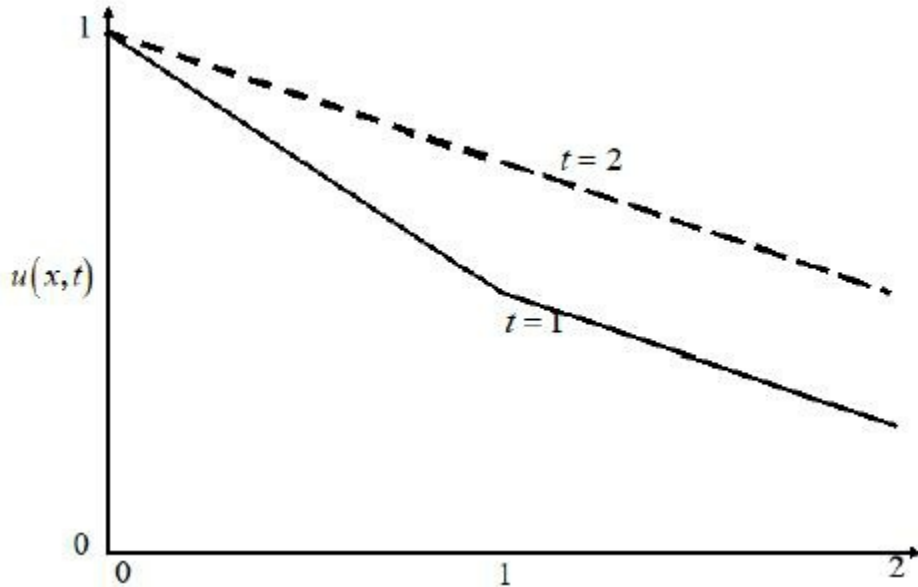


Figure 6.7: Solution to finite difference approximation after two time steps have been executed.

6.3.2 Example of Second Order Accurate Finite Difference Approximation in Space

We will now consider the possibility of representing the space derivative in the example problem using a different spatial approximation. Consider once again the type equation presented above, namely

$$\frac{\partial u(x,t)}{\partial t} + \frac{\partial u(x,t)}{\partial x} = 0 \quad x \in [0,2], \quad t \in [0,2] \quad (6.37)$$

and impose the following initial and boundary conditions

$$u(x,0) = 0 \quad (6.38)$$

$$u(0,t) = 1. \quad (6.39)$$

We now need to select a suitable finite difference approximation to this equation. Let us use the following centered in space and backward in time approximations

$$\frac{u_{i,j+1} - u_{i,j}}{\Delta t} + \frac{u_{i+1,j+1} - u_{i-1,j+1}}{2\Delta x} = 0 \quad (6.40)$$

or, upon rearranging so the unknown information is on the left hand side of the equation and the known information is on the right,

$$\frac{u_{i,j+1}}{\Delta t} + \frac{u_{i+1,j+1} - u_{i-1,j+1}}{2\Delta x} = \frac{u_{i,j}}{\Delta t}. \quad (6.41)$$

The template we wish to use for this equation is shown in [Fig. 6.8](#) where the unknown values are located at filled in circles and the known information is at the open circles.

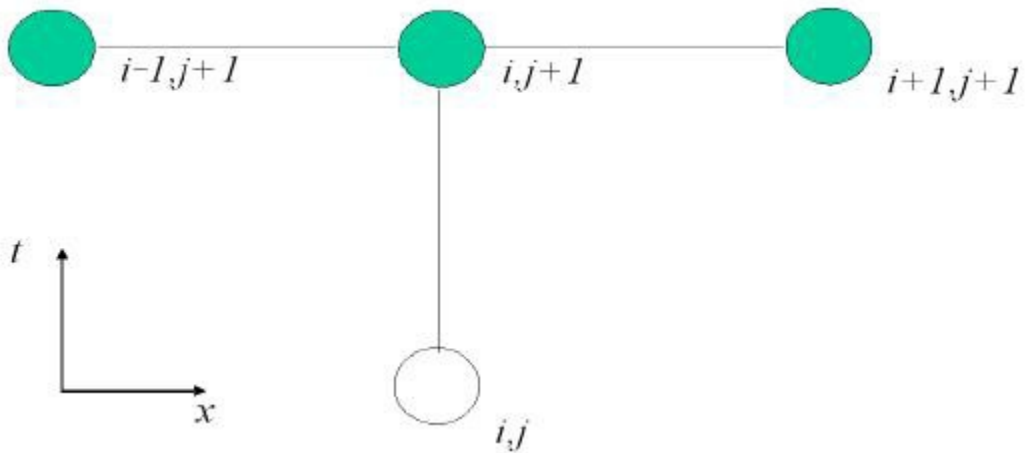


Figure 6.8: Computational molecule for an implicit first-order accurate finite difference calculation in time and a second-order accurate approximation in space.

Step One

Write [Eq. \(6.41\)](#) for the special case of $\Delta t = 1$ and $\Delta x = 1$ as shown in [Fig. 6.6](#) which is

$$\frac{u_{1,1}}{1} + \frac{u_{2,1} - u_{0,1}}{2(1)} = \frac{u_{1,0}}{1}. \quad (6.42)$$

Now we use the fact that the initial condition is given as $u(x, 0) = 0$ which implies that $u_{1,0} = 0$. Because the boundary condition is $u(0, t) = 1$ we have $u_{0,1} = 1$. Introducing these conditions into [Eq. \(6.42\)](#) we obtain

$$\frac{u_{1,1}}{1} + \frac{u_{2,1} - 1}{2(1)} = \frac{0}{1}. \quad (6.43)$$

We cannot write an equation for node 2, 1 using [Eq. \(6.41\)](#) because there is no node to the right of node 2, 1 which would be the $i + 1$ node in [Eq. \(6.41\)](#). Thus [Eq. \(6.43\)](#) has two unknowns. How can we solve this equation? One strategy is to use a backward in space approximation for node 2, 1. A suitable expression is

$$\frac{u_{2,1} - u_{2,0}}{\Delta t} + \frac{u_{2,1} - u_{1,1}}{\Delta x} = 0 \quad (6.44)$$

which for our specific case becomes

$$\frac{u_{2,1} - u_{2,0}}{1} + \frac{u_{2,1} - u_{1,1}}{1} = 0. \quad (6.45)$$

This would provide two equations in two unknowns which can be solved. This is an **implicit approximation**. An implicit approximation requires that one solve an equation involving both the current state of the system and the later one. This requires the solution of a system of equations, and in this case the set of equations can be written in matrix form as

$$\underbrace{\begin{bmatrix} -\frac{1}{\Delta x} & \frac{1}{\Delta x} & 0 \\ -\frac{1}{2\Delta x} & 0 & \frac{1}{2\Delta x} \\ 0 & -\frac{1}{\Delta x} & \frac{1}{\Delta x} \end{bmatrix}}_A \begin{bmatrix} u(0,0) \\ u_{1,1} \\ u_{2,1} \end{bmatrix} = - \underbrace{\begin{bmatrix} \frac{1}{\Delta t} & 0 & 0 \\ 1 & \frac{1}{\Delta t} & 0 \\ 1 & 0 & \frac{1}{\Delta t} \end{bmatrix}}_B \left[\begin{Bmatrix} u(0,0) \\ u_{1,1} \\ u_{2,1} \end{Bmatrix} - \begin{Bmatrix} u(0,0) \\ u(1,0) \\ u(2,0) \end{Bmatrix} \right] \quad (6.46)$$

where the term identified as A represents the space derivative and the term identified as B the time derivative. Note that some unknown variables have been replaced by known values represented by initial and boundary conditions.

Since $u(0, t)$ is known to be 1, the first column can be multiplied by this value and the first row and column now contain no useful information and can be eliminated to give

$$\underbrace{\begin{bmatrix} 0 & \frac{1}{2\Delta x} \\ -\frac{1}{\Delta x} & \frac{1}{\Delta x} \end{bmatrix}}_A \begin{bmatrix} u_{1,1} \\ u_{2,1} \end{bmatrix} = - \underbrace{\begin{bmatrix} \frac{1}{\Delta t} & 0 \\ 0 & \frac{1}{\Delta t} \end{bmatrix}}_B \left\{ \begin{bmatrix} u_{1,1} \\ u_{2,1} \end{bmatrix} - \begin{bmatrix} u_{1,0} \\ u_{02,0} \end{bmatrix} \right\} + \underbrace{\begin{bmatrix} \left(\frac{1}{2\Delta x}\right)(1) \\ 0 \end{bmatrix}}_C. \quad (6.47)$$

The term C now contains the information about the boundary condition.

Now collect all the information regarding the unknown vector $\begin{bmatrix} u_{1,1} \\ u_{2,1} \end{bmatrix}$ to the left-hand side of the equation to yield

$$\left\{ \begin{bmatrix} 0 & \frac{1}{2\Delta x} \\ -\frac{1}{\Delta x} & \frac{1}{\Delta x} \end{bmatrix} + \begin{bmatrix} \frac{1}{\Delta t} & 0 \\ 0 & \frac{1}{\Delta t} \end{bmatrix} \right\} \begin{bmatrix} u_{1,1} \\ u_{2,1} \end{bmatrix} = \begin{bmatrix} \frac{1}{\Delta t} & 0 \\ 0 & \frac{1}{\Delta t} \end{bmatrix} \begin{bmatrix} u_{1,0} \\ u_{2,0} \end{bmatrix} + \begin{bmatrix} \left(\frac{1}{2\Delta x}\right)(1) \\ 0 \end{bmatrix}. \quad (6.48)$$

We now substitute the specific values for $\Delta x = 1$ and $\Delta t = 1$ to give

$$\left\{ \begin{bmatrix} 0 & \frac{1}{2} \\ -\frac{1}{1} & \frac{1}{1} \end{bmatrix} + \begin{bmatrix} \frac{1}{1} & 0 \\ 0 & \frac{1}{1} \end{bmatrix} \right\} \begin{bmatrix} u_{1,1} \\ u_{2,1} \end{bmatrix} = \begin{bmatrix} \frac{1}{1} & 0 \\ 0 & \frac{1}{1} \end{bmatrix} \begin{bmatrix} u_{1,0} \\ u_{2,0} \end{bmatrix} + \begin{bmatrix} \left(\frac{1}{2}\right)(1) \\ 0 \end{bmatrix} \quad (6.49)$$

which reduces to

$$\begin{bmatrix} 1 & 0.5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} u_{1,1} \\ u_{2,1} \end{bmatrix} = \begin{bmatrix} \frac{1}{1} & 0 \\ 0 & \frac{1}{1} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix} \quad (6.50)$$

or

$$\begin{bmatrix} 1 & 0.5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} u_{1,1} \\ u_{2,1} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix} \quad (6.51)$$

which yields

$$\begin{bmatrix} u_{1,1} \\ u_{2,1} \end{bmatrix} = \frac{1}{2.5} \begin{bmatrix} 2 & -0.5 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.2 \end{bmatrix}. \quad (6.52)$$

Now let us calculate the next time step. This is done by substituting the newly calculated solution into [Eq. \(6.50\)](#) to get

$$\begin{bmatrix} 1 & 0.5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} u_{1,1} \\ u_{2,1} \end{bmatrix} = \begin{bmatrix} \frac{1}{1} & 0 \\ 0 & \frac{1}{1} \end{bmatrix} \begin{bmatrix} 0.4 \\ 0.2 \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix} \quad (6.53)$$

which simplifies to

$$\begin{bmatrix} 1 & 0.5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} u_{1,1} \\ u_{2,1} \end{bmatrix} = \begin{bmatrix} 0.9 \\ 0.2 \end{bmatrix}. \quad (6.54)$$

We can solve this equation by taking the inverse of the coefficient matrix and we obtain for the unknown values $u_{1,1}$ and $u_{2,1}$

$$\begin{bmatrix} u_{1,1} \\ u_{2,1} \end{bmatrix} = \frac{1}{2.5} \begin{bmatrix} 2 & -0.5 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0.9 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 0.68 \\ 0.44 \end{bmatrix}. \quad (6.55)$$

This ends the calculation for the first time step and the process is repeated until the target time period is completed.

In the above example we have introduced two ideas we will use again later. The first is the concept of introducing the specified value boundary condition (first or Dirichlet type) by using a value of 1 on the diagonal of the first row. The second is reducing the matrix rank by removing rows and columns through the process of the accommodation of the boundary condition.

6.4 Stability of Finite Difference Approximations

Several approaches exist to determine the stability of a numerical scheme that involves derivatives in space and time. The approach we will employ is called the **Matrix Method**. Consider the equation

$$\frac{\partial u(x, t)}{\partial t} - \frac{\partial^2 u(x, t)}{\partial x^2} = 0 \quad x \in [0, L] \quad t \in [0, \infty) \quad (6.56)$$

with boundary and initial conditions (note that this is second order in space, so we will need two boundary conditions in x as well as an initial condition)

$$u(x, 0) = u_0(x) = 0 \quad (6.57)$$

$$u(0, t) = u_1(t) = 0 \quad (6.58)$$

$$u(L, t) = 0. \quad (6.59)$$

In this development we will employ, as an example, an explicit forward difference approximation in time and a centered finite difference approximation in space. Introduction of these approximations yields the finite difference equation

$$\frac{1}{\Delta t} \left[u(x_i, t_{j+1}) - u(x_i, t_j) \right] = \frac{1}{\Delta x^2} \left[u(x_{i+1}, t_j) - 2u(x_i, t_j) + u(x_{i-1}, t_j) \right]. \quad (6.60)$$

To simplify notation let

$$r \equiv \frac{\Delta t}{\Delta x^2} \quad (6.61)$$

and rewrite the finite difference formula as

$$u_{i,j+1} = ru_{i+1,j} + (1 - 2r)u_{i,j} + ru_{i-1,j} \quad i = 0, 1, \dots, N. \quad (6.62)$$

Now we write the matrix form of [Eq. \(6.62\)](#) for the $j + 1$ time step. After imposing the boundary conditions (remember the boundary conditions are zero so no term appears on the right hand side to identify them) we have

$$\begin{bmatrix} u_{1,j+1} \\ u_{2,j+1} \\ u_{3,j+1} \\ \vdots \\ \vdots \\ u_{N-1,j+1} \end{bmatrix} = \begin{bmatrix} (1-2r) & r & 0 & 0 & 0 & 0 & 0 \\ r & (1-2r) & r & 0 & 0 & 0 & 0 \\ 0 & r & (1-2r) & r & 0 & 0 & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & r & (1-2r) \end{bmatrix} \begin{bmatrix} u_{1,j} \\ u_{2,j} \\ u_{3,j} \\ \vdots \\ \vdots \\ u_{N-1,j} \end{bmatrix}. \quad (6.63)$$

We can write [Eq. \(6.63\)](#) symbolically as

$$\{u\}_{j+1} = [A] \{u\}_j. \quad (6.64)$$

Assume an exact solution, free of round-off error is given by $\{\bar{u}\}_{j+1}$. Note that this vector contains the values of $u_{i,j+1}$ at all nodes, except the Dirichlet boundary nodes, at time $(j+1)\Delta t$. Since $\{\bar{u}\}_{j+1}$ is a solution to our equation we can write

$$\{\bar{u}\}_{j+1} = [A] \{\bar{u}\}_j. \quad (6.65)$$

Now subtract [Eq. \(6.65\)](#) from [Eq. \(6.64\)](#) to obtain a measure of the error or difference between the exact and computed values, that is

$$\{u - \bar{u}\}_{j+1} = [A] \{u - \bar{u}\}_j. \quad (6.66)$$

Next define this error as

$$\{\varepsilon\}_{j+1} = \{u - \bar{u}\}_{j+1}. \quad (6.67)$$

We can now rewrite [Eq. \(6.66\)](#) as

$$\{\varepsilon\}_{j+1} = [A] \{\varepsilon\}_j. \quad (6.68)$$

Assuming no further errors are committed, other than the round-off errors that exist at u_0 we have

$$\{\varepsilon\}_1 = [A] \{\varepsilon\}_0 \quad (6.69)$$

$$\{\varepsilon\}_2 = [A] \{\varepsilon\}_1 = [A]^2 \{\varepsilon\}_0 \quad (6.70)$$

$$\vdots \quad (6.71)$$

$$\vdots$$

$$\vdots$$

$$\{\varepsilon\}_j = [A]^j \{\varepsilon\}_0$$

which illustrates how the error $\{\varepsilon\}_j$ changes as the solution advances through time.

In the next step we analyze the behavior of this scheme through the use of *eigenvalues*. Assume: (1) there exist $N - 1$ different eigenvalues of $[A]$, that is λ_s , $s = 1, 2, \dots, N - 1$ and (2) there exist a set of linearly independent eigenvectors, $\{v\}_s$, $s = 1, \dots, N - 1$. Using these eigenvectors as a kind of basis, we can write

$$\{\varepsilon\}_0 = \sum_{s=1}^{N-1} c_s \{v\}_s \quad (6.72)$$

where c_s , $s = 1, \dots, N - 1$ are numbers. Note also that by definition

$$[A] \{v\}_s = \lambda_s \{v\}_s. \quad (6.73)$$

Now substitute [Eq. \(6.72\)](#) in [Eq. \(6.71\)](#) to obtain

$$\{\varepsilon\}_j = [A]^j \{\varepsilon\}_0 = [A]^j \sum_{s=1}^{N-1} c_s \{v\}_s. \quad (6.74)$$

We now proceed through a series of apparently unrelated steps in pursuit of a specific result. We begin with the modification of [Eq. \(6.74\)](#), that is

$$\begin{aligned} \{\varepsilon\}_j &= [A]^j \sum_{s=1}^{N-1} c_s \{v\}_s \\ &= [A]^{j-1} [A] \sum_{s=1}^{N-1} c_s \{v\}_s \\ &= [A]^{j-1} \sum_{s=1}^{N-1} c_s [A] \{v\}_s. \end{aligned} \quad (6.75)$$

Note that we can use [Eq. \(6.73\)](#) to modify the last expression, that is

$$\begin{aligned}
[A]^{j-1} \sum_{s=1}^{N-1} c_s [A] \{v\}_s &= [A]^{j-1} \sum_{s=1}^{N-1} c_s \lambda_s \{v\}_s & (6.76) \\
&= [A]^{j-2} [A] \sum_{s=1}^{N-1} c_s \lambda_s \{v\}_s \\
&= [A]^{j-2} \sum_{s=1}^{N-1} c_s \lambda_s [A] \{v\}_s \\
&= [A]^{j-2} \sum_{s=1}^{N-1} c_s \lambda_s^2 \{v\}_s \\
&\cdot \\
&\cdot \\
&\cdot \\
&= [A]^0 \sum_{s=1}^{N-1} c_s \lambda_s^j \{v\}_s .
\end{aligned}$$

So we have

$$\{\varepsilon\}_j = [A]^0 \sum_{s=1}^{N-1} c_s \lambda_s^j \{v\}_s . \quad (6.77)$$

Examination of [Eq. \(6.77\)](#) indicates that the error $\{\varepsilon\}_j$ will grow if the eigenvalue λ_s is greater than 1.

The matrix $[A]$ is known to have the eigenvalues

$$\lambda_s = 1 - 4 r \sin^2 \left(\frac{s\pi}{2N'} \right) \quad (6.78)$$

where N' is the maximum number of nodes. Thus to assure stability

$$\lambda_s \leq 1 \quad (6.79)$$

or

$$\left| 1 - 4 r \sin^2 \left(\frac{s\pi}{2N'} \right) \right| \leq 1 \quad (6.80)$$

which translates into

$$-1 \leq 1 - 4 r \sin^2 \left(\frac{s\pi}{2N'} \right) \leq 1. \quad (6.81)$$

Because $\sin^2 \left(\frac{s\pi}{2N'} \right)$ is always positive, the upper bound is always satisfied. For the lower bound we have

$$-1 \leq 1 - 4 r \sin^2 \left(\frac{s\pi}{2N'} \right) \quad (6.82)$$

$$\Rightarrow -2 \leq -4 r \sin^2 \left(\frac{s\pi}{2N'} \right) \quad (6.83)$$

$$\Rightarrow 2 \geq 4 r \sin^2 \left(\frac{s\pi}{2N'} \right) \quad (6.84)$$

$$\Rightarrow \frac{2}{4 \sin^2 \left(\frac{s\pi}{2N'} \right)} \geq r \quad (6.85)$$

$$\Rightarrow r \leq \frac{1}{2 \sin^2 \left(\frac{s\pi}{2N'} \right)}. \quad (6.86)$$

The worst case scenario is when

$$\sin^2 \left(\frac{s\pi}{2N'} \right) = 1. \quad (6.87)$$

Then the constraint on stability is, according to [Eq. \(6.86\)](#)

$$r \leq \frac{1}{2}. \quad (6.88)$$

Recalling the definition of r (see [Eq. \(6.61\)](#)) we obtain

$$\frac{\Delta t}{\Delta x^2} \leq \frac{1}{2}. \quad (6.89)$$

Equation [\(6.89\)](#) states that to assure stability, the time step must be related to the space

increment such that $\Delta t \leq \frac{\Delta x^2}{2}$.

6.4.1 Example of Stability

Case of $r = 1$

Consider the approximation for

$$\frac{\partial u(x,t)}{\partial t} - \frac{\partial^2 u(x,t)}{\partial x^2} = 0 \quad x = [0,1] \quad t = [0,2/9] \quad (6.90)$$

provided in [Eq. \(6.62\)](#) for the case of four nodes in space, namely

$$u_{i,j+1} = ru_{i+1,j} + (1 - 2r)u_{i,j} + ru_{i-1,j} \quad i = 0,1,2,3 \quad x \in [0,1] \quad (6.91)$$

with boundary and initial conditions

$$u(x,0) = 0 \quad (6.92)$$

$$u(0,t) = 1 \quad (6.93)$$

$$u(1,t) = 0. \quad (6.94)$$

The problem is defined in [Fig. 6.9](#)

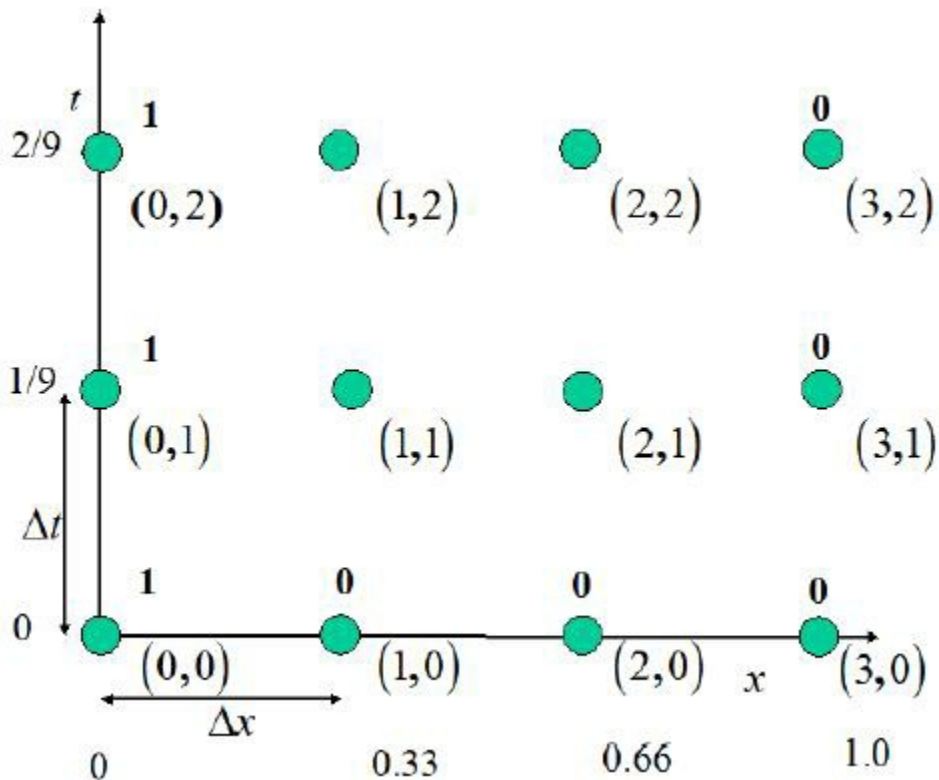


Figure 6.9: Nodal array for the finite difference solution to [Eqs. \(6.90\)–\(6.94\)](#).

The template or computational molecule for the finite difference approximation to [Eq. \(6.90\)](#) given by [Eq. \(6.91\)](#) is shown in [Fig. 6.10](#). As the calculations proceed, only the shaded node is unknown, that is the clear nodes are known from the preceding time level.

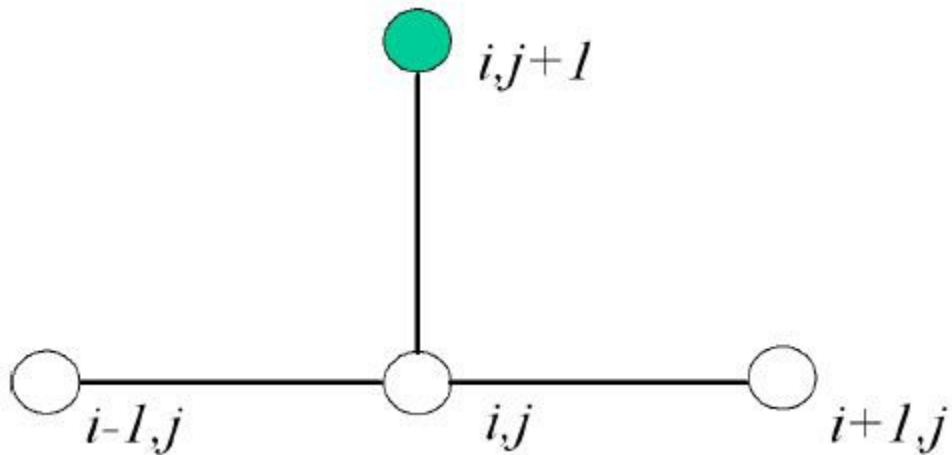


Figure 6.10: Computational molecule for the Euler (forward difference) approximation to [Eq. \(6.90\)](#) given by [Eq. \(6.91\)](#).

Step One ($j = 0$)

Substitution of $j = 0$ into [Eq. \(6.91\)](#) yields

$$u_{1,1} = (1) \times \underbrace{u_{2,0}}_0 - (1) \times \underbrace{u_{1,0}}_0 + (1) \times \underbrace{u_{0,0}}_1 = 1 \quad (6.95)$$

$$u_{2,1} = (1) \times \underbrace{u_{3,0}}_0 - (1) \times \underbrace{u_{2,0}}_0 + (1) \times \underbrace{u_{1,0}}_0 = 0 \quad (6.96)$$

where the values under the variable names indicate the current value of those variables.

Step Two ($j = 1$)

$$u_{1,2} = (1) \times \underbrace{u_{2,1}}_0 - (1) \times \underbrace{u_{1,1}}_1 + (1) \times \underbrace{u_{0,1}}_1 = 0 \quad (6.97)$$

$$u_{2,2} = (1) \times \underbrace{u_{3,1}}_0 - (1) \times \underbrace{u_{2,1}}_0 + (1) \times \underbrace{u_{1,1}}_1 = 1. \quad (6.98)$$

Step Three ($j = 2$)

$$u_{1,3} = (1) \times \underbrace{u_{2,2}}_1 - (1) \times \underbrace{u_{1,2}}_0 + (1) \times \underbrace{u_{0,2}}_1 = 2 \quad (6.99)$$

$$u_{2,3} = (1) \times \underbrace{u_{3,2}}_0 - (1) \times \underbrace{u_{2,2}}_1 + (1) \times \underbrace{u_{1,2}}_0 = -1. \quad (6.100)$$

The solution for the first three time steps ($j = 0, 1, 2$) is shown in [Fig. 6.11](#). It is clearly unstable.

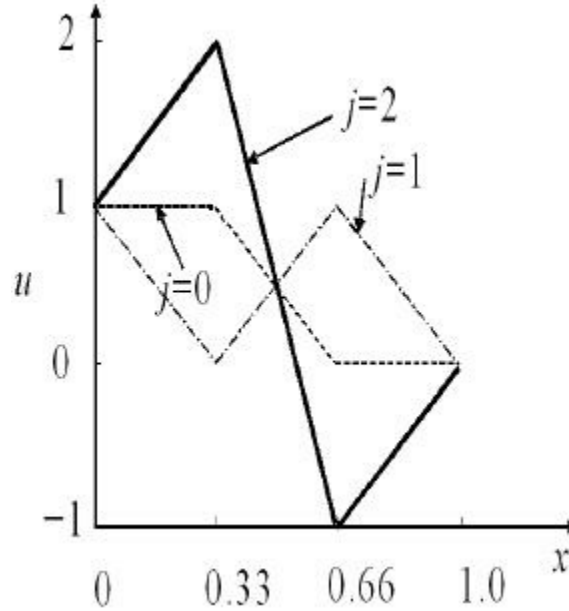


Figure 6.11: Explicit unstable solution for time steps $j = 0, 1, 2$.

Case of $r = 1/4$

Now we reduce the value of r to $r = 1/4$. The calculations follow:

Step One ($j = 0$)

$$u_{1,1} = \left(\frac{1}{4}\right) \times \underbrace{u_{2,0}}_0 + \left(\frac{1}{2}\right) \times \underbrace{u_{1,0}}_0 + \left(\frac{1}{4}\right) \times \underbrace{u_{0,0}}_1 = \frac{1}{4} \quad (6.101)$$

$$u_{2,1} = \left(\frac{1}{4}\right) \times \underbrace{u_{3,0}}_0 + \left(\frac{1}{2}\right) \times \underbrace{u_{2,0}}_0 + \left(\frac{1}{4}\right) \times \underbrace{u_{1,0}}_0 = 0. \quad (6.102)$$

Step Two ($j = 1$)

$$u_{1,2} = \left(\frac{1}{4}\right) \times \underbrace{u_{2,1}}_0 + \left(\frac{1}{2}\right) \times \underbrace{u_{1,1}}_{\frac{1}{4}} + \left(\frac{1}{4}\right) \times \underbrace{u_{0,1}}_1 = \frac{3}{8} \quad (6.103)$$

$$u_{2,2} = \left(\frac{1}{4}\right) \times \underbrace{u_{3,1}}_0 + \left(\frac{1}{2}\right) \times \underbrace{u_{2,1}}_0 + \left(\frac{1}{4}\right) \times \underbrace{u_{1,1}}_{\frac{1}{4}} = \frac{1}{16}. \quad (6.104)$$

Step Three ($j = 2$)

$$u_{1,3} = \left(\frac{1}{4}\right) \underbrace{u_{2,2}}_{\frac{1}{16}} + \left(\frac{1}{2}\right) \underbrace{u_{1,2}}_{\frac{3}{8}} + \left(\frac{1}{4}\right) \underbrace{u_{0,2}}_1 = \frac{31}{64} \quad (6.105)$$

$$u_{2,3} = \left(\frac{1}{4}\right) \underbrace{u_{3,2}}_0 + \left(\frac{1}{2}\right) \underbrace{u_{2,2}}_{\frac{1}{16}} + \left(\frac{1}{4}\right) \underbrace{u_{1,2}}_{\frac{3}{8}} = \frac{1}{8}. \quad (6.106)$$

The solution for the first three time steps is provided in [Fig. 6.12](#)

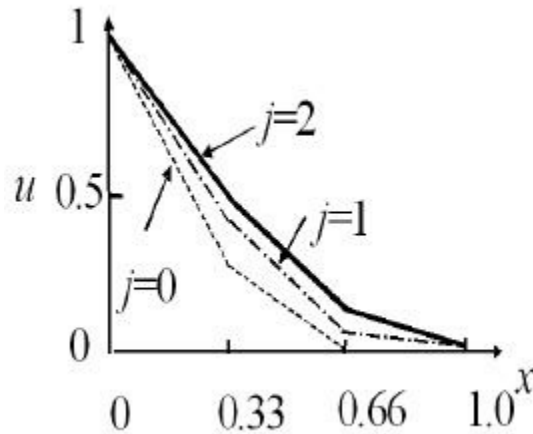


Figure 6.12: Solution for first three time steps for case of $r = \frac{1}{2}$.

It is clear from [Fig. 6.12](#) that the solution is stable.

While these two examples do not prove that the constraints on r are necessary, they do suggest that to be the case.

6.4.2 Example Simulation

Let us assume we have the following equation and boundary conditions

$$D \frac{\partial^2 c(x, t)}{\partial x^2} = \frac{\partial c(x, t)}{\partial t} \quad (6.107)$$

with

$$\begin{aligned} c(0, t) &= c_{0,j} = 0 \quad \text{for all } j \\ \frac{\partial c(L, t)}{\partial x} &= \frac{\partial c_{3,j}}{\partial x} = 0 \quad \text{for all } j \\ c(x, 0) &= c_{i,0} = 0 \quad \text{for all } i, \quad i \neq 0. \end{aligned} \quad (6.108)$$

A centered in space, backwards in time finite difference expression for this equation is

$$D_i \frac{c_{i+1,j+1} - 2c_{i,j+1} + c_{i-1,j+1}}{(\Delta x)^2} = \frac{c_{i,j+1} - c_{i,j}}{\Delta t}. \quad (6.109)$$

We assume that there are four nodes equally separated by distance Δx . The second through fourth row of the matrix equation for this problem are obtained using [Eq. \(6.109\)](#) for nodes 1, and 2. The first row is used to accommodate the first type (Dirichlet) boundary condition and the last for the second type (Neumann) boundary condition which is obtained by using a backward difference in space approximation for the expression appearing in [Eq. \(6.108\)](#), that is

$$\begin{aligned} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ D_i \frac{1}{(\Delta x)^2} & D_i \frac{-2}{(\Delta x)^2} & D_i \frac{1}{(\Delta x)^2} & 0 \\ 0 & D_i \frac{1}{(\Delta x)^2} & D_i \frac{-2}{(\Delta x)^2} & D_i \frac{1}{(\Delta x)^2} \\ 0 & 0 & \frac{-1}{\Delta x} & \frac{1}{\Delta x} \end{bmatrix} \begin{bmatrix} c_{0,j+1} \\ c_{1,j+1} \\ c_{2,j+1} \\ c_{3,j+1} \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\Delta t} & 0 & 0 \\ 0 & 0 & \frac{1}{\Delta t} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ c_{1,j+1} - c_{1,j} \\ c_{2,j+1} - c_{2,j} \\ 0 \end{bmatrix}. \end{aligned} \quad (6.110)$$

Collecting the coefficients multiplying the unknown vector at time $(j + 1)\Delta t$ on the left, hand side of the equation we obtain

$$\begin{aligned}
 & \begin{bmatrix} 1 & 0 & 0 & 0 \\ D_i \frac{1}{(\Delta x)^2} & D_i \frac{-2}{(\Delta x)^2} - \frac{1}{\Delta t} & D_i \frac{1}{(\Delta x)^2} & 0 \\ 0 & D_i \frac{1}{(\Delta x)^2} & D_i \frac{-2}{(\Delta x)^2} - \frac{1}{\Delta t} & D_i \frac{1}{(\Delta x)^2} \\ 0 & 0 & \frac{-1}{\Delta x} & \frac{1}{\Delta x} \end{bmatrix} \begin{bmatrix} c_{0,j+1} \\ c_{1,j+1} \\ c_{2,j+1} \\ c_{3,j+1} \end{bmatrix} \\
 & = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\Delta t} & 0 & 0 \\ 0 & 0 & \frac{1}{\Delta t} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ c_{1,j} \\ c_{2,j} \\ 0 \end{bmatrix}
 \end{aligned} \tag{6.111}$$

where the first and last rows provide for the boundary conditions. Now we begin calculating. Since $c_{i,j}$ is zero for $j = 0$ we have

$$\begin{aligned}
 & \begin{bmatrix} 1 & 0 & 0 & 0 \\ D_i \frac{1}{(\Delta x)^2} & D_i \frac{-2}{(\Delta x)^2} - \frac{1}{\Delta t} & D_i \frac{1}{(\Delta x)^2} & 0 \\ 0 & D_i \frac{1}{(\Delta x)^2} & D_i \frac{-2}{(\Delta x)^2} - \frac{1}{\Delta t} & D_i \frac{1}{(\Delta x)^2} \\ 0 & 0 & \frac{-1}{\Delta x} & \frac{1}{\Delta x} \end{bmatrix} \begin{bmatrix} c_{0,1} \\ c_{1,1} \\ c_{2,1} \\ c_{3,1} \end{bmatrix} \\
 & = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\Delta t} & 0 & 0 \\ 0 & 0 & \frac{1}{\Delta t} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ c_{1,0} \\ c_{2,0} \\ 0 \end{bmatrix}.
 \end{aligned} \tag{6.112}$$

The initial conditions can now be employed to replace $c_{1,0}$ and $c_{2,0}$ with known values to yield

$$\begin{aligned}
 & \begin{bmatrix} 1 & 0 & 0 & 0 \\ D_i \frac{1}{(\Delta x)^2} & D_i \frac{-2}{(\Delta x)^2} - \frac{1}{\Delta t} & D_i \frac{1}{(\Delta x)^2} & 0 \\ 0 & D_i \frac{1}{(\Delta x)^2} & D_i \frac{-2}{(\Delta x)^2} - \frac{1}{\Delta t} & D_i \frac{1}{(\Delta x)^2} \\ 0 & 0 & \frac{-1}{\Delta x} & \frac{1}{\Delta x} \end{bmatrix} \begin{bmatrix} c_{0,1} \\ c_{1,1} \\ c_{2,1} \\ c_{3,1} \end{bmatrix} \\
 & = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\Delta t} & 0 & 0 \\ 0 & 0 & \frac{1}{\Delta t} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.
 \end{aligned} \tag{6.113}$$

Assume that we solve these equations for time step 1 and proceed to time step 2. The resulting equation is the same as the above but for the updated ones in time concentrations:

$$\begin{aligned}
& \begin{bmatrix} 1 & 0 & 0 & 0 \\ D_i \frac{1}{(\Delta x)^2} & D_i \frac{-2}{(\Delta x)^2} - \frac{1}{\Delta t} & D_i \frac{1}{(\Delta x)^2} & 0 \\ 0 & D_i \frac{1}{(\Delta x)^2} & D_i \frac{-2}{(\Delta x)^2} - \frac{1}{\Delta t} & D_i \frac{1}{(\Delta x)^2} \\ 0 & 0 & \frac{-1}{\Delta x} & \frac{1}{\Delta x} \end{bmatrix} \begin{bmatrix} c_{0,2} \\ c_{1,2} \\ c_{2,2} \\ c_{3,2} \end{bmatrix} \\
& = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\Delta t} & 0 & 0 \\ 0 & 0 & \frac{1}{\Delta t} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ c_{1,1} \\ c_{2,1} \\ 0 \end{bmatrix}.
\end{aligned} \tag{6.114}$$

Repeating this process one more time we obtain for time step 3

$$\begin{aligned}
& \begin{bmatrix} 1 & 0 & 0 & 0 \\ D_i \frac{1}{(\Delta x)^2} & D_i \frac{-2}{(\Delta x)^2} + \frac{1}{\Delta t} & D_i \frac{1}{(\Delta x)^2} & 0 \\ 0 & D_i \frac{1}{(\Delta x)^2} & D_i \frac{-2}{(\Delta x)^2} + \frac{1}{\Delta t} & D_i \frac{1}{(\Delta x)^2} \\ 0 & 0 & \frac{-1}{\Delta x} & \frac{1}{\Delta x} \end{bmatrix} \begin{bmatrix} c_{0,3} \\ c_{1,3} \\ c_{2,3} \\ c_{3,3} \end{bmatrix} \\
& = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\Delta t} & 0 & 0 \\ 0 & 0 & \frac{1}{\Delta t} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ c_{1,2} \\ c_{2,2} \\ 0 \end{bmatrix}
\end{aligned} \tag{6.115}$$

and so on.

6.5 Galerkin Finite Element Approximations in Time

In this section we extend the ideas introduced in [Section 6.3](#) to consider a time-dependant Galerkin finite-element formulation. We will consider the same equation as [6.107](#), but assume $D = 1$, that is

$$\frac{\partial u(x, t)}{\partial t} - \frac{\partial^2 u(x, t)}{\partial x^2} = 0 \quad x \in [0, 1] \tag{6.116}$$

with initial and boundary conditions

$$u(x, 0) = u_1(x) = 0 \tag{6.117}$$

$$u(0, t) = \bar{u}_0(t) \tag{6.118}$$

$$u(1, t) = \bar{u}_1(t). \tag{6.119}$$

We begin with the definition of the approximating function for one element; this is similar to Eq. (5.58) on page 95, but in this case $\hat{u}(x, t)$ is a function of both space x and

time t . We first represent $u(x, t)$ as

$$u(x, t) = \hat{u}(x, t) + E(x, t) \quad (6.120)$$

and then define $\hat{u}(x, t)$ as follows:

$$\hat{u}(x, t) = \sum_{j=0}^n u_j(t) \ell_j^n(x). \quad (6.121)$$

Note the structure of this expression. *Where we have previously had constant coefficients u_j , we now have functions of time $u_j(t)$.* We could have used another strategy in which we represent the time dimension using a Lagrange polynomial $\ell_j^n(t)$ and our approximating expression would be

$$\hat{u}(x, t) = \sum_{j=0}^n u_j \ell_j^n(t) \ell_j^n(x), \quad (6.122)$$

but the strategy presented in [6.121](#) is a more commonly used approach.

As earlier, we now apply Galerkin's method of weighted residuals for one element, again using the basis function as the weighting function. We obtain

$$\int_x \left(\frac{\partial \hat{u}(x, t)}{\partial t} - \frac{\partial^2 \hat{u}(x, t)}{\partial x^2} \right) \ell_i^n(x) dx = 0 \quad i = 0, 1, \dots, n. \quad (6.123)$$

Using integration by parts on the second-order derivative results in

$$\int_x \frac{\partial \hat{u}(x, t)}{\partial x} \frac{d\ell_i^n(x)}{dx} dx + \int_x \frac{\partial \hat{u}(x, t)}{\partial t} \ell_i^n(x) dx - \frac{\partial \hat{u}}{\partial x} \ell_i(x) \Big|_0^1 = 0, \quad i = 0, 1, \dots, n. \quad (6.124)$$

We now substitute [Eq. \(6.121\)](#) into [Eq. \(6.124\)](#) to obtain, once again for one element,

$$\sum_{j=0}^n u_j(t) \int_x \frac{\partial \ell_j^n}{\partial x} \frac{d\ell_i^n}{dx} dx + \sum_{j=0}^n \frac{du_j(t)}{dt} \int_x \ell_j^n \ell_i^n dx - \frac{\partial \hat{u}}{\partial x} \ell_i(x) \Big|_0^1 = 0, \quad i = 0, 1, \dots, n. \quad (6.125)$$

Examination of [Eq. \(6.125\)](#) reveals that it is similar to earlier Galerkin finite element expressions but for one important detail. Both the function $u_j(t)$ and the time derivative of this function, $\frac{du_j(t)}{dt}$, appear in this equation. In other words we have a set of $n + 1$ ordinary differential equations in time. If we assume a value for n , say $n = 1$ for linear Lagrange functions, we can look at multiple elements by changing the notation to

(6.126)

$$\sum_{j=0}^N u_j(t) \int_x \frac{\partial \ell_j}{\partial x} \frac{d\ell_i}{dx} dx + \sum_{j=0}^N \frac{du_j(t)}{dt} \int_x \ell_j \ell_i dx - \frac{\partial \bar{u}}{\partial x} \ell_i(x) \Big|_0^1 = 0, \quad i = 0, 1, \dots, N$$

where N is the total number of nodes and $\ell_j(x)$ is the node-wise defined Lagrange polynomial; that is each $\ell_j(x)$ ranges over multiple elements (in this case, two). Since [Eq. \(6.126\)](#) now represents a system of elements with $N + 1$ equations in $N + 1$ unknowns, we can write it in matrix form

$$[A] \{u(t)\} + [B] \left\{ \frac{du(t)}{dt} \right\} + \{f\} = 0 \quad (6.127)$$

where the coefficients of matrices $[A]$ and $[B]$ are defined as

$$\begin{aligned} & \begin{bmatrix} a_{11} & a_{12} & 0 & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} & 0 & 0 & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & a_{N-1,N-2} & a_{N-1,N-1} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \cdot \\ \cdot \\ \cdot \\ u_{N-1} \end{bmatrix} + \\ & \begin{bmatrix} b_{11} & b_{12} & 0 & 0 & 0 & 0 & 0 \\ b_{21} & b_{22} & b_{23} & 0 & 0 & 0 & 0 \\ 0 & b_{32} & b_{33} & b_{34} & 0 & 0 & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & b_{N-1,N-2} & b_{N-1,N-1} \end{bmatrix} \begin{bmatrix} \frac{du_1}{dt} \\ \frac{du_2}{dt} \\ \frac{du_3}{dt} \\ \cdot \\ \cdot \\ \cdot \\ \frac{du_{N-1}}{dt} \end{bmatrix} + \\ & \begin{bmatrix} a_{10} \bar{u}_0 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ a_{N-1,N} \bar{u}_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \end{aligned} \quad (6.128)$$

where, as we saw earlier,

$$a_{ij} = \int_0^1 \frac{d\ell_j}{dx} \frac{d\ell_i}{dx} dx \quad (6.129)$$

$$b_{ij} = \int_0^1 \ell_j \ell_i dx \quad (6.130)$$

$$u_i = u_i(t). \quad (6.131)$$

Where there is one or two second type or Neumann boundary conditions we also have

$$f_i = -\frac{d\hat{u}}{dx} \ell_i \Big|_0^1 \quad (6.132)$$

defined as non-zero in either the first or last row of the vector of known values, f_i , depending upon whether the Neumann condition is specified at the first or last node, respectively. Otherwise f_i contains information on Dirichlet boundary conditions (such as in the above example).

So the question arises as to how to efficiently treat the time derivative. We will consider two possibilities.

6.5.1 Strategy One: Forward Difference Approximation

Let us approximate $\frac{du}{dt}$ by a forward difference approximation

$$\left\{ \frac{du(t)}{dt} \Big|_k \right\} = \frac{\{u\}_{k+1} - \{u\}_k}{\Delta t}. \quad (6.133)$$

Then substitute [Eq. \(6.133\)](#) into [Eq. \(6.127\)](#) to obtain from

$$[A] \{u(t)\} + [B] \left\{ \frac{du(t)}{dt} \right\} + \{f\} = 0 \quad (6.134)$$

the expression

$$[A] \{u\}_k + [B] \frac{\{u\}_{k+1} - \{u\}_k}{\Delta t} + \{f\} = 0. \quad (6.135)$$

Note that the spacial approximation involving $[A]\{u\}_k$ is evaluated at the time $t = k\Delta t$.

We now rearrange this equation to give

$$[B] \{u\}_{k+1} = ([B] - \Delta t [A]) \{u\}_k - \Delta t \{f\}. \quad (6.136)$$

Because the coefficients

$$b_{ij} = \int_{x=0}^{x=1} \ell_j \ell_i dx \quad (6.137)$$

create a tridiagonal $[B]$ matrix, [Eq. \(6.136\)](#) constitutes a system of equations. The solution of the problem is written formally as

$$\{u\}_{k+1} = [B]^{-1} ([B] - \Delta t [A] \{u\}_k - \Delta t \{f\}). \quad (6.138)$$

An important observation is that although we are using a forward difference equation, which in the case of finite difference approximations yielded one equation in one unknown, here, as noted, we have a matrix equation. This is due to the fact that the matrix $[B]$ is no longer diagonal. In other words, the three elements in each row of the matrix $[B]$ tie adjacent nodes together. The consequence of this from a practical perspective is that, although the method is explicit, with its stability limitations, it is nevertheless necessary to take an inverse of matrix $[B]$ to get a solution.

6.5.2 Strategy Two: Backward Difference Approximation

In this approach we employ a backward difference in time approximation

$$\left\{ \frac{du}{dt} \right\} \Big|_{(k+1)} = \frac{\{u\}_{k+1} - \{u\}_k}{\Delta t}. \quad (6.139)$$

Then from [Eq. \(6.127\)](#) we have

$$[A] \{u(t)\} + [B] \left\{ \frac{du(t)}{dt} \right\} + \{f\} = 0 \quad (6.140)$$

which becomes

$$[A] \{u\}_{k+1} + [B] \frac{\{u\}_{k+1} - \{u\}_k}{\Delta t} + \{f\} = 0. \quad (6.141)$$

We can rearrange this equation to give

$$(\Delta t [A] + [B]) \{u\}_{k+1} = [B] \{u\}_k - \Delta t \{f\}. \quad (6.142)$$

We can rewrite this expression as

$$[C] \{u\}_{k+1} = [B] \{u\}_k - \Delta t \{f\} \quad (6.143)$$

where

$$[C] = (\Delta t [A] + [B]). \quad (6.144)$$

One can formally obtain as a solution for $\{u\}_{k+1}$, that is

$$\{u\}_{k+1} = [C]^{-1} ([B] \{u\}_k - \Delta t \{f\}). \quad (6.145)$$

As noted earlier, the formulation that generates [Eq. \(6.138\)](#) is explicit. That is the space derivative is at the old time level j and the solution is computed at the new time level $j + 1$. This formulation is conditionally stable, that is there is a restriction on the value of $r \equiv \frac{\Delta t}{\Delta x^2}$, although we do not prove it here. On the other hand, the formulation that

generates [Eq. \(6.145\)](#) is implicit. The space derivative approximation is at the new time level. It is unconditionally stable.

If the $[B]$ matrix were to be diagonal, that is there are non-zero values only along the diagonal, then, as before, the explicit formulation would be very efficient, involving one equation in one unknown for each row of the matrix equation. However, it is not, so a complete matrix solution is required.

On the other hand, the implicit formulation shown in [Eq. \(6.145\)](#), which requires a complete matrix equation solution, is unconditionally stable and therefore one can move forward in time with larger time steps. In essence, the advantage of the explicit approximation in the finite difference formulation (where we have one equation in one unknown for each node), is normally **not transferable to the finite element case**. There is a procedure called mass lumping which we discuss later in Section 10 on page 273 that allows for a more efficient solution to the explicit formulation equations.

6.6 Chapter Summary

In this chapter we introduce the extension of spatial approximations to include time dependence. Using polynomial approximation theory we create finite difference formulae that address transient one space- dimensional partial-differential equations. This is followed by a discussion of stability including illustrative examples of how an instability can be exhibited. The Galerkin finite element method is then reformulated to consider time,dependent problems in which both forward and backward in time finite difference methods are presented.

6.7 Problems

1. Consider the segment of the grid shown in gray in [Figure 6.3](#). This area can be expressed as

$$[a, b] = [x_i, x_{i+1}] \quad (6.146)$$

$$[c, d] = [t_j, t_{j+1}]. \quad (6.147)$$

Obtain a finite,difference approximation to $\frac{\partial}{\partial x} f(x, t_j)$ over this rectangular area.

Begin with [Eq. \(6.10\)](#)

$$\frac{\partial}{\partial t} f(x, t) = \sum_{i=0}^n \sum_{j=0}^m \frac{\partial \ell_i^n(x)}{\partial t} \ell_j^m(t) f(x_i, t_j) + \frac{\partial}{\partial t} E_{nm} \quad (6.148)$$

and customize it for this use. Then use the definitions of the Lagrange polynomials to obtain the finite-difference form.

2. Consider the type equation

$$\frac{\partial^2 u(x, t)}{\partial x^2} - \frac{\partial u(x, t)}{\partial t} = 0 \quad (6.149)$$

and impose the following initial and boundary conditions

$$u(x, 0) = 0 \quad (6.150)$$

$$u(0, t) = 1 \quad (6.151)$$

$$u(2, t) = 0. \quad (6.152)$$

Now approximate [Eq. \(6.28\)](#) at location $i, j + 1$ using a backward difference in time. Let $i = 0, 1, 2$, and $j = 0, 1, 2$, where $x = i\Delta x$ and $t = j\Delta t$. Also assume $\Delta x = \Delta t = 1$, such that $\rho \equiv \frac{\Delta t}{\Delta x^2} = 1$ The finite difference expression for [Eq. \(6.28\)](#) using the node x_i, t_{j+1} as the point of reference is

$$\frac{u_{i,j+1} - u_{i,j}}{\Delta t} - \frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{\Delta x^2} = 0. \quad (6.153)$$

The computational molecule for this expression is found in [Figure 6.13](#). Calculate the values of u at time levels Δt and $2\Delta t$. Here you have an ambiguity at $u_{i-1,j}$. Assume the initial condition as the boundary condition at this node to resolve this. Should this scheme be stable? How do you know?

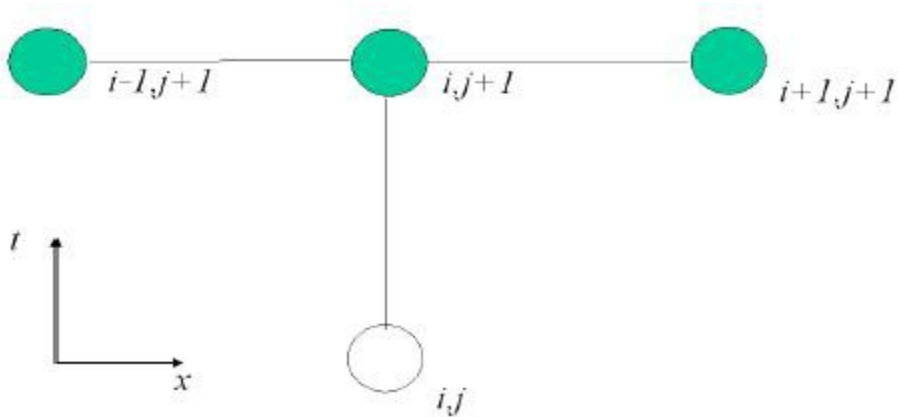


Figure 6.13: Computational molecule for [Eq. \(6.153\)](#).

3. Consider the type equation

$$\frac{\partial^2 u(x, t)}{\partial x^2} - \frac{\partial u(x, t)}{\partial t} = 0 \quad (6.154)$$

- obtain a finite difference form of [Eq. \(6.154\)](#) at location $i, j + 1/2$, that is, evaluate $\frac{\partial^2}{\partial x^2} u(x_i, t_{j+1/2})$
- obtain the finite difference form of $\frac{\partial}{\partial t} u(x_i, t_{j+1/2})$.

Write the finite difference equation at the point $(i, j + 1/2)$ (see [Fig. 6.14](#)).

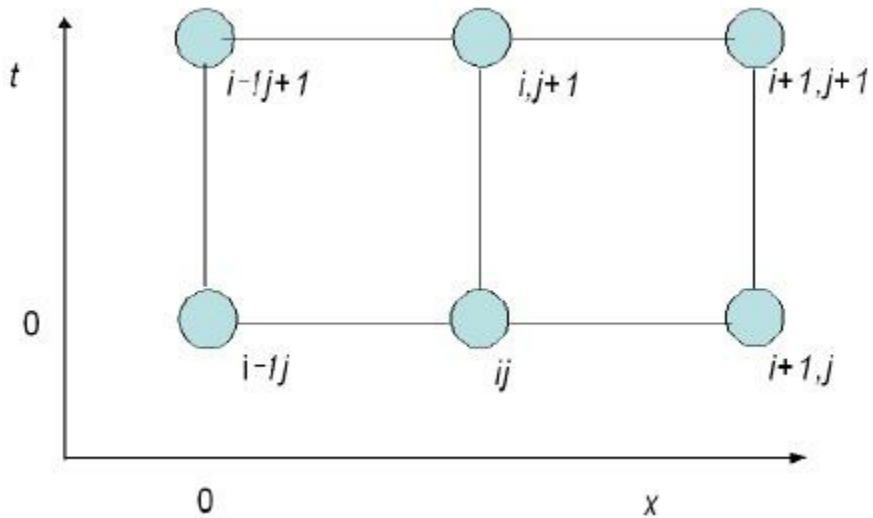


Figure 6.14: Computational molecule for problem 6.14.

4. Consider the equation

$$\frac{\partial u(x, t)}{\partial t} - \frac{\partial^2}{\partial x^2} u(x, t) = 0 \quad (6.155)$$

with boundary and initial conditions

$$u(x, 0) = u_0(x) = 0 \quad (6.156)$$

$$u(0, t) = u_1(t) = 0 \quad (6.157)$$

$$u(L, t) = 0 \quad (6.158)$$

If you employ an implicit backward difference approximation in time and a centered finite difference approximation in space, you will obtain

$$\frac{1}{\Delta t} [u(x_i, t_{j+1}) - u(x_i, t_j)] = \frac{1}{\Delta x^2} [u(x_{i+1}, t_{j+1}) - 2u(x_i, t_{j+1}) - u(x_{i-1}, t_{j+1})] + O(\Delta t, \Delta x^2). \quad (6.159)$$

To simplify notation let

$$r \equiv \frac{\Delta t}{\Delta x^2}. \quad (6.160)$$

The task is to show that the above formulation is unconditionally stable. To achieve this note also that given a matrix $[A]$, eigenvalues λ_s and eigenvectors $\{v\}_s$ the following relationships hold:

$$[A] \{v\}_s = \lambda_s \{v\}_s \quad (6.161)$$

$$[A]^{-1} \{v\}_s = \frac{1}{\lambda_s} \{v\}_s \quad (6.162)$$

and that

$$\lambda_s = 1 + 4 r \sin^2 \left(\frac{s\pi}{2N'} \right). \quad (6.163)$$

5. If you were to approximate the equation

$$\frac{\partial u}{\partial t} + v \frac{\partial u}{\partial x} = 0 \quad (6.164)$$

using a backward in time, backward in space approximation, you require one boundary defined at $x = 0$ and one initial condition. What would happen if you were to apply a second boundary condition at the terminal end of the interval? What would happen if you used a central difference approximation for the space derivative and a boundary condition at the terminal end of the interval. To make these determinations, generate a nodal array with one element in the vertical (time) dimension and three in the horizontal (space) dimension and the computational molecule for both formulations and use this information in your explanation.

6. Approximate the equation

$$\frac{\partial u}{\partial t} + a \frac{\partial^2 u}{\partial x^2} = 0 \quad (6.165)$$

using three nodes and a centered finite difference approximation in space and a backward finite difference equation in time for the central spatial node. Now do the same problem using a linear finite element in space and backward difference in time formulation for the central spatial node. What is the fundamental difference in the algebraic equations and describe what you think this means in terms of how the time derivative is handled.

7. Write the finite difference approximation at the location $x = i\Delta x$ and $t = \Delta t/2$. for the equation

$$\frac{\partial u}{\partial t} + a \frac{\partial^2 u}{\partial x^2} = 0. \quad (6.166)$$

Hint: you will need the space approximation at both the $t = 0$ and $t = \Delta t$ levels.

Bibliography

Allen, M.B III and E.L. Isaacson (1998) *Numerical Analysis for Applied Science*, John Wiley & Sons, Hoboken.

Anderson, J.D. Jr. (1994) *Computational Fluid Dynamics: The Basics with Applications*, McGraw Hill, New York.

Botha, J.F. and G.F. Pinder (1983) *Fundamental Concepts in the Numerical Solution of Differential Equations*, John Wiley and Sons, New York.

Isaacson, E. and H.B. Keller (1966) *Analysis of Numerical Methods*, John Wiley and Sons, New York.

Richtmeyer, R.D. and K.W. Morton (1967) *Difference Methods for Initial Value Problems*, Wiley Interscience, New York.

Smith, G.D. (1985) *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, 3rd ed. Oxford University Press, Oxford.

von Rosenberg, D.U. (1969) *Methods for the Numerical Solution of Partial Differential Equations*, American Elsevier, New York.

Chapter 7

Finite Difference Methods in Two Space

In this chapter we will consider the finite difference approximation to a steady-state (time-independent) problem a two-space dimensional system. The “type” equation we will initially consider is the Poisson’s equation

$$\frac{\partial^2 u(x, y)}{\partial x^2} + \frac{\partial^2 u(x, y)}{\partial y^2} = Q(x, y). \tag{7.1}$$

Since this equation is second order in both x and y , we need two boundary conditions for each coordinate direction. Thus we have, for example,

$$u(x, y) = u_0(x, y) \quad x, y \in \partial\Omega_x \tag{7.2}$$

$$\frac{\partial}{\partial x} u(x, y) = \frac{\partial}{\partial x} u_1(x, y) \quad x, y \in \partial\Omega_y \tag{7.3}$$

$$\partial\Omega_x + \partial\Omega_y = \partial\Omega \tag{7.4}$$

where $u_0(x, y)$ and $\frac{\partial}{\partial x} u_1(x, y)$ are known functions and the notation $\partial\Omega_x$ and $\partial\Omega_y$ denote the portion of the perimeter as indicated in [Fig. 7.1](#).

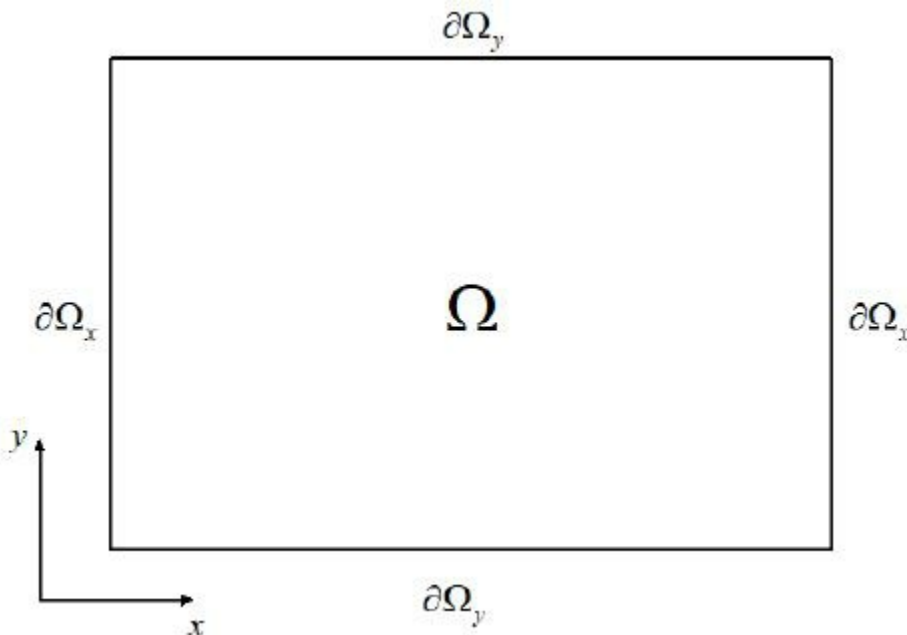


Figure 7.1 Domain of interest regarding the boundary conditions specified in [Eq. \(7.2\)](#)

through 7.4.

For the case of rectangular elements and a finite difference approximation, the two-space dimensional case is analogous to the two-dimensional space-time case see Eq. (6.10). One simply replaces the time coordinate with a second space coordinate. The approximating function over three nodes becomes

$$u(x, y) = \sum_{i=-1}^1 \sum_{j=-1}^1 u_{i,j} \ell_i^n(x) \ell_j^m(y) + E_{nm} \quad (7.5)$$

which is illustrated in Fig. 7.2 for the case when the superscripts n and m (the degrees of the Lagrange polynomials in x and y , respectively) are both 1. Figure 7.3 presents the quadratic basis function defined over one element. Notice that we have hijacked the subscript i . Earlier it was identified in general with a node, but in the case of weighted residuals, also the weighting function. Now it represents the x coordinate for the node i only, that is $x = i\Delta x$ but there is no reference to its use as an index for a weighting function. We will only use i in this way in this section and redefine it in the following section. Note that along both the x and y directions the function is linear, but in the diagonal directions it is a quadratic because of the product $\ell_i^1(x)\ell_j^1(y)$. We can use these functions to represent the first derivative, but for the second derivative we need to use at least a quadratic-second degree Lagrange polynomial. If we differentiate Eq. (7.5) twice when $n = m = 2$ and evaluate the result at point $x_i = i\Delta x$ and $y_j = j\Delta y$ we obtain

$$\left. \frac{\partial^2 u}{\partial x^2} \right|_{(x_i, y_j)} = \frac{1}{\Delta x^2} [u_{i-1,j} - 2u_{i,j} + u_{i+1,j}] + \frac{\partial^2}{\partial x^2} E_{i,j} \quad (7.6)$$

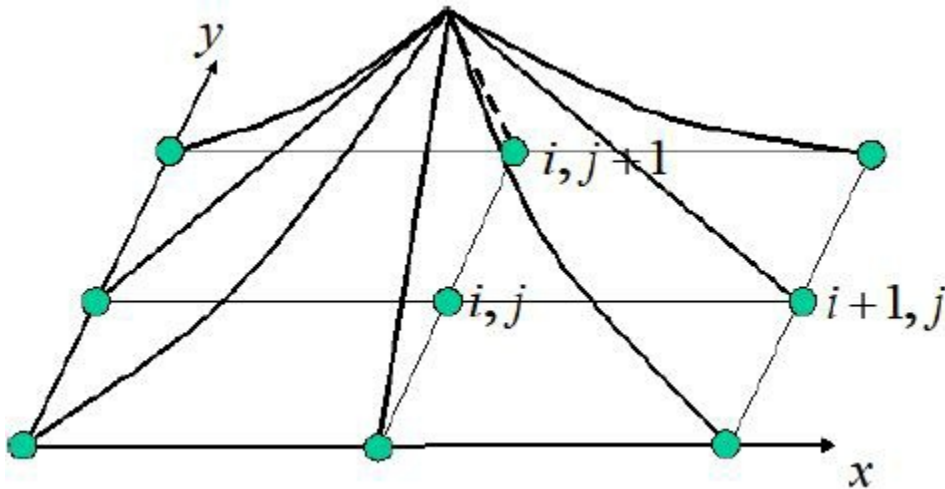


Figure 7.2 Diagrammatic representation of the function $\ell_i^1(x)\ell_j^1(y)$.

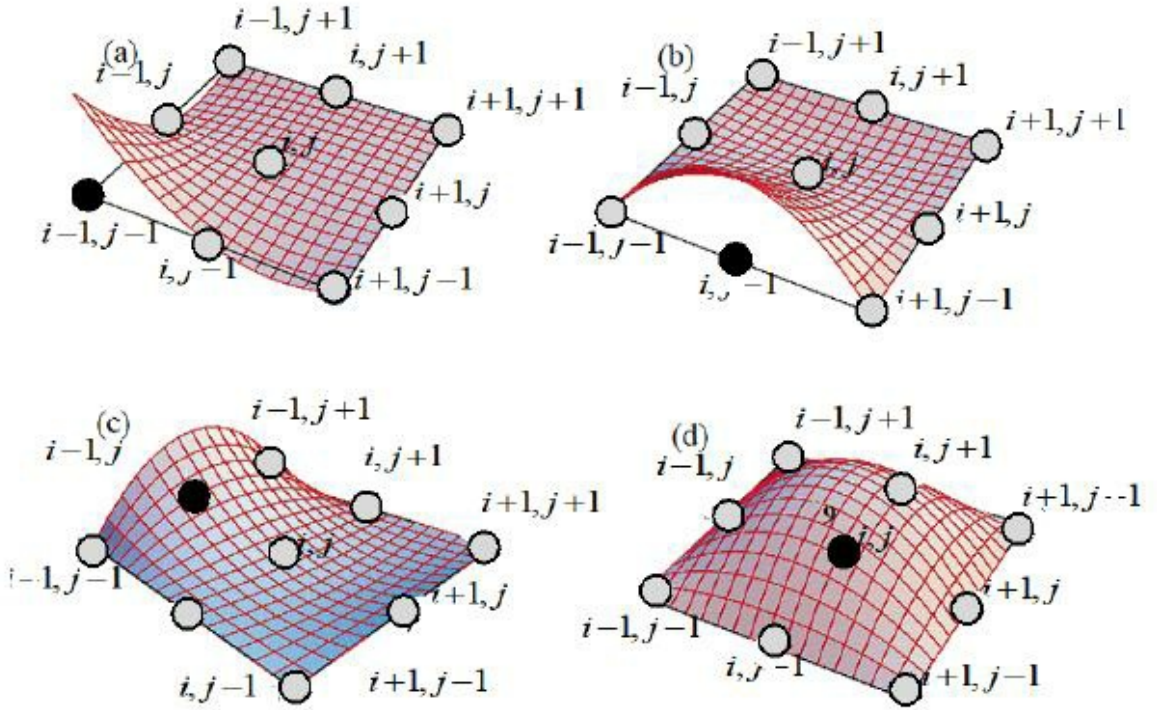


Figure 7.3 Quadratic polynomials in two dimensions. The circles are the nodes and the black circles indicate the location for which the polynomial is defined, that is, the location at which it equals unity (modified from [Felippa, 2015]).

and

$$\left. \frac{\partial^2 u}{\partial y^2} \right|_{(x_i, y_j)} = \frac{1}{\Delta y^2} [u_{i, j-1} - 2u_{i, j} + u_{i, j+1}] + \frac{\partial^2}{\partial y^2} E_{i, j}. \quad (7.7)$$

Substitution of [Eqs. \(7.6\)](#) and [\(7.7\)](#) into [Eq. \(7.1\)](#) yields our two-dimensional finite-difference approximation of [Eq. \(7.1\)](#) which is given by

(7.8)

$$\begin{aligned} & \frac{1}{\Delta x^2} [u_{i-1, j} - 2u_{i, j} + u_{i+1, j}] + \frac{1}{\Delta y^2} [u_{i, j-1} - 2u_{i, j} + u_{i, j+1}] \\ & + \frac{\partial^2}{\partial y^2} E_{i, j} + \frac{\partial^2}{\partial x^2} E_{i, j} = Q(x_i, y_j). \end{aligned}$$

If we neglect the error $E_{i, j}$ this equation becomes

(7.9)

$$\frac{1}{\Delta x^2} [u_{i-1, j} - 2u_{i, j} + u_{i+1, j}] + \frac{1}{\Delta y^2} [u_{i, j-1} - 2u_{i, j} + u_{i, j+1}] = Q(x_i, y_j)$$

and the computational molecule for this equation is given in [Fig. 7.4](#). In this figure, the weights for each of the nodal values of $u_{i,u}$ etc. are shown assuming $\Delta x = \Delta y = 1$

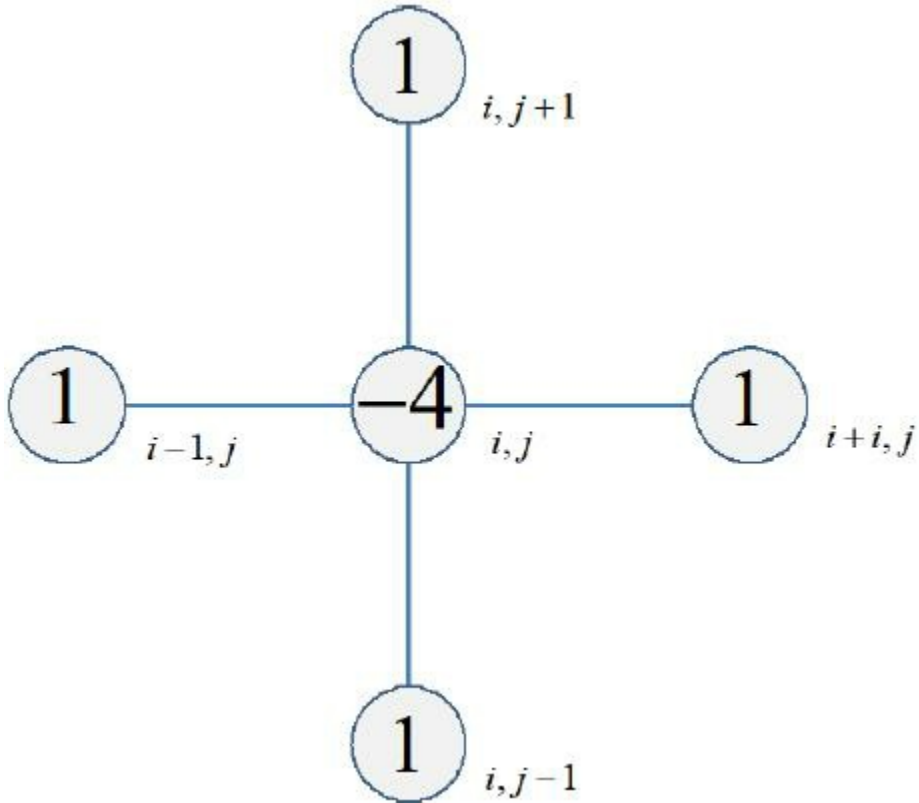


Figure 7.4 Computational molecule for [Eq. 7.9](#) where $\Delta x = \Delta y = 1$.

We will now consider an example to illustrate how this equation is used when a cross-derivative is involved. Consider the generalization of [Eq. \(7.1\)](#). This type of equation arises when the coefficients in the equation are tensors, that is

$$a_{xx} \frac{\partial^2 u(x, y)}{\partial x^2} + a_{xy} \frac{\partial}{\partial x} \left(\frac{\partial u(x, y)}{\partial y} \right) + a_{yx} \frac{\partial}{\partial y} \left(\frac{\partial u(x, y)}{\partial x} \right) + a_{yy} \frac{\partial^2 u(x, y)}{\partial y^2} = 0. \quad (7.10)$$

where coefficients are known constants. Now the finite difference approximation is a little more complicated. We begin by writing the second degree Lagrange approximation for $\hat{u}(x, y)$ viz.

$$\hat{u}(x, y) = \sum_{i=0}^n \sum_{j=0}^m u(x_i, y_j) \ell_i^2(x) \ell_j^2(y). \quad (7.11)$$

Now we differentiate this expression with respect to x and then y to get

$$\frac{\partial \hat{u}(x, y)}{\partial x} = \sum_{i=0}^n \sum_{j=0}^m u(x_i, y_j) \frac{\partial \ell_i^2(x)}{\partial x} \ell_j^2(y) \quad (7.12)$$

$$\frac{\partial \hat{u}(x, y)}{\partial y} = \sum_{i=0}^n \sum_{j=0}^m u(x_i, y_j) \ell_i^2(x) \frac{\partial \ell_j^2(y)}{\partial y}. \quad (7.13)$$

Next differentiate [Eq. \(7.12\)](#) with respect to x and [Eq. \(7.13\)](#) with respect to y . Performing these actions we get

$$\frac{\partial^2 \hat{u}(x, y)}{\partial x^2} = \sum_{i=0}^n \sum_{j=0}^m u(x_i, y_j) \frac{\partial^2 \ell_i^2(x)}{\partial x^2} \ell_j^2(y) \quad (7.14)$$

$$\frac{\partial^2 \hat{u}(x, y)}{\partial y^2} = \sum_{i=0}^n \sum_{j=0}^m u(x_i, y_j) \ell_i^2(x) \frac{\partial^2 \ell_j^2(y)}{\partial y^2}. \quad (7.15)$$

We can now reverse the procedure and differential [Eq. \(7.12\)](#) with respect to y and [Eq. \(7.13\)](#) with respect to x . When this is done one obtains

$$\frac{\partial^2 \hat{u}(x, y)}{\partial y \partial x} = \sum_{i=0}^n \sum_{j=0}^m u(x_i, t_j) \frac{\partial \ell_i^2(x)}{\partial x} \frac{\partial \ell_j^2(y)}{\partial y} \quad (7.16)$$

$$\frac{\partial^2 \hat{u}(x, y)}{\partial x \partial y} = \sum_{i=0}^n \sum_{j=0}^m u(x_i, t_j) \frac{\partial \ell_i^2(x)}{\partial x} \frac{\partial \ell_j^2(y)}{\partial y}. \quad (7.17)$$

We have evaluated these derivatives above in [Eqs. \(7.6\)](#) and [\(7.7\)](#) and we found

$$\frac{d\hat{u}^2(x_i, y_j)}{dx^2} = \frac{1}{\Delta x^2} [u_{i+1, j} - 2u_{i, j} + u_{i-1, j}] \quad (7.18)$$

$$\frac{d\hat{u}^2(x_i, y_j)}{dy^2} = \frac{1}{\Delta y^2} [u_{i, j+1} - 2u_{i, j} + u_{i, j-1}]. \quad (7.19)$$

The cross derivatives are a little tricky. One can show by expanding the derivatives found in [Eqs. \(7.16\)](#) and [\(7.17\)](#) that the complete expansion will consist of nine terms, one for

each of the nine nodes found in the Fig. [?]. However, all but four of them vanish because of the nature of the functions. They are zero on at least two sides of the square and, therefore their derivatives are zero along the tangent to those sides. What remains are the terms

$$\frac{\partial^2 \hat{u}(x, y)}{\partial x \partial y} = \tag{7.20}$$

$$u(x_{i-1}, y_{j-1}) \frac{\partial \ell_{i-1}^2(x)}{\partial x} \frac{\partial \ell_{j-1}^2(y)}{\partial y} + u(x_{i-1}, y_{j+1}) \frac{\partial \ell_{i-1}^2(x)}{\partial x} \frac{\partial \ell_{j+1}^2(y)}{\partial y} +$$

$$u(x_{i+1}, y_{j+1}) \frac{\partial \ell_{i+1}^2(x)}{\partial x} \frac{\partial \ell_{j+1}^2(y)}{\partial y} + u(x_{i+1}, y_{j-1}) \frac{\partial \ell_{i+1}^2(x)}{\partial x} \frac{\partial \ell_{j-1}^2(y)}{\partial y}$$

The various derivatives found in this equation evaluated at (x_i, y_j) are as follows:

$$\begin{aligned} \frac{\partial}{\partial x} \ell_{i-1}^2(x) \Big|_{x_i} &= \left[\frac{\partial}{\partial x} \left(\frac{(x - x_i)(x - x_{i+1})}{(x_{i-1} - x_i)(x_{i-1} - x_{i+1})} \right) \right] \Big|_{x_i} \tag{7.21} \\ &= \left[\frac{(x - x_i) + (x - x_{i+1})}{(x_{i-1} - x_i)(x_{i-1} - x_{i+1})} \right] \Big|_{x_i} \\ &= \frac{x_i - x_{i+1}}{(x_{i-1} - x_i)(x_{i-1} - x_{i+1})} \\ &= \frac{-\Delta x}{2(\Delta x^2)} \\ &= \frac{-1}{2\Delta x} \end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial x} \ell_{i+1}^2(x) \Big|_{x_i} &= \left[\frac{\partial}{\partial x} \left(\frac{(x - x_{i-1})(x - x_i)}{(x_{i+1} - x_{i-1})(x_{i+1} - x_i)} \right) \right]_{x_i} & (7.22) \\
&= \left[\frac{(x - x_{i-1}) + (x - x_i)}{(x_{i+1} - x_{i-1})(x_{i+1} - x_i)} \right]_{x_i} \\
&= \frac{x_i - x_{i-1}}{(x_{i+1} - x_{i-1})(x_{i+1} - x_i)} \\
&= \frac{\Delta x}{2(\Delta x^2)} \\
&= \frac{1}{2\Delta x}.
\end{aligned}$$

Similarly, for the y direction we get the following using a similar reasoning

$$\frac{\partial}{\partial y} \ell_{j-1}^2(y) \Big|_{y_j} = \frac{-1}{2\Delta y} \quad (7.23)$$

$$\frac{\partial}{\partial y} \ell_{j+1}^2(y) \Big|_{y_j} = \frac{1}{2\Delta y}. \quad (7.24)$$

Substituting the above derivative approximations into [Eq. \(7.20\)](#) we get

$$\begin{aligned}
\frac{\partial^2 \hat{u}(x, y)}{\partial x \partial y} \Big|_{x_i, y_j} &= & (7.25) \\
&u(x_{i-1}, y_{j-1}) \left(\frac{-1}{2\Delta x} \right) \left(\frac{-1}{2\Delta y} \right) + u(x_{i-1}, y_{j+1}) \left(\frac{-1}{2\Delta x} \right) \left(\frac{1}{2\Delta y} \right) \\
&+ u(x_{i+1}, y_{j+1}) \left(\frac{1}{2\Delta x} \right) \left(\frac{1}{2\Delta y} \right) + u(x_{i+1}, y_{j-1}) \left(\frac{1}{2\Delta x} \right) \left(\frac{-1}{2\Delta y} \right)
\end{aligned}$$

$$\frac{\partial^2 \hat{u}(x, y)}{\partial x \partial y} \Big|_{x_i, y_j} = \frac{1}{4\Delta x \Delta y} \left(u(x_{i-1}, y_{j-1}) - u(x_{i-1}, y_{j+1}) + u(x_{i+1}, y_{j+1}) - u(x_{i+1}, y_{j-1}) \right) \quad (7.26)$$

To get the final form of the approximation we combine [Eqs. \(7.19\)](#), [\(7.18\)](#) and [\(7.26\)](#) to get

(7.27)

(7.28)

$$\begin{aligned} & \frac{a_{xx}}{\Delta x^2} [u_{i+1,j} - 2u_{i,j} + u_{i-1,j}] + \frac{a_{yy}}{\Delta y^2} [u_{i,j+1} - 2u_{i,j} + u_{i,j-1}] + \\ & \frac{a_{xy}}{4\Delta x\Delta y} [u(x_{i-1}, y_{j-1}) - u(x_{i-1}, y_{j+1}) + u(x_{i+1}, y_{j+1}) - u(x_{i+1}, y_{j-1})] + \\ & \frac{a_{yx}}{4\Delta y\Delta x} [u(x_{i+1}, y_{j+1}) - u(x_{i+1}, y_{j-1}) + u(x_{i-1}, y_{j-1}) - u(x_{i-1}, y_{j+1})] \\ & = 0. \end{aligned}$$

If we assume $\Delta x = \Delta y = a_{xx} = a_{xy} = a_{yx} = a_{yy} = 1$ the computational molecule for this cross-derivative bearing equations is given in [Fig. 7.5](#). Notice that now there are nine, not five, nodes in the molecule. The cross derivative added four. From a computational point of view this is not a good thing. The matrix used to solve a problem with this formulation will have four extra bands parallel to the diagonal and this means significantly more effort is needed to solve the resulting set of equations when standard matrix algebra methods are used.

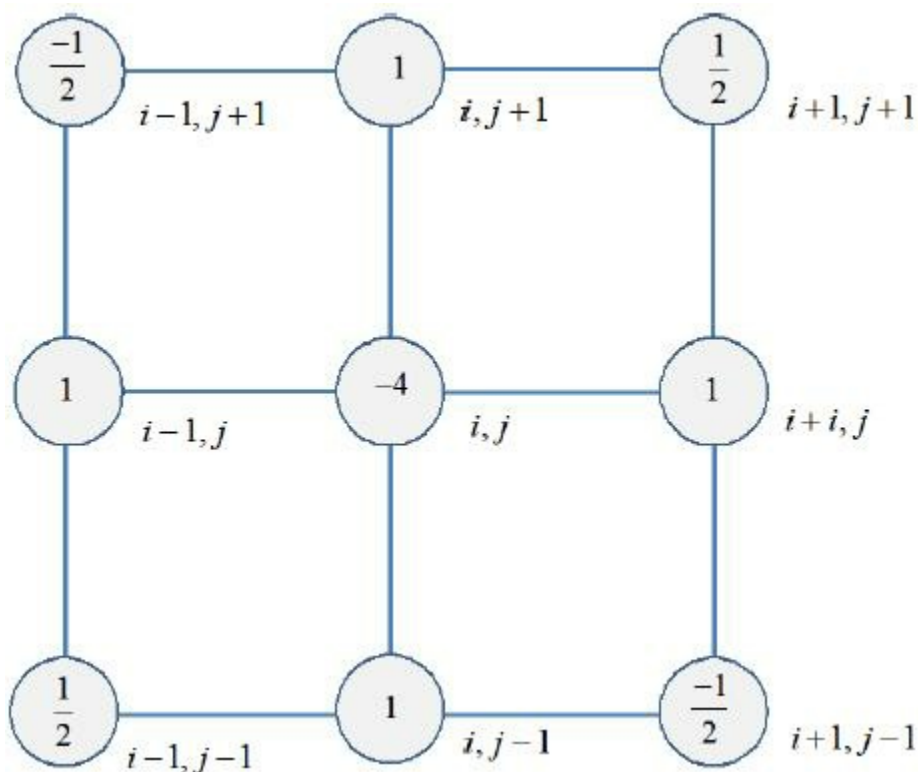


Figure 7.5 Computational molecule for [Eq. 7.28](#) when we assume $\Delta x = \Delta y = a_{xx} = a_{xy} = a_{yx} = a_{yy} = 1$.

7.1 Example Problem

Consider the problem defined by equation

$$\frac{\partial^2}{\partial x^2} u(x, y) + \frac{\partial^2}{\partial y^2} u(x, y) = 1 \quad (7.29)$$

shown in [Fig. 7.6](#).

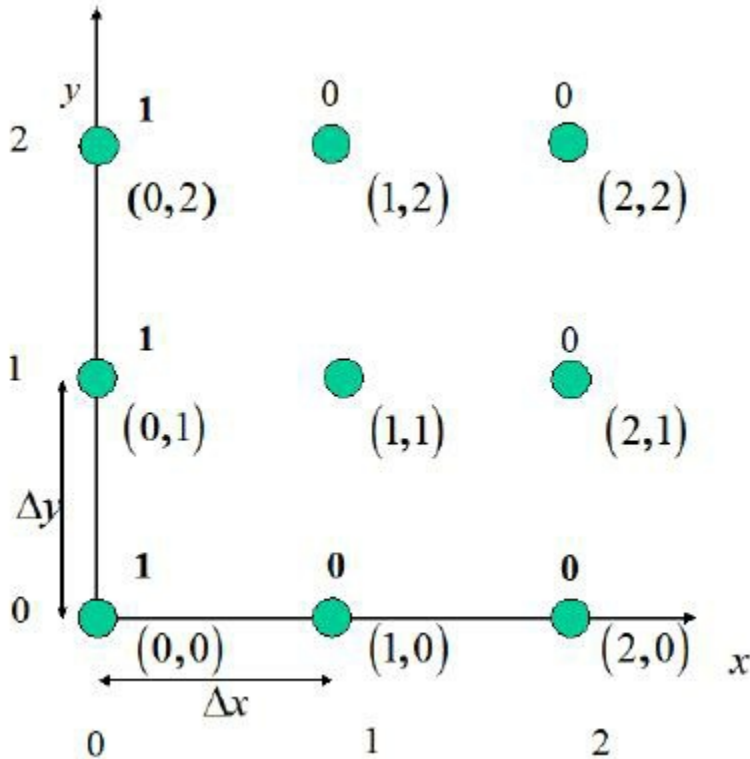


Figure 7.6 Grid for the example considered in [Section 7.6](#).

Assume boundary conditions

$$u(0, y) = 1 \quad (7.30)$$

$$u(2, y) = 0 \quad (7.31)$$

$$u(x, 0) = 0 \quad (7.32)$$

$$u(x, 2) = 0. \quad (7.33)$$

The finite difference expression for this problem is obtained directly from [Eq. \(7.8\)](#) by assigning the appropriate nodal identifiers. Because only the node (1, 1) is unknown in this problem (the rest are known from boundary conditions), the difference equation is

$$\frac{1}{\Delta x^2} [u_{0,1} - 2u_{1,1} + u_{2,1}] + \frac{1}{\Delta y^2} [u_{1,0} - 2u_{1,1} + u_{1,2}] = 1. \quad (7.34)$$

Substitution of values known from boundary conditions yields

$$\frac{1}{12} [1 - 2u_{1,1} + 0] + \frac{1}{12} [0 - 2u_{1,1} + 0] = 1 \quad (7.35)$$

which simplifies to

$$-4u_{1,1} = 0 \quad (7.36)$$

and finally

$$u_{1,1} = 0. \quad (7.37)$$

If we assume there is no flux, that is, $Q = 0$ we get

$$\frac{1}{12} [1 - 2u_{1,1} + 0] + \frac{1}{12} [0 - 2u_{1,1} + 0] = 0 \quad (7.38)$$

which reduces to

$$-4u_{1,1} = -1 \quad (7.39)$$

and

$$u_{1,1} = 1/4. \quad (7.40)$$

7.2 Chapter Summary

In this brief chapter, we introduce the concept of representing a two-dimensional problem using finite difference approximations and present an elementary example.

7.3 Problems

1. Consider the finite difference mesh indicated in Fig. 7.7. The equation of interest is

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad (7.41)$$

with boundary conditions

$$u(0, y) = 1 \quad (7.42)$$

$$\frac{\partial u(4, y)}{\partial x} = 2 \quad (7.43)$$

$$\frac{\partial u(x, 0)}{\partial y} = 0 \quad (7.44)$$

$$\frac{\partial u(x, 4)}{\partial y} = 0 \quad (7.45)$$

Do the following:

- a. Write the finite difference expression for [Eq. \(7.41\)](#). You should have five terms in your approximation.
 - b. Write a finite difference approximation for the Neumann boundary conditions (Eqs. [\(7.43\)](#)-[\(7.45\)](#)).
 - c. Write the finite difference equation for [Eq. \(7.42\)](#). You now have enough equations.
 - d. Substitute the information from steps 1b and 1c into step 1a.
 - e. Solve the resulting equation for the unknown nodes.
2. The approximation of u for the case of the two-dimensional element shown in [Fig 7.7](#) is written as

$$u(x, y) \approx \hat{u}(x, y) = \sum_{j=0}^2 \sum_{i=0}^2 u_{i,j} \ell_i(x) \ell_j(y) \quad (7.46)$$

where the $\ell_i(x)$ and $\ell_j(y)$ are quadratic Lagrange polynomials defined as

$$\ell_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} \quad (7.47)$$

$$\ell_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} \quad (7.48)$$

$$\ell_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} \quad (7.49)$$

with a similar set of definitions for $\ell_j(y)$, $j = 0, 1, 2$.

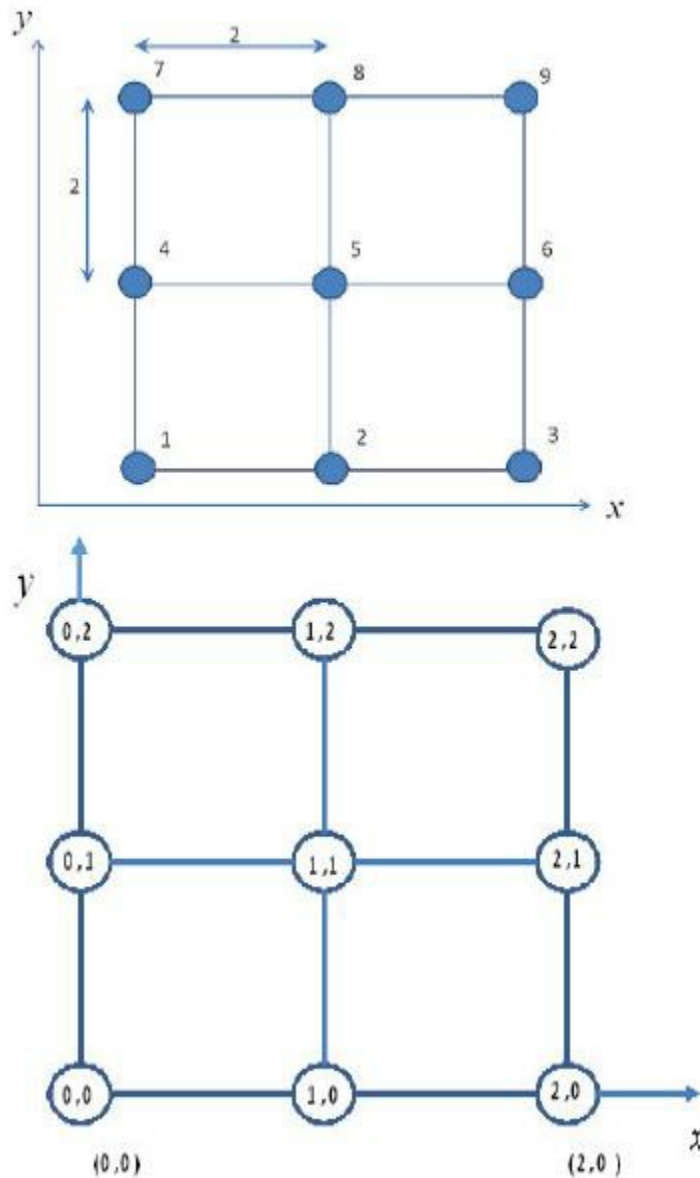


Figure 7.7 Mesh and nodal arrangement for problems 1 (panel a) and 2 (panel b).

3. You are to obtain the finite difference representation of

$$\frac{\partial^2 \hat{u}}{\partial x \partial y}. \quad (7.50)$$

Please proceed in the following steps:

- a. Differentiate [Eq. \(7.46\)](#) first with respect to y and then the result with respect to x . You should have nine terms. of the general form

$$u_{i,j} \frac{\partial l_i(x)}{\partial x} \frac{\partial l_j(y)}{\partial y}. \quad (7.51)$$

- b. Place these into [Eq. \(7.46\)](#) after differentiation and evaluate the nine terms using [Eq. \(7.47\)](#)
- c. Replace the various spatial increments with the values from the above figure, that is $x_1 - x_0 = 1$ etc.; be careful with your signs.
- d. Now evaluate the result at the center node $u(1, 1)$, since that is where you are writing the approximation
- e. The result should be

$$(u_{2,2} - u_{0,2} - u_{2,0} + u_{0,0}) / 4 \quad (7.52)$$

$$(u_{2,2} - u_{0,2} - u_{2,0} + u_{0,0}) / 4. \quad (7.53)$$

Bibliography

Allen, M.B. III and E.L. Isaacson (1988) *Numerical Analysis for Applied Science*, John Wiley and Sons, New York.

Celia, M.A. and W.G. Gray (1992) *Numerical Methods for Differential Equations: Fundamental Concepts for Scientific and Engineering Applications*, Prentice-Hall, Englewood Cliffs.

Collatz, L. (1966) *The Numerical Treatment of Differential Equations*, Springer Verlag, New York.

Felippa, C.A. (2015) *Introduction to Finite Element Methods (ASEN 5007)*, Chapter 18, Shape Function Magic, University of Colorado.

Lapidus, L. and G.F. Pinder (1982) *Numerical Solution of Partial Differential Equations in Science and Engineering*, Wiley-Interscience, John Wiley and Sons.

Li, R., Z. Chen, and W. Wu (2000) *Generalized Difference Methods for Differential Equations*, Pure and Applied Mathematics, Marcel Dekker, New York, 2000.

Morton, K.W. and D.F. Mayers (2005) *Numerical Solution of Partial Differential Equations, An Introduction*, Cambridge University Press.

Saulev, V.K. (1957) "On a Class of Elliptic Equations Solvable by the Method of Finite Differences," *Vycisl. Mat.*, 1, 81–86.

Smith, G.D. (1985) *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, 3rd ed., Oxford University Press.

Chapter 8

Finite Element Methods in Two Space

8.1 Finite Element Approximations over Rectangles

In this section we extend the one-dimensional finite element method of approximation to two space dimensions. We begin with rectangular elements because our earlier theory lends itself most readily to this approach. We begin by once again considering the Poisson's equation on a rectangular domain

$$\frac{\partial^2 u(x, y)}{\partial x^2} + \frac{\partial^2 u(x, y)}{\partial y^2} = Q(x, y) \tag{8.1}$$

$$u(x, y) = u_0(x, y) \quad x, y \in \partial\Omega_x \tag{8.2}$$

$$u(x, y) = u_1(x, y) \quad x, y \in \partial\Omega_y \tag{8.3}$$

$$\partial\Omega_x + \partial\Omega_y = \partial\Omega. \tag{8.4}$$

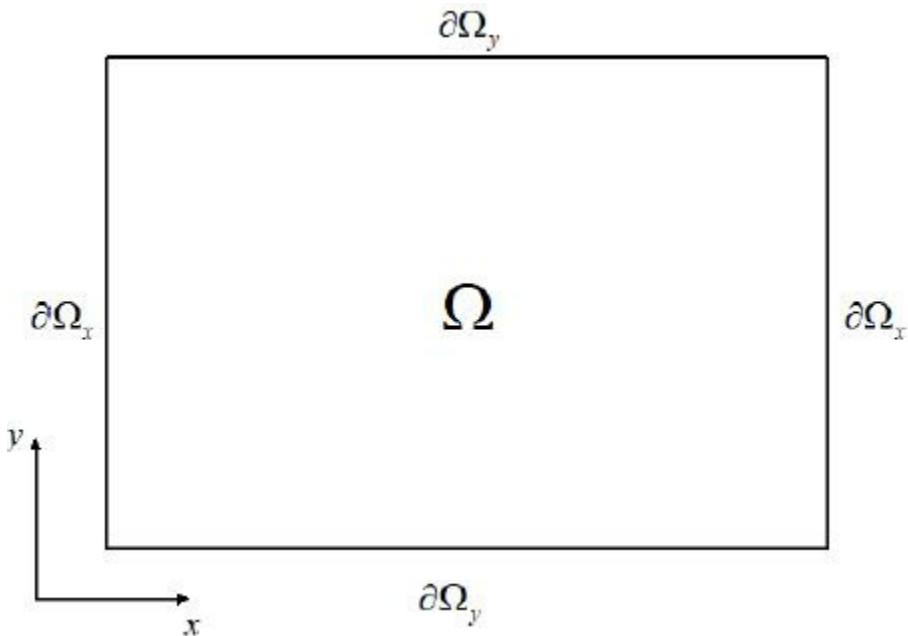


Figure 8.1: Definition sketch of the region of interest for [Eq. 8.1](#).

We now digress briefly to introduce some notation. As we have done earlier, let us begin with the approximation of $u(x, y)$ by $\hat{u}(x, y)$. We will use the same point of departure as we used for the finite difference approximations in Chapter 7, that is

$$\hat{u}(x, y) = \sum_{i=0}^n \sum_{j=0}^m u_{i,j} \ell_i^n(x) \ell_j^m(y)$$

where we assume we are working with Lagrange polynomials of degree n and m . But now we make a change in notation which, while not critical for our current discussion of rectangular elements, will be essential when we talk about triangular elements in the next section. The strategy is to replace the i, j notation for a nodal location in two-dimensional problems with a single index.

Earlier, in the case of problems defined in one space dimension, there was no ambiguity; the i, j index played the role of both nodal identifier and nodal location through the relationship $x = i\Delta x$ and $y = j\Delta y$. However this is all about to change. To make a long story short, we will replace the double i, j subscript notation *for a node location* (that is based on the idea of i associated with a row and j associated with a column) with a new p_α index where α in this instance could take on the value i or j ... The i, j notation was convenient because *it identified the node by the value i, j and also identified its location $x = i\Delta x$ and $y = j\Delta y$* (assuming Δx and Δy are constant everywhere).

Note that p_α does not, in and of itself, tell you where the node p_α is located. We need additional information. For example, we need to know that node number p_α is located at $x_{p_\alpha}, y_{p_\alpha}$. In other words we need a table that relates our node number p_α to the its location in terms of x_α and y_α . The index p_α is simply a nodal number and, in general, $x \neq p_\alpha\Delta x$ and $y \neq p_\alpha\Delta y$.

To solidify this idea, consider the node located at p_5 in [Fig. 8.2](#). Unless we know that p_5 is located at location $(x_5, y_5) = (1, 1)$ the knowledge that $p_\alpha = p_5$ is meaningless. Whereas, using i, j notation we would know both that the node is identified by the unique values of $i = 1$ and $j = 1$, that is if we have $\ell_1^n(x)\ell_1^m(y)$ we know it is located at $(1, 1)$. Somehow using the single index does not seem like progress, but let us go a little further.

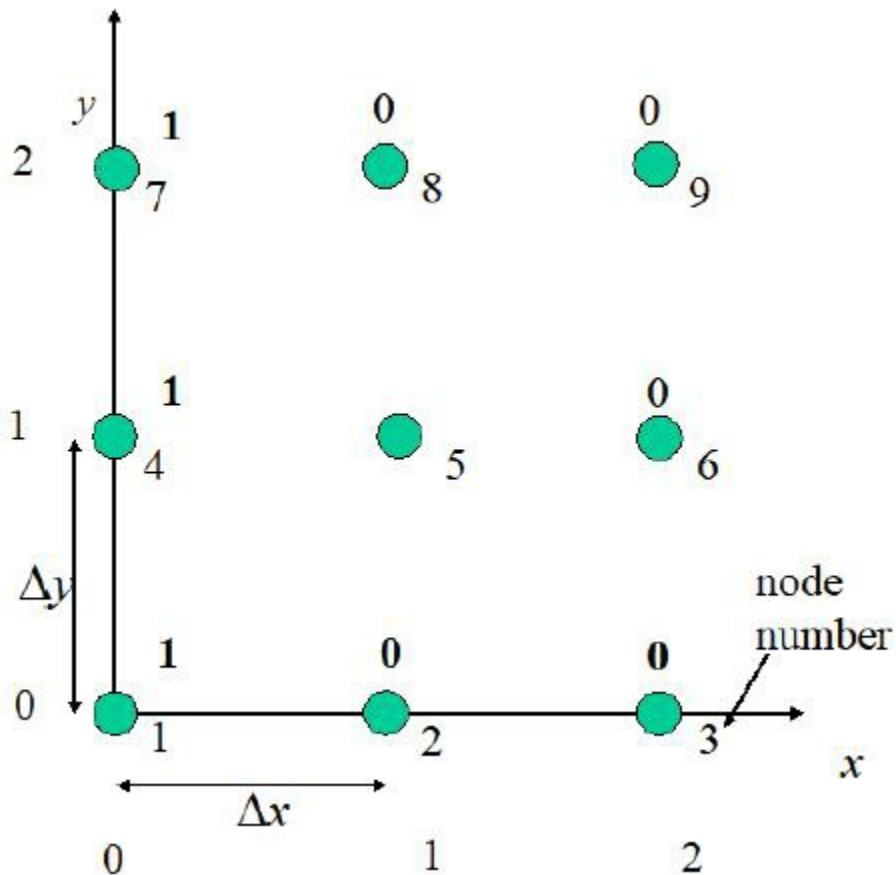


Figure 8.2: Single digit numbering for the finite element method of solution.

Let us take a look at how we would describe a Lagrangian polynomial using this notation. First of all, the Lagrangian polynomial now needs only one subscript to be identified uniquely. Because any two-dimensional Lagrangian is the product of a one dimensional Lagrangian in x and another in y , we can write

$$\ell_i(x, y) = \ell_i(x) \ell_i(y). \quad (8.6)$$

This holds as long as the index i refers to a node, irrespective of where that node is located.

To make the value of [Eq. \(8.6\)](#) unique, we need to know where the node i is located. To illustrate our point, assume function $\ell_i(x, y)$ is identified with a node located at $p_5 = (x_5, y_5)$. Then we know from our earlier discussion that the Lagrange polynomial is identified with a node located at $(x_5, y_5) = (1, 1)$. If we know this information and the mesh spacing, we have a unique representation of $\ell_5(x, y)$ provided we have similar information about all the nodes in the neighborhood (because all connecting node locations are part of the definition of $\ell_5(x, y)$). For the case in point, those are all the nodes appearing in [Fig. 8.2](#).

In summary, if you have a Lagrange polynomial (or its derivatives) with subscript i , you need to know where node i is located. A table will tell you that node i is located at location p_i and that location p_i is x_{p_i}, y_{p_i} . You will see that you will need to have this

location information when you need to evaluate $\ell_i(x, y)$ when, for example, you need to calculate an integral to obtain a coefficient.

It is helpful at this point to compare [Fig. 8.2](#) with Fig. 7.6 to see the correspondence between the two notations. Employing the double index notation, the approximating function $\hat{u}(x, y)$ shown in Fig. 7.6 is

$$\hat{u}(x, y) = \sum_{i=1}^{i=3} \sum_{j=1}^{j=3} u_{ij} \ell_i(x) \ell_j(y) \quad (8.7)$$

and for the grid shown in [Fig. 8.2](#), which uses only one index we have

$$\hat{u}(x, y) = \sum_{k=1}^{k=9} u_k \ell_k(x, y). \quad (8.8)$$

We have used the index k in [Eq. \(8.8\)](#) to avoid confusion with the use of the index i in [Eq. \(8.7\)](#). These two expressions, [Eq. \(8.7\)](#) and [\(8.8\)](#), represent the same information in the special case of rectangular elements, provided the relationship between the index k and the location p_k is provided. So, in the future, when you see an index i or j you must identify each with a unique location p_i or p_j in order to proceed. Note that both indices i and j can take on the same p_α value since either i or j (or both) may be identified with location $(x_{p_\alpha}, y_{p_\alpha})$.

With the notation issue behind us, let us proceed by stating the residual, which is obtained by replacing $u(x, y)$ with $\hat{u}(x, y)$ in [Eq. \(8.1\)](#), that is,

$$\frac{\partial^2}{\partial x^2} \hat{u}(x, y) + \frac{\partial^2}{\partial y^2} \hat{u}(x, y) - Q(x, y) = R(x, y). \quad (8.9)$$

The appropriate weighted residual formulation is

$$\int_{\Omega} R(x, y) \ell_i(x, y) d\Omega = 0 \quad i = 1, 2, \dots, N \quad (8.10)$$

where N is the number of nodes. Now, substitute for $R(x, y)$ from [Eq. \(8.9\)](#).

$$\int_{\Omega} \left[\frac{\partial^2}{\partial x^2} \hat{u}(x, y) + \frac{\partial^2}{\partial y^2} \hat{u}(x, y) - Q(x, y) \right] \ell_i(x, y) = 0 \quad i = 1, 2, \dots, N \quad (8.11)$$

Next employ Green's Theorem (two-dimensional version of integration by parts) to the second-order terms. We get

(8.12)

$$\int_{\Omega} - \left[\frac{\partial}{\partial x} \hat{u}(x, y) \frac{\partial}{\partial x} \ell_i(x, y) + \frac{\partial}{\partial y} \hat{u}(x, y) \frac{\partial}{\partial y} \ell_i(x, y) + Q(x, y) \ell_i(x, y) \right] d\Omega$$

$$+ \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_i(x, y) dl = 0 \quad i = 1, 2, \dots, N$$

where dl is taken along the side $\partial\Omega$ in a counterclockwise direction. We will explain how to handle the term

$$\int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_i(x, y) dl$$

later on page 191. Substitution for $\hat{u}(x, y)$ yields

(8.13)

$$\int_{\Omega} - \left[\sum_{j=1}^N u_j \left\{ \frac{\partial}{\partial x} \ell_j(x, y) \frac{\partial}{\partial x} \ell_i(x, y) + \frac{\partial}{\partial y} \ell_j(x, y) \frac{\partial}{\partial y} \ell_i(x, y) \right\} \right.$$

$$\left. + Q(x, y) \ell_i(x, y) d\Omega \right] + \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_i(x, y) dl = 0 \quad i = 1, 2, \dots, N.$$

Keep in mind that we are now using single valued indices so *the subscripts i and j are now referring to basis functions and weighting functions located at nodes i and j .*

Let us assume we are working with the same problem presented earlier and reproduced as [Fig. 8.2](#). We now write [Eq. \(8.13\)](#) for this problem

(8.14)

$$\int_{\Omega} - \left[\sum_{j=1}^9 u_j \left\{ \frac{\partial}{\partial x} \ell_j(x, y) \frac{\partial}{\partial x} \ell_i(x, y) + \frac{\partial}{\partial y} \ell_j(x, y) \frac{\partial}{\partial y} \ell_i(x, y) \right\} + Q(x, y) \ell_i(x, y) \right] d\Omega$$

$$+ \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_i(x, y) dl = 0 \quad i = 1, 2, \dots, 9.$$

The matrix form of this equation is

$$\begin{bmatrix} a_{11} & a_{12} & 0 & a_{14} & a_{15} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & 0 & a_{35} & a_{36} & 0 & 0 & 0 \\ a_{41} & a_{42} & 0 & a_{44} & a_{45} & 0 & a_{47} & a_{48} & 0 \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} & a_{58} & a_{59} \\ 0 & a_{62} & a_{63} & 0 & a_{65} & a_{66} & 0 & a_{68} & a_{69} \\ 0 & 0 & 0 & a_{74} & a_{75} & 0 & a_{77} & a_{78} & 0 \\ 0 & 0 & 0 & a_{84} & a_{85} & a_{86} & a_{87} & a_{88} & a_{89} \\ 0 & 0 & 0 & 0 & a_{95} & a_{96} & 0 & a_{98} & a_{99} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \\ u_7 \\ u_8 \\ u_9 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{bmatrix} \quad (8.15)$$

or in more simplified notation.

$$[A] \{u\} = \{f\}. \quad (8.16)$$

The elements of the coefficient matrix $[A]$ are defined as follows

$$a_{ij} = - \int_{\Omega} \left(\frac{\partial}{\partial x} \ell_j(x, y) \frac{\partial}{\partial x} \ell_i(x, y) + \frac{\partial}{\partial y} \ell_j(x, y) \frac{\partial}{\partial y} \ell_i(x, y) \right) d\Omega \quad (8.17)$$

$$f_i = \int_{\Omega} Q(x, y) \ell_i(x, y) d\Omega - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_i(x, y) d\ell. \quad (8.18)$$

The information appearing on the right-hand side will be known from a specification of $Q(x, y)$ or, as will be seen later, by boundary conditions. Now consider the evaluation of the integral appearing in [Eq. \(8.17\)](#). In our earlier work (see Section 5.2 on page 94) we performed the integrations over the elements. We will do the same thing here. We rewrite [Eq. \(8.17\)](#) as

$$a_{ij} = - \sum_{e=1}^{e=E} \int_{\Omega_e} \left(\frac{\partial}{\partial x} \ell_j(x, y) \frac{\partial}{\partial x} \ell_i(x, y) + \frac{\partial}{\partial y} \ell_j(x, y) \frac{\partial}{\partial y} \ell_i(x, y) \right) d\Omega_e \quad (8.19)$$

where the integration is now over element Ω_e .

Consider a typical element such as shown in [Fig. 8.3](#); in our case it happens to be a unit square, but that is not necessary. Note that element numbers are shown circled. The contribution of this element for node (i, j) is

$$- \int_{\Omega_1} \left(\frac{\partial}{\partial x} \ell_j(x, y) \frac{\partial}{\partial x} \ell_i(x, y) + \frac{\partial}{\partial y} \ell_j(x, y) \frac{\partial}{\partial y} \ell_i(x, y) \right) d\Omega_1 \quad i = 1, 2, 5, 4, \quad j = 1, 2, 5, 4. \quad (8.20)$$

We can expand the definitions of the basis and weighting functions using [Eq. \(8.6\)](#) to give

$$- \int_{\Omega_1} \left[\frac{\partial}{\partial x} (\ell_j(x) \ell_j(y)) \frac{\partial}{\partial x} (\ell_i(x) \ell_i(y)) + \frac{\partial}{\partial y} (\ell_j(y) \ell_j(x)) \frac{\partial}{\partial y} (\ell_i(y) \ell_i(x)) \right] d\Omega_1 \quad (8.21)$$

$i = 1, 2, 5, 4, \quad j = 1, 2, 5, 4.$

This looks like a backwards step, but, as we will see in a moment, it actually simplifies our development.

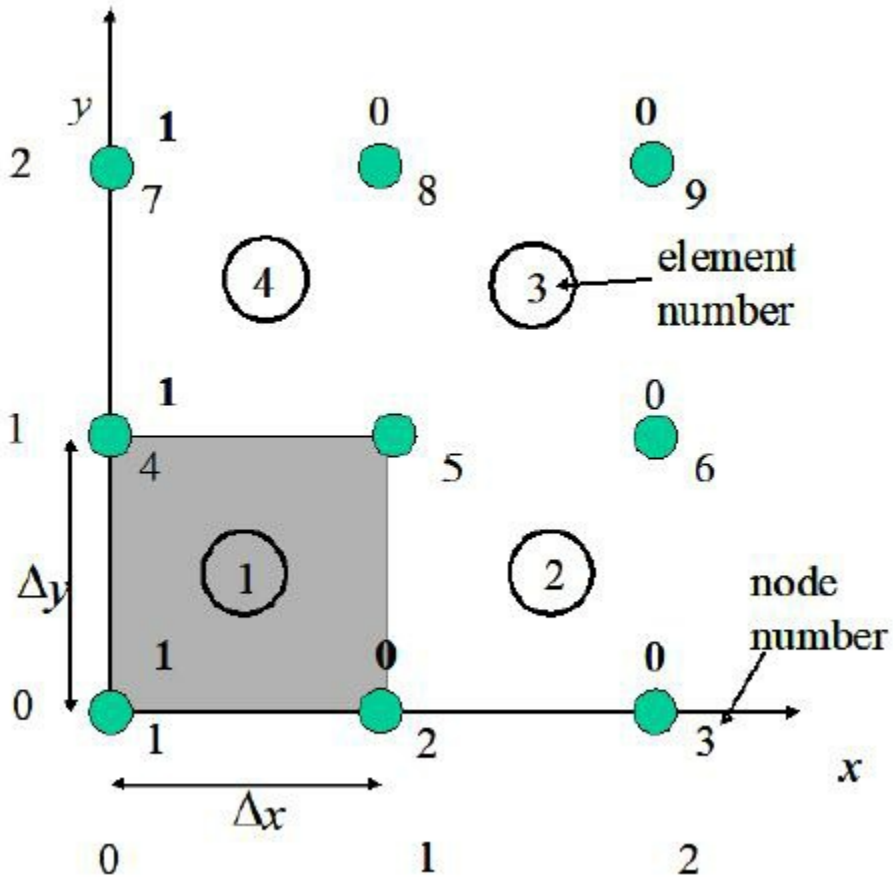


Figure 8.3: Finite difference grid of two-dimensional problem in space.

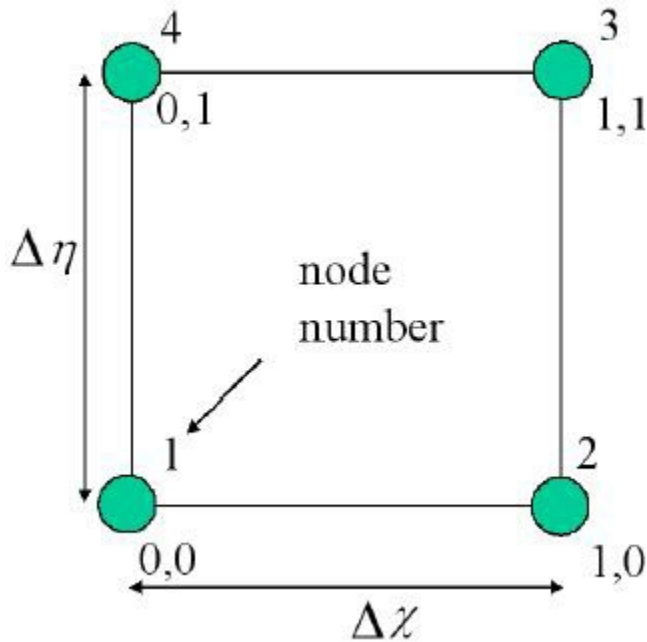


Figure 8.4: One local element defined in ξ, η local coordinates.

To set the stage for what follows, we will change direction and consider the case of a generic element with node numbering 1, 2, 5 4 shown in Fig. 8.3. There are a lot of numbers in this figure, so let us take a minute to clarify them. The ones and zeros around the outside of the array of nodes are the boundary condition values. The numbers to the lower right of the nodes are the node numbers; they go from 1 to 9. The numbers in the circles are the element numbers. There are four of them and the one in the lower left of the figure is shaded. At the far left and bottom of the figure are the distances along y and x , respectively. The ranges are $x \in [0, 2]$ and $y = [0, 2]$. Then the translation in numbering between our **actual or 'global' element** and our **generic or 'local element'** is given in the following table. Take a moment to study this little table. In the first row are the nodes defined in the local coordinate system (ξ, η) . In the second row are the corresponding values in the global coordinate system. Thus, for example, if you look in the second column you see that the local node 1 corresponds to the global node 1 but in the case of column 4 the local node 3 corresponds to the global node 5.

The integrations we need to perform to employ the Galerkin finite element method can be accomplished for this element using local coordinates χ for x and η for y (see Section 1.8 on page 14 for an introduction to local coordinates):

Local	1	2	3	4	(8.22)
Global	1	2	5	4	

$$\begin{aligned}
& - \int_0^1 \int_0^1 \left(\frac{\partial}{\partial \chi} [\ell_J(\chi) \ell_J(\eta)] \frac{\partial \chi}{\partial x} \right) \left(\frac{\partial}{\partial \chi} [\ell_I(\chi) \ell_I(\eta)] \frac{\partial \chi}{\partial x} \right) \frac{dx}{d\chi} d\chi \frac{dy}{d\eta} d\eta \\
& - \int_0^1 \int_0^1 \left(\frac{\partial}{\partial \eta} [\ell_J(\chi) \ell_J(\eta)] \frac{\partial \eta}{\partial y} \right) \left(\frac{\partial}{\partial \eta} [\ell_I(\chi) \ell_I(\eta)] \frac{\partial \eta}{\partial y} \right) \frac{dx}{d\chi} d\chi \frac{dy}{d\eta} d\eta \\
& J = 1, 2, 3, 4 \quad I = 1, 2, 3, 4.
\end{aligned} \tag{8.23}$$

Expansion of the derivatives and noting that since χ and η are orthogonal so that terms such as $\frac{\partial}{\partial \eta} [\ell_J(\chi)] = 0$, yields

$$\begin{aligned}
& - \int_0^1 \int_0^1 \left(\frac{\partial}{\partial \chi} [\ell_J(\chi)] \ell_J(\eta) \frac{\partial \chi}{\partial x} \right) \left(\frac{\partial}{\partial \chi} [\ell_I(\chi)] \ell_I(\eta) \frac{\partial \chi}{\partial x} \right) \frac{dx}{d\chi} d\chi \frac{dy}{d\eta} d\eta \\
& - \int_0^1 \int_0^1 \left(\frac{\partial}{\partial \eta} [\ell_J(\eta)] \ell_J(\chi) \frac{\partial \eta}{\partial y} \right) \left(\frac{\partial}{\partial \eta} [\ell_I(\eta)] \ell_I(\chi) \frac{\partial \eta}{\partial y} \right) \frac{dx}{d\chi} d\chi \frac{dy}{d\eta} d\eta \\
& J = 1, 2, 3, 4 \quad I = 1, 2, 3, 4.
\end{aligned} \tag{8.24}$$

Notice the subtle difference between [Eqs. \(8.23\)](#) and [\(8.24\)](#).

Consider, as an example, the term associated with $I = 1, J = 1$, that is

$$\begin{aligned}
& - \int_0^1 \int_0^1 \left(\frac{\partial}{\partial \chi} [\ell_1(\chi)] \ell_1(\eta) \frac{\partial \chi}{\partial x} \right) \left(\frac{\partial}{\partial \chi} [\ell_1(\chi)] \ell_1(\eta) \frac{\partial \chi}{\partial x} \right) \frac{dx}{d\chi} d\chi \frac{dy}{d\eta} d\eta \\
& - \int_0^1 \int_0^1 \left(\frac{\partial}{\partial \eta} [\ell_1(\eta)] \ell_1(\chi) \frac{\partial \eta}{\partial y} \right) \left(\frac{\partial}{\partial \eta} [\ell_1(\eta)] \ell_1(\chi) \frac{\partial \eta}{\partial y} \right) \frac{dx}{d\chi} d\chi \frac{dy}{d\eta} d\eta.
\end{aligned} \tag{8.25}$$

From the definitions of the Lagrange polynomials in local coordinates (see Eq. 5.15 on page 85) we have, after performing the required differentiations,

$$\begin{aligned}
& - \int_0^1 \int_0^1 \left(\frac{-1}{1} \ell_1(\eta) \frac{1}{1} \right) \left(\frac{-1}{1} \ell_1(\eta) \frac{1}{1} \right) \frac{1}{1} d\chi \frac{1}{1} d\eta \\
& - \int_0^1 \int_0^1 \left(\frac{-1}{1} \ell_1(\chi) \frac{1}{1} \right) \left(\frac{-1}{1} \ell_1(\chi) \right) \frac{1}{1} d\chi \frac{1}{1} d\eta
\end{aligned} \tag{8.26}$$

or, on further substitution

$$\begin{aligned}
& - \int_0^1 \int_0^1 \left(\frac{-1}{1} (1 - \eta) \frac{1}{1} \right) \left(\frac{-1}{1} (1 - \eta) \frac{1}{1} \right) \frac{1}{1} d\chi \frac{1}{1} d\eta \\
& - \int_0^1 \int_0^1 \left(\frac{-1}{1} (1 - \chi) \frac{1}{1} \right) \left(\frac{-1}{1} (1 - \chi) \right) \frac{1}{1} d\chi \frac{1}{1} d\eta
\end{aligned} \tag{8.27}$$

which reduces to

$$\begin{aligned}
& - \int_0^1 \int_0^1 (1 - \eta) (1 - \eta) d\chi d\eta \\
& - \int_0^1 \int_0^1 (1 - \chi) (1 - \chi) d\chi d\eta.
\end{aligned} \tag{8.28}$$

Evaluation of the integrals gives the (1, 1) entry of the local coefficient matrix, that is

$$\left(-\frac{1}{3} \right) + \left(-\frac{1}{3} \right) = -\frac{2}{3}. \tag{8.29}$$

As a second example, consider $I = 1, J = 2$. which produces an off-diagonal coefficient. From [Eq. \(8.24\)](#) we get

$$\begin{aligned}
& - \int_0^1 \int_0^1 \left(\frac{\partial}{\partial \chi} [\ell_2(\chi)] \ell_2(\eta) \frac{\partial \chi}{\partial x} \right) \left(\frac{\partial}{\partial \chi} [\ell_1(\chi)] \ell_1(\eta) \frac{\partial \chi}{\partial x} \right) \frac{dx}{d\chi} d\chi \frac{dy}{d\eta} d\eta \\
& - \int_0^1 \int_0^1 \left(\frac{\partial}{\partial \eta} [\ell_2(\eta)] \ell_2(\chi) \frac{\partial \eta}{\partial y} \right) \left(\frac{\partial}{\partial \eta} [\ell_1(\eta)] \ell_1(\chi) \frac{\partial \eta}{\partial y} \right) \frac{dx}{d\chi} d\chi \frac{dy}{d\eta} d\eta.
\end{aligned} \tag{8.30}$$

Again we substitute for the definitions of the Lagrange polynomials and differentiate to give

$$\begin{aligned}
& - \int_0^1 \int_0^1 \left(\frac{1}{1} \ell_2(\eta) \frac{1}{1} \right) \left(\frac{-1}{1} \ell_1(\eta) \frac{1}{1} \right) \frac{1}{1} d\chi \frac{1}{1} d\eta \\
& - \int_0^1 \int_0^1 \left(\frac{1}{1} \ell_2(\chi) \frac{1}{1} \right) \left(\frac{-1}{1} \ell_1(\chi) \frac{1}{1} \right) \frac{1}{1} d\chi \frac{1}{1} d\eta
\end{aligned} \tag{8.31}$$

which upon further substitution yields,

$$\begin{aligned}
& - \int_0^1 \int_0^1 \left(\frac{1}{1} (1 - \eta) \frac{1}{1} \right) \left(\frac{-1}{1} (1 - \eta) \frac{1}{1} \right) \frac{1}{1} d\chi \frac{1}{1} d\eta \\
& - \int_0^1 \int_0^1 \left(\frac{-1}{1} (\chi) \frac{1}{1} \right) \left(\frac{-1}{1} (1 - \chi) \right) \frac{1}{1} d\chi \frac{1}{1} d\eta
\end{aligned} \tag{8.32}$$

which simplifies to

$$\begin{aligned}
& \int_0^1 \int_0^1 (1 - \eta) (1 - \eta) d\chi d\eta \\
& - \int_0^1 \int_0^1 (\chi) (1 - \chi) d\chi d\eta
\end{aligned} \tag{8.33}$$

and finally yields, upon integration,

$$\left(\frac{1}{3} \right) + \left(-\frac{1}{6} \right) = \frac{1}{6}. \tag{8.34}$$

If we perform similar integrations for the remaining nodal combinations in this element we obtain the local coefficient matrix which we provide as a table. For convenience in this table we have identified the nodes in row 2 and also in column 5

J					
1	2	3	4		
$-\frac{2}{3}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$	1	I
$\frac{1}{6}$	$-\frac{2}{3}$	$\frac{1}{6}$	$\frac{1}{3}$	2	
$\frac{1}{3}$	$\frac{1}{6}$	$-\frac{2}{3}$	$\frac{1}{6}$	3	
$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$	$-\frac{2}{3}$	4	

(8.35)

The form of the matrix is

$$\begin{bmatrix}
-\frac{2}{3} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\
\frac{1}{6} & -\frac{2}{3} & \frac{1}{6} & \frac{1}{3} \\
\frac{1}{3} & \frac{1}{6} & -\frac{2}{3} & \frac{1}{6} \\
\frac{1}{6} & \frac{1}{3} & \frac{1}{6} & -\frac{2}{3}
\end{bmatrix}. \tag{8.36}$$

All the element matrices will be the same in this case so we can construct the global matrix using the local-global coefficient transformation table, that is the table shown as [Eq. \(8.37\)](#). To be sure we understand this table, let us take a minute to examine it. The second row contains the nodal locations expressed in terms of global coordinates (see

[Fig. 8.5](#)). The last column on the right contains the element numbers. The third row contains the local coordinate node numbers in element 1 corresponding to the same nodes defined in the global numbering system. For example, global node 5 shown in the second row corresponds to the local node 3 identified in the third row with the local coordinate system for element 1. Since this mapping from global node numbers to local node numbers is a basic concept, *the reader is encouraged to review this paragraph until it is clearly understood.*

Local	Global									Element
	1	2	3	4	5	6	7	8	9	
	1	2		4	3					1
		1	2		4	3				2
					1	2		4	3	3
				1	2		4	3		4

(8.37)

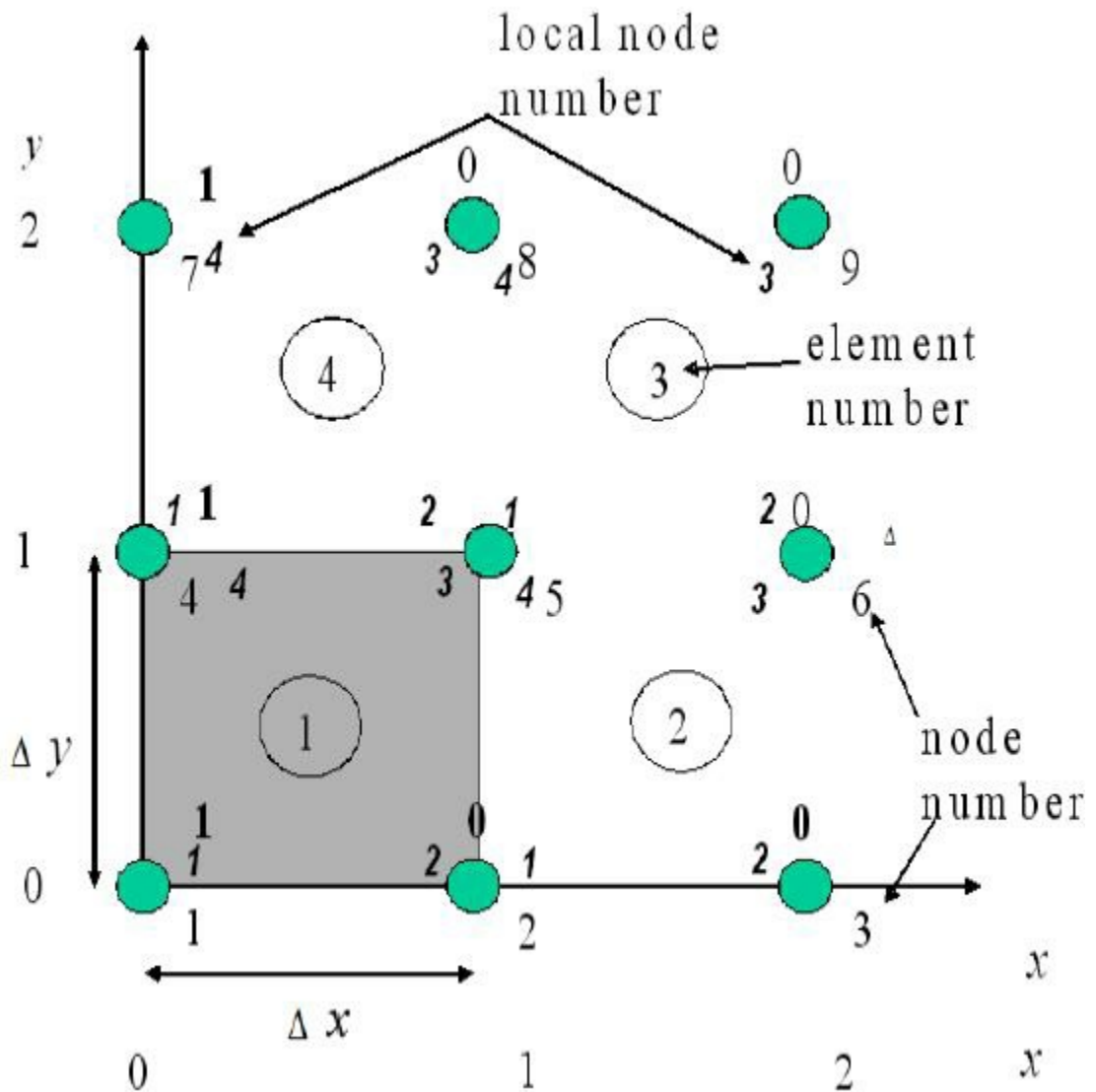


Figure 8.5: Figure used to relate the location of nodes in the local and global coordinate systems.

The global matrix is shown in [Eq. \(8.38\)](#). Let us see how it is obtained. Consider the (1, 1) coefficient in [Fig. \(8.38\)](#). This coefficient is associated only with global node 1. We first locate the element in which global node 1 is located (please refer to [Fig. 8.5](#)). We see that it is in element number 1. Next find the local node that corresponds to the global node (1, 1) in local element 1; it is local node 1. So the value we seek to fill location (1, 1) in the global matrix is the coefficient (1, 1) in local element matrix 1. Equation [Eq. \(8.35\)](#) or [Eq. \(8.36\)](#) tell us that the value we seek is $-\frac{2}{3}$. Since node 1 in the global matrix appears in only one element, the information in location (1, 1) is all that is available. Thus placing the value $-\frac{2}{3}$ in the global matrix at location (1, 1) completes the operation to fill this global coefficient location.

The global location (2, 2) is different because it involves information from both elements 1 and 2. This is because the weighting function for this node spans both elements. Thus,

to obtain this coefficient, it is necessary to use the table in [Eq. \(8.37\)](#) for element 1 and 2. For the first element we see that global node 2 corresponds to local node 2. Therefore we need to take the value (2, 2) from the element coefficient matrix and this is again $-\frac{2}{3}$ and place it in global matrix location (2, 2).

But we are not finished. We now need to obtain the contribution from element 2. Global node 2 in element 2 is local node 1 as seen from the table in row 4, column 3. Going to [Eq. \(8.36\)](#) we again find in element (1, 1) a value of $-\frac{2}{3}$. We need to add this to the global coefficient location (2, 2). Thus we have $-\frac{2}{3}$ from element 1 and $-\frac{2}{3}$ from element 2 which gives a total of $\left(-\frac{2}{3} - \frac{2}{3} = -\frac{4}{3}\right)$ which we found in the global matrix at location (2, 2). It would be prudent for the reader to try and obtain the global coefficient (2, 5), which involves yet another concept:.

$$[A] = - \begin{bmatrix} \frac{2}{3} & -\frac{1}{6} & 0 & -\frac{1}{6} & -\frac{1}{3} & 0 & 0 & 0 & 0 \\ -\frac{1}{6} & \frac{4}{3} & -\frac{1}{6} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 \\ 0 & -\frac{1}{6} & \frac{2}{3} & 0 & -\frac{1}{3} & -\frac{1}{6} & 0 & 0 & 0 \\ -\frac{1}{6} & -\frac{1}{3} & 0 & \frac{4}{3} & -\frac{1}{3} & 0 & -\frac{1}{6} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & \frac{8}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & -\frac{1}{3} & -\frac{1}{6} & 0 & -\frac{1}{3} & \frac{4}{3} & 0 & -\frac{1}{3} & -\frac{1}{6} \\ 0 & 0 & 0 & -\frac{1}{6} & -\frac{1}{3} & 0 & \frac{2}{3} & -\frac{1}{6} & 0 \\ 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{6} & \frac{4}{3} & -\frac{1}{6} \\ 0 & 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}. \quad (8.38)$$

Consider now the integral associated with the right-hand side of our matrix equation the includes the source term $Q(x, y)$, namely

$$f_i = \int_{\Omega} Q(x, y) \ell_i(x, y) d\Omega \quad (8.39)$$

which, assuming $Q(x, y)$ constant over the element, is for our local element

$$f_I = Q \int_0^1 \int_0^1 \ell_I(\chi, \eta) \frac{dx}{d\chi} d\chi \frac{dy}{d\eta} d\eta. \quad (8.40)$$

Consider the case of $I = 1$

$$f_1 = Q \int_0^1 \int_0^1 \ell_1(\chi, \eta) \frac{dx}{d\chi} d\chi \frac{dy}{d\eta} d\eta \quad (8.41)$$

or, substituting for $\ell_1(x, y)$, we obtain

$$f_1 = Q \int_0^1 \int_0^1 \ell_1(\chi) \ell_1(\eta) \frac{dx}{d\chi} d\chi \frac{dy}{d\eta} d\eta. \quad (8.42)$$

Upon substitution of the definitions of the Lagrange polynomials we get

$$f_1 = Q \int_0^1 \int_0^1 (1 - \chi) (1 - \eta) \frac{dx}{d\chi} d\chi \frac{dy}{d\eta} d\eta \quad (8.43)$$

or

$$f_1 = Q \left[\int_0^1 (1 - \chi) \frac{dx}{d\chi} d\chi \right] \left[\int_0^1 (1 - \eta) \frac{dy}{d\eta} d\eta \right]. \quad (8.44)$$

Integration of this equation provides our final result, namely

$$f_1 = Q \left(\frac{1}{2} \right) \left(\frac{1}{2} \right) = \frac{Q}{4}. \quad (8.45)$$

In general we also have a term of the form (see [Eq. \(8.18\)](#)) which represents a Neumann, second type or flux boundary condition, that is,

$$\int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_i(x, y) d\ell. \quad (8.46)$$

Writing this integral in local coordinates we obtain for integration along the η direction

$$\frac{\partial \hat{u}}{\partial n} \int_0^1 \ell_I(\chi, \eta) \frac{dy}{d\eta} d\eta \quad (8.47)$$

or

$$\frac{\partial \hat{u}}{\partial n} \int_0^1 \ell_I(\chi) \ell_I(\eta) \frac{dy}{d\eta} d\eta. \quad (8.48)$$

Let us select $I = 1$ to provide a concrete example

(8.49)

$$\begin{aligned}
& \frac{\partial \hat{u}}{\partial n} \ell_1(\chi) \int_0^1 \ell_1(\eta) \frac{dy}{d\eta} d\eta \\
&= \frac{\partial \hat{u}}{\partial n} \ell_1(\chi) \int_0^1 (1-\eta) \frac{1}{1} \frac{dy}{d\eta} d\eta \\
&= \frac{\partial \hat{u}}{\partial n} \ell_1(\chi) \left(\eta - \frac{\eta^2}{2} \right) \Big|_0^1 \frac{dy}{d\eta} \\
&= \frac{\partial \hat{u}}{\partial n} (1) \left(\frac{1}{2} \right).
\end{aligned}$$

Note that the value of $\ell_1(\chi)$ at $\chi = 0$ is unity and thus the (1) in the last equation. [Eq. \(8.49\)](#) states that one half of the flux entering the side of the element of length 1 is allocated to node 1. The other half would be allocated to the adjacent node along the external face of the element being considered. The $f(\chi, \eta)$ vector is obtained by examining all the nodes in the global system, that is

$$\{f\} = \begin{bmatrix} 0 \\ 4 \\ 0 \\ 2 \\ 0 \\ 4 \\ 0 \\ 2 \\ 0 \\ 1 \\ 0 \\ 2 \\ 0 \\ 4 \\ 0 \\ 2 \\ 0 \\ 4 \end{bmatrix} - \begin{bmatrix} \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_1(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_2(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_3(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_4(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_5(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_6(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_7(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_8(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \ell_9(x, y) d\ell \end{bmatrix}. \quad (8.50)$$

If we combine the information provided above about our approximation of [Eq. \(8.1\)](#), we obtain the matrix equation

$$\begin{aligned}
& - \begin{bmatrix} \frac{2}{3} & -\frac{1}{6} & 0 & -\frac{1}{6} & -\frac{1}{3} & 0 & 0 & 0 & 0 \\ -\frac{1}{6} & \frac{4}{3} & -\frac{1}{6} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 \\ 0 & -\frac{1}{3} & \frac{2}{6} & 0 & -\frac{1}{3} & -\frac{1}{6} & 0 & 0 & 0 \\ -\frac{1}{6} & -\frac{1}{3} & 0 & \frac{4}{3} & -\frac{1}{3} & 0 & -\frac{1}{6} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & \frac{8}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & -\frac{1}{3} & -\frac{1}{6} & 0 & -\frac{1}{3} & \frac{4}{3} & 0 & -\frac{1}{3} & -\frac{1}{6} \\ 0 & 0 & 0 & -\frac{1}{6} & -\frac{1}{3} & 0 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{6} & \frac{4}{3} & -\frac{1}{6} \\ 0 & 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & \frac{2}{3} & \frac{2}{6} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \\ u_7 \\ u_8 \\ u_9 \end{bmatrix} \\
& = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{2} \end{bmatrix} - \begin{bmatrix} \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_1(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_2(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_3(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_4(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_5(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_6(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_7(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_8(x, y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_9(x, y) d\ell \end{bmatrix}.
\end{aligned} \tag{8.51}$$

Application of the boundary conditions gives

$$\begin{aligned}
 & - \begin{bmatrix} \frac{2}{3} & -\frac{1}{6} & 0 & -\frac{1}{6} & -\frac{1}{3} & 0 & 0 & 0 & 0 \\ -\frac{1}{6} & \frac{4}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 \\ 0 & -\frac{1}{6} & \frac{2}{6} & 0 & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 \\ -\frac{1}{6} & -\frac{1}{6} & 0 & \frac{4}{6} & -\frac{1}{3} & 0 & -\frac{1}{6} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & \frac{8}{3} & -\frac{1}{3} & -\frac{1}{6} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & -\frac{1}{3} & -\frac{1}{6} & 0 & -\frac{1}{3} & \frac{4}{3} & 0 & -\frac{1}{3} & -\frac{1}{6} \\ 0 & 0 & 0 & -\frac{1}{6} & -\frac{1}{3} & 0 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{6} & \frac{4}{3} & -\frac{1}{6} \\ 0 & 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & \frac{2}{3} & \frac{2}{6} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ u_5 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{8.52} \\
 & = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_1(x,y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_2(x,y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_3(x,y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_4(x,y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_5(x,y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_6(x,y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_7(x,y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_8(x,y) d\ell \\ \int_{\partial\Omega} \frac{\partial u}{\partial n} \ell_9(x,y) d\ell \end{bmatrix}
 \end{aligned}$$

which simplifies for this simple case to

$$-8u_5 = -1 - 1 - 1 + 3Q/1 \tag{8.53}$$

where the surface integral associated with this equation (Eq. (8.53)) is zero because $\ell_5(x, y)$ is zero along the external boundary.

Assuming $Q = 1$

$$u_5 = 0. \tag{8.54}$$

This is the same result we obtained with the finite difference method.

If we assume $Q = 0$ we have

$$u_5 = \frac{3}{8} \tag{8.55}$$

which compares with $u_5 = \frac{1}{4}$ obtained using the finite difference method. The secret as to why these solutions are different lies in the way that the source term Q is handled. In essence we use a point value at the center node for the finite difference representation and an averaged value in the finite element formulation. This will be better understood

after reading the following discussion on how the finite difference and finite element methods differ for this problem.

To see how the finite difference and finite element approximations are different we expand the equation for u_5 to obtain

$$\frac{1}{3} (u_1 + u_2 + u_3 + u_4 - 8u_5 + u_6 + u_7 + u_8) = Q\Delta x\Delta y. \quad (8.56)$$

We now modify this equation through addition and extraction of these eight values to give something that resembles a series of second derivative approximations, that is

$$\begin{aligned} & \frac{1}{6} [(u_1 - 2u_2 + u_3) + 4(u_4 - 2u_5 + u_6) + (u_7 - 2u_8 + u_9)] \\ & + \frac{1}{6} [(u_1 - 2u_4 + u_7) + 4(u_2 - 2u_5 + u_8) - (u_3 - 2u_6 + u_9)] \\ & = -Q\Delta x\Delta y. \end{aligned} \quad (8.57)$$

Next we multiply and divide the first term by Δx^2 and the second by Δy^2 . The result is

$$\begin{aligned} & \frac{\Delta x^2}{6} \left[\frac{(u_1 - 2u_2 + u_3)}{\Delta x^2} + \frac{4(u_4 - 2u_5 + u_6)}{\Delta x^2} + \frac{(u_7 - 2u_8 + u_9)}{\Delta x^2} \right] \\ & + \frac{\Delta y^2}{6} \left[\frac{(u_1 - 2u_4 + u_7)}{\Delta y^2} + 4\frac{(u_2 - 2u_5 + u_8)}{\Delta y^2} - \frac{(u_3 - 2u_6 + u_9)}{\Delta y^2} \right] \\ & = -Q\Delta x\Delta y \end{aligned} \quad (8.58)$$

The next step is to divide the equation through by $\Delta x\Delta y$ to yield

$$\begin{aligned} & \frac{1}{\Delta x\Delta y} \frac{\Delta x^2}{6} \left[\frac{(u_1 - 2u_2 + u_3)}{\Delta x^2} + \frac{4(u_4 - 2u_5 + u_6)}{\Delta x^2} + \frac{(u_7 - 2u_8 + u_9)}{\Delta x^2} \right] \\ & + \frac{1}{\Delta x\Delta y} \frac{\Delta y^2}{6} \left[\frac{(u_1 - 2u_4 + u_7)}{\Delta y^2} + 4\frac{(u_2 - 2u_5 + u_8)}{\Delta y^2} - \frac{(u_3 - 2u_6 + u_9)}{\Delta y^2} \right] \\ & = -Q. \end{aligned} \quad (8.59)$$

We now simplify and rearrange the terms multiplying the contents of the square brackets. The result is

(8.60)

$$\begin{aligned}
& \frac{1}{2\Delta y} \frac{\Delta x}{3} \left[\frac{(u_1 - 2u_2 + u_3)}{\Delta x^2} + \frac{4(u_4 - 2u_5 + u_6)}{\Delta x^2} + \frac{(u_7 - 2u_8 + u_9)}{\Delta x^2} \right] \\
& + \frac{1}{2\Delta x} \frac{\Delta y}{3} \left[\frac{(u_1 - 2u_4 + u_7)}{\Delta y^2} + 4 \frac{(u_2 - 2u_5 + u_8)}{\Delta y^2} - \frac{(u_3 - 2u_6 + u_9)}{\Delta y^2} \right] \\
& = -Q.
\end{aligned}$$

Finally, we recognize that the multipliers of the second derivative approximations are those used in Simpson's rule. We take advantage of this relationship to give the following equation.

(8.61)

$$\begin{aligned}
& \frac{1}{2\Delta y} \frac{\Delta x}{3} \left[\frac{(u_1 - 2u_2 + u_3)}{\Delta x^2} + \frac{4(u_4 - 2u_5 + u_6)}{\Delta x^2} + \frac{(u_7 - 2u_8 + u_9)}{\Delta x^2} \right] \\
& + \frac{1}{2\Delta x} \frac{\Delta y}{3} \left[\frac{(u_1 - 2u_4 + u_7)}{\Delta y^2} + 4 \frac{(u_2 - 2u_5 + u_8)}{\Delta y^2} - \frac{(u_3 - 2u_6 + u_9)}{\Delta y^2} \right] \\
& + Q \\
& \approx \frac{1}{2\Delta y} \int_0^{2\Delta y} \frac{\partial^2 u}{\partial x^2} dy + \frac{1}{2\Delta x} \int_0^{2\Delta x} \frac{\partial^2 u}{\partial y^2} dx + Q.
\end{aligned}$$

[Equation 8.61](#) states that the finite element formulation for the problem analyzed produces a discrete equation that uses the finite difference approximations to the second derivative terms defined in one coordinate direction averaged over the coordinate direction orthogonal to that of the derivatives.

The template for the finite element algorithm described above is found in [Fig. 8.6](#). It is made up of the template for the x -derivative ([Fig. 8.7](#)) added to the template of the y -derivative ([Fig. 8.8](#)).

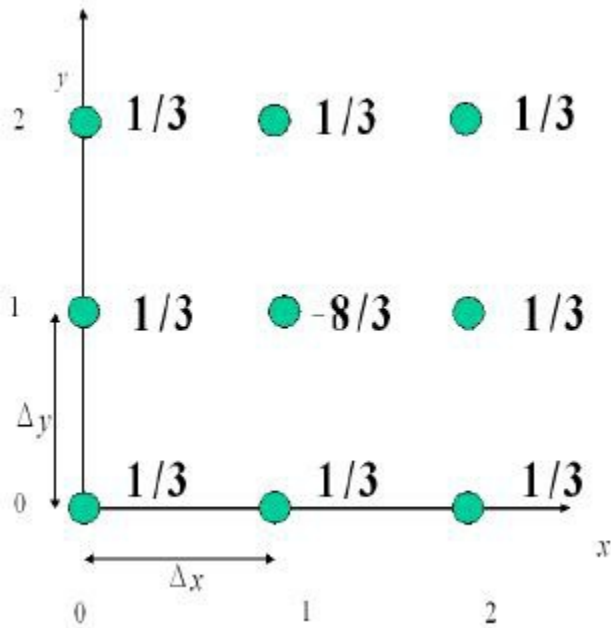


Figure 8.6: Finite element mesh with weights.

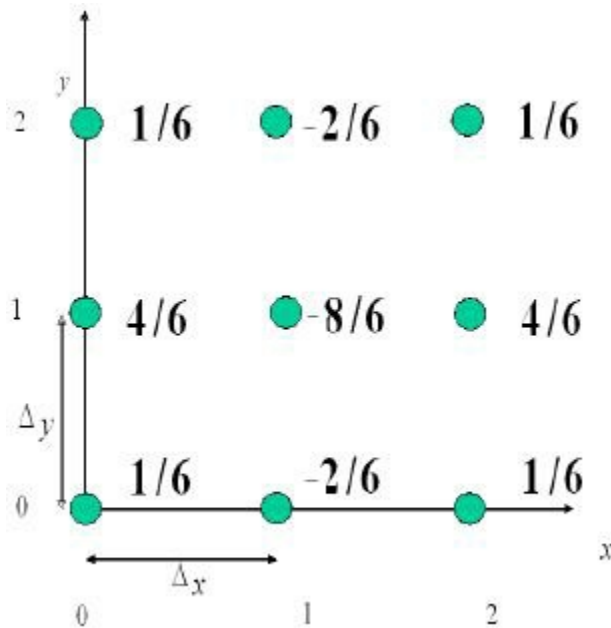


Figure 8.7: Finite element template for x -derivative.

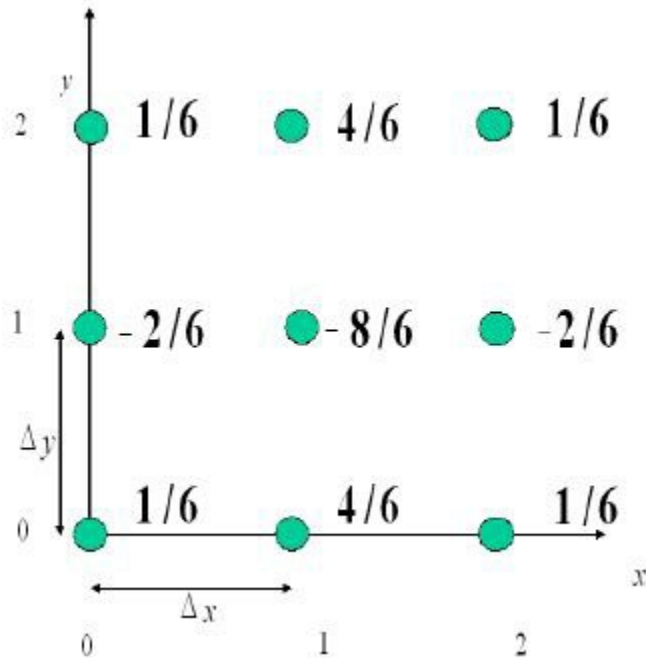


Figure 8.8: Finite element template for y -derivative.

8.2 Finite Element Approximations over Triangles

Consider once again the equation

$$\frac{\partial^2}{\partial x^2} u(x, y) + \frac{\partial^2}{\partial y^2} u(x, y) = Q(x, y) \quad x, y \in \Omega \quad (8.62)$$

defined on the irregular domain illustrated in [Fig. 8.9](#). As earlier, the interior of the domain is defined as Ω and the perimeter as $\partial\Omega$

$$u(x, y) = u_0(x, y) \quad x, y \in \partial\Omega_1 \quad (8.63)$$

$$u(x, y) = u_1(x, y) \quad x, y \in \partial\Omega_2 \quad (8.64)$$

$$\partial\Omega_1 + \partial\Omega_2 = \partial\Omega. \quad (8.65)$$

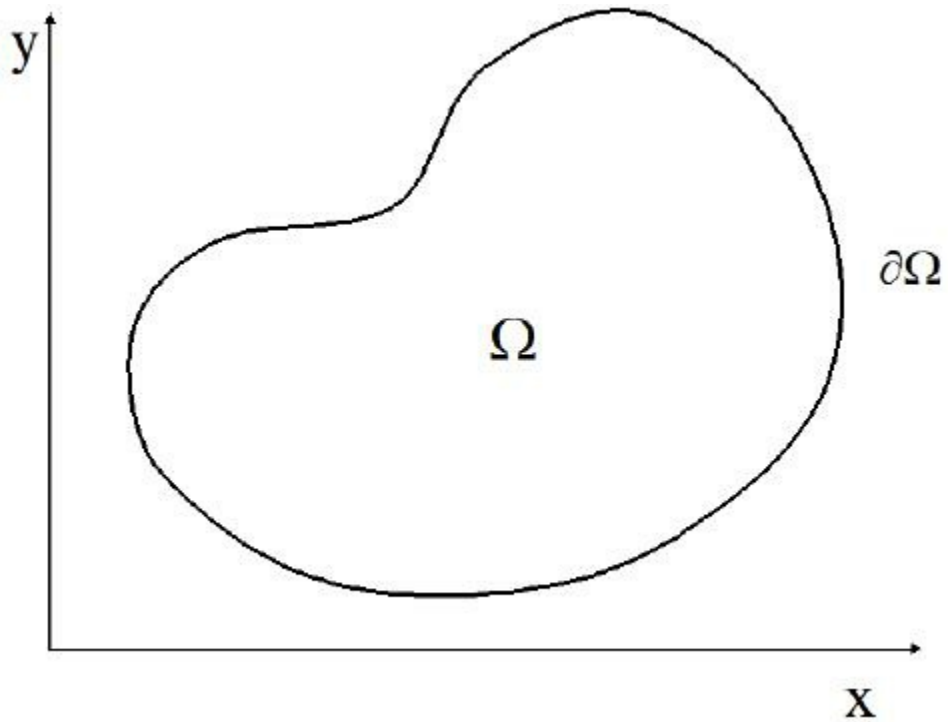


Figure 8.9: Region of interest for the triangular finite element method. The area is Ω and the perimeter is $\partial\Omega$.

We begin the analysis by defining the approximating function applicable to this problem, that is

$$\hat{u}(x, y) = \sum_{j=1}^N u_j \phi_j(x, y). \quad (8.66)$$

In this expression we have replaced our Lagrangian basis functions with a new functional form $\phi_j(x, y)$. This is a polynomial basis function defined on a triangular subspace which we denote as a triangular finite element. The triangular element discretization of the domain shown in [Fig. 8.9](#) is shown in [Fig. 8.10](#).

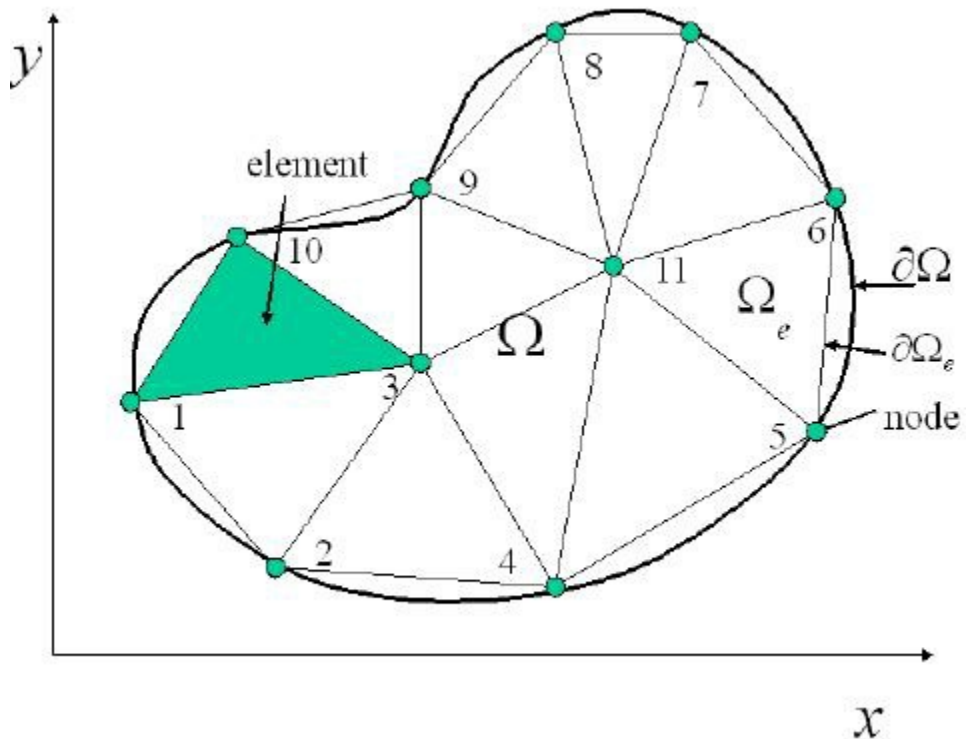


Figure 8.10: Array of finite elements which discretize the area Ω . Note that the boundary of Ω , denoted as $\partial\Omega$, is subdivided into linear elements $\partial\Omega_e$ that are also the sides of the interior triangles. Observe that nodes are located at the vertices of each element.

The triangular basis function for a typical node is shown in [Fig. 8.11](#). Note that, as usual, the basis function is unity at the node for which it is defined and zero at all other nodes. In [Fig. 8.12](#) we show the set of three basis functions for the three nodes in the triangle. Note that everywhere they sum to unity. *It is important that the elements be numbered in the counterclockwise order for the formulae to be presented below for integration to provide correct answers.*

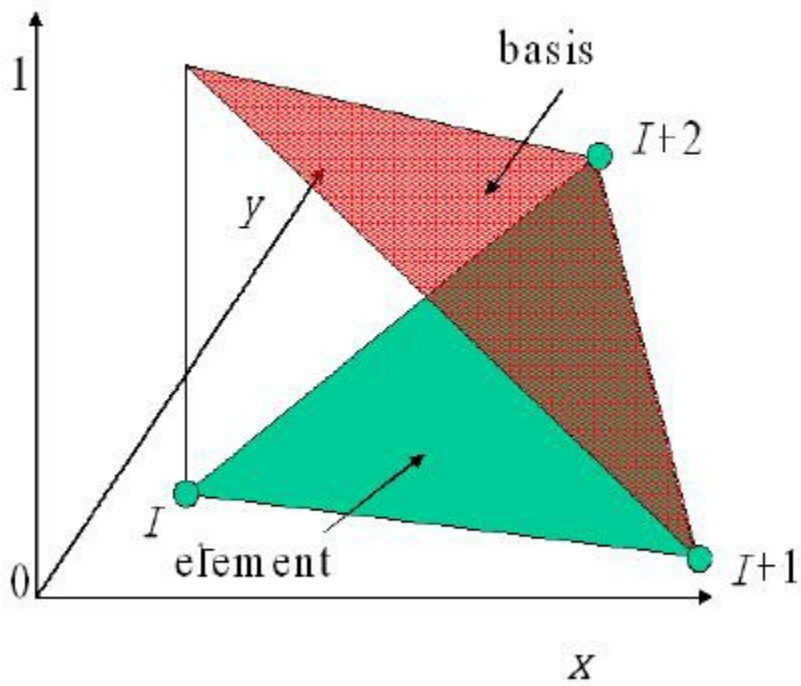


Figure 8.11: Triangular basis function for node I. The patterned inclined plane is the function denoted as the 'basis' and the corresponding element is the solid surface in the x - y plane identified as the 'element'.

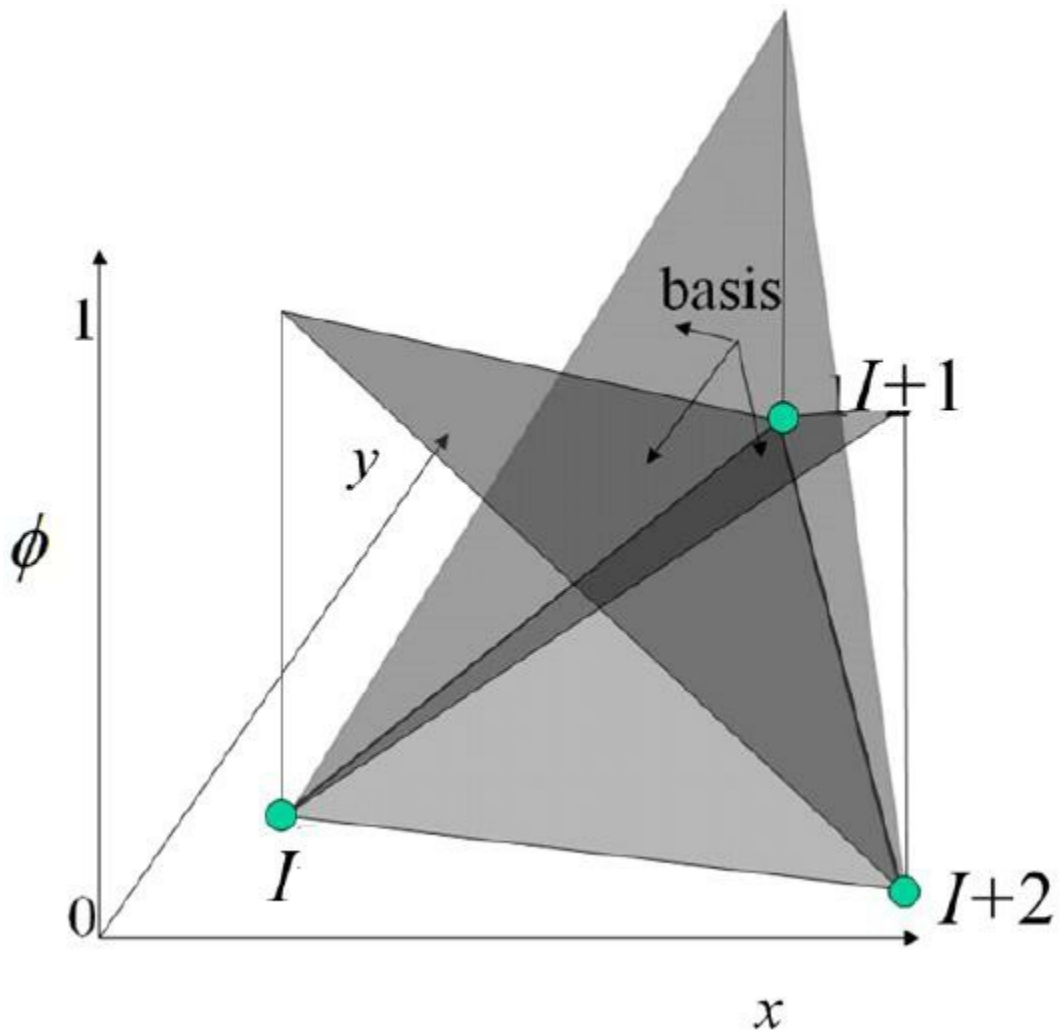


Figure 8.12: Diagrammatic representation of three triangular basis function, one for each node of the triangle. Note that the values of the triangles sum to unity anywhere on the triangle.

8.2.1 Formulation of Triangular Basis Functions

Consider one basis function, for example the one located at node I in [Fig. 8.11](#). Note that this triangle employs the local indices $I, I + 1$ and $I + 2$. Thus the index I can take on any of the nodal values in **global coordinates** shown in [Fig. 8.10](#). To be more specific, if we consider the shaded element in [Fig. 8.10](#) we could choose I to be node number 1, 3, or 10. If we chose $I = 1$, then $I + 1$ would be 3 and $I + 2$ would be 10.

In developing the equations for the triangular basis functions we will require, as noted above, that they be unity at the node for which they are defined and zero at the other two nodes in the triangle. In other words we have for the basis function at node I

$$\phi_I(x_I, y_I) = 1 \quad (8.67)$$

$$\phi_I(x_{I+1}, y_{I+1}) = 0 \quad (8.68)$$

$$\phi_I(x_{I+2}, y_{I+2}) = 0 \quad (8.69)$$

where the notation $\phi_I(x_I, y_I)$ denotes the value of ϕ_I at the nodal location (x_I, y_I) .

To begin our development, note that the equation of a plane defined in the $x - y$ coordinate system is

$$\phi_I(x, y) = ax + by + c. \quad (8.70)$$

If we now evaluate this expression at each node, we have

$$\phi_I(x_I, y_I) = ax_I + by_I + c = 1 \quad (8.71)$$

$$\phi_I(x_{I+1}, y_{I+1}) = ax_{I+1} + by_{I+1} + c = 0 \quad (8.72)$$

$$\phi_I(x_{I+2}, y_{I+2}) = ax_{I+2} + by_{I+2} + c = 0 \quad (8.73)$$

which can be written in matrix form as

$$\begin{bmatrix} x_I & y_I & 1 \\ x_{I+1} & y_{I+1} & 1 \\ x_{I+2} & y_{I+2} & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad (8.74)$$

or, written symbolically

$$[P] \{q\} = \{g\}. \quad (8.75)$$

The solution to this equation gives the expression of $\phi_I(x, y)$:

$$\phi_I(x, y) = \frac{x(y_{I+1} - y_{I+2}) + y(x_{I+2} - x_{I+1}) + (x_{I+1}y_{I+2} - x_{I+2}y_{I+1})}{\det [P]} \quad (8.76)$$

where $\det [P]$ is the determinant of the matrix $[P]$. A similar formulation leads to equations for $\phi_{I+1}(x, y)$ and $\phi_{I+2}(x, y)$, that is

$$\phi_{I+1}(x, y) = \frac{x(y_{I+2} - y_I) + y(x_I - x_{I+2}) + (x_{I+2}y_I - x_I y_{I+2})}{\det [P]} \quad (x, y) \in [\Omega_e] \quad (8.77)$$

and

(8.78)

$$\phi_{I+2}(x, y) = \frac{x(y_I - y_{I+1}) + y(x_{I+1} - x_I) + (x_I y_{I+1} - x_{I+1} y_I)}{\det [P]} \quad (x, y) \in [\Omega_e].$$

It turns out that there is a simple integration formula for triangular elements as defined above; the formula is:

(8.79)

$$\int_{\Omega_e} \phi_I^{m_I} \phi_{I+1}^{m_{I+1}} \phi_{I+2}^{m_{I+2}} d\Omega_e = 2A \frac{m_I! m_{I+1}! m_{I+2}!}{(m_I + m_{I+1} + m_{I+2} + 2)!}$$

where m_I is the **power** to which the ϕ_I polynomial is raised. For example in the integral

$$\int_{\Omega_e} \phi_I \phi_{I+1} \phi_{I+2} d\Omega_e \quad (8.80)$$

m_I, m_{I+1} , and m_{I+2} are all unity, while in the integral

$$\int_{\Omega_e} \phi_I \phi_I \phi_{I+2} d\Omega_e \quad (8.81)$$

or, written slightly differently,

$$\int_{\Omega_e} \phi_I^2 \phi_{I+1}^0 \phi_{I+2}^1 d\Omega_e \quad (8.82)$$

m_I, m_{I+1} , and m_{I+2} are 2, 0, 1 respectively.

The area of the triangle A is given by

$$A = \frac{1}{2} \det [P]. \quad (8.83)$$

Integration Example

Consider the following integral

$$\text{Integral} = \int_{\Omega_e = A_e} \phi_I(x, y) d\Omega_e. \quad (8.84)$$

Since the function ϕ_I appears alone in the integral, $m_I = 1$, $m_{I+1} = 0$, and $m_{I+2} = 0$. Using

[Eq. \(8.79\)](#) and recalling that the factorial of zero is defined to be 1, we obtain

$$\begin{aligned}
 & \int_{\Omega_e=A_e} \phi_I(x, y) d\Omega_e & (8.85) \\
 & = 2A_e \frac{1 \cdot 1 \cdot 1}{(1 + 0 + 0 + 2)!} \\
 & = \frac{2}{6} A_e \\
 & = \frac{A_e}{3}.
 \end{aligned}$$

8.2.2 Example Problem of Finite Element Approximation over Triangles

The equation of interest is, once again,

$$\frac{\partial^2}{\partial x^2} u(x, y) + \frac{\partial^2}{\partial y^2} u(x, y) = 1 \quad (8.86)$$

where (see [Fig. 8.13](#))

$$u(x, y) = 1 - 0.763x \quad x, y \in \partial\Omega_1 \quad (8.87)$$

$$u(x, y) = 1 - 0.38x \quad x, y \in \partial\Omega_2 \quad (8.88)$$

$$u(x, y) = 0 \quad x, y \in \partial\Omega_3. \quad (8.89)$$

These boundary conditions were selected to simplify the calculations by having the area of each triangle be unity and while letting the Dirichlet boundary values vary linearly along the sides of the triangle.

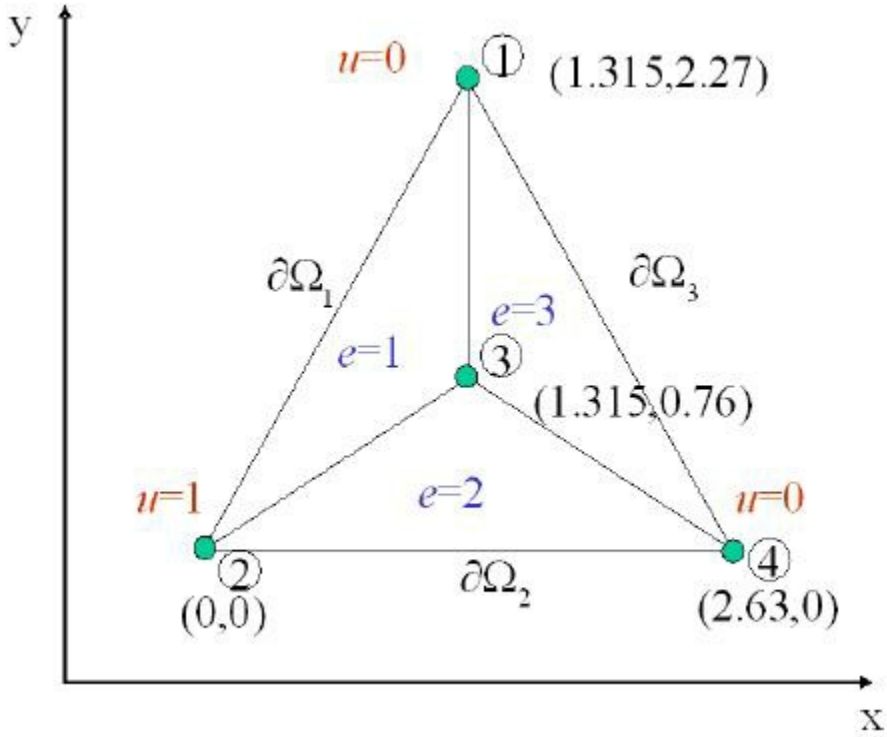


Figure 8.13: Illustration of triangle used as example in this subsection.

We begin by selecting our approximation for $u(x, y)$ as

$$\hat{u}(x, y) = \sum_{j=1}^4 u_j \phi_j(x, y) \quad (8.90)$$

where the basis functions $\phi_j(x, y)$ are defined using [Eqs. \(8.76\) – \(8.78\)](#). As earlier, we now define the residual by substituting the approximation $\hat{u}(x, y)$ into the governing equation, in this case [Eq. \(8.86\)](#) and obtain the residual $R(x, y)$

$$\frac{\partial^2}{\partial x^2} \hat{u}(x, y) + \frac{\partial^2}{\partial y^2} \hat{u}(x, y) - 1 = R(x, y). \quad (8.91)$$

Using Galerkin's method we weight the residual by the same set of functions we are using for the basis functions, that is $\phi_i(x, y)$ $i = 1, \dots, 4$, and integrate each of the four products over $\Omega = \Omega_1 + \Omega_2 + \Omega_3$. The result is

$$\int_{\Omega} R(x, y) \phi_i(x, y) d\Omega = 0 \quad i = 1, \dots, 4. \quad (8.92)$$

Substitution of [Eqs. \(8.90\)](#) and [\(8.91\)](#) into [\(8.92\)](#) yields

(8.93)

$$\int_{\Omega} \left\{ \sum_{j=1}^4 u_j \left(\frac{\partial^2}{\partial x^2} \phi_j(x, y) + \frac{\partial^2}{\partial y^2} \phi_j(x, y) \right) - 1 \right\} \phi_i(x, y) d\Omega = 0 \quad i = 1, \dots, 4.$$

We now apply Green's theorem to [Eq. \(8.93\)](#) to obtain:

(8.94)

$$\begin{aligned} & - \sum_{j=1}^4 u_j \int_{\Omega} \left(\frac{\partial}{\partial x} \phi_j(x, y) \frac{\partial}{\partial x} \phi_i(x, y) + \frac{\partial}{\partial y} \phi_j(x, y) \frac{\partial}{\partial y} \phi_i(x, y) \right) d\Omega \\ & = \int_{\Omega} (1) \phi_i(x, y) d\Omega - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_i(x, y) dl \quad i = 1, \dots, 4. \end{aligned}$$

The equations represented by [Eq. \(8.93\)](#) are of the form

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} \quad (8.95)$$

or

$$[A] \{u\} = \{f\} \quad (8.96)$$

where the coefficients in [Eq. \(8.95\)](#) are defined using [Eq. \(8.94\)](#) as

$$a_{ij} = - \sum_{e=1}^3 \int_{\Omega_e} \left(\frac{\partial \phi_j(x, y)}{\partial x} \frac{\partial \phi_i(x, y)}{\partial x} + \frac{\partial \phi_j(x, y)}{\partial y} \frac{\partial \phi_i(x, y)}{\partial y} \right) d\Omega_e \quad (8.97)$$

$$f_i = - \sum_{e=1}^3 \int_{\partial\Omega_e} \frac{\partial \hat{u}(x, y)}{\partial n} \phi_i(x, y) dl + \sum_{e=1}^3 \int_{\Omega_e} (1) \phi_i(x, y) d\Omega_e. \quad (8.98)$$

Note that Ω is now represented explicitly using the individual elements denoted by the index e . The basis functions are identified with the element over which the integration is being performed. For example when integrating the basis function defined for node 1 over element 3, $\phi_i(x, y)$ would appear as $\phi_1(x, y)$ defined over element 3.

As in our earlier example on page 189 we need a transformation table. Using the I local index notation introduced for triangles we have

		Global				
	1	2	3	4	Element	
	I	I+1	I+2		<i>I</i>	

Local	I+1	I	I+2	2
	I	I+1	I+2	3

Again take a moment to review this table. The global node numbers are provided in the second row and go from 1 to 4. From the third row we find that the I^{th} node for triangle 1 was selected as the global node 1. Once this decision is made, the other two local nodes, namely $I + 1$ and $I + 2$, are automatically assigned to global node numbers 2 and 3, respectively. We know this order to be the case not only because of the table, but also because we know that we must number the nodes in each element in counterclockwise order.

Element Coefficient Matrix

Because we are using triangles, as indicated in the above table, there will be three nodes for each element and therefore the element coefficient matrix will have nine entries, that is

$$\begin{bmatrix} a_{I,I} & a_{I,I+1} & a_{I,I+2} \\ a_{I+1,I} & a_{I+1,I+1} & a_{I+1,I+2} \\ a_{I+2,I} & a_{I+2,I+1} & a_{I+2,I+2} \end{bmatrix}. \quad (8.99)$$

The corresponding coefficient matrix for element 1, **using global numbering** would be

$$A_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad (8.100)$$

as can also be seen from the above table. The integral associated with coefficient a_{11} in the matrix defined by [Eq. \(8.100\)](#) is obtained from [Eq. \(8.97\)](#) using the local I numbering system:

$$- \int_{\Omega_1} \left(\frac{\partial}{\partial x} \phi_I(x, y) \frac{\partial}{\partial x} \phi_I(x, y) + \frac{\partial}{\partial y} \phi_I(x, y) \frac{\partial}{\partial y} \phi_I(x, y) \right) d\Omega. \quad (8.101)$$

The next step is to use [Eq. \(8.76\)](#), reproduced below, to obtain the derivatives needed:

$$\phi_I(x, y) = \frac{x(y_{I+1} - y_{I+2}) + y(x_{I+2} - x_{I+1}) + (x_{I+1}y_{I+2} - x_{I+2}y_{I+1})}{\det [P]} \quad x, y \in [\Omega_e] \quad (8.102)$$

from which we obtain upon differentiation

$$\frac{\partial}{\partial x} \phi_I(x, y) = \frac{(y_{I+1} - y_{I+2})}{\det [P]} \quad (8.103)$$

and

$$\frac{\partial}{\partial y} \phi_I(x, y) = \frac{(x_{I+2} - x_{I+1})}{\det [P]} \quad (8.104)$$

where

$$[P] = \begin{bmatrix} x_I & y_I & 1 \\ x_{I+1} & y_{I+1} & 1 \\ x_{I+2} & y_{I+2} & 1 \end{bmatrix}. \quad (8.105)$$

We now substitute [Eqs. \(8.103\)](#) and [\(8.104\)](#) into [\(8.101\)](#) to get

$$\begin{aligned} & - \int_{\Omega_1} \left(\frac{\partial}{\partial x} \phi_I(x, y) \frac{\partial}{\partial x} \phi_I(x, y) + \frac{\partial}{\partial y} \phi_I(x, y) \frac{\partial}{\partial y} \phi_I(x, y) \right) d\Omega \\ & = \frac{-1}{(\det [P])^2} \int_{\Omega_1} ((y_{I+1} - y_{I+2})(y_{I+1} - y_{I+2}) + (x_{I+2} - x_{I+1})(x_{I+2} - x_{I+1})) d\Omega. \end{aligned} \quad (8.106)$$

Using our global-local coordinate table, and recalling that

$$\det [P] = 2A_e \quad (8.107)$$

this term becomes

$$\begin{aligned} & - \frac{1}{(\det [P])^2} \int_{\Omega_1} ((y_{I+1} - y_{I+2})(y_{I+1} - y_{I+2}) + (x_{I+2} - x_{I+1})(x_{I+2} - x_{I+1})) d\Omega \\ & = - \frac{1}{4A_e^2} \int_{\Omega_1} ((y_2 - y_3)(y_2 - y_3) + (x_3 - x_2)(x_3 - x_2)) d\Omega \\ & = - \frac{1}{4A_e^2} ((y_2 - y_3)(y_2 - y_3) + (x_3 - x_2)(x_3 - x_2)) \int_{\Omega_1} d\Omega \\ & = - \frac{1}{4A_e} ((y_2 - y_3)(y_2 - y_3) + (x_3 - x_2)(x_3 - x_2)) \\ & = - \frac{1}{4} ((-0.76)(-0.76) + (1.31)(1.31)) \\ & = -0.573. \end{aligned} \quad (8.108)$$

As a second example of element integration, let us evaluate coefficient a_{21} of element 1.

The required integral becomes

$$- \int_{\Omega_1} \left(\frac{\partial}{\partial x} \phi_I(x, y) \frac{\partial}{\partial x} \phi_{I+1}(x, y) + \frac{\partial}{\partial y} \phi_I(x, y) \frac{\partial}{\partial y} \phi_{I+1}(x, y) \right) d\Omega. \quad (8.109)$$

From [Eqs. \(8.76\)](#) and [\(8.77\)](#) on page 198 we have

$$\frac{\partial}{\partial x} \phi_I(x, y) = \frac{(y_{I+1} - y_{I+2})}{\det [P]} \quad (8.110)$$

$$\frac{\partial}{\partial y} \phi_I(x, y) = \frac{(x_{I+2} - x_{I+1})}{\det [P]} \quad (8.111)$$

$$\frac{\partial}{\partial x} \phi_{I+1}(x, y) = \frac{(y_{I+2} - y_I)}{\det [P]} \quad (8.112)$$

$$\frac{\partial}{\partial y} \phi_{I+1}(x, y) = \frac{(x_I - x_{I+2})}{\det [P]}. \quad (8.113)$$

The required integral becomes (recalling that the area of each triangle is, for convenience, designed to be unity)

$$\begin{aligned} & - \int_{\Omega_1} \left(\frac{\partial}{\partial x} \phi_I(x, y) \frac{\partial}{\partial x} \phi_{I+1}(x, y) + \frac{\partial}{\partial y} \phi_I(x, y) \frac{\partial}{\partial y} \phi_{I+1}(x, y) \right) d\Omega \\ &= - \frac{1}{(\det [P])^2} \int_{\Omega_1} ((y_{I+1} - y_{I+2})(y_{I+2} - y_I) + (x_{I+2} - x_{I+1})(x_I - x_{I+2})) d\Omega \\ &= - \frac{1}{(\det [P])^2} ((y_{I+1} - y_{I+2})(y_{I+2} - y_I) + (x_{I+2} - x_{I+1})(x_I - x_{I+2})) \int_{\Omega_1} d\Omega \\ &= - \frac{1}{4A_1} ((y_{I+1} - y_{I+2})(y_{I+2} - y_I) + (x_{I+2} - x_{I+1})(x_I - x_{I+2})) \\ &= - \frac{1}{4A_1} ((y_2 - y_3)(y_3 - y_1) + (x_3 - x_2)(x_1 - x_3)) \\ &= - \frac{1}{4} ((-0.76)(-1.51) + (1.31)(0)) \\ &= -0.286. \end{aligned} \quad (8.114)$$

Global Coefficient Matrix

When calculations such as shown above are completed for each of the nodes and each of the elements, the following element coefficient matrices are generated

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}}_{\text{element 1}} = - \begin{bmatrix} 0.573 & 0.286 & -0.85 \\ 0.286 & 0.57 & -0.856 \\ -0.85 & -0.856 & 1.71 \end{bmatrix} \quad (8.115)$$

$$\underbrace{\begin{bmatrix} a_{33} & a_{32} & a_{34} \\ a_{23} & a_{22} & a_{24} \\ a_{43} & a_{42} & a_{44} \end{bmatrix}}_{\text{element 2}} = - \begin{bmatrix} 1.73 & -0.868 & -0.868 \\ -0.868 & 0.573 & 0.286 \\ -0.868 & 0.286 & 0.573 \end{bmatrix} \quad (8.116)$$

$$\underbrace{\begin{bmatrix} a_{11} & a_{13} & a_{14} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{bmatrix}}_{\text{element 3}} = - \begin{bmatrix} 0.573 & -0.849 & 0.276 \\ -0.849 & 1.71 & -0.868 \\ 0.276 & -0.868 & 0.573 \end{bmatrix}. \quad (8.117)$$

The next step is to assemble the global matrix by taking the coefficients for each node from all the elements and adding them together as illustrated below

$$\begin{aligned} & \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \\ &= - \begin{bmatrix} 0.573 + 0.573 & 0.286 & -0.859 - 0.849 & 0.276 \\ 0.286 & 0.570 + 0.573 & -0.856 - 0.868 & 0.286 \\ -0.859 - 0.849 & -0.856 - 0.868 & 1.71 + 1.73 + 1.71 & -0.868 - 0.868 \\ 0.276 & 0.286 & -0.868 - 0.868 & 0.573 + 0.573 \end{bmatrix} \\ &= - \begin{bmatrix} 1.14 & 0.286 & -1.70 & 0.276 \\ 0.286 & 1.14 & -1.72 & 0.286 \\ -1.70 & -1.72 & 5.16 & -1.74 \\ 0.276 & 0.286 & -1.74 & 1.15 \end{bmatrix}. \end{aligned} \quad (8.118)$$

Before we can solve the problem, we need to evaluate the forcing function

$$\int_{\Omega} (1) \phi_i(x, y) d\Omega - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_i(x, y) dl \quad (8.119)$$

which is the same as

$$\begin{aligned} & \int_{\Omega} \phi_i(x, y) d\Omega - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_i(x, y) dl \\ &= \sum_{e=1}^3 \int_{\Omega_e} \phi_i^e(x, y) d\Omega - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_i(x, y) dl. \end{aligned} \tag{8.120}$$

From [Eq. \(8.85\)](#) we have

$$\int_{\Omega_e} \phi_i^e(x, y) d\Omega = \frac{A_e}{3} \tag{8.121}$$

so we can rewrite [Eq. \(8.120\)](#) for each node (again keeping in mind that the area of each element in this problem is unity).

$$\begin{aligned} f_1 &= \sum_{e=1}^3 \int_{\Omega_e} \phi_1^e(x, y) d\Omega - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_1(x, y) dl \\ &= \frac{A_1}{3} + \frac{A_3}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_1(x, y) dl \\ &= \frac{2}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_1(x, y) dl \end{aligned} \tag{8.122}$$

$$\begin{aligned} f_2 &= \sum_{e=1}^3 \int_{\Omega_e} \phi_2^e(x, y) d\Omega - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_2(x, y) dl \\ &= \frac{A_1}{3} + \frac{A_2}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_2(x, y) dl \\ &= \frac{2}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_2(x, y) dl \end{aligned} \tag{8.123}$$

$$\begin{aligned}
f_3 &= \sum_{e=1}^3 \int_{\Omega_e} \phi_3^e(x, y) d\Omega - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_3(x, y) dl & (8.124) \\
&= \frac{A_1}{3} + \frac{A_2}{3} + \frac{A_3}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_3(x, y) dl \\
&= \frac{3}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_3(x, y) dl.
\end{aligned}$$

$$\begin{aligned}
f_4 &= \sum_{e=1}^3 \int_{\Omega_e} \phi_4^e(x, y) d\Omega - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_4(x, y) dl & (8.125) \\
&= \frac{A_2}{3} + \frac{A_3}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_4(x, y) dl \\
&= \frac{2}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_4(x, y) dl.
\end{aligned}$$

Thus, the right-hand side vector becomes

$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_1(x, y) dl \\ \frac{2}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_2(x, y) dl \\ \frac{3}{3} + 0 \\ \frac{2}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_4(x, y) dl \end{bmatrix}. \quad (8.126)$$

You may be wondering what we will do with the normal derivative appearing on the right-hand side of [Eq. \(8.126\)](#). As you will see, if we need it, it is provided as a boundary condition. If we do not have a derivative-type boundary condition, this term will not be needed.

Solution of Equations

The matrix equation now becomes

(8.127)

$$- \begin{bmatrix} 1.14 & 0.286 & -1.70 & 0.276 \\ 0.286 & 1.14 & -1.72 & 0.286 \\ -1.70 & -1.72 & 5.16 & -1.74 \\ 0.276 & 0.286 & -1.74 & 1.15 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_1(x, y) dl \\ \frac{2}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_2(x, y) dl \\ \frac{2}{3} + 0 \\ \frac{2}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_4(x, y) dl \end{bmatrix}.$$

Application of boundary conditions provides the following:

(8.128)

$$- \begin{bmatrix} 1.14 & 0.286 & -1.70 & 0.276 \\ 0.286 & 1.14 & -1.72 & 0.286 \\ -1.70 & -1.72 & 5.16 & -1.74 \\ 0.276 & 0.286 & -1.74 & 1.15 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ u_3 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_1(x, y) dl \\ \frac{2}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_2(x, y) dl \\ \frac{2}{3} + 0 \\ \frac{2}{3} - \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_4(x, y) dl \end{bmatrix}$$

which can be reduced to

$$- (5.16) u_3 = [(-1.70)(0) - (1.72)(1) - (1.74)(0)] + 1 \quad (8.129)$$

or

$$u_3 = \frac{-0.72}{-5.16} = 0.14. \quad (8.130)$$

Notice that the equations containing the derivative on the right-hand side have been eliminated by virtue of the specification of the function value of $u(x, y)$. Thus, because we did not have any Neumann boundary conditions, the equations containing this information disappeared by virtue of the alternative specification of the Dirichlet condition. We will see in the next subsection what to do when there is a Neumann condition specified. The complete approximation for $u(x, y)$ for this problem becomes

(8.131)

$$\begin{aligned} \hat{u}(x, y) &= \sum_{j=1}^4 u_j \phi_j(x, y) \\ &= (0) \phi_1(x, y) + (1) \phi_2(x, y) + (0.14) \phi_3(x, y) + (0) \phi_4(x, y). \end{aligned}$$

8.2.3 Second Type or Neumann Boundary-Value Problem

In this section we will modify our earlier problem presented in [Section 8.2.2](#) to consider a Neumann boundary condition. The problem definition is provided below:

$$\frac{\partial^2}{\partial x^2} u(x, y) + \frac{\partial^2}{\partial y^2} u(x, y) = 1 \quad (8.132)$$

where (see [Fig. 8.13](#))

$$\frac{\partial u}{\partial n}(x, y) = 1 \quad x, y \in \partial\Omega_1 \quad (8.133)$$

$$u(x, y) = 0 \quad x, y \in \partial\Omega_2 \quad (8.134)$$

$$u(x, y) = 0 \quad x, y \in \partial\Omega_3 \quad (8.135)$$

where we have assumed that nodes 1 and 2 are associated with $\partial\Omega_1$. The only changes we need to make to our earlier formulation is to modify the right-hand side to accommodate the new boundary conditions. The new boundary specifications are shown in [Fig. 8.14](#). We first write the earlier formulation

$$- \begin{bmatrix} 1.14 & 0.286 & -1.70 & 0.276 \\ 0.286 & 1.14 & -1.72 & 0.286 \\ -1.70 & -1.72 & 5.16 & -1.74 \\ 0.276 & 0.286 & -1.74 & 1.15 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} - \int_{\partial\Omega} \Omega \frac{\partial u}{\partial n} \phi_1(x, y) dl \\ \frac{2}{3} - \int_{\partial\Omega} \Omega \frac{\partial u}{\partial n} \phi_2(x, y) dl \\ \frac{3}{3} + 0 \\ \frac{2}{3} - \int_{\partial\Omega} \Omega \frac{\partial u}{\partial n} \phi_4(x, y) dl \end{bmatrix} \quad (8.136)$$

The terms of interest are in the boxes in [Eq. \(8.136\)](#).

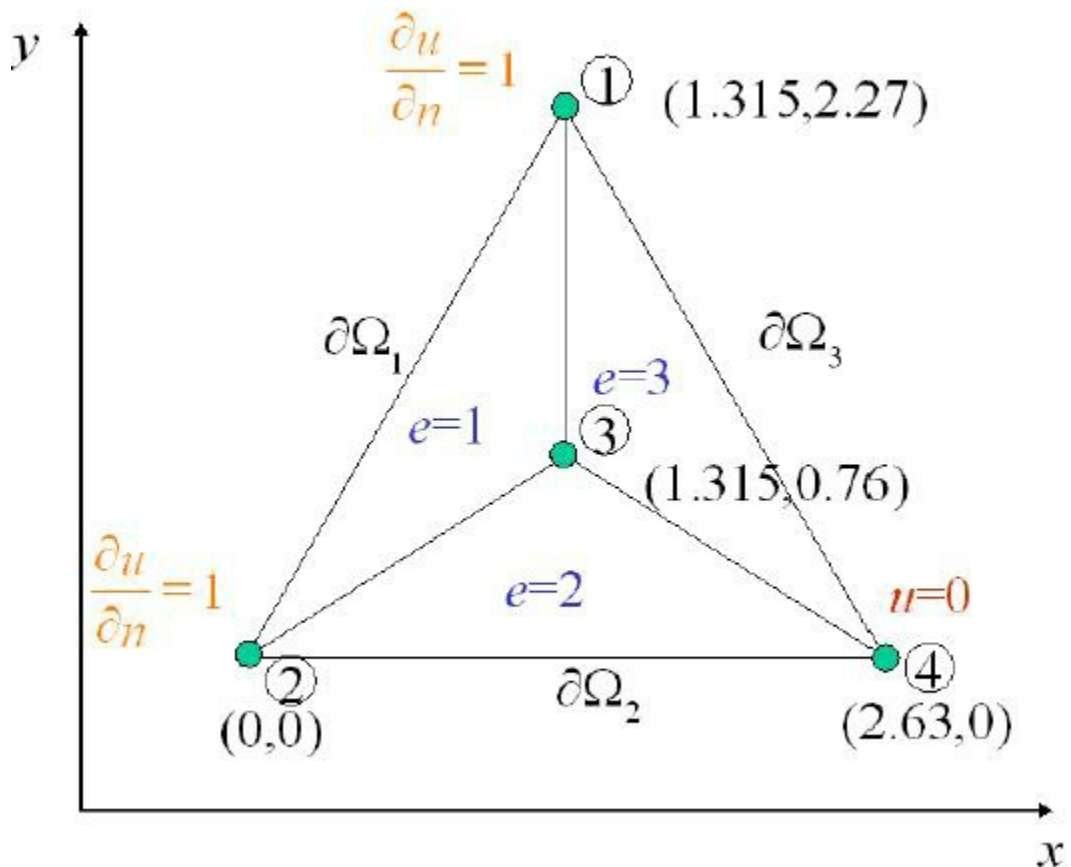


Figure 8.14: Repeated for reference, illustration of triangle used as example in this

subsection, that is [Section 8.2.2](#).

Consider the term associated with the first row, that is the equation identified with node 1 and therefore weighting function 1. The first thing to note is that the triangular basis function shown in [Fig. 8.11](#) is a linear Lagrange polynomial along the side of the triangle. Thus the following holds

$$\int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_1(x, y) dl = \int_{\partial\Omega_1} \frac{\partial \hat{u}}{\partial n} \phi_1(x, y) dl. \quad (8.137)$$

If we assume $\frac{\partial \hat{u}}{\partial n}$ constant along the boundary segment $\partial\Omega_1$, which is the normal procedure in linear triangular elements, it can be taken out of the integral and we obtain

$$\int_{\partial\Omega_1} \frac{\partial \hat{u}}{\partial n} \phi_1(x, y) dl = \frac{\partial \hat{u}}{\partial n} \int_{\partial\Omega_1} \phi_1(x, y) dl. \quad (8.138)$$

The function $\phi_1(x, y)$, which is defined only along the boundary $\partial\Omega_1$ of the triangle, will be a linear Lagrange polynomial. We use the symbol σ to denote distance along this side. In other words σ is the one-dimensional coordinate defined along the line $\partial\Omega_1$. Using this notation we obtain

$$\begin{aligned} & \frac{\partial \hat{u}}{\partial n} \int_{\partial\Omega_1} \phi_1(\sigma) dl \\ &= \frac{\partial \hat{u}}{\partial n} \int_{\partial\Omega_1} \frac{\sigma - \sigma_2}{\sigma_1 - \sigma_2} dl \end{aligned} \quad (8.139)$$

where σ_1 and σ_2 are the locations of nodes 1 and 2 in the σ coordinate system. In [Fig. 8.11](#) we see, for example, that σ_1 and σ_2 could correspond to nodes I and $I + 1$ and the integration would be performed along the line connecting them. Integration now yields

$$\begin{aligned}
& \frac{\partial \hat{u}}{\partial n} \int_{\partial \Omega_1} \frac{\sigma_2 - \sigma}{\sigma_2 - \sigma_1} dl \\
&= \frac{\partial \hat{u}}{\partial n} \left(\frac{\sigma_2 \sigma - \sigma^2/2}{\sigma_2 - \sigma_1} \right) \Big|_{\sigma_1}^{\sigma_2} \\
&= \frac{\partial \hat{u}}{\partial n} \left(\frac{2\sigma_2 \sigma - \sigma^2}{2(\sigma_2 - \sigma_1)} \right) \Big|_{\sigma_1}^{\sigma_2} \\
&= \frac{\partial \hat{u}}{\partial n} \left(\frac{2(\sigma_2)^2 - \sigma_2^2}{2(\sigma_2 - \sigma_1)} \right) - \left(\frac{2(\sigma_2 \sigma_1) - \sigma_1^2}{2(\sigma_2 - \sigma_1)} \right) \\
&= \frac{\partial \hat{u}}{\partial n} \left(\frac{(\sigma_2)^2 - 2(\sigma_2 \sigma_1) + \sigma_1^2}{2(\sigma_2 - \sigma_1)} \right) \\
&= \frac{\partial \hat{u}}{\partial n} \left(\frac{(\sigma_2 - \sigma_1)^2}{2(\sigma_2 - \sigma_1)} \right) \\
&= \frac{\partial \hat{u}}{\partial n} \left(\frac{(\sigma_2 - \sigma_1)}{2} \right).
\end{aligned}$$

The interpretation of [Eq. \(8.140\)](#) is that the the portion of the contribution of $\frac{\partial \hat{u}}{\partial n}$ associated with node 1 is $\frac{\partial \hat{u}}{\partial n} \frac{\sigma_2}{2}$. Thus the integral in the first row on the right-hand-side of [Eq. \(8.136\)](#) will take on the value presented in the last line of [Eq. 8.140](#). The coordinate σ_2 is determined by inspection from [Fig. 8.14](#) to be 2.63 length units (see the length of the side of the triangle along x) and $\frac{\partial \hat{u}}{\partial n} = 1$. Thus we have

$$\frac{\partial \hat{u}}{\partial n} \frac{\sigma_2}{2} = (1) \frac{2.63}{2} = 1.32. \quad (8.141)$$

A similar argument can be used to obtain $\int_{\partial \Omega_1} \frac{\partial \hat{u}}{\partial n} \phi_2(x, y) dl$ and this turns out to be the same value, namely 1.32.

We are now in a position to modify our matrix to accommodate the boundary conditions. We obtain

(8.142)

$$- \begin{bmatrix} 1.14 & 0.286 & -1.70 & 0.276 \\ 0.286 & 1.14 & -1.72 & 0.286 \\ -1.70 & -1.72 & 5.16 & -1.74 \\ 0.276 & 0.286 & -1.74 & 1.15 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ 0 \end{bmatrix} = \begin{bmatrix} -1.32 \\ -1.32 \\ 0 \\ \frac{2}{3} - \int_{\partial\Omega} \frac{\partial u}{\partial n} \phi_4(x, y) dl \end{bmatrix}.$$

You will notice that the integral in the last row remains unevaluated. We will see in the next step that the reason for this is that it will not be needed. Using the information from the Dirichlet boundary condition that $u_4 = 0$, as we implicitly did in the last example [Section 8.2.2](#), we can multiply this value by the last column of the matrix, put the resulting information on the right-hand side of the matrix equation, and eliminate the last row and last column of the matrix. Note that we now have three unknowns since nodes 1 and 2 are not specified $u(x, y)$ values. Simplifying [Eq. \(8.142\)](#) we have

$$- \begin{bmatrix} 1.14 & 0.286 & -1.70 \\ 0.286 & 1.14 & -1.72 \\ -1.70 & -1.72 & 5.16 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} - 1.32 + (0.276)(0) \\ \frac{2}{3} - 1.32 + (0.286)(0) \\ \frac{3}{3} + 0 - (1.74)(0) \end{bmatrix} \quad (8.143)$$

or

$$- \begin{bmatrix} 1.14 & 0.286 & -1.70 \\ 0.286 & 1.14 & -1.72 \\ -1.70 & -1.72 & 5.16 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} -0.654 \\ -0.654 \\ 1 \end{bmatrix}. \quad (8.144)$$

Taking the inverse of the coefficient matrix, that is

$$\begin{bmatrix} 1.14 & 0.286 & -1.70 \\ 0.286 & 1.14 & -1.72 \\ -1.70 & -1.72 & 5.16 \end{bmatrix}^{-1} = \begin{bmatrix} 2.25 & 1.10 & 1.11 \\ 1.10 & 2.30 & 1.13 \\ 1.11 & 1.13 & 0.934 \end{bmatrix} \quad (8.145)$$

we can solve [Eq. \(8.144\)](#) directly as follows:

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = - \begin{bmatrix} 2.25 & 1.10 & 1.11 \\ 1.10 & 2.30 & 1.13 \\ 1.11 & 1.13 & 0.934 \end{bmatrix} \begin{bmatrix} -0.654 \\ -0.654 \\ 1 \end{bmatrix} \\ = \begin{bmatrix} 1.09 \\ 1.09 \\ 0.53 \end{bmatrix}. \quad (8.146)$$

Thus the solution to our problem is $[u_1, u_2, u_3]^T = [1.09, 1.09, 0.53]^T$ where T denotes, transpose.

To understand the results, one must remember that the vector \mathbf{n} associated with $\frac{\partial u}{\partial n}$ is

outward directed. Thus, if u is a state variable, such as mass, there is mass attempting to enter the system along-side $\partial\Omega_1$. Thus we would expect that the values of u near $\partial\Omega_1$ would be higher than further away, say, at node 3. However, we also have a flux $Q = 1$ exiting the system that must be satisfied, in part through inward flow along the boundary. Node 4 must remain at zero because that is the fixed boundary condition value for u at that node.

If we assume that $Q = 0$, that is there is no source term in the equation, we get

$$-\begin{bmatrix} 1.14 & 0.286 & -1.70 \\ 0.286 & 1.14 & -1.72 \\ -1.70 & -1.72 & 5.16 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 - 1.32 + (0.276)(0) \\ 0 - 1.32 + (0.286)(0) \\ 0 + 0 - (1.73)(0) \end{bmatrix} \quad (8.147)$$

or

$$-\begin{bmatrix} 1.14 & 0.29 & -1.70 \\ 0.29 & 1.14 & -1.72 \\ -1.70 & -1.72 & 5.16 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} -1.32 \\ -1.32 \\ 0 \end{bmatrix} \quad (8.148)$$

or

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = -\begin{bmatrix} 2.25 & 1.10 & 1.11 \\ 1.10 & 2.30 & 1.13 \\ 1.11 & 1.13 & 0.934 \end{bmatrix} \begin{bmatrix} -1.32 \\ -1.32 \\ 0 \end{bmatrix} \quad (8.149)$$

or

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 4.42 \\ 4.48 \\ 2.95 \end{bmatrix}. \quad (8.150)$$

One now observes that the slope of the surface has changed. The high side along $\partial\Omega_1$ is even higher since this is the side from which the mass flows inwards. If the domain were rectangular, we would expect a uniform slope across the domain from the left to the right. However, the domain is triangular and so the surface is not perfectly flat, but it is close.

As a variant on this example, let us see what happens if we again remove the uniform flux, that is let $Q = 0$, but also reverse the sense of the flux boundary condition, that is set

$$\frac{\partial}{\partial n} u(x, y) = -1 \quad x, y \in \partial\Omega_1 \quad (8.151)$$

$$u(x, y) = 0 \quad x, y \in \partial\Omega_2 \quad (8.152)$$

$$u(x, y) = 0 \quad x, y \in \partial\Omega_3. \quad (8.153)$$

The matrix equation becomes

$$-\begin{bmatrix} 1.14 & 0.286 & -1.70 \\ 0.286 & 1.14 & -1.72 \\ -1.70 & -1.72 & 5.16 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 + 1.32 + (0.276)(0) \\ 0 + 1.32 + (0.286)(0) \\ 0 - (1.73)(0) \end{bmatrix} \quad (8.154)$$

for which the solution is

$$\begin{aligned} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} &= -\begin{bmatrix} 2.25 & 1.10 & 1.11 \\ 1.10 & 2.30 & 1.13 \\ 1.11 & 1.13 & 0.934 \end{bmatrix} \begin{bmatrix} 1.32 \\ 1.32 \\ 0 \end{bmatrix} \\ &= -\begin{bmatrix} 4.42 \\ 4.49 \\ 2.96 \end{bmatrix} \end{aligned} \quad (8.155)$$

or $(u_1, u_2, u_3)^T = -(4.42, 4.49, 2.96)^T$. One now observes that the slope of the surface has changed. The low side is along $\partial\Omega_1$ since this is the side from which the mass flows outwards. If the domain were rectangular, we would expect a uniform positive slope across the domain from the left to the right.

8.3 Isoparametric Finite Element Approximation

In this section we will introduce the concept of isoparametric finite elements. These elements are characteristically irregular in their geometry. They are essentially deformed rectangles. However the concept can be further extended to allow for sides that have curvature. They are useful in some specific engineering applications. A necessary tool for understanding this concept is the idea of a 'natural coordinate system' which is a special form of local coordinate system and which we will now introduce.

8.3.1 Natural Coordinate Systems

One Dimensional Natural Coordinate Systems

A natural coordinate system is a local coordinate system (see page 16) which is special in that it permits the specification of a point within an element by a *set of dimensionless numbers whose magnitudes never exceed unity*. Consider the information presented in [Fig. 8.15](#). The variable x is a global coordinate and L is a **local coordinate** which is also a **natural coordinate**. Notice that any point along the indicated interval $[x_1, x_2]$ can be specified by a number whose absolute magnitude does not exceed unity in the L coordinate system.

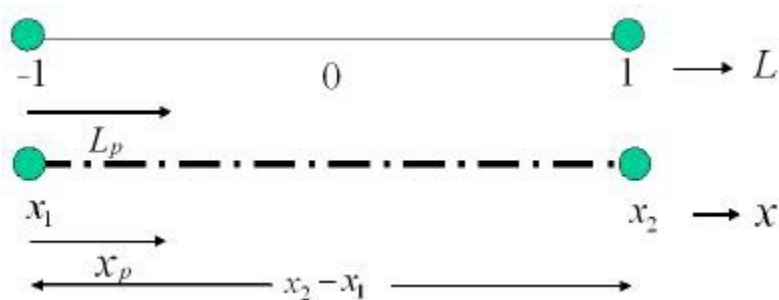


Figure 8.15: Illustration of the natural coordinate system, L .

Our first task is to determine how to relate the natural and local coordinate systems one to the other. From [Fig. 8.15](#) we observe the following: a point x_p somewhere within the interval (x_1, x_2) is related to a point L_p located in the interval $(-1, 1)$ by the expression

$$\begin{aligned} \frac{x_p}{x_2 - x_1} &= \frac{L_p}{1 - (-1)} \\ &= \frac{L_p}{2}. \end{aligned} \quad (8.156)$$

In general, for any x , we have

$$\frac{x - x_1}{x_2 - x_1} = \frac{L - (-1)}{2}. \quad (8.157)$$

We can solve for x in [Eq. \(8.157\)](#) and obtain

$$x = x_1 + \frac{L + 1}{2} (x_2 - x_1). \quad (8.158)$$

A little algebra gives

$$x = \frac{1}{2} [2x_1 - Lx_1 - x_1 + (L + 1)x_2]. \quad (8.159)$$

We now rearrange the terms in this expression with malice of forethought to give

$$x = \frac{1}{2} (1 - L)x_1 + \frac{1}{2} (1 + L)x_2. \quad (8.160)$$

[Eq. \(8.160\)](#) can also be written as

$$x = \sum_{i=1}^2 x_i \phi_i(L) \quad (8.161)$$

where

$$\phi_1(L) = \frac{1-L}{2} \quad (8.162)$$

$$\phi_2(L) = \frac{1+L}{2}. \quad (8.163)$$

Let us now look more carefully at the function $\phi_i(L)$. We observe the following:

$$\phi_1(L = -1) = 1 \quad (8.164)$$

$$\phi_1(L = 1) = 0 \quad (8.165)$$

$$\phi_2(L = -1) = 0 \quad (8.166)$$

$$\phi_2(L = 1) = 1. \quad (8.167)$$

It appears that the values of $\phi_i(L)$ satisfy the requirements of basis functions for the interpolation of the coordinate x in terms of the end points of the interval (we did something similar in Section 1.5 on page 4)!

Two-Dimensional Natural Coordinate Systems

The one-dimensional transformation just presented was rather straight-forward. However, the two-dimensional version is a little more complicated. The local and global elements are shown in [Figs. 8.16](#) and [Fig. 8.17](#), respectively. Now let us examine the relationship between them. To be a one-to-one transformation we will require the following:

Node Number	→	Local Coordinate
1	→	(-1, -1)
2	→	(1, -1)
3	→	(1, 1)
4	→	(-1, 1)

(8.168)

where the symbol \rightarrow is to be interpreted as ‘corresponds to.’ Building upon what we learned from our one-dimensional example, namely [Eq. \(8.161\)](#), we want to be able to determine the coordinates of x and y within the global element using a relationship of the form:

$$x = \sum_{i=1}^4 x_i \phi_i(\xi, \eta) \quad (8.169)$$

$$y = \sum_{i=1}^4 y_i \phi_i(\xi, \eta).$$

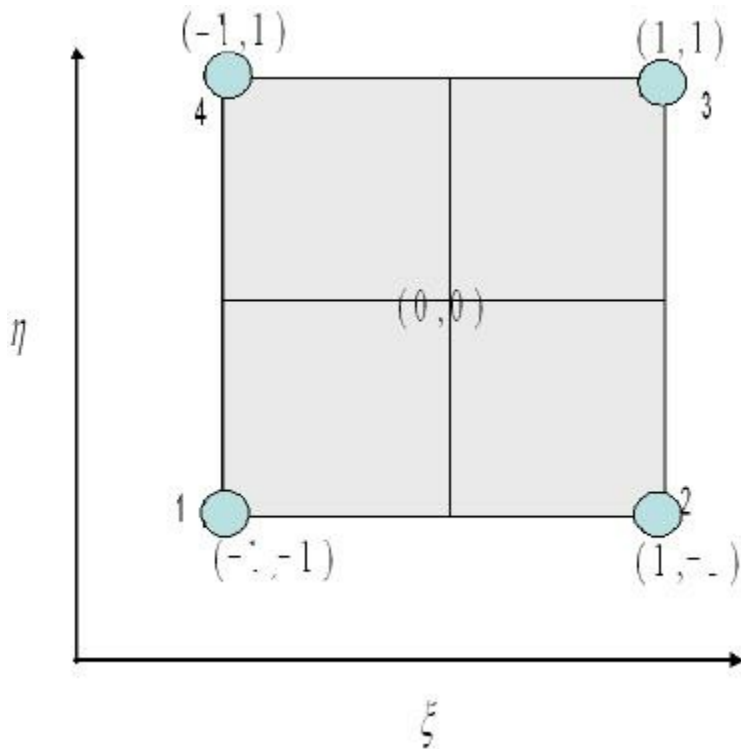


Figure 8.16: Local element in local $\xi - \eta$ coordinate system.

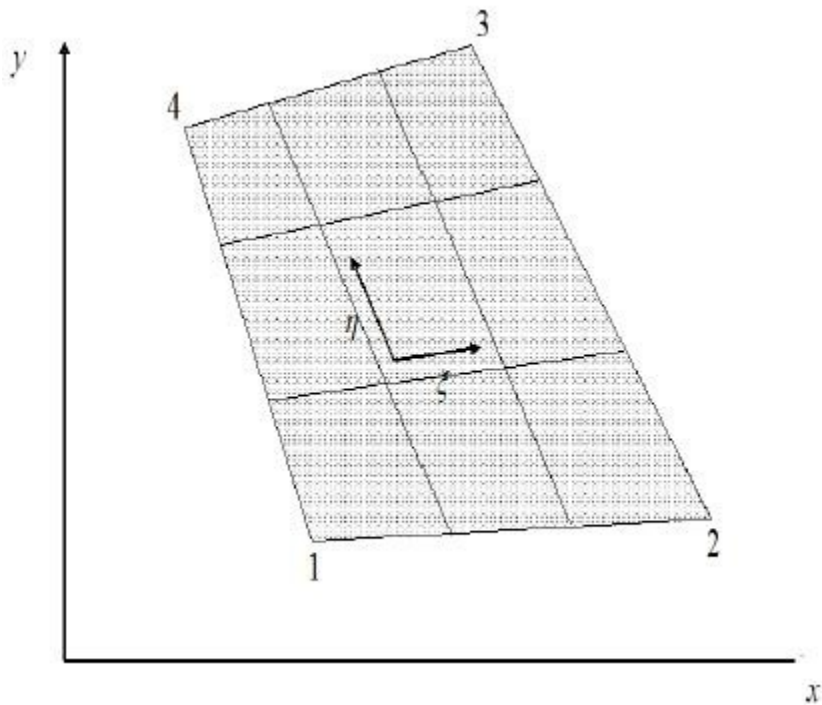


Figure 8.17: Isoparametric element defined in local (ξ, η) and global (x, y) coordinates.

We will return to this relationship shortly. However, for now let us consider a different strategy for representing x .

Let us impose the constraint that a line that forms an element side that is a straight line in the global coordinate system is also a straight line in the local coordinate system. If this is to be true, we need to be able to have x vary linearly along ξ and y vary linearly along η . A function of ξ and η that varies linearly along ξ will have the form

$$x = \alpha_1 + \alpha_2\xi + \alpha_3\eta + \alpha_4\xi\eta. \quad (8.171)$$

To see that this equation satisfies the requirement of being linear along ξ , substitute $\eta = 1$. This gives

$$x = \alpha_1 + \alpha_2\xi + \alpha_3 + \alpha_4\xi. \quad (8.172)$$

which is a linear function of ξ . Similarly if we substitute $\eta = -1$ we have

$$x = \alpha_1 + \alpha_2\xi - \alpha_3 - \alpha_4\xi. \quad (8.173)$$

Both [Eqs. \(8.172\)](#) and [\(8.173\)](#) provide a linear change in x with a linear change in ξ , thus satisfying the requirement that an element side that is linear in the global system is linear in the local system.

One can, using a similar argument, show that the appropriate form of the interpolator for y is

$$y = \beta_1 + \beta_2\xi + \beta_3\eta + \beta_4\xi\eta. \quad (8.174)$$

It can be shown that when ξ is ± 1 , y is a linear function of η .

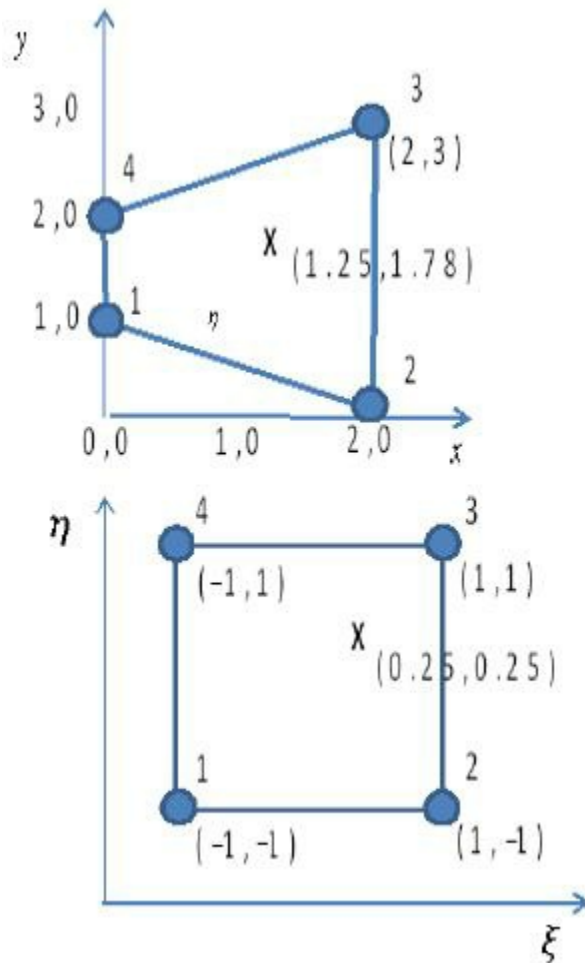
We would like to find out what the values of α and β are. To achieve this we impose on [Eqs. \(8.173\)](#) and [\(8.174\)](#) the requirement of one-to-one correspondence between the nodal locations in the global and local coordinate systems as shown in the above table. We obtain by substituting the corresponding values of x , ξ , and η into [Eq. \(8.173\)](#) the following matrix equation for the unknown values of $a_i, i = 1, \dots, 4$:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & \xi_1 & \eta_1 & \xi_1\eta_1 \\ 1 & \xi_2 & \eta_2 & \xi_2\eta_2 \\ 1 & \xi_3 & \eta_3 & \xi_3\eta_3 \\ 1 & \xi_4 & \eta_4 & \xi_4\eta_4 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} \quad (8.175)$$

Substitution of the local coordinate values for ξ and η we yields

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}. \quad (8.176)$$

Inverting the coefficient matrix provides the solution for $\{\alpha\}^T$, that is,



$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}. \quad (8.177)$$

Recall from [Eq. \(8.169\)](#) that we can express x as a function of the nodal coordinates via the expression

$$x = [\phi_1(\xi, \eta), \phi_2(\xi, \eta), \phi_3(\xi, \eta), \phi_4(\xi, \eta)] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \quad (8.178)$$

and from [Eq. \(8.171\)](#)

$$x = [1, \xi, \eta, \xi\eta] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}. \quad (8.179)$$

But, we have another representation of the α vector from [Eq. \(8.177\)](#), so we can substitute that relationship into [Eq. \(8.179\)](#) to obtain

$$x = \frac{1}{4} [1, \xi, \eta, \xi\eta] \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}. \quad (8.180)$$

We now examine [Eqs. \(8.178\)](#) and [\(8.180\)](#) and see that both of these equations describe the behavior of x . Thus we can equate their right-hand-sides to give the relationship

$$[\phi_1(\xi, \eta), \phi_2(\xi, \eta), \phi_3(\xi, \eta), \phi_4(\xi, \eta)] = \frac{1}{4} [1, \xi, \eta, \xi\eta] \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}. \quad (8.181)$$

To obtain further insight into this relationship, consider the functions $\varphi_i(\xi, \eta)$. We have from [Eq. \(8.181\)](#)

$$\begin{aligned} \phi_1(\xi, \eta) &= \frac{1}{4} [(1)(1) - (1)(\xi) - (1)(\eta) + \xi\eta] & (8.182) \\ &= \frac{1}{4} (1 - \xi - \eta + \xi\eta) \\ &= \left(\frac{1 - \xi}{2}\right) \left(\frac{1 - \eta}{2}\right) \end{aligned}$$

$$\begin{aligned} \phi_2(\xi, \eta) &= \frac{1}{4} (1 + \xi - \eta - \xi\eta) & (8.183) \\ &= \left(\frac{1 + \xi}{2}\right) \left(\frac{1 - \eta}{2}\right) \end{aligned}$$

$$\begin{aligned}\phi_3(\xi, \eta) &= \frac{1}{4}(1 + \xi + \eta + \xi\eta) \\ &= \left(\frac{1 + \xi}{2}\right) \left(\frac{1 + \eta}{2}\right)\end{aligned}\quad (8.184)$$

$$\begin{aligned}\phi_4(\xi, \eta) &= \frac{1}{4}[1 - \xi + \eta - \xi\eta] \\ &= \left(\frac{1 - \xi}{2}\right) \left(\frac{1 + \eta}{2}\right).\end{aligned}\quad (8.185)$$

The above functions are called **Coordinate Transformation Functions**.

Example Problem

Consider the point located with the X in the lower panel of [Fig. 8.18](#). The goal is to find the global coordinate location for this point. The first step is to write the general form of the global-local transformation. One obtains global coordinates in terms of the global coordinate node locations and the local coordinate transformation functions as

$$x = \sum_{i=1}^4 x_i \phi_i(\xi, \eta) \quad (8.186)$$

$$y = \sum_{i=1}^4 y_i \phi_i(\xi, \eta). \quad (8.187)$$

Next substitute for the x_i and y_i values to get

$$x = 0[\phi_1(\xi, \eta)] + 2[\phi_2(\xi, \eta)] + 2[\phi_3(\xi, \eta)] + 0[\phi_4(\xi, \eta)] \quad (8.188)$$

which gives, using [Eqs. \(8.182\)–\(8.185\)](#)

$$\begin{aligned}x|_{\xi=0.25, \eta=0.25} &= 2 \left[\frac{(1 + 0.25)(1 - 0.25)}{4} \right] + 2 \left[\frac{(1 + 0.25)(1 + 0.25)}{4} \right] \\ &= 0.468 + 0.781 = 1.25.\end{aligned}\quad (8.189)$$

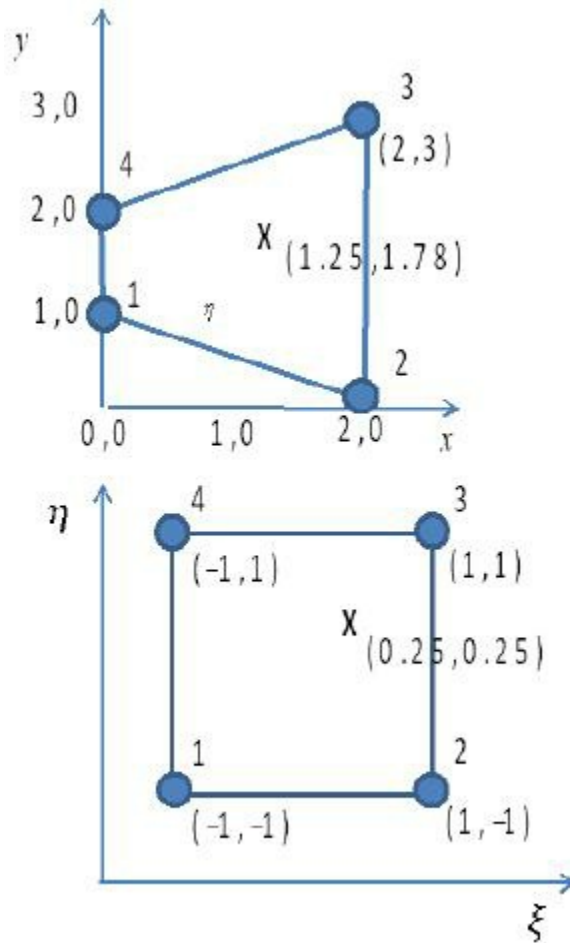


Figure 8.18: Element in local and global coordinates used for example problem. The ξ in the lower panel maps to the X in the upper panel via [Eqs. 8.186](#) and [8.187](#).

Similarly for the y_1 coordinate location

(8.190)

$$y = 1 [\phi_1 (\xi, \eta)] + 0 [\phi_2 (\xi, \eta)] + 3 [\phi_3 (\xi, \eta)] + 2 [\phi_4 (\xi, \eta)]$$

(8.191)

$$y|_{\xi=0.25, \eta=0.25} = 1 \left[\frac{(1 - 0.25)(1 - 0.25)}{4} \right] + 3 \left[\frac{(1 + 0.25)(1 + 0.25)}{4} \right] + 2 \left[\frac{(1 - 0.25) \times (1 + 0.25)}{4} \right]$$

$$= 0.141 + 1.17 + 0.469 = 1.78.$$

So, the calculations indicate that the location $(\xi, \eta) = (0.25, 0.25)$ maps to $(x, y) = (1.25, 1.78)$, as shown in the upper panel of [Fig. 8.18](#). Now that we have the coordinate transformation functions, let us consider the basis functions.

8.3.2 Basis Functions

As usual, we are seeking to represent our approximating function $\hat{u}(x, y)$ by a series employing the product of coefficients and basis functions. Consider for example the approximation of $u(x, y)$ over one element

$$\hat{u}(x, y) = \sum_{k=1}^4 u_k \psi_k(\xi, \eta) \quad (8.192)$$

where $\psi_k(\xi, \eta)$ are the basis functions. To ensure that a function is approximated linearly along a linear side (see [Eq. \(8.172\)](#)) we consider the relationship

$$\psi_i = \alpha_{1i} + \alpha_{2i}\xi + \alpha_{3i}\eta + \alpha_{4i}\xi\eta \quad i = 1, 2, 3, 4. \quad (8.193)$$

Consider as an example the conditions that apply to basis functions in general and $\psi_1(\xi, \eta)$ in particular, namely that

$$\psi_1(\xi_1, \eta_1) = 1 \quad (8.194)$$

and

$$\psi_1(\xi_p, \eta_p) = 0 \quad p = 2, 3, 4. \quad (8.195)$$

Combination of these constraints with the definition provided in [Eq. \(8.193\)](#) yields, for $\psi_1(\xi, \eta)$,

$$\begin{bmatrix} \psi_1(\xi_1, \eta_1) \\ \psi_1(\xi_2, \eta_2) \\ \psi_1(\xi_3, \eta_3) \\ \psi_1(\xi_4, \eta_4) \end{bmatrix} = \begin{bmatrix} 1 & \xi_1 & \eta_1 & \xi_1\eta_1 \\ 1 & \xi_2 & \eta_2 & \xi_2\eta_2 \\ 1 & \xi_3 & \eta_3 & \xi_3\eta_3 \\ 1 & \xi_4 & \eta_4 & \xi_4\eta_4 \end{bmatrix} \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \\ \alpha_{41} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (8.196)$$

or

$$\begin{bmatrix} \psi_1(\xi_1, \eta_1) \\ \psi_1(\xi_2, \eta_2) \\ \psi_1(\xi_3, \eta_3) \\ \psi_1(\xi_4, \eta_4) \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \\ \alpha_{41} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (8.197)$$

Solving for the α vector we obtain

$$\begin{aligned} \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \\ \alpha_{41} \end{bmatrix} &= \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{4} \\ -\frac{1}{4} \\ -\frac{1}{4} \\ \frac{1}{4} \end{bmatrix}. \end{aligned} \quad (8.198)$$

From [Eq. \(8.193\)](#) we have for $\psi_1(\xi, \eta)$

$$\psi_1 = \alpha_{11} + \alpha_{21}\xi + \alpha_{31}\eta + \alpha_{41}\xi\eta \quad (8.199)$$

or using [Eq. \(8.198\)](#)

$$\psi_1 = \frac{1}{4} - \frac{1}{4}\xi - \frac{1}{4}\eta + \frac{1}{4}\xi\eta. \quad (8.200)$$

In order to demonstrate a concept, we now rearrange this equation to give

$$\begin{aligned} \psi_1 &= \frac{1}{4} (1 - \xi - \eta + \xi\eta) \\ &= \frac{1}{2} (1 - \xi) \frac{1}{2} (1 - \eta). \end{aligned} \quad (8.201)$$

Thus we see by comparison of [Eq. \(8.201\)](#) to [Eq. \(8.182\)](#) that $\varphi_1(\xi, \eta)$ and $\psi_1(\xi, \eta)$ are identical. Thus $\varphi_i(\xi, \eta)$ (or $\psi_i(\xi, \eta)$) is both the coordinate transformation function and the basis function.

8.3.3 Calculation of the Jacobian

To see how to use the isoparametric element we will address an example problem. Consider the equation

$$\frac{\partial^2}{\partial x^2} u(x, y) + \frac{\partial^2}{\partial y^2} u(x, y) - \frac{\partial}{\partial x} u(x, y) - \frac{\partial}{\partial y} u(x, y) = 0. \quad (8.202)$$

Using our standard approach we can write an approximation for $u(x, y)$ as

$$\hat{u}(x, y) = \sum_{j=1}^n u_j \phi_j(x, y). \quad (8.203)$$

From Galerkin's method we have

$$\int_{\Omega} \left(\frac{\partial^2}{\partial x^2} \hat{u}(x, y) + \frac{\partial^2}{\partial y^2} \hat{u}(x, y) - \frac{\partial}{\partial x} \hat{u}(x, y) - \frac{\partial}{\partial y} \hat{u}(x, y) \right) \phi_i(x, y) d\Omega = 0 \quad i = 1, \dots, n. \quad (8.204)$$

Application of Green's theorem to the second-order terms yields

$$\begin{aligned} & \int_{\Omega} \left(-\frac{\partial}{\partial x} \hat{u}(x, y) \frac{\partial}{\partial x} \phi_i(x, y) - \frac{\partial}{\partial y} \hat{u}(x, y) \frac{\partial}{\partial y} \phi_i(x, y) \right) d\Omega - \\ & \int_{\Omega} \left(\frac{\partial}{\partial x} \hat{u}(x, y) \phi_i(x, y) + \frac{\partial}{\partial y} \hat{u}(x, y) \phi_i(x, y) \right) d\Omega \\ & + \int_{\partial\Omega} \frac{\partial \hat{u}(x, y)}{\partial n} \phi_i(x, y) = 0 \quad i = 1, \dots, n. \end{aligned} \quad (8.205)$$

Substituting [Eq. \(8.203\)](#) into [Eq. \(8.204\)](#), and introducing the local indices I and J , we get for each element that consists of four nodes

$$\begin{aligned} & - \sum_{J=1}^4 u_J \int_{\Omega_\varepsilon} \left(\frac{\partial}{\partial x} \phi_J(\xi, \eta) \frac{\partial}{\partial x} \phi_I(\xi, \eta) + \frac{\partial}{\partial y} \phi_J(\xi, \eta) \frac{\partial}{\partial y} \phi_I(\xi, \eta) \right. \\ & \left. + \frac{\partial}{\partial x} \phi_J(\xi, \eta) \phi_I(\xi, \eta) + \frac{\partial}{\partial y} \phi_J(\xi, \eta) \phi_I(\xi, \eta) \right) d\Omega \\ & + \int_{\partial\Omega} \frac{\partial}{\partial n} \hat{u} \phi_I(\xi, \eta) dl = 0 \quad I = 1, \dots, 4 \end{aligned} \quad (8.207)$$

where n is now the outward-directed normal to the element face. Notice that this is not an equation equal to zero because we are considering only one element and we would need to sum the information contributed by all the elements associated with node I before we could write an equation for node I .

Typical integrals in [Eq. \(8.207\)](#) are of the form

$$\int_{\Omega_c} \left(\frac{\partial}{\partial x} \phi_J(\xi, \eta) \frac{\partial}{\partial x} \phi_I(\xi, \eta) \right) d\Omega \quad (8.208)$$

and

$$\int_{\Omega_c} \left(\frac{\partial}{\partial x} \phi_J(\xi, \eta) \phi_I(\xi, \eta) \right) d\Omega. \quad (8.209)$$

Let us see how we evaluate these integrals. Since

$$\xi = \xi(x, y) \quad (8.210)$$

and

$$\eta = \eta(x, y) \quad (8.211)$$

from the chain rule we have

$$\frac{\partial}{\partial \xi} \phi_J(\xi, \eta) = \frac{\partial \phi_J(\xi, \eta)}{\partial x} \frac{\partial x}{\partial \xi} + \frac{\partial \phi_J(\xi, \eta)}{\partial y} \frac{\partial y}{\partial \xi} \quad (8.212)$$

$$\frac{\partial}{\partial \eta} \phi_J(\xi, \eta) = \frac{\partial \phi_J(\xi, \eta)}{\partial x} \frac{\partial x}{\partial \eta} + \frac{\partial \phi_J(\xi, \eta)}{\partial y} \frac{\partial y}{\partial \eta}. \quad (8.213)$$

We can assemble these equations into matrix form to obtain

$$\begin{bmatrix} \frac{\partial \phi_J(\xi, \eta)}{\partial \xi} \\ \frac{\partial \phi_J(\xi, \eta)}{\partial \eta} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \end{bmatrix} \begin{bmatrix} \frac{\partial \phi_J(\xi, \eta)}{\partial x} \\ \frac{\partial \phi_J(\xi, \eta)}{\partial y} \end{bmatrix}. \quad (8.214)$$

By convention we define the Jacobian matrix $[J]$ as

$$[J] = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \end{bmatrix} \quad (8.215)$$

such that

$$\begin{bmatrix} \frac{\partial \phi_J(\xi, \eta)}{\partial \xi} \\ \frac{\partial \phi_J(\xi, \eta)}{\partial \eta} \end{bmatrix} = [J] \begin{bmatrix} \frac{\partial \phi_J(\xi, \eta)}{\partial x} \\ \frac{\partial \phi_J(\xi, \eta)}{\partial y} \end{bmatrix}. \quad (8.216)$$

However we see in [Eqs. \(8.212\)](#) and [\(8.213\)](#) that we need terms of the form

$$\frac{\partial}{\partial x} \phi_J(\xi, \eta) \quad (8.217)$$

and

$$\frac{\partial}{\partial y} \phi_J (\xi, \eta) . \quad (8.218)$$

To obtain this form, we need to take the inverse of $[J]$ to yield

$$\begin{bmatrix} \frac{\partial \phi_J (\xi, \eta)}{\partial x} \\ \frac{\partial \phi_J (\xi, \eta)}{\partial y} \end{bmatrix} = [J]^{-1} \begin{bmatrix} \frac{\partial \phi_J (\xi, \eta)}{\partial \xi} \\ \frac{\partial \phi_J (\xi, \eta)}{\partial \eta} \end{bmatrix} . \quad (8.219)$$

Computing $[J]^{-1}$ we have

$$[J]^{-1} = \frac{1}{\det [J]} \begin{bmatrix} \frac{\partial y}{\partial \eta} & -\frac{\partial y}{\partial \xi} \\ -\frac{\partial x}{\partial \eta} & \frac{\partial x}{\partial \xi} \end{bmatrix} . \quad (8.220)$$

Combination of [Eqs. \(8.219\)](#) and [\(8.220\)](#) yields

$$\begin{bmatrix} \frac{\partial \phi_J (\xi, \eta)}{\partial x} \\ \frac{\partial \phi_J (\xi, \eta)}{\partial y} \end{bmatrix} = \frac{1}{\det [J]} \begin{bmatrix} \frac{\partial y}{\partial \eta} & -\frac{\partial y}{\partial \xi} \\ -\frac{\partial x}{\partial \eta} & \frac{\partial x}{\partial \xi} \end{bmatrix} \begin{bmatrix} \frac{\partial \phi_J (\xi, \eta)}{\partial \xi} \\ \frac{\partial \phi_J (\xi, \eta)}{\partial \eta} \end{bmatrix} . \quad (8.221)$$

The next problem is to determine how to get the Jacobian matrix $[J]$?

Recall that

$$x = \sum_{I=1}^4 x_I \phi_I (\xi, \eta) \quad (8.222)$$

and

$$y = \sum_{I=1}^4 y_I \phi_I (\xi, \eta) . \quad (8.223)$$

We can now use the relationships in [Eqs. \(8.222\)](#) and [\(8.223\)](#) and write

$$\frac{\partial x}{\partial \xi} = \sum_{I=1}^4 x_I \frac{\partial}{\partial \xi} \phi_I(\xi, \eta) \quad (8.224)$$

$$= x_1 \left[\frac{1}{4} (-1 + \eta) \right] + x_2 \left[\frac{1}{4} (1 - \eta) \right]$$

$$+ x_3 \left[\frac{1}{4} (1 + \eta) \right] + x_4 \left[\frac{1}{4} (-1 - \eta) \right] \quad (8.225)$$

$$\frac{\partial x}{\partial \eta} = \sum_{I=1}^4 x_I \frac{\partial}{\partial \eta} \phi_I(\xi, \eta) \quad (8.226)$$

$$= x_1 \left[\frac{1}{4} (-1 + \xi) \right] + x_2 \left[\frac{1}{4} (-1 - \xi) \right]$$

$$+ x_3 \left[\frac{1}{4} (1 + \xi) \right] + x_4 \left[\frac{1}{4} (1 - \xi) \right]$$

$$\frac{\partial y}{\partial \xi} = \sum_{I=1}^4 y_I \frac{\partial}{\partial \xi} \phi_I(\xi, \eta) \quad (8.227)$$

$$= y_1 \left[\frac{1}{4} (-1 + \eta) \right] + y_2 \left[\frac{1}{4} (1 - \eta) \right]$$

$$+ y_3 \left[\frac{1}{4} (1 + \eta) \right] + y_4 \left[\frac{1}{4} (-1 - \eta) \right]$$

$$\begin{aligned}
\frac{\partial y}{\partial \eta} &= \sum_{I=1}^4 y_I \frac{\partial}{\partial \eta} \phi_I(\xi, \eta) \\
&= y_1 \left[\frac{1}{4} (\xi - 1) \right] + y_2 \left[\frac{1}{4} (-1 - \xi) \right] + \\
&\quad y_3 \left[\frac{1}{4} (1 + \xi) \right] + y_4 \left[\frac{1}{4} (1 - \xi) \right].
\end{aligned}
\tag{8.228}$$

We can now evaluate the Jacobian. We have

$$\begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} (-1 + \eta) & (1 - \eta) & (1 + \eta) & (-1 - \eta) \\ (\xi - 1) & (-1 - \xi) & (1 + \xi) & (1 - \xi) \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ x_4 & y_4 \end{bmatrix}.
\tag{8.229}$$

One can also show that

$$dxdy = \det [J] d\xi d\eta
\tag{8.230}$$

and

$$\det [J] = \left[\frac{\partial y}{\partial \eta} \frac{\partial x}{\partial \xi} - \frac{\partial x}{\partial \eta} \frac{\partial y}{\partial \xi} \right].
\tag{8.231}$$

We are now in a position to evaluate the integrals of interest. Consider as an example

$$\int_{\Omega_\varepsilon} \left(\frac{\partial}{\partial x} \phi_J(\xi, \eta) \frac{\partial}{\partial x} \phi_I(\xi, \eta) + \frac{\partial}{\partial y} \phi_J(\xi, \eta) \frac{\partial}{\partial y} \phi_I(\xi, \eta) \right) dxdy.
\tag{8.232}$$

We can write the integral in natural coordinates using [Eqs. \(8.212\)–\(8.221\)](#) and [Eqs. \(8.230\)–\(8.231\)](#) as

(8.233)

$$\begin{aligned}
& \int_{\Omega_e} \left(\frac{\partial}{\partial x} \phi_J(\xi, \eta) \frac{\partial}{\partial x} \phi_I(\xi, \eta) + \frac{\partial}{\partial y} \phi_J(\xi, \eta) \frac{\partial}{\partial y} \phi_I(\xi, \eta) \right) dx dy \\
&= \int_{-1}^1 \int_{-1}^1 \frac{1}{\det [J]} \left(\frac{\partial}{\partial \xi} \phi_J(\xi, \eta) \frac{\partial y}{\partial \eta} - \frac{\partial}{\partial \eta} \phi_J(\xi, \eta) \frac{\partial y}{\partial \xi} \right) \\
& \frac{1}{\det [J]} \left(\frac{\partial}{\partial \xi} \phi_I(\xi, \eta) \frac{\partial y}{\partial \eta} - \frac{\partial}{\partial \eta} \phi_I(\xi, \eta) \frac{\partial y}{\partial \xi} \right) \\
& + \frac{1}{\det [J]} \left(-\frac{\partial}{\partial \xi} \phi_J(\xi, \eta) \frac{\partial x}{\partial \eta} + \frac{\partial}{\partial \eta} \phi_J(\xi, \eta) \frac{\partial x}{\partial \xi} \right) \\
& \frac{1}{\det [J]} \left(-\frac{\partial}{\partial \xi} \phi_I(\xi, \eta) \frac{\partial x}{\partial \eta} + \frac{\partial}{\partial \eta} \phi_I(\xi, \eta) \frac{\partial x}{\partial \xi} \right) \det [J] d\xi d\eta.
\end{aligned}$$

The challenge now is to evaluate the integrals appearing on the right-hand side of [Eq. \(8.233\)](#). This is accomplished using Gauss–Legendre Quadrature introduced in Section 3.1.6 on page 59.

8.3.4 Example of Isoparametric Formulation

Let us consider the problem of solving for the temperature distribution in the $x - y$ plane. The problem is defined in [Fig. 8.19](#). The governing equation is

$$K \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right) = 0 \quad (8.234)$$

with boundary conditions

$$T(x, y) = 6 \quad x = 0 \quad y \in [0, 5] \quad (8.235)$$

$$T(x, y) = 12 \quad x = 20 \quad y \in [0, 5] \quad (8.236)$$

$$\frac{\partial T}{\partial n} = 0 \quad y = 0 \quad x \in (0, 20) \quad (8.237)$$

$$\frac{\partial T}{\partial n} = 0 \quad y = 5 \quad x \in (0, 20). \quad (8.238)$$

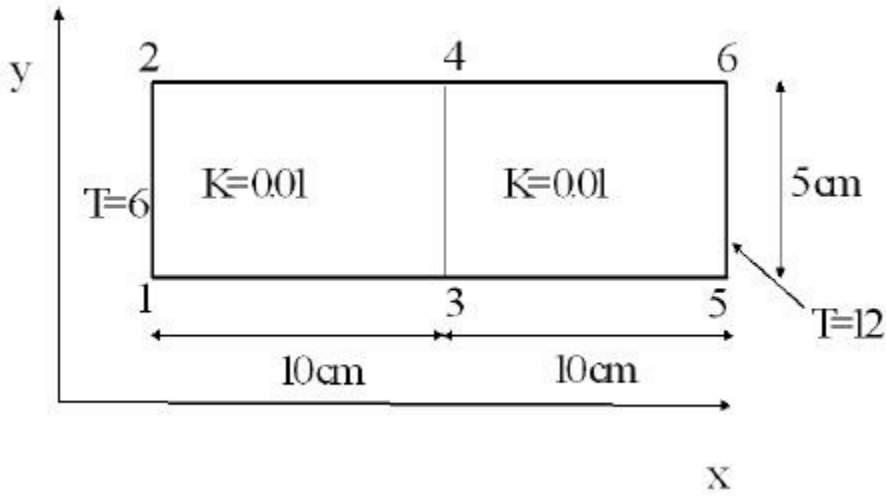


Figure 8.19: Finite element mesh for the example problem considered in Section 8.3.4.

Galerkin Formulation

Let us assume the temperature can be represented by the series

$$\hat{T}(x, y) = \sum_{j=1}^6 T_j \phi_j(x, y). \quad (8.239)$$

Define the residual

$$\frac{\partial^2 \hat{T}}{\partial x^2} + \frac{\partial^2 \hat{T}}{\partial y^2} = R(x, y). \quad (8.240)$$

From the Galerkin method we have

$$\int_{\Omega} R(x, y) \phi_i(x, y) d\Omega = 0 \quad i = 1, \dots, 6. \quad (8.241)$$

Expanding $R(x, y)$ we have

$$\int_{\Omega} \left(\frac{\partial^2 \hat{T}}{\partial x^2} + \frac{\partial^2 \hat{T}}{\partial y^2} \right) \phi_i(x, y) d\Omega = 0 \quad i = 1, \dots, 6. \quad (8.242)$$

Substitution of [8.239](#) into [8.242](#) gives

$$\sum_{j=1}^6 T_j \int_{\Omega} \left(\frac{\partial^2 \phi_j(x, y)}{\partial x^2} \phi_i(\xi, \eta) + \frac{\partial^2 \phi_j(x, y)}{\partial y^2} \phi_i(\xi, \eta) \right) d\Omega = 0 \quad i = 1, \dots, 6 \quad (8.243)$$

We can now apply Green's theorem to modify the second derivative terms and obtain

$$-\sum_{j=1}^6 T_j \int_{\Omega} \left(\frac{\partial \phi_j(x, y)}{\partial x} \frac{\partial \phi_i(x, y)}{\partial x} + \frac{\partial \phi_j(x, y)}{\partial y} \frac{\partial \phi_i(x, y)}{\partial y} \right) d\Omega + \int_{\partial\Omega} \frac{\partial \hat{T}}{\partial n} \phi_i(x, y) dl = 0 \quad i = 1, \dots, 6. \quad (8.244)$$

This set of equations can be written in global coordinates as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ 0 & 0 & a_{53} & a_{54} & a_{55} & a_{56} \\ 0 & 0 & a_{63} & a_{64} & a_{65} & a_{66} \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \end{bmatrix}. \quad (8.245)$$

Since we know from the boundary conditions that

$$T_1 = T_2 = 6 \quad (8.246)$$

and

$$T_5 = T_6 = 12 \quad (8.247)$$

we can reduce the coefficient matrix to give

$$\begin{bmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} T_3 \\ T_4 \end{bmatrix} = \begin{bmatrix} f_3 \\ f_4 \end{bmatrix} - \begin{bmatrix} a_{31}T_1 + a_{32}T_2 + a_{35}T_5 + a_{36}T_6 \\ a_{41}T_1 + a_{42}T_2 + a_{45}T_5 + a_{46}T_6 \end{bmatrix}. \quad (8.248)$$

Although it may not be obvious, we can show from symmetry that

$$a_{34} = a_{43} \quad (8.249)$$

and from geometry that

$$a_{33} = a_{44} \quad (8.250)$$

$$a_{31} = a_{42} \quad (8.251)$$

$$a_{35} = a_{46} \quad (8.252)$$

$$a_{36} = a_{45} \quad (8.253)$$

$$a_{32} = a_{41} \quad (8.254)$$

$$a_{35} = a_{31} \quad (8.255)$$

$$a_{42} = a_{46} \quad (8.256)$$

$$a_{36} = a_{32} \quad (8.257)$$

With these considerations in mind, we are left with the coefficients

$$a_{31}, a_{32}, a_{33}, a_{34} \quad (8.258)$$

Consider coefficient a_{31} :

$$a_{31} = \int_{\Omega} \left(\frac{\partial \phi_1}{\partial x} \frac{\partial \phi_3}{\partial x} + \frac{\partial \phi_1}{\partial y} \frac{\partial \phi_3}{\partial y} \right) d\Omega. \quad (8.259)$$

The integration for this coefficient is restricted to element 1 because nodes 1 and 3 are only connected in element 1 as seen in [Fig. 8.19](#). We now do the integration over only element 1 and we do this in the local coordinate system. The process is as follows.

Transform Integrals to Natural Coordinate System

The relationship between the natural and global coordinate systems for the first element are provided in the following table:

Local	Global
1 or $(-1, -1)$	1
2 or $(1, -1)$	3
3 or $(1, 1)$	4
4 or $(-1, 1)$	2

(8.260)

Consider, for element 1 in natural coordinates, the integral (see [Section 8.3.3](#) on page 219 above for insight into how this is done)

$$\int_{\Omega_{e=1}} \left(\frac{\partial \phi_1(\xi, \eta)}{\partial x} \frac{\partial \phi_2(\xi, \eta)}{\partial x} + \frac{\partial \phi_1(\xi, \eta)}{\partial y} \frac{\partial \phi_2(\xi, \eta)}{\partial y} \right) d\Omega_e. \quad (8.261)$$

We need functions of the form $\frac{\partial \phi_1(\xi, \eta)}{\partial \xi}$ defined in natural coordinates. We start with the relationship derived from the chain rule, that is,

$$\begin{bmatrix} \frac{\partial \phi_j(\xi, \eta)}{\partial \xi} \\ \frac{\partial \phi_j(\xi, \eta)}{\partial \eta} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \end{bmatrix} \begin{bmatrix} \frac{\partial \phi_j(\xi, \eta)}{\partial x} \\ \frac{\partial \phi_j(\xi, \eta)}{\partial y} \end{bmatrix}. \quad (8.262)$$

From [Eq. \(8.229\)](#) on page 222, reproduced below, we can write the coefficient matrix in [Eq. \(8.262\)](#) as

$$\begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} (-1 + \eta) & (1 - \eta) & (1 + \eta) & (-1 - \eta) \\ (-1 + \xi) & (-1 - \xi) & (1 + \xi) & (1 - \xi) \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ x_4 & y_4 \end{bmatrix}. \quad (8.263)$$

So, for the particular case of $\varphi_1(\xi, \eta)$ we substitute [Eq. \(8.263\)](#) into [Eq. \(8.262\)](#) to obtain

(8.264)

$$\begin{bmatrix} \frac{\partial \phi_1(\xi, \eta)}{\partial \xi} \\ \frac{\partial \phi_1(\xi, \eta)}{\partial \eta} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} (-1 + \eta) & (1 - \eta) & (1 + \eta) & (-1 - \eta) \\ (-1 + \xi) & (-1 - \xi) & (1 + \xi) & (1 - \xi) \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 10 & 0 \\ 10 & 5 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} \frac{\partial \phi_1(\xi, \eta)}{\partial x} \\ \frac{\partial \phi_1(\xi, \eta)}{\partial y} \end{bmatrix}$$

or (8.265)

$$\begin{bmatrix} \frac{\partial \phi_1(\xi, \eta)}{\partial \xi} \\ \frac{\partial \phi_1(\xi, \eta)}{\partial \eta} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} (1 - \eta)(10) + (1 + \eta)(10) & (1 + \eta)(5) + (-1 - \eta)(5) \\ (-1 - \xi)(10) + (1 + \xi)(10) & (1 + \xi)(5) + (1 - \xi)(5) \end{bmatrix} \begin{bmatrix} \frac{\partial \phi_1(\xi, \eta)}{\partial x} \\ \frac{\partial \phi_1(\xi, \eta)}{\partial y} \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial \phi_1(\xi, \eta)}{\partial \xi} \\ \frac{\partial \phi_1(\xi, \eta)}{\partial \eta} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 20 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} \frac{\partial \phi_1(\xi, \eta)}{\partial x} \\ \frac{\partial \phi_1(\xi, \eta)}{\partial y} \end{bmatrix}. \quad (8.266)$$

Similarly we have

$$\begin{bmatrix} \frac{\partial \phi_2(\xi, \eta)}{\partial \xi} \\ \frac{\partial \phi_2(\xi, \eta)}{\partial \eta} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 20 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} \frac{\partial \phi_2(\xi, \eta)}{\partial x} \\ \frac{\partial \phi_2(\xi, \eta)}{\partial y} \end{bmatrix}. \quad (8.267)$$

Taking the inverse of the Jacobian matrix we have

$$\begin{aligned} \begin{bmatrix} \frac{\partial \phi_1(\xi, \eta)}{\partial x} \\ \frac{\partial \phi_1(\xi, \eta)}{\partial y} \end{bmatrix} &= \frac{16}{200} \begin{bmatrix} \frac{10}{4} & 0 \\ 0 & \frac{20}{4} \end{bmatrix} \begin{bmatrix} \frac{\partial \phi_1(\xi, \eta)}{\partial \xi} \\ \frac{\partial \phi_1(\xi, \eta)}{\partial \eta} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{5} & 0 \\ 0 & \frac{2}{5} \end{bmatrix} \begin{bmatrix} \frac{\partial \phi_1(\xi, \eta)}{\partial \xi} \\ \frac{\partial \phi_1(\xi, \eta)}{\partial \eta} \end{bmatrix} \end{aligned} \quad (8.268)$$

and.

$$\begin{bmatrix} \frac{\partial \phi_2(\xi, \eta)}{\partial x} \\ \frac{\partial \phi_2(\xi, \eta)}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{1}{5} & 0 \\ 0 & \frac{2}{5} \end{bmatrix} \begin{bmatrix} \frac{\partial \phi_2(\xi, \eta)}{\partial \xi} \\ \frac{\partial \phi_2(\xi, \eta)}{\partial \eta} \end{bmatrix}. \quad (8.269)$$

This is the relationship we need to transform our derivatives in [Eq. \(8.261\)](#) so they are in the form we need.

To see how this is done, we return to our integral of interest, that is

$$\int_{\Omega_{\varepsilon=1}} \left(\frac{\partial \phi_1(\xi, \eta)}{\partial x} \frac{\partial \phi_2(\xi, \eta)}{\partial x} + \frac{\partial \phi_1(\xi, \eta)}{\partial y} \frac{\partial \phi_2(\xi, \eta)}{\partial y} \right) d\Omega_e \quad (8.270)$$

which becomes, using [Eqs. \(8.268\)](#) and [\(8.269\)](#) as follows:

$$\begin{aligned} & \int_{\Omega_{\varepsilon=1}} \left(\frac{\partial \phi_1(\xi, \eta)}{\partial x} \frac{\partial \phi_2(\xi, \eta)}{\partial x} + \frac{\partial \phi_1(\xi, \eta)}{\partial y} \frac{\partial \phi_2(\xi, \eta)}{\partial y} \right) d\Omega_e \\ &= \int_{-1}^1 \int_{-1}^1 \left[\left(\frac{1}{5} \frac{\partial \phi_1(\xi, \eta)}{\partial \xi} \right) \left(\frac{1}{5} \frac{\partial \phi_2(\xi, \eta)}{\partial \xi} \right) \right. \\ & \quad \left. + \left(\frac{2}{5} \frac{\partial \phi_1(\xi, \eta)}{\partial \eta} \right) \left(\frac{2}{5} \frac{\partial \phi_2(\xi, \eta)}{\partial \eta} \right) \right] \det [J] d\xi d\eta \end{aligned} \quad (8.271)$$

and $\det [J]$ is given by

$$\begin{aligned} & \det \frac{1}{4} \begin{bmatrix} 20 & 0 \\ 0 & 10 \end{bmatrix} \\ &= \det \begin{bmatrix} 5 & 0 \\ 0 & \frac{5}{2} \end{bmatrix} \\ &= \frac{25}{2}. \end{aligned} \quad (8.272)$$

Perform Integrations

We now use Gauss Quadrature (see Section 3.1.6 on page 59) to evaluate integrals of the form found in [Eq. \(8.271\)](#). The Gauss points are at

$$b_i(\xi) = \pm \frac{1}{\sqrt{3}}, \quad b_i(\eta) = \pm \frac{1}{\sqrt{3}} \quad (8.273)$$

where $\frac{1}{\sqrt{3}} = 0.5773$, and the weighting function is $H_i = 1.0$. Consider the first term in the integral on the right-hand side of [Eq. \(8.271\)](#). From the definitions of $\varphi_1(\xi, \eta)$ and $\varphi_2(\xi, \eta)$ given by

$$\phi_1 = \left[\left(\frac{1 - \xi}{2} \right) \left(\frac{1 - \eta}{2} \right) \right] \quad (8.274)$$

and

$$\phi_2 = \left[\left(\frac{1 + \xi}{2} \right) \left(\frac{1 - \eta}{2} \right) \right] \quad (8.275)$$

we have for the first term in [Eq. \(8.271\)](#),

$$(8.276)$$

$$\begin{aligned} & \int_{-1}^1 \int_{-1}^1 \left(\frac{1}{5} \frac{\partial \phi_1(\xi, \eta)}{\partial \xi} \right) \left(\frac{1}{5} \frac{\partial \phi_2(\xi, \eta)}{\partial \xi} \right) \det [J] d\xi d\eta \\ &= \sum_{\xi = \pm \frac{1}{\sqrt{3}}} \sum_{\eta = \pm \frac{1}{\sqrt{3}}} \left[\left(\frac{1}{5} \left[- \left(\frac{1 - \eta}{4} \right) \right] \right) \left(\frac{1}{5} \left[\left(\frac{1 - \eta}{4} \right) \right] \right) \frac{25}{2} \right] \\ &= 2 \sum_{\xi = \pm \frac{1}{\sqrt{3}}} - \left(\frac{1}{25} \right) \left(\frac{1}{16} \right) (1 - 2\eta + \eta^2) \left(\frac{25}{2} \right) \\ &= 2 \left\{ - \left(\frac{1}{25} \right) \left(\frac{1}{16} \right) \left[1 - 2 \frac{1}{\sqrt{3}} + \left(\frac{1}{\sqrt{3}} \right)^2 \right] - \left(\frac{1}{25} \right) \left(\frac{1}{16} \right) \left[1 - 2 \frac{-1}{\sqrt{3}} + \left(\frac{-1}{\sqrt{3}} \right)^2 \right] \right\} \left(\frac{25}{2} \right) \\ &= -2 \left(\frac{1}{25} \right) \left(\frac{1}{16} \right) \left(\left[1 - \frac{2}{\sqrt{3}} + \frac{1}{3} \right] + \left[1 + \frac{2}{\sqrt{3}} + \frac{1}{3} \right] \right) \left(\frac{25}{2} \right) \\ &= -2 \left(\frac{1}{25} \right) \left(\frac{1}{16} \right) \left(2 + \frac{2}{3} \right) \left(\frac{25}{2} \right) \\ &= -\frac{2}{16} \left(\frac{8}{3} \right) \left(\frac{1}{2} \right) \\ &= -\frac{2}{12} \\ &= -\frac{1}{6}. \end{aligned}$$

Note that the notation $\sum_{\xi = \pm \frac{1}{\sqrt{3}}} (\cdot)$ means that you substitute in the values $\xi = \pm \frac{1}{\sqrt{3}}$ into

the argument sequentially and sum the result. Thus there will be four terms in the summation on the right-hand side of the above equation. Now consider the second term on the right-hand-side of [Eq. \(8.271\)](#).

(8.277)

$$\begin{aligned}
 & \int_{-1}^1 \int_{-1}^1 \left[\left(\frac{2}{5} \frac{\partial \phi_1(\xi, \eta)}{\partial \eta} + \frac{2}{5} \frac{\partial \phi_2(\xi, \eta)}{\partial \eta} \right) \right] \det [J] d\xi d\eta \\
 &= \sum_{\eta=\pm \frac{1}{\sqrt{3}}} \sum_{\xi=\pm \frac{1}{\sqrt{3}}} \left[\left(\frac{2}{5} \left[- \left(\frac{1-\xi}{4} \right) \right] \right) \left(\frac{2}{5} \left[\left(-\frac{1+\xi}{4} \right) \right] \right) \frac{25}{2} \right] \\
 &= 2 \sum_{\eta=\pm \frac{1}{\sqrt{3}}} \left(\frac{4}{25} \right) \left(\frac{1}{16} \right) (1 - \xi^2) \left(\frac{25}{2} \right) \\
 &= 2 \left(\frac{4}{25} \right) \left(\frac{1}{16} \right) \left[\left(1 - \frac{1}{3} \right) + \left(1 - \frac{1}{3} \right) \right] \left(\frac{25}{2} \right) \\
 &= 2 \left(\frac{4}{16} \right) \left(\frac{4}{3} \right) \left(\frac{1}{2} \right) \\
 &= \frac{2}{6}.
 \end{aligned}$$

The total term is the sum of $-\frac{1}{6} + \frac{2}{6} = \frac{1}{6}$. We now have the value for the coefficient (13) in the local coefficient matrix for element 1.

Matrix Assembly

If we evaluate all the integral terms for both elements, we can obtain the element coefficient matrices.

To obtain the global coefficient matrix we need a transformation table for the global and local nodal numbers. In the top row we have the global node numbers. In the second row, the local element node numbers of each node value for element 1 is provided. For example, the local node number in element 1 for global node 2 is 4. The same information for element 2 is found in the last row.

	Global						(8.278)
Local	1	2	3	4	5	6	Element
	1	4	2	3			1
			1	4	2	3	2

Element 1 and element 2 have the same coefficient values since they have the same geometry and properties, that is

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} 5/6 & 1/6 & -5/12 & -7/12 \\ 1/6 & 5/6 & -7/12 & -5/12 \\ -5/12 & -7/12 & 5/6 & 1/6 \\ -7/12 & -5/12 & 1/6 & 5/6 \end{bmatrix}. \quad (8.279)$$

The global matrix is obtained by combining information from the two element matrices and the transformation table. As an example, consider the global coefficient (b_{11}). This coefficient is according to the table, the (a_{11}) coefficient of the element coefficient matrix for element 1; this value is $5/6$ and this value appears as the (b_{11}) coefficient in the global matrix given in [Eq. \(8.280\)](#) below. Now let us consider a more interesting case. Node 3 in [Fig. 8.19](#) lies at the interface of the two elements. Thus this coefficient in the global matrix draws information from both of these elements. We see this in the table because global node 3 is the same as local node 2 in element 1 it is the same as local node 1 in element 2. So, to build the global coefficient (b_{33}) we need to add the (a_{22}) coefficient from element 1 to the (a_{11}) coefficient from element 2. We obtain $5/6 + 5/6$ as is seen in the (b_{33}) location of the matrix in [Eq. \(8.280\)](#).

As a final example we take the most complicated situation, the formation of global element coefficient (b_{34}). This coefficient involves information from two nodes, each located at the interface of two elements. From the table we see that in the first element, global node (b_{34}) is equivalent to element node (a_{23}) in element 1 and ($a_{1,4}$) in element 2. Thus we need to add these two values together to obtain the global coefficient (b_{34}) and so we obtain $-7/12 - 7/12$ as seen in coefficient location (b_{34}).

Note that for this second-order equation both the global [B] and local [A] coefficient matrices are symmetric and the rows and columns sum to zero.

$$\begin{aligned}
 & \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} & b_{16} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} & b_{26} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} & b_{36} \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} & b_{46} \\ b_{51} & b_{52} & b_{53} & b_{54} & b_{55} & b_{56} \\ b_{61} & b_{62} & b_{63} & b_{64} & b_{65} & b_{66} \end{bmatrix} & (8.280) \\
 & = \begin{bmatrix} \frac{5}{6} & -\frac{7}{12} & \frac{1}{6} & -\frac{5}{12} & 0 & 0 \\ -\frac{7}{12} & \frac{5}{6} & -\frac{5}{12} & \frac{1}{6} & 0 & 0 \\ \frac{1}{6} & -\frac{5}{12} & \frac{5}{6} + \frac{5}{6} & -\frac{7}{12} - \frac{7}{12} & \frac{1}{6} & -\frac{5}{12} \\ -\frac{5}{12} & \frac{1}{6} & -\frac{7}{12} - \frac{7}{12} & \frac{5}{6} + \frac{5}{6} & -\frac{5}{12} & \frac{1}{6} \\ 0 & 0 & \frac{1}{6} & -\frac{5}{12} & \frac{5}{6} & -\frac{7}{12} \\ 0 & 0 & -\frac{5}{12} & \frac{1}{6} & -\frac{7}{12} & \frac{5}{6} \end{bmatrix}.
 \end{aligned}$$

The complete matrix equation is

$$- \begin{bmatrix} \frac{5}{6} & -\frac{7}{12} & \frac{1}{6} & -\frac{5}{12} & 0 & 0 \\ -\frac{7}{12} & \frac{5}{6} & -\frac{5}{12} & \frac{1}{6} & 0 & 0 \\ \frac{1}{6} & -\frac{5}{12} & \frac{5}{6} + \frac{5}{6} & -\frac{7}{12} - \frac{7}{12} & \frac{1}{6} & -\frac{5}{12} \\ -\frac{5}{12} & \frac{1}{6} & -\frac{7}{12} - \frac{7}{12} & \frac{5}{6} + \frac{5}{6} & -\frac{5}{12} & \frac{1}{6} \\ 0 & 0 & \frac{1}{6} & -\frac{5}{12} & \frac{5}{6} & -\frac{7}{12} \\ 0 & 0 & -\frac{5}{12} & \frac{1}{6} & -\frac{7}{12} & \frac{5}{6} \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \end{bmatrix} \quad (8.281)$$

where

$$f_i = - \int_{\partial\Omega} \frac{\partial \hat{T}}{\partial n} \Phi_i(x, y) dl. \quad (8.282)$$

Introducing the boundary conditions we have

$$- \begin{bmatrix} \frac{5}{6} & -\frac{7}{12} & \frac{1}{6} & -\frac{5}{12} & 0 & 0 \\ -\frac{7}{12} & \frac{5}{6} & -\frac{5}{12} & \frac{1}{6} & 0 & 0 \\ \frac{1}{6} & -\frac{5}{12} & \frac{5}{6} + \frac{5}{6} & -\frac{7}{12} - \frac{7}{12} & \frac{1}{6} & -\frac{5}{12} \\ -\frac{5}{12} & \frac{1}{6} & -\frac{7}{12} - \frac{7}{12} & \frac{5}{6} + \frac{5}{6} & -\frac{5}{12} & \frac{1}{6} \\ 0 & 0 & \frac{1}{6} & -\frac{5}{12} & \frac{5}{6} & -\frac{7}{12} \\ 0 & 0 & -\frac{5}{12} & \frac{1}{6} & -\frac{7}{12} & \frac{5}{6} \end{bmatrix} \begin{bmatrix} 6 \\ 6 \\ T_3 \\ T_4 \\ 12 \\ 12 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \end{bmatrix} \quad (8.283)$$

which after bringing information to the right-hand side of the equation simplifies to

$$- \begin{bmatrix} 5/6 + 5/6 & -7/12 - 7/12 \\ -7/12 - 7/12 & 5/6 + 5/6 \end{bmatrix} \begin{bmatrix} T_3 \\ T_4 \end{bmatrix} = \begin{bmatrix} (1/6)(6) + (-5/12)(6) + (1/6)(12) + (-5/12)(12) \\ (-5/12)(6) + (1/6)(6) + (-5/12)(12) + (1/6)(12) \end{bmatrix} \quad (8.284)$$

$$- \begin{bmatrix} 5/3 & -7/6 \\ -7/6 & 5/3 \end{bmatrix} \begin{bmatrix} T_3 \\ T_4 \end{bmatrix} = - \begin{bmatrix} 9/2 \\ 9/2 \end{bmatrix}. \quad (8.285)$$

Solving for the unknown vector we get (on inversion of the coefficient matrix)

$$\begin{aligned} \begin{bmatrix} T_3 \\ T_4 \end{bmatrix} &= \frac{1}{(5/3)(5/3) - (7/6)(7/6)} \begin{bmatrix} 5/3 & 7/6 \\ 7/6 & 5/3 \end{bmatrix} \begin{bmatrix} 9/2 \\ 9/2 \end{bmatrix} & (8.286) \\ &= \frac{36}{51} \begin{bmatrix} (5/3 + 7/6)(9/2) \\ (7/6 + 5/3)(9/2) \end{bmatrix} \\ &= \frac{3}{51} \begin{bmatrix} (17/1)(9/1) \\ (17/1)(9/1) \end{bmatrix} \\ &= \begin{bmatrix} 9 \\ 9 \end{bmatrix} \end{aligned}$$

which is the exact solution.

8.4 Chapter Summary

In this chapter we introduce the Galerkin finite element method. We begin with rectangular elements. The approach used for rectangles is next extended to triangular finite elements, a method that is somewhat more abstract than the case of rectangular elements. The rectangular finite element formulation was next generalized to the case of isoparametric finite element that allowed for elements that are squares in a local coordinate system but can be deformed rectangles or trapezoids in the global coordinate system. The integrations inherent in the finite element method are easily performed in this local coordinate system.

8.5 Problems

1. Given a triangle with nodes at $(0, 0)$, $(2, 0)$ and $(1, 2)$ as shown below in [Fig. 8.20](#) and the values $u(0, 0) = 1$, $u(2, 0) = 2$, and $u(1, 2) = 0$, determine the value of $u(1, 1)$ and show that it is 1.0 This can be determined from the following two relationships:

$$\hat{u}(x, y) = \sum_{i=1}^3 u_i L_i(x_i, y_i) \quad (8.287)$$

$$\sum_{i=1}^3 L_i(x, y) = 1 \quad (8.288)$$

where $L_i(x, y)$ is a triangular basis function.

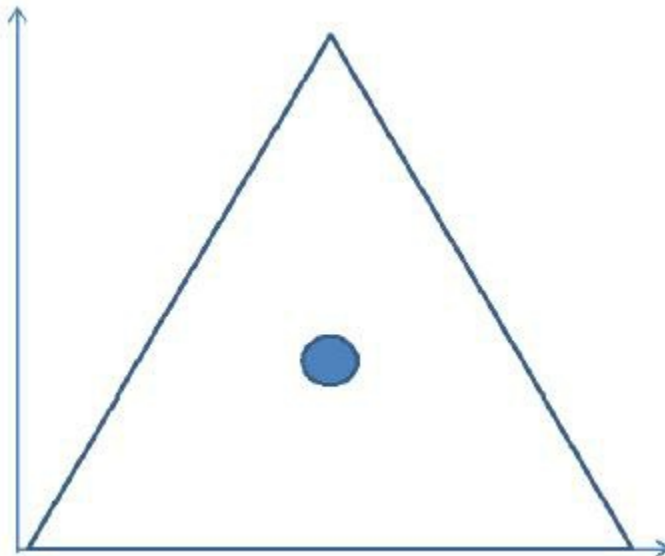


Figure 8.20: Triangular element and target location for problem 1.

2. Given the following equation that is used in the isoparametric formulation,

$$\begin{bmatrix} \frac{\partial \phi_j(\xi, \eta)}{\partial \xi} \\ \frac{\partial \phi_j(\xi, \eta)}{\partial \eta} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \end{bmatrix} \begin{bmatrix} \frac{\partial \phi_j(\xi, \eta)}{\partial x} \\ \frac{\partial \phi_j(\xi, \eta)}{\partial y} \end{bmatrix} \quad (8.289)$$

and the information in [Figs. 8.21](#) and [8.22](#), calculate the value of $\frac{\partial x}{\partial \xi}$, $\frac{\partial y}{\partial \xi}$, $\frac{\partial x}{\partial \eta}$ and $\frac{\partial y}{\partial \eta}$ and show that they are as follows:

$$\begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 2.5 \end{bmatrix} \quad (8.290)$$

Hint: You can do this problem just through careful observation of the figures below and the properties of basis functions, but just in case you need them here are the equations for the basis functions: $\varphi_j(x, y) = \varphi_j(x)\varphi_j(y)$ where $\phi_j(x) = \frac{x_{j+1}-x}{\Delta x}$ and $\varphi_j(\xi, \eta) = \varphi_j(\xi)\varphi_j(\eta)$ where $\phi_j(\xi) = \frac{\xi_{j+1}-\xi}{2}$.

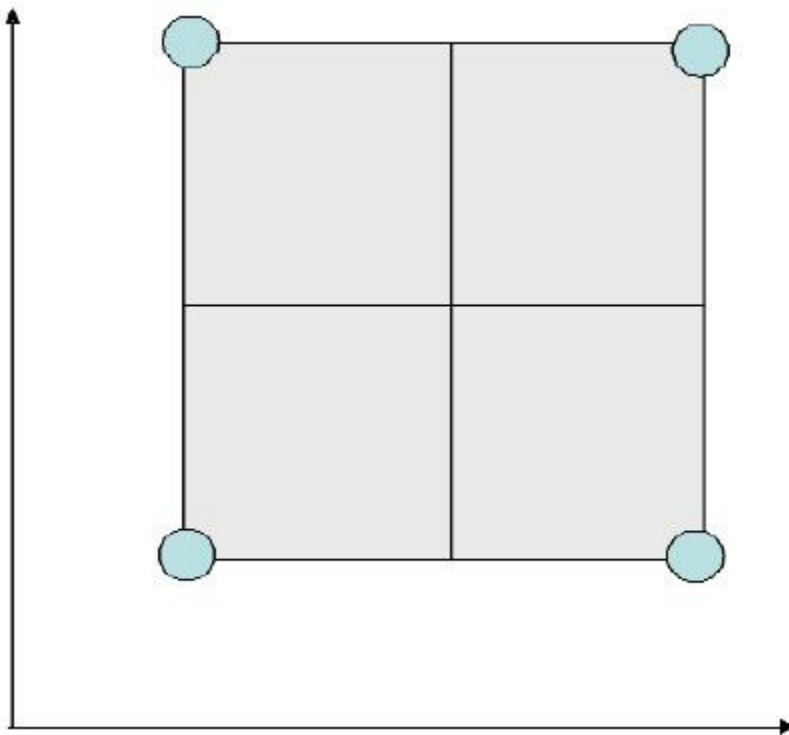


Figure 8.21: Explanation sketch for problem 2.

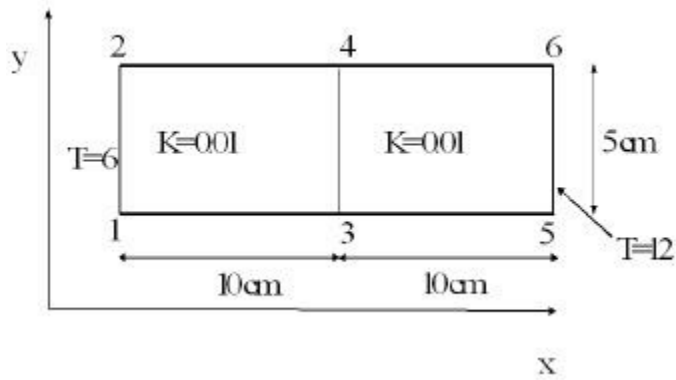


Figure 8.22: Explanation sketch for problem 2.

3. Consider the finite element array found in Fig. 8.5 the following figure. Assume the global matrix for this problem is

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & 3 & 2 & 5 \\ 0 & 2 & 2 & 4 \\ 2 & 5 & 4 & 4 \end{bmatrix}$$

and the element matrix for element $e = 1$ is

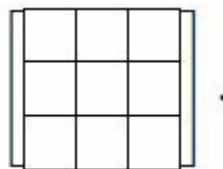
$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 2 \\ 2 & 2 & 1 \end{bmatrix}$$

(8.291)

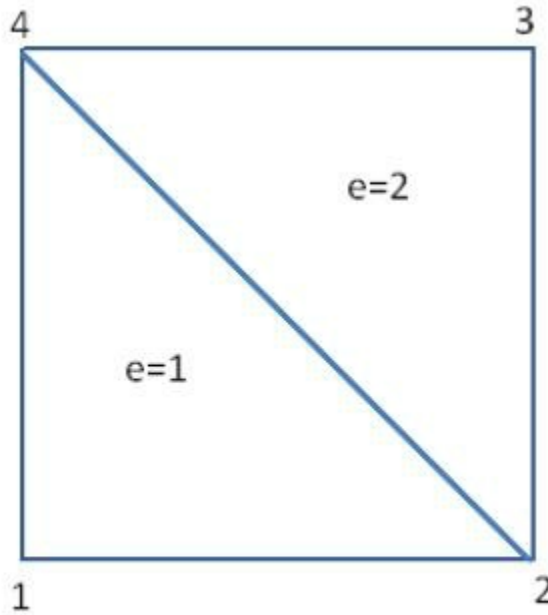
The transformation table is

Global	1	2	3	4
Element one	1	2		3
Element two		1	2	3

Please fill in the values for element matrix $e = 2$



(8.292)



Definition diagram for problem 3.

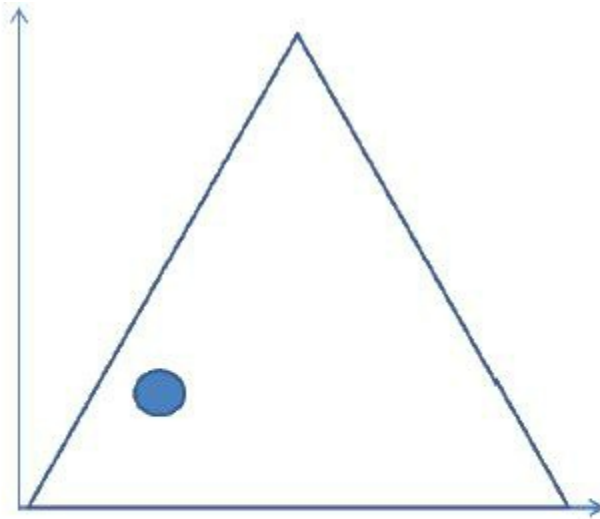
4. Given a triangle as illustrated in [Fig. 8.5](#), with nodes at $(0, 0)$, $(2, 0)$, and $(1, 2)$ as shown below and the values $u(0, 0) = 1$, $u(2, 0) = 2$, and $u(1, 2) = 0$, and the basis functions

$$L_I(x, y) = \frac{x(y_{I+1} - y_{I+2}) + y(x_{I+2} - x_{I+1}) + (x_{I+1}y_{I+2} - x_{I+2}y_{I+1})}{\det[P]} \quad (8.293)$$

where

$$A = \frac{1}{2} \det[P] \quad (8.294)$$

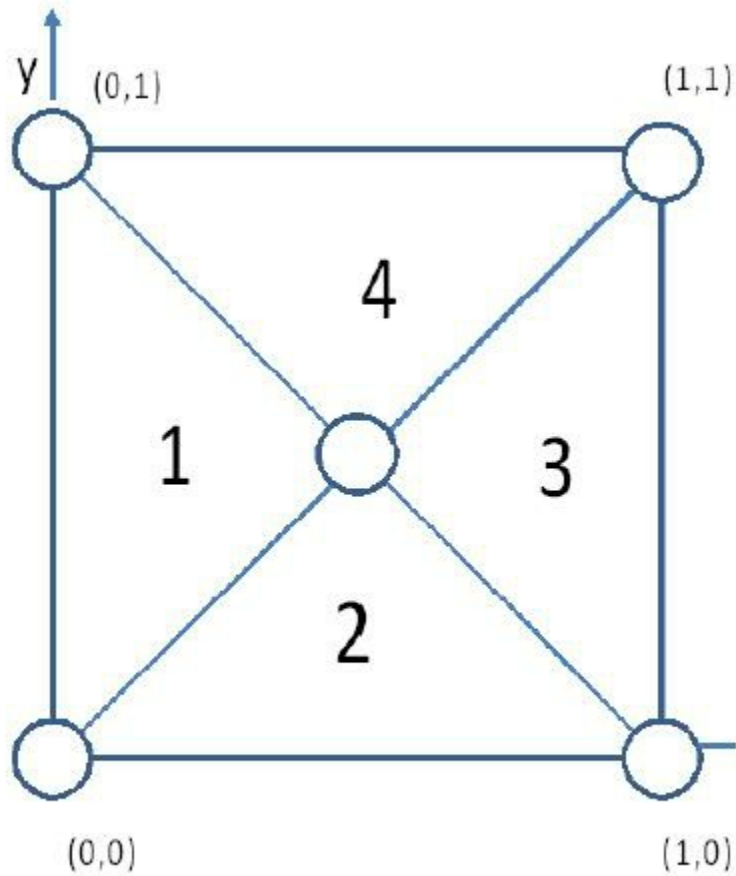
and $A = 2$ is the area of the triangle, show, **using the above basis functions**, that the value of $\frac{\partial u(x, y)}{\partial x}$ at the location $(x, y) = (0.5, 0.5)$ is 0.5.



Information pertinent to problem 8.5.

5. Given the finite element mesh in [Fig. 8.5](#), locate and place the global nodal numbers on the figure (in the circles). In the following table, the global numbers are in bold and the element numbers are in italics. Remember that local element numbering (the $I, I+1, I+2$) is counter-clockwise.

	1	2	3	4	5
<i>1</i>	<i>I</i>	<i>I+1</i>			<i>I+2</i>
<i>2</i>	<i>I</i>	<i>I+2</i>	<i>I+1</i>		
<i>3</i>		<i>I+2</i>	<i>I</i>	<i>I+1</i>	
<i>4</i>		<i>I</i>		<i>I+1</i>	<i>I+2</i>



Illustrative figure associated with problem 8.5.

7. Consider the equation

$$\frac{\partial}{\partial x} u + \frac{\partial}{\partial y} u = 0 \quad (8.295)$$

defined in the one-element domain shown in [Fig. 8.23](#) with boundary conditions

$$u(0, 0) = 1 \quad (8.296)$$

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

Solve for the value at node 3, namely $u(0, 1)$, or equivalently u_3 , using a Galerkin finite element method.

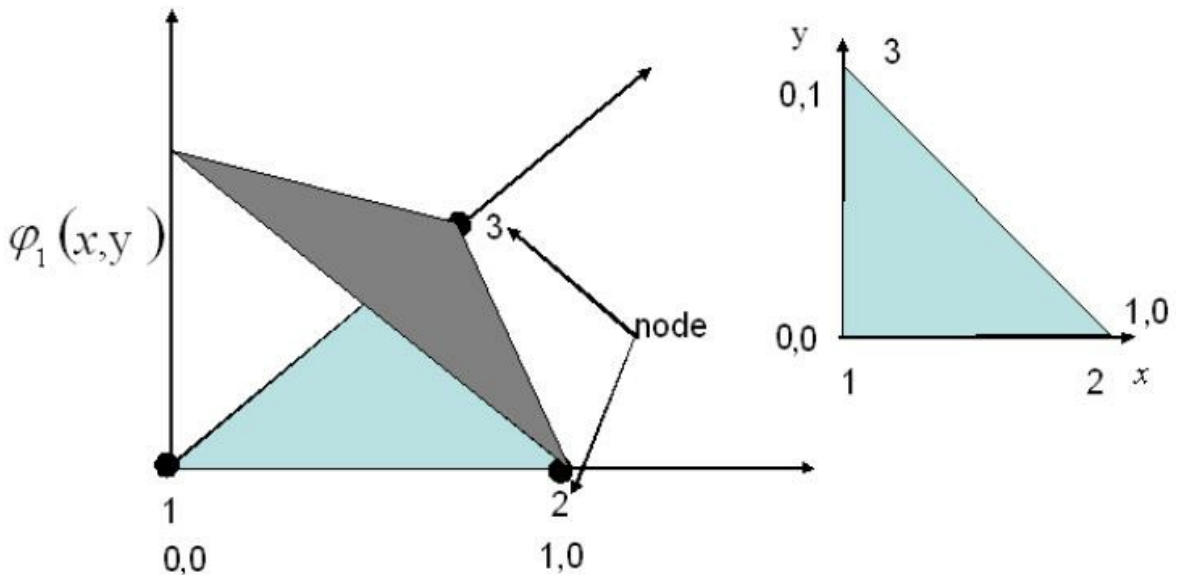


Figure 8.23: One-element domain with the basis function at node one illustrated.

Steps

1. Write the approximation for $u(x, y)$, that is, $\hat{u}(x, y) = \sum_{j=1}^3 u_j \phi_j(x, y)$.
2. Generate the weighted residual equations (there are three).
3. Substitute for $\hat{u}(x, y)$ in the weighted residual equation.
4. Write the matrix equation symbolically, that is the global matrix (3×3) in terms of the integral such as $\int_{\Omega_e} \frac{\partial \phi_i}{\partial x} \phi_i da$, the unknown vector of values of u_j and the right-hand side (which at this stage is zero).
5. Impose the two boundary conditions (you will now have one unknown, u_3) and note that you will now have integrals in the (1×1) global matrix and on the right-hand side (two of them).
6. Evaluate the integrals using the geometry shown in the figure to define the derivatives of ϕ_j and the relationship $\int_{\Omega_e} \phi_i da = \frac{A}{3}$ (you actually don't need this last relationship, but in case you don't see why, here it is).
7. You now have one equation in one unknown, u_3 . Solve for it.
8. Check your answer by determining whether the derivatives of $u(x, y)$ satisfy the original differential equation.

Bibliography

Allen, M.B. III and E.L. Isaacson (1988) *Numerical Analysis for Applied Science*, John Wiley and Sons, New York.

Bathe, K.J. (1976) *Numerical Methods in Finite Element Analysis*, Prentice-Hall.

Becker, E.B., G.F. Carey and J.T. Oden (1981) *Finite Elements An Introduction*,

Prentice-Hall, Englewood Cliffs.

Celia, M.A. and W.G. Gray (1992) *Numerical Methods for Differential Equations: Fundamental Concepts for Scientific and Engineering Applications*, Prentice-Hall, Englewood Cliffs.

Collatz, L. (1966) "The Numerical Treatment of Differential Equations," Springer Verlag, New York.

Finlayson, B.A. (1972) *The Method of Weighted Residuals and Variational Principles*, Academic Press, New York.

Li, R., Z. Chen, and W. Wu (2000) "Generalized difference methods for differential equations," *Pure and Applied Mathematics*, Marcel Dekker, New York.

Reddy, J.N. (2006) *An Introduction to the Finite Element Method*, Mcgraw Hill, New York.

Strang, G. and G. Fix (1973) *Analysis of the Finite Element Method*, Prentice-Hall, Englewood Cliffs.

Zienkiewicz, O.C., R.L. Taylor and J.Z. Zhu (2005) *The Finite Element Method: Its Basis and Fundamentals*, (6th ed.) Butterworth-Heinemann.

Chapter 9

Finite Volume Approximation in Two Space

9.1 Finite Volume Formulation

In this section we will extend the finite volume formulation we introduced in Section 5.1 on page 84 to two-dimensional problems. The system we will consider is one that allows for a cross derivative such as $\frac{\partial^2 u(x,y)}{\partial x \partial y}$, namely

$$a_{xx} \frac{\partial^2 u(x,y)}{\partial x^2} + a_{yy} \frac{\partial^2 u(x,y)}{\partial y^2} + a_{xy} \frac{\partial^2 u(x,y)}{\partial x \partial y} + a_{yx} \frac{\partial^2 u(x,y)}{\partial y \partial x} \quad (9.1)$$

$$+ \kappa u(x,y) - Q = 0 \quad x, y \in \Omega \quad (9.2)$$

$$u(x,y) = u_0(x,y) \quad x, y \in \partial\Omega_D \quad (9.3)$$

$$\frac{\partial u(x,y)}{\partial n} = u_1(x,y) \quad x, y \in \partial\Omega_N \quad (9.4)$$

$$\partial\Omega_D + \partial\Omega_N = \partial\Omega \quad (9.5)$$

where, as earlier, $\partial\Omega_D$ and $\partial\Omega_N$ are the segments of the boundary $\partial\Omega$ of the domain Ω along which Dirichlet and Neumann boundary conditions are defined respectively.

The formulation of the finite volume method in two space dimensions can be done using more than one approach. The one we have selected is chosen because it is consistent with the method of weighted residuals which we have used as the intellectual glue to hold together the various methods we have introduced thus far. With this in mind, as in the case of earlier methods, we begin with a representation of $u(x,y)$ as

$$u(x,y) \approx \hat{u}(x,y) = \sum_{j=1}^N u_j \phi_j(x,y) \quad (9.6)$$

where, as in Section 8.2 on page 195, we use basis functions defined on triangular elements.

A typical finite element arrangement is seen in [Fig. 9.1](#). The black dots are the nodes. The small circles are the circumcenters of the triangles. The circumcenter is the location of the intersection of the perpendicular bisectors of each side. It is also the center of a circle that passes through all three of the nodes of the triangular element. We will make

reference to circumcenters in the next paragraph.

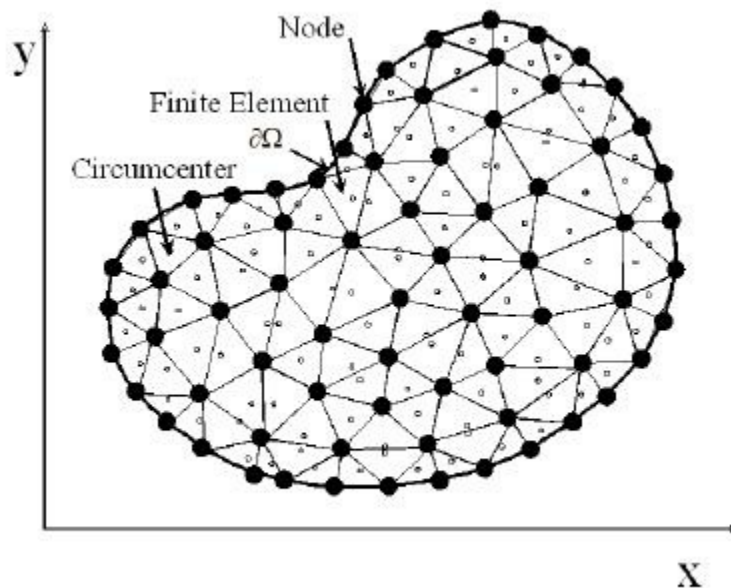


Figure 9.1: Domain of interest subdivided into elements as part of the finite volume formulation presentation. The small circles are the approximate locations of the circumcenters of the triangles.

To begin the development of the finite volume approximation, a series of polygons is constructed within the array of elements as seen in [Fig. 9.2](#). The polygons are created by connecting the circumcenters of each triangle. Note that the sides of the polygons are orthogonal to the sides of the triangles and are continuous across them. The polygons, in the absence of the triangles, are shown in [Fig. 9.3](#). We will now discuss the rationale for the various steps just presented and we will explain, via two examples, how to employ the resulting finite-volume formulation.

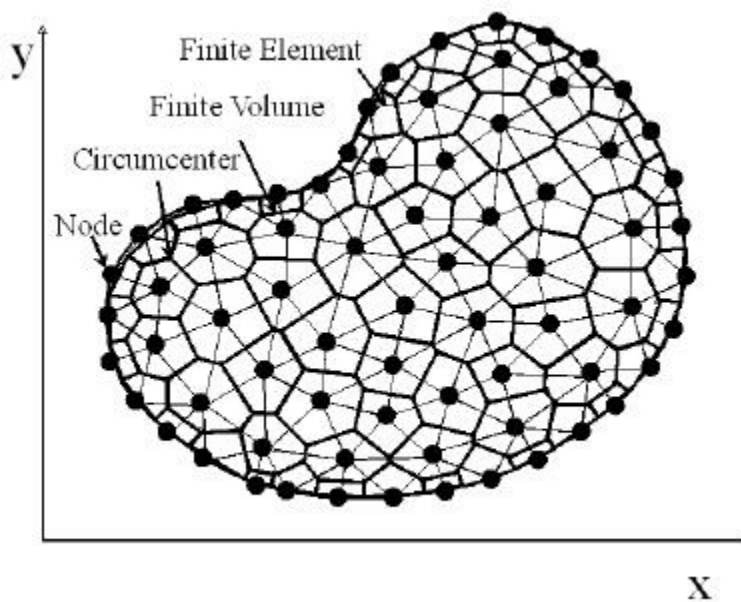


Figure 9.2: Polygonal finite volumes generated from the finite element array shown in [Fig. 9.1](#).

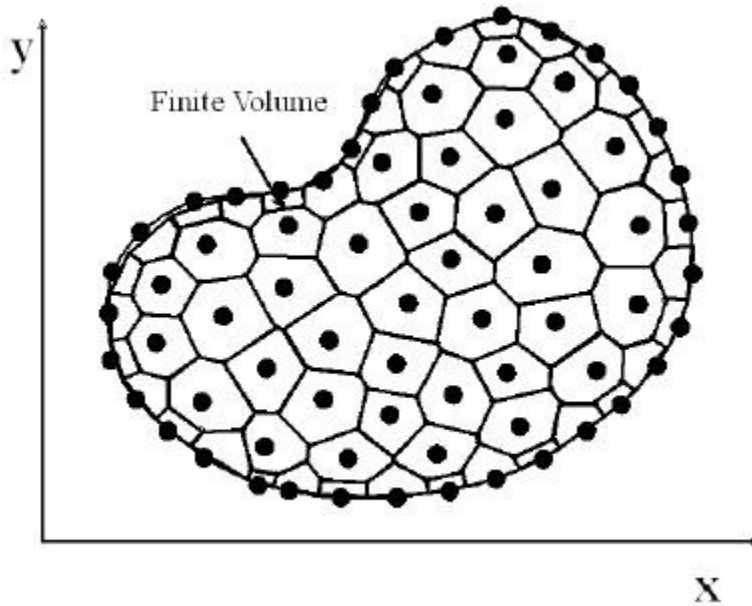


Figure 9.3: Finite volume mesh after removal of the triangular elements.

Once we have the finite element basis functions defined, as presented in [Eq. \(9.6\)](#) and discussed in more detail later, the next step is to write the weighted residual formulation, that is

$$\int_{\Omega} R(x, y) w_i(x, y) d\Omega = 0 \quad i = 1, 2, \dots, N \quad (9.7)$$

which contains the weighting function $w_i(x, y)$. The choice of the weighting function for the two-dimensional finite-volume method is a direct extension of the one-dimensional case discussed in Section 5.4 on page 112.

To visualize our choice, imagine extending a line of unit length vertically at the circumcenter of each triangle as denoted by the letter C in [Fig. 9.5](#). Connect the tops of the resulting lines to form the polygon at a height of 1 unit (this is the surface B in [Fig. 9.5](#)). Finally stretch a hypothetical membrane over the polygon on the top and then enclose each of the vertical lines just constructed with the membrane. The resulting ‘tower’ is the needed weighting function $w_i(x, y)$ and is consistent with the finite element finite-volume array shown in [Fig. 9.6](#). The projection of the weighting function onto the x - y plane is illustrated in [Figs. 9.5](#) and [9.6](#). We will learn why we have selected this functional form for the weighting function a little later.

As we have done in earlier examples, we now expand the weighted residual formulation of [Eq. \(9.1\)](#) to give

$$\begin{aligned}
& \int_{\Omega} \left(a_{xx} \frac{\partial^2 \hat{u}(x, y)}{\partial x^2} + a_{yy} \frac{\partial^2 \hat{u}(x, y)}{\partial y^2} \right) d\Omega + \\
& \int_{\Omega} \left(a_{xy} \frac{\partial^2 \hat{u}(x, y)}{\partial x \partial y} + a_{yx} \frac{\partial^2 \hat{u}(x, y)}{\partial y \partial x} + \kappa \hat{u}(x, y) - Q w_i(x, y) \right) d\Omega \\
& = 0 \quad i = 1, 2, \dots, N.
\end{aligned} \tag{9.8}$$

Now we apply Green's theorem to the second-order terms in [Eq. \(9.8\)](#) and we get

$$\begin{aligned}
& - \int_{\Omega} \left(a_{xx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{xy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) \frac{\partial w_i(x, y)}{\partial x} d\Omega + \\
& - \int_{\Omega} \left(a_{yx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{yy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) \frac{\partial w_i(x, y)}{\partial y} d\Omega \\
& + \int_{\Omega} (\kappa \hat{u}(x, y) - Q) w_i(x, y) d\Omega \\
& + \int_{\partial\Omega} \left(a_{xx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{xy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) n_x w_i(x, y) ds \\
& + \int_{\partial\Omega} \left(a_{yx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{yy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) n_y w_i(x, y) ds \\
& = 0 \quad i = 1, 2, \dots, N
\end{aligned} \tag{9.9}$$

where n_x and n_y are the components of the vector \mathbf{n} that is normal to Ω in the x and y directions respectively.

Consider now the information provided in [Fig. 9.4](#). Here you see a typical finite element with the coordinates x, y, n and τ defined. Note that n is orthogonal to the side of the finite-volume and τ is tangential to it.

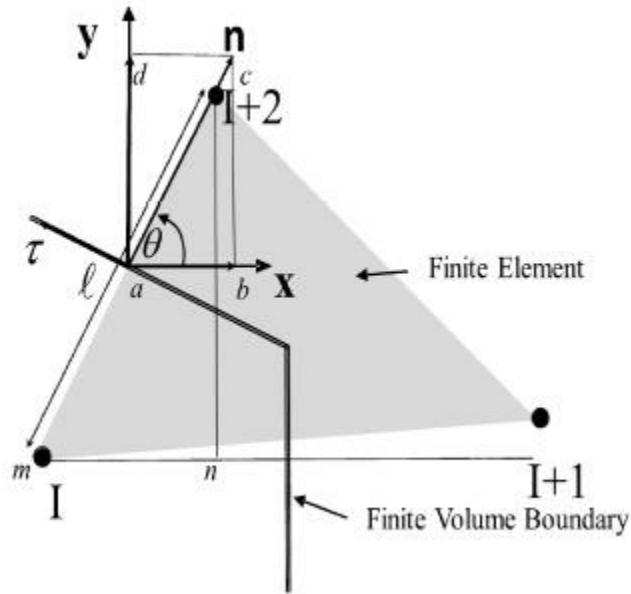


Figure 9.4: Coordinate system relationship associated with finite-volume formulation. Note that the positive angle for θ is counter-clockwise.

Next we chain out the w_i derivatives with respect to τ and n . To see how this is done, recall that by definition

$$\frac{\partial}{\partial x} w_i(x, y) = \nabla w_i(x, y) \cdot \mathbf{i} \quad (9.10)$$

and

$$\frac{\partial}{\partial y} w_i(x, y) = \nabla w_i(x, y) \cdot \mathbf{j}. \quad (9.11)$$

The unit vectors in the x and y coordinate directions are given by \mathbf{i} and \mathbf{j} respectively. We can also use the definition of the gradient operator to write it in the normal-tangential coordinate system

$$\frac{\partial}{\partial n} w_i(x, y) = \nabla w_i(x, y) \cdot \mathbf{n} \quad (9.12)$$

and

$$\frac{\partial}{\partial \tau} w_i(x, y) = \nabla w_i(x, y) \cdot \boldsymbol{\tau} \quad (9.13)$$

where, in this case

$$\nabla w_i(x, y) = \frac{\partial}{\partial n} w_i(x, y) \mathbf{n} + \frac{\partial}{\partial \tau} w_i(x, y) \boldsymbol{\tau} \quad (9.14)$$

and \mathbf{n} and $\boldsymbol{\tau}$ are the unit vectors in the n , and τ coordinate directions respectively.

Now we expand this relationship to give

$$\frac{\partial}{\partial x} w_i(x, y) = \left(\frac{\partial}{\partial n} w_i(x, y) \mathbf{n} + \frac{\partial}{\partial \tau} w_i(x, y) \boldsymbol{\tau} \right) \cdot \mathbf{i}. \quad (9.15)$$

We can now rewrite this expression as

$$\frac{\partial}{\partial x} w_i(x, y) = \left(\frac{\partial}{\partial n} w_i(x, y) \mathbf{n} \cdot \mathbf{i} + \frac{\partial}{\partial \tau} w_i(x, y) \boldsymbol{\tau} \cdot \mathbf{i} \right). \quad (9.16)$$

Similarly, for $\frac{\partial}{\partial y} w_i(x, y)$ we obtain

$$\frac{\partial}{\partial y} w_i(x, y) = \left(\frac{\partial}{\partial n} w_i(x, y) \mathbf{n} \cdot \mathbf{j} + \frac{\partial}{\partial \tau} w_i(x, y) \boldsymbol{\tau} \cdot \mathbf{j} \right). \quad (9.17)$$

As noted earlier and illustrated in [Figs. 9.5](#) and [9.6](#), the weighting function $w_i(x, y)$ is a constant in the τ direction such that the terms of the form of $\frac{\partial}{\partial \tau} w_i(x, y)$ are zero. Using this property, the above equations become

$$\frac{\partial}{\partial x} w_i(x, y) = \frac{\partial}{\partial n} w_i(x, y) \mathbf{n} \cdot \mathbf{i} = -\delta(x, y) |_{\partial\Omega_i} \mathbf{n} \cdot \mathbf{i} \quad (9.18)$$

and

$$\frac{\partial}{\partial y} w_i(x, y) = \frac{\partial}{\partial n} w_i(x, y) \mathbf{n} \cdot \mathbf{j} = -\delta(x, y) |_{\partial\Omega_i} \mathbf{n} \cdot \mathbf{j} \quad (9.19)$$

where $\delta(x, y) |_{\partial\Omega_i}$ is the Dirac delta function defined along $\partial\Omega_i$.

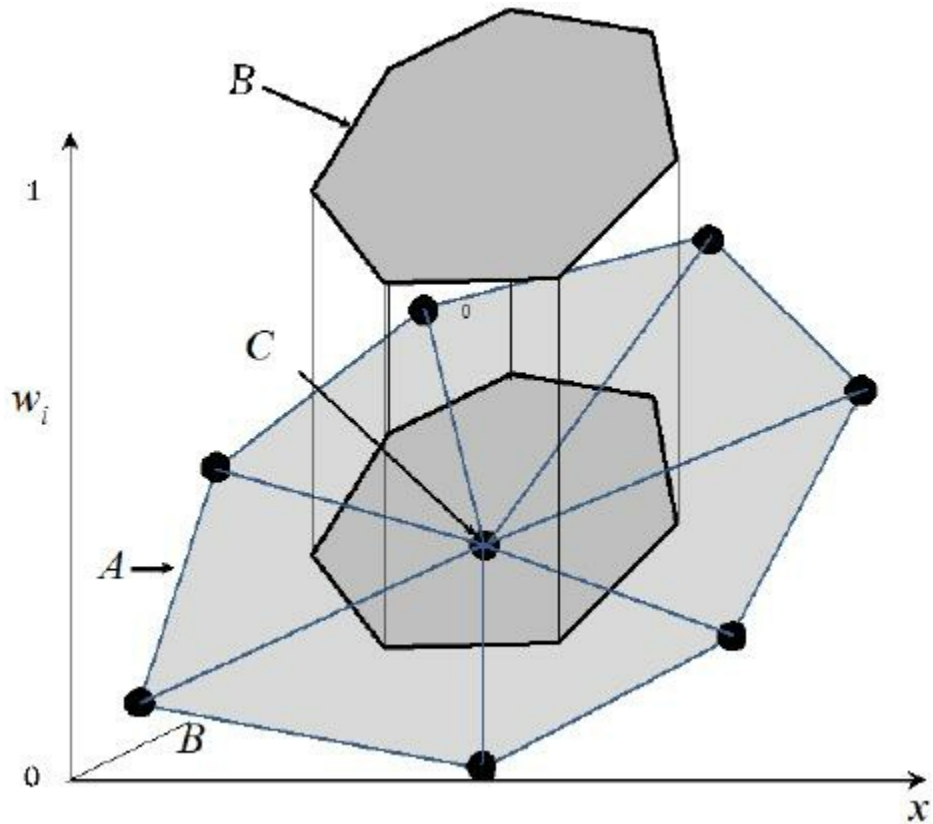


Figure 9.5: The weighting function used in the finite-volume formulation. The top of the function is denoted by the dark-grey polygon.

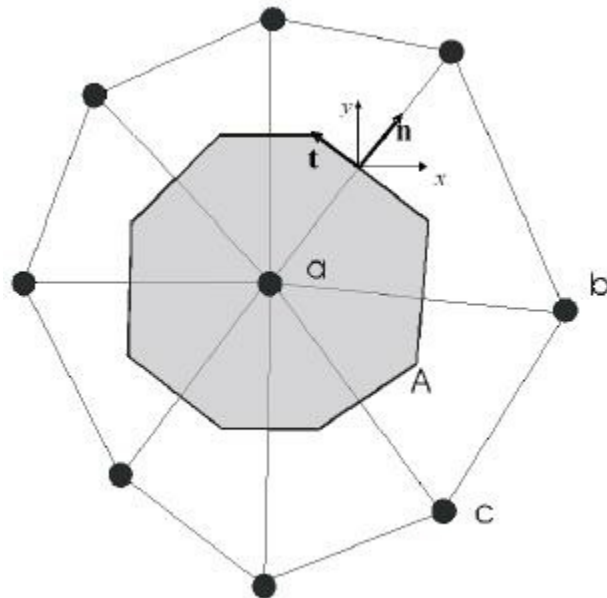


Figure 9.6: Finite volume (gray) and finite elements with illustration of normal \mathbf{n} and tangential $\boldsymbol{\tau}$ coordinate directions.

Now we need to determine the product $\mathbf{n} \cdot \mathbf{j}$. Referring to [Fig. 9.4](#) we see that the vector

from a to c in the n direction can be expressed as

$$\mathbf{n}_{ac} = (x_c - x_a, y_c - y_a). \quad (9.20)$$

To normalize this vector to obtain the unit vector \mathbf{n} , we divide by the length of \mathbf{n}_{ac} , that is

$$\mathbf{n} = \frac{(x_c - x_a, y_c - y_a)}{\sqrt{(x_c - x_a)^2 + (y_c - y_a)^2}}. \quad (9.21)$$

Thus we find the unit vector coordinates for \mathbf{n}_x to be

$$\mathbf{n} \cdot \mathbf{i} \equiv n_x = \frac{x_c - x_a}{\sqrt{(x_c - x_a)^2 + (y_c - y_a)^2}} \quad (9.22)$$

and for \mathbf{n}_y to be

$$\mathbf{n} \cdot \mathbf{j} \equiv n_y = \frac{y_c - y_a}{\sqrt{(x_c - x_a)^2 + (y_c - y_a)^2}}. \quad (9.23)$$

Note that the above calculations could be performed using the coordinates at the locations c , m , and n in lieu of c , a , and b with the same results.

We return to [Eq. \(9.10\)](#) and introduce the relationships developed above to obtain

$$\begin{aligned} & - \int_{\Omega} \left[\left(a_{xx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{xy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) (-\delta(x, y) |_{\partial\Omega_i}) \mathbf{n} \cdot \mathbf{i} \right. \\ & \left. + \left(a_{yx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{yy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) (-\delta(x, y) |_{\partial\Omega_i}) \mathbf{n} \cdot \mathbf{j} \right] d\Omega \\ & + \int_{\Omega} (\kappa \hat{u}(x, y) - Q) w_i(x, y) d\Omega \\ & + \int_{\partial\Omega} \left[\left(a_{xx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{xy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) n_x \right. \\ & \left. + \left(a_{yx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{yy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) n_y \right] w_i(x, y) ds = 0 \\ & i = 1, 2, \dots, N \end{aligned} \quad (9.24)$$

where $\hat{u}(x, y)$ is presented in [Eq. \(9.6\)](#). From the definition of the Dirac delta function, this equation can be rewritten as

(9.25)

$$\begin{aligned}
& \int_{\partial\Omega_i} \left[\left(a_{xx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{xy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) n_x + \left(a_{yx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{yy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) n_y \right] ds \\
& + \int_{\Omega} (\kappa \hat{u}(x, y) - Q) w_i(x, y) d\Omega \\
& + \int_{\partial\Omega} \left[\left(a_{xx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{xy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) n_x \right. \\
& \left. + \left(a_{yx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{yy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) n_y \right] w_i(x, y) ds = 0 \\
& i = 1, 2, \dots, N.
\end{aligned}$$

Notice that integration is now over the finite-volume perimeter and that there are two uses of the notation n_x and n_y . One is associated with the finite volume i and the other with the region Ω .

As a final step, we introduce the triangular basis functions $\phi_j(x, y)$ to give the final form we will use in the following examples

$$\begin{aligned}
& \sum_{e=1}^{E_i} \sum_{j=1}^N \int_{\partial\Omega_e^i} u_j \left(a_{xx} \frac{\partial \phi_j(x, y)}{\partial x} + a_{xy} \frac{\partial \phi_j(x, y)}{\partial y} \right) n_x ds + \\
& \sum_{e=1}^{E_i} \sum_{j=1}^N \int_{\partial\Omega_e^i} u_j \left(a_{yx} \frac{\partial \phi_j(x, y)}{\partial x} + a_{yy} \frac{\partial \phi_j(x, y)}{\partial y} \right) n_y ds + \\
& \sum_{e=1}^{E_i} \sum_{j=1}^N u_j \int_{\Omega_e} (\kappa \phi_j(x, y)) w_i(x, y) d\Omega - \int_{\Omega} Q w_i(x, y) d\Omega + \\
& \int_{\partial\Omega} \left(a_{xx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{xy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) n_x w_i(x, y) ds + \\
& \int_{\partial\Omega} \left(a_{yx} \frac{\partial \hat{u}(x, y)}{\partial x} + a_{yy} \frac{\partial \hat{u}(x, y)}{\partial y} \right) n_y w_i(x, y) ds \\
& = 0 \quad i = 1, 2, \dots, N.
\end{aligned} \tag{9.26}$$

The above development may seem a bit abstract at this point, but how we use this information in application will be clearer after we work through the following two examples:

9.2 Finite Volume Example Problem 1

9.2.1 Problem Definition

Consider the system of equations presented above modified to be descriptive of the

problem presented in [Fig. 9.7](#), namely

$$2 \frac{\partial^2 u(x, y)}{\partial x^2} + \frac{\partial^2 u(x, y)}{\partial y^2} + \kappa u(x, y) - Q = 0 \quad x, y \in \Omega, \quad (9.27)$$

$$u(x, y) = 1 \quad x = 0, \quad y \in [0, 2], \quad (9.28)$$

$$u(x, y) = 0 \quad x = 2, \quad y \in [0, 2], \quad (9.29)$$

$$\frac{\partial u(x, y)}{\partial n} = 0 \quad x \in (0, 2), \quad y = 0, \quad (9.30)$$

$$\frac{\partial u(x, y)}{\partial n} = 0 \quad x \in (0, 2), \quad y = 2. \quad (9.31)$$

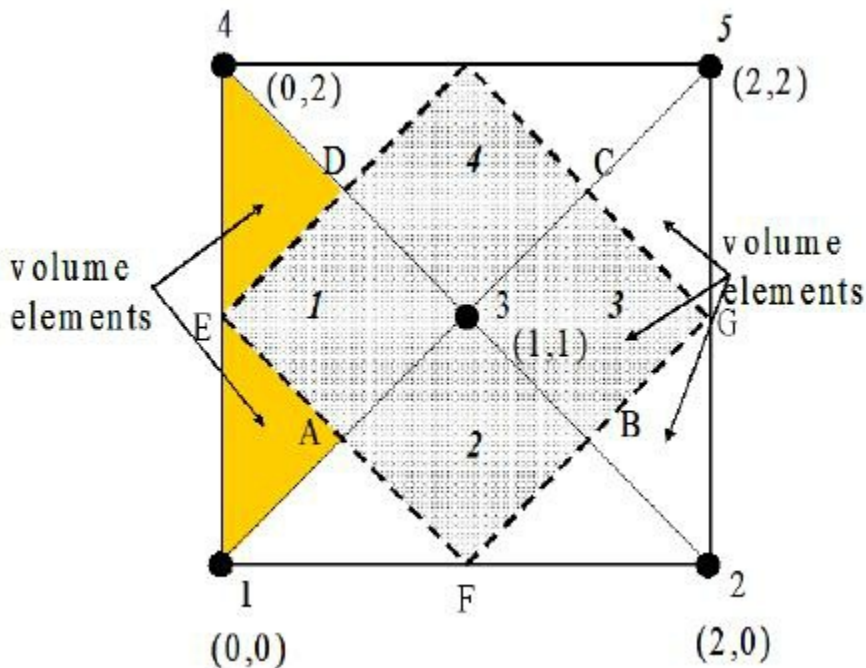


Figure 9.7: Definition drawing of finite volume problem presented as ‘example one’ in [Section 9.2](#).

9.2.2 Weighted Residual Formulation

The domain of interest presented in [Fig. 9.7](#) has **four large triangular elements** and **five finite volumes**. The finite volumes consist of the square of area 2 in the middle and the four triangles in the corners. In other words, **there is one square volume element and four small triangular volume elements**.

We begin by employing [Eq. \(9.26\)](#). We have for our problem

$$\begin{aligned}
& \sum_{e=1}^{E_i} \sum_{j=1}^5 u_j \int_{\partial\Omega_i^e} \left(2 \frac{\partial\phi_j(x,y)}{\partial x} n_x + \frac{\partial\phi_j(x,y)}{\partial y} n_y \right) ds \\
& + \sum_{e=1}^{E_i} \sum_{j=1}^5 u_j \int_{\Omega_e} \left(\kappa\phi_j^e(x,y) \right) w_i(x,y) d\Omega - \int_{\Omega} Q w_i(x,y) d\Omega + \\
& \int_{\partial\Omega} \left[2 \frac{\partial\hat{u}(x,y)}{\partial x} n_x + \frac{\partial\hat{u}(x,y)}{\partial y} n_y \right] w_i(x,y) ds = 0 \quad i = 1, 2, \dots, 5
\end{aligned} \tag{9.32}$$

where the number of elements identified with (attached to) each node i , where i identifies the finite volume of interest, is given by E_i . If we write [Eq. \(9.32\)](#) in matrix form we have

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & 0 \\ a_{21} & a_{22} & a_{23} & 0 & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & 0 & a_{43} & a_{44} & a_{45} \\ 0 & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \end{bmatrix}. \tag{9.33}$$

The number of rows in the coefficient matrix in [Eq. \(9.33\)](#) is five, which corresponds to the number of finite volumes. The number of columns is also five and represents the number of unknown values of u_j at the nodes.

At this point we realize that we actually have only one unknown, namely the value of u_j at node 3, because the values of u_j at other nodes are all provided by the boundary conditions. Thus, due to the imposition of Dirichlet boundary conditions, the rows associated with these known values can be eliminated. As a result this system of equations reduces to

$$\begin{bmatrix} a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ u_3 \\ 1 \\ 0 \end{bmatrix} = [f_3] \tag{9.34}$$

or

$$(a_{33}) (u_3) = f_3 - (a_{31}) (1) - (a_{34}) (1). \tag{9.35}$$

9.2.3 Element Coefficient Matrices

The first step in evaluating the coefficients is to recognize that $e = 1, 2, 3, 4$ and therefore we have four 3×3 element coefficient matrices of the form.

$$A = \begin{bmatrix} a_{I,I} & a_{I,I+1} & a_{I,I+2} \\ a_{I+1,I} & a_{I+1,I+1} & a_{I+1,I+2} \\ a_{I+2,I} & a_{I+2,I+1} & a_{I+2,I+2} \end{bmatrix} \quad (9.36)$$

where I are local coordinates (the concept of local coordinates for triangles was introduced in Section 8.2.1 on page 196). Recall once again that the rows represent the nodes defining where the finite volumes are centered and the columns with the nodes that define the location of the triangular basis functions.

The local coefficient matrix for element e is (where we again assume $\varphi_I(x, y)$ can be represented by φ_I)

$$\begin{bmatrix} \left(\int_{\partial\Omega_I^e} \left(2 \frac{\partial\phi_I}{\partial x} n_x + \frac{\partial\phi_I}{\partial y} n_y \right) ds \right) + \int_{\Omega_I^e} (\kappa\phi_I) w_I d\Omega & \left(\int_{\partial\Omega_I^e} \left(2 \frac{\partial\phi_{I+1}}{\partial x} n_x + \frac{\partial\phi_{I+1}(x,y)}{\partial y} n_y \right) ds \right) + \int_{\Omega_I^e} (\kappa\phi_{I+1}) w_I d\Omega \\ \left(\int_{\partial\Omega_{I+1}^e} \left(2 \frac{\partial\phi_I}{\partial x} n_x + \frac{\partial\phi_I}{\partial y} n_y \right) ds \right) + \int_{\Omega_{I+1}^e} (\kappa\phi_I) w_{I+1} d\Omega & \left(\int_{\partial\Omega_{I+1}^e} \left(2 \frac{\partial\phi_{I+1}}{\partial x} n_x + \frac{\partial\phi_{I+1}}{\partial y} n_y \right) ds \right) + \int_{\Omega_{I+1}^e} (\kappa\phi_{I+1}) w_{I+1} d\Omega \\ \left(\int_{\partial\Omega_{I+2}^e} \left(2 \frac{\partial\phi_I}{\partial x} n_x + \frac{\partial\phi_I}{\partial y} n_y \right) ds \right) + \int_{\Omega_{I+2}^e} (\kappa\phi_I) w_{I+2} d\Omega & \left(\int_{\partial\Omega_{I+2}^e} \left(2 \frac{\partial\phi_{I+1}}{\partial x} n_x + \frac{\partial\phi_{I+1}}{\partial y} n_y \right) ds \right) + \int_{\Omega_{I+2}^e} (\kappa\phi_{I+1}) w_{I+2} d\Omega \\ \left(\int_{\partial\Omega_I^e} \left(2 \frac{\partial\phi_{I+2}}{\partial x} n_x + \frac{\partial\phi_{I+2}}{\partial y} n_y \right) ds \right) + \int_{\Omega_I^e} (\kappa\phi_{I+2}) w_I d\Omega & \\ \left(\int_{\partial\Omega_{I+1}^e} \left(2 \frac{\partial\phi_{I+2}}{\partial x} n_x + \frac{\partial\phi_{I+2}}{\partial y} n_y \right) ds \right) + \int_{\Omega_{I+1}^e} (\kappa\phi_{I+2}) w_{I+1} d\Omega & \\ \left(\int_{\partial\Omega_{I+2}^e} \left(2 \frac{\partial\phi_{I+2}}{\partial x} n_x + \frac{\partial\phi_{I+2}}{\partial y} n_y \right) ds \right) + \int_{\Omega_{I+2}^e} (\kappa\phi_{I+2}) w_{I+2} d\Omega & \end{bmatrix} \quad (9.37)$$

The table relating the global and local numbering for the problem shown in [Fig. 9.7](#):

If the values in the above table are introduced into [Eq. \(9.37\)](#) one obtains for the global coefficient matrix elements

$$a_{31} = \sum_{e=1}^{E_3} \int_{\partial\Omega_3^e} \left(2 \frac{\partial\phi_1(x,y)}{\partial x} n_x + \frac{\partial\phi_1(x,y)}{\partial y} n_y \right) ds + \int_{\Omega_3^e} (\kappa\phi_1(x,y)) w_3(x,y) d\Omega \quad (9.38)$$

$$a_{33} = \sum_{e=1}^{E_3} \int_{\partial\Omega_3^e} \left(2 \frac{\partial\phi_3(x,y)}{\partial x} n_x + \frac{\partial\phi_3(x,y)}{\partial y} n_y \right) ds + \int_{\Omega_3^e} (\kappa\phi_3(x,y)) w_3(x,y) d\Omega \quad (9.39)$$

$$a_{34} = \sum_{e=1}^{E_3} \int_{\partial\Omega_3^e} \left(2 \frac{\partial\phi_4(x,y)}{\partial x} n_x + \frac{\partial\phi_4(x,y)}{\partial y} n_y \right) ds + \int_{\Omega_3^e} (\kappa\phi_4(x,y)) w_3(x,y) d\Omega \quad (9.40)$$

$$f_3 = \int_{\Omega_3^e} Qw_i(x,y) d\Omega \quad (9.41)$$

where the subscripts on the $\phi_i(x,y)$ values are in global coordinates.

9.2.4 Evaluation of the Line Integral

The global coefficient a_{31} contributed by element 1 is calculated for basis function ϕ_1 and weighting function w_3 as follows. First we substitute the values of $I = 3$ and $J = 1$ to give

$$\begin{aligned} & \int_{\partial\Omega_3^1} \left(2 \frac{\partial\phi_1}{\partial x} n_x + \frac{\partial\phi_1}{\partial y} n_y \right) ds \quad (9.42) \\ &= \int_{\overline{DE}} \left(2 \frac{\partial\phi_1}{\partial x} n_x|_{\overline{DE}} + \frac{\partial\phi_1}{\partial y} n_y|_{\overline{DE}} \right) ds \\ &+ \int_{\overline{EA}} \left(2 \frac{\partial\phi_1}{\partial x} n_x|_{\overline{EA}} + \frac{\partial\phi_1}{\partial y} n_y|_{\overline{EA}} \right) ds \end{aligned}$$

where $|_{EA}$ is the segment of side $\partial\Omega_3^e$ that is composed of the line \overline{EA} . and $|_{DE}$ is the segment along the line \overline{DE} . Now we recall the definition of the the basis and obtain their derivatives, that is

$$\phi_I(x,y) = \frac{x(y_{I+1} - y_{I+2}) + y(x_{I+2} - x_{I+1}) + (x_{I+1}y_{I+2} - x_{I+2}y_{I+1})}{\det[P]} \quad (9.43)$$

from which

$$\frac{\partial\phi_I}{\partial x} = \frac{(y_{I+1} - y_{I+2})}{\det[P]} \quad (9.44)$$

and

$$\frac{\partial \phi_I}{\partial y} = \frac{(x_{I+2} - x_{I+1})}{\det[P]} \quad (9.45)$$

Therefore, [Eq. \(9.42\)](#) becomes

$$\begin{aligned} & \int_{\overline{DE}} \frac{2(y_3 - y_4) n_x \Big|_{\overline{DE}} + (x_4 - x_3) n_y \Big|_{\overline{DE}}}{\det[P]} ds \\ & + \int_{\overline{AE}} \frac{2(y_3 - y_4) n_x \Big|_{\overline{EA}} + (x_4 - x_3) n_y \Big|_{\overline{EA}}}{\det[P]} ds. \end{aligned} \quad (9.46)$$

Next observe in [Fig. 9.7](#), the location of the line that defines the side of the triangle connecting the location of node 3, which identifies the volume of interest, and the node 1 which identifies the element basis function of interest. If we now imagine being at the node 3 and direct our attention towards node 1, we observe that the resulting line is given by $(x_3 - x_1, y_3 - y_1)$. We now need the n_x and n_y values to complete the evaluation associated with [Eq. \(9.46\)](#). From [Eq. \(9.22\)](#) we see that the value of n_x is given by

$$\begin{aligned} n_x \Big|_{\overline{DE}} &= \frac{x_4 - x_3}{\sqrt{(x_3 - x_4)^2 + (y_3 - y_4)^2}} \\ &= \frac{0 - 1}{\sqrt{(1)^2 + (1)^2}} \\ &= \frac{-1}{\sqrt{2}}. \end{aligned} \quad (9.47)$$

Similarly for the evaluation of n_y we have

$$\begin{aligned} n_y \Big|_{\overline{DE}} &= \frac{y_4 - y_3}{\sqrt{(x_3 - x_4)^2 + (y_3 - y_4)^2}} \\ &= \frac{2 - 1}{\sqrt{(1)^2 + (1)^2}} \\ &= \frac{1}{\sqrt{2}}. \end{aligned} \quad (9.48)$$

Now we can combine this information to give

$$\begin{aligned}\mathbf{n}|_{\overline{DE}} &= (n_x|_{\overline{DE}}, n_y|_{\overline{DE}}) \\ &= \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right).\end{aligned}\tag{9.49}$$

But we still need to have the information relevant to side \overline{EA} . Following the strategy presented above we get

$$\begin{aligned}n_x|_{\overline{EA}} &= \frac{x_1 - x_3}{\sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2}} \\ &= \frac{0 - 1}{\sqrt{(1)^2 + (1)^2}} \\ &= \frac{-1}{\sqrt{2}}.\end{aligned}\tag{9.50}$$

and

$$\begin{aligned}n_y|_{\overline{EA}} &= \frac{y_1 - y_3}{\sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2}} \\ &= \frac{0 - 1}{\sqrt{(1)^2 + (1)^2}} \\ &= -\frac{1}{\sqrt{2}}.\end{aligned}\tag{9.51}$$

such that

$$\begin{aligned}\mathbf{n}|_{\overline{EA}} &= (n_x|_{\overline{EA}}, n_y|_{\overline{EA}}) \\ &= \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right).\end{aligned}\tag{9.52}$$

We now substitute these values into [Eq. \(9.42\)](#) to give

$$\begin{aligned}
& \int_{\overline{DE}} \frac{2(y_3 - y_4) n_x|_{\overline{DE}} + (x_4 - x_3) n_y|_{\overline{DE}}}{\det[P]} ds & (9.53) \\
& + \int_{\overline{EA}} \frac{2(y_3 - y_4) n_x|_{\overline{EA}} + (x_4 - x_3) n_y|_{\overline{EA}}}{\det[P]} ds \\
& = \int_{\overline{DE}} \frac{2(y_3 - y_4) \frac{-1}{\sqrt{2}} + (x_4 - x_3) \frac{1}{\sqrt{2}}}{\det[P]} ds \\
& + \int_{\overline{EA}} \frac{2(y_3 - y_4) \frac{-1}{\sqrt{2}} + (x_4 - x_3) \frac{-1}{\sqrt{2}}}{\det[P]} ds.
\end{aligned}$$

After substituting the coordinate locations of x_i, y_i we get

$$\begin{aligned}
& \int_{\overline{DE}} \frac{2(y_3 - y_4) \frac{-1}{\sqrt{2}} + (x_4 - x_3) \frac{1}{\sqrt{2}}}{\det[P]} ds & (9.54) \\
& + \int_{\overline{EA}} \frac{2(y_3 - y_4) \frac{-1}{\sqrt{2}} + (x_4 - x_3) \frac{-1}{\sqrt{2}}}{\det[P]} ds \\
& = \int_{\overline{DE}} \frac{2(1 - 2) \frac{-1}{\sqrt{2}} + (0 - 1) \frac{1}{\sqrt{2}}}{\det[P]} ds \\
& + \int_{\overline{EA}} \frac{2(1 - 2) \frac{-1}{\sqrt{2}} + (0 - 1) \frac{-1}{\sqrt{2}}}{\det[P]} ds \\
& = \int_{\overline{DE}} \frac{2 - 1}{\sqrt{2} \det[P]} ds \\
& + \int_{\overline{EA}} \frac{2 + 1}{\sqrt{2} \det[P]} ds.
\end{aligned}$$

We can calculate the $\det[P]$ from the coordinate matrix which, for this element, is defined using the above table as

$$[P] = \begin{bmatrix} x_I & y_I & 1 \\ x_{I+1} & y_{I+1} & 1 \\ x_{I+2} & y_{I+2} & 1 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 \\ x_3 & y_3 & 1 \\ x_4 & y_4 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1.4 & 0 & 1 \\ 1.4 & 1.4 & 1 \end{bmatrix} \quad (9.55)$$

We can determine $\det[P]$ either by calculation from [Eq. \(9.55\)](#), or from the known relationship that $\det[P] = 2A$, presented earlier in Section 8.2.1 on page 196. In either event we determine that $\det[P] = 2$ and [Eq. \(9.54\)](#) becomes

$$\begin{aligned}
& \int_{\overline{DE}} \frac{2-1}{2\sqrt{2}} ds + \int_{\overline{EA}} \frac{2+1}{2\sqrt{2}} ds & (9.56) \\
& \int_{\overline{DE}} \frac{1}{2\sqrt{2}} ds + \int_{\overline{EA}} \frac{2+1}{2\sqrt{2}} ds \\
& = \left(\frac{1}{2\sqrt{2}} \right) |\overline{DE}| + \left(\frac{2+1}{2\sqrt{2}} \right) |\overline{EA}| \\
& = \left(\frac{1}{2\sqrt{2}} \right) \sqrt{2} + \left(\frac{2+1}{2\sqrt{2}} \right) \sqrt{2} \\
& = \frac{4}{4} = 1.
\end{aligned}$$

Thus the line integral component of the coefficient for $I + 1 = 3, J = 1$ for element 1 is 1.0.

As a second example consider $I + 1 = 3, J + 2 = 1$ and substitute this information in [Eq. \(9.42\)](#) for element 2. We obtain

$$\begin{aligned}
& \int_{\partial\Omega_2^2} \left(2 \frac{\partial\phi_1}{\partial x} n_x + \frac{\partial\phi_1}{\partial y} n_y \right) ds & (9.57) \\
& = \int_{\overline{AF}} \left(2 \frac{\partial\phi_1}{\partial x} n_x|_{\overline{AF}} + \frac{\partial\phi_1}{\partial y} n_y|_{\overline{AF}} \right) ds \\
& \quad + \int_{\overline{FB}} \left(2 \frac{\partial\phi_1}{\partial x} n_x|_{\overline{FB}} + \frac{\partial\phi_1}{\partial y} n_y|_{\overline{FB}} \right) ds \\
& = \left(2 \frac{\partial\phi_1}{\partial x} n_x|_{\overline{AF}} + \frac{\partial\phi_1}{\partial y} n_y|_{\overline{AF}} \right) |\overline{AF}| \\
& \quad + \left(2 \frac{\partial\phi_1}{\partial x} n_x|_{\overline{FB}} + \frac{\partial\phi_1}{\partial y} n_y|_{\overline{FB}} \right) |\overline{FB}| \\
& = \left[2 \left(-\frac{1}{2} \right) \left(-\frac{\sqrt{2}}{2} \right) + \left(-\frac{1}{2} \right) \left(-\frac{\sqrt{2}}{2} \right) \right] \frac{\sqrt{2}}{2} \\
& \quad + \left[2 \left(-\frac{1}{2} \right) \left(\frac{\sqrt{2}}{2} \right) + \left(-\frac{1}{2} \right) \left(-\frac{\sqrt{2}}{2} \right) \right] \frac{\sqrt{2}}{2} \\
& = \left[\left(\frac{3}{4} \right) \sqrt{2} - \frac{1}{4} \sqrt{2} \right] \frac{\sqrt{2}}{2} \\
& = 0.5
\end{aligned}$$

so the coefficient for $I + 1 = 3, J + 2 = 2$ is 0.5 in element 2.

In this, the third example, we consider the corner finite volume, that is the one associated with node 2, element 3, and the basis function identified with node 2, that is $I + 2 = 2$ and $J + 2 = 2$. Substitution in [Eq. \(9.42\)](#) gives

$$\int_{\partial\Omega_2^3} \left(2 \frac{\partial\phi_2}{\partial x} n_x + \frac{\partial\phi_2}{\partial y} n_y \right) ds. \quad (9.58)$$

The finite volume perimeter of interest is

$$\partial\Omega_2^3 = \overline{2G} + \overline{GB}. \quad (9.59)$$

so [Eq. \(9.58\)](#) becomes, where $\overline{2G}$ is the segment connecting node 2 and point G,

$$\int_{\partial\Omega_2^3} \left(2 \frac{\partial\phi_2}{\partial x} n_x + \frac{\partial\phi_2}{\partial y} n_y \right) ds \quad (9.60)$$

$$= \int_{\overline{2G}} \left(2 \frac{\partial\phi_2}{\partial x} n_x + \frac{\partial\phi_2}{\partial y} n_y \right) ds$$

$$+ \int_{\overline{GB}} \left(2 \frac{\partial\phi_2}{\partial x} n_x + \frac{\partial\phi_2}{\partial y} n_y \right) ds. \quad (9.61)$$

The normal vector to $\overline{2G}$ is given by (see [Eqs. \(9.47\)](#) to [\(9.49\)](#))

$$\begin{aligned} \mathbf{n}|_{\overline{2G}} &= \left(n_x|_{\overline{2G}}, n_y|_{\overline{2G}} \right) \\ &= (0, 1) \end{aligned} \quad (9.62)$$

and

$$\begin{aligned} \mathbf{n}|_{\overline{GB}} &= \left(n_x|_{\overline{GB}}, n_y|_{\overline{GB}} \right) \\ &= \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right). \end{aligned} \quad (9.63)$$

The next step is to obtain the derivatives of the basis function for $\phi_2(x, y)$ which we obtain from [Eq. \(9.44\)](#) and [9.45](#), that is

$$\begin{aligned} \frac{\partial\phi_2}{\partial x} &= \frac{y_5 - y_3}{\det[P]} \\ &= \frac{1}{2} \end{aligned} \quad (9.64)$$

and

$$\begin{aligned}\frac{\partial\phi_2}{\partial y} &= \frac{x_3 - x_5}{\det[P]} \\ &= -\frac{1}{2}.\end{aligned}\tag{9.65}$$

We can now evaluate the equation of interest, namely

$$\begin{aligned}& \int_{\overline{2G}} \left(2 \frac{\partial\phi_2}{\partial x} n_x + \frac{\partial\phi_2}{\partial y} n_y \right) ds \\ & + \int_{\overline{GB}} \left(2 \frac{\partial\phi_2}{\partial x} n_x + \frac{\partial\phi_2}{\partial y} n_y \right) ds \\ &= \left(2 \frac{\partial\phi_2}{\partial x} n_x \Big|_{\overline{2G}} + \frac{\partial\phi_2}{\partial y} n_y \Big|_{\overline{2G}} \right) |\overline{2G}| \\ & + \left(2 \frac{\partial\phi_2}{\partial x} n_x \Big|_{\overline{GB}} + \frac{\partial\phi_2}{\partial y} n_y \Big|_{\overline{GB}} \right) |\overline{GB}| \\ &= \left[2 \left(\frac{1}{2} \right) (1) + \left(-\frac{1}{2} \right) (0) \right] (1) \\ & + \left[2 \left(\frac{1}{2} \right) \left(-\frac{\sqrt{2}}{2} \right) + \left(-\frac{1}{2} \right) \left(\frac{\sqrt{2}}{2} \right) \right] \left(\frac{\sqrt{2}}{2} \right) \\ &= 1 - \left(\frac{3\sqrt{2}}{4} \right) \left(\frac{\sqrt{2}}{2} \right) \\ &= 0.25.\end{aligned}\tag{9.66}$$

Thus, the coefficient located at $I + 2 = 2, J + 2 = 2$ in element 3 is 0.25.

Finally, we consider the case of the contribution to coefficient a_{33} attributable to element 3, that is $I + 1 = 3$ and $J + 1 = 3$. The two segments of the boundary of the finite volume of interest are

$$\partial\Omega_3^3 = \overline{BG} + \overline{GC}.\tag{9.67}$$

The term of interest is

(9.68)

$$\begin{aligned}
& \int_{\partial\Omega_3^3} \left(2 \frac{\partial\phi_3}{\partial x} n_x + \frac{\partial\phi_3}{\partial y} n_y \right) ds \\
&= \int_{\overline{BG}} \left(2 \frac{\partial\phi_3}{\partial x} n_x + \frac{\partial\phi_3}{\partial y} n_y \right) ds \\
&+ \int_{\overline{GC}} \left(2 \frac{\partial\phi_3}{\partial x} n_x + \frac{\partial\phi_3}{\partial y} n_y \right) ds \\
&= \left(2 \frac{\partial\phi_3}{\partial x} n_x|_{\overline{BG}} + \frac{\partial\phi_3}{\partial y} n_y|_{\overline{BG}} \right) |\overline{BG}| \\
&+ \left(2 \frac{\partial\phi_3}{\partial x} n_x|_{\overline{GC}} + \frac{\partial\phi_3}{\partial y} n_y|_{\overline{GC}} \right) |\overline{GC}|.
\end{aligned}$$

The derivatives of the basis function $\phi_3(x, y)$ are

$$\frac{\partial\phi_3}{\partial x} = \frac{y_2 - y_5}{\det[P]} \quad (9.69)$$

and

$$\frac{\partial\phi_3}{\partial y} = \frac{x_5 - x_2}{\det[P]}. \quad (9.70)$$

The values of $n_x|_{\overline{BG}}$ and $n_x|_{\overline{GC}}$ are computed as follows:

$$\begin{aligned}
\mathbf{n}|_{\overline{BG}} &= \left(n_x|_{\overline{BG}}, n_y|_{\overline{GC}} \right) \\
&= \frac{1}{\sqrt{(x_2 - x_3)^2 + (y_2 - y_3)^2}} ((x_2 - x_3), (y_2 - y_3)) \\
&= \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right)
\end{aligned} \quad (9.71)$$

$$\begin{aligned}
\mathbf{n}|_{\overline{GC}} &= (n_x|_{\overline{GC}}, n_y|_{\overline{GC}}) & (9.72) \\
&= \frac{1}{\sqrt{(x_5 - x_3)^2 + (y_5 - y_3)^2}} ((x_5 - x_3), (y_5 - y_3)) \\
&= \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right).
\end{aligned}$$

We can now evaluate our coefficient

$$\begin{aligned}
&\left(2 \frac{\partial \phi_3}{\partial x} n_x|_{\overline{BG}} + \frac{\partial \phi_3}{\partial y} n_y|_{\overline{BG}} \right) |\overline{BG}| & (9.73) \\
&+ \left(2 \frac{\partial \phi_3}{\partial x} n_x|_{\overline{GC}} + \frac{\partial \phi_3}{\partial y} n_y|_{\overline{GC}} \right) |\overline{GC}| \\
&= \left[2(-1) \left(\frac{\sqrt{2}}{2} \right) \right] \frac{\sqrt{2}}{2} + \left[2(-1) \left(\frac{\sqrt{2}}{2} \right) \frac{\sqrt{2}}{2} \right] \\
&= -1 - 1 = -2.
\end{aligned}$$

Thus the coefficient $i = 3, j = 3$ of element 3 is -2

If we evaluate all the coefficients in each element coefficient matrix we obtain for the line-integral component:

Element 1

$$\begin{bmatrix} 0.25 & -1.0 & 0.75 \\ 1.0 & -2.0 & 1.0 \\ 0.75 & -1.0 & 0.25 \end{bmatrix} \quad (9.74)$$

Element 2

$$\begin{bmatrix} -0.25 & -0.50 & 0.75 \\ 0.50 & -1.0 & 0.50 \\ 0.75 & -0.50 & -0.25 \end{bmatrix} \quad (9.75)$$

Element 3

$$\begin{bmatrix} 0.25 & -1.0 & 0.75 \\ 1.0 & -2.0 & 1.0 \\ 0.75 & -1.0 & 0.25 \end{bmatrix} \quad (9.76)$$

Element 4

$$\begin{bmatrix} -0.25 & -0.50 & 0.75 \\ 0.50 & -1.0 & 0.50 \\ 0.75 & -0.50 & -0.25 \end{bmatrix}. \quad (9.77)$$

9.2.5 Evaluation of the Area Integral

We need to also evaluate integrals in [Eq. \(9.32\)](#) of the form

$$\int_{\Omega_e} (\kappa \phi_j(x, y) d\Omega) w_i(x, y) d\Omega = \int_{\Omega_e} \kappa \phi_j(x, y) d\Omega$$

for the system shown in [Fig. 9.7](#). As an example of how this can be done consider element 1 shown in [Fig. 9.8](#) which is the same as element 1 in [Fig. 9.7](#) rotated clockwise 45 degrees so that the sides of the element line up with the x and y axes. The value of the integral does not change by doing this, but the presentation is simplified. Let values of i and j both be 3.

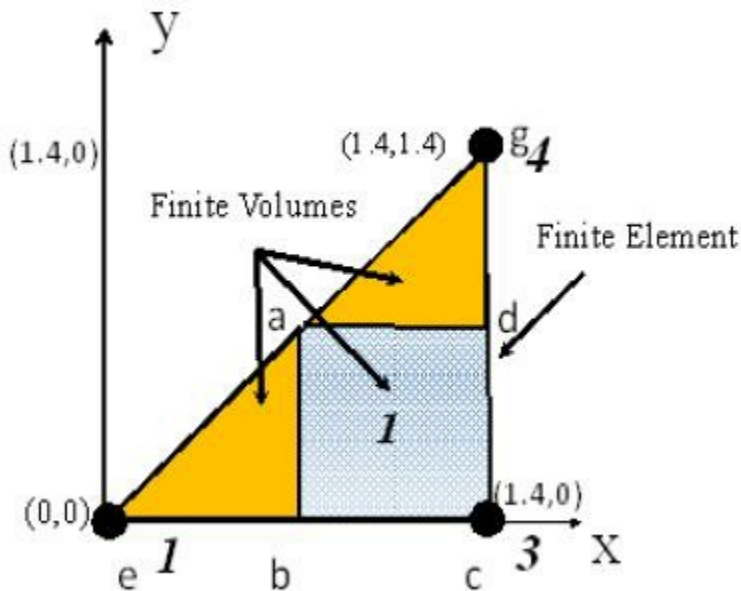


Figure 9.8: Representation of the finite element configuration for calculation of the area integral.

The finite volume of interest is the shaded square area in [Fig. 9.8](#). In this example, the length of the line \overline{bc} is one half the length of the side \overline{ec} . The basis function in element 1 identified with node 3 has a value of 1 at node 3 and zero at nodes 1 and 4. Thus, at the point c the value of ϕ_3 is 1.0, at the points b and d the values are 0.5, and at e and g the values are 0.

Two approaches to obtaining this integral are presented. The first is based upon geometric considerations and the second on upon numerical integration.

Geometric Approach

To obtain the integral of ϕ_3 located under the shaded square, one first calculates the area

under the triangle Δecg . This is obtained from the equation given in Eq. (8.83) and the integration formula 8.121 Using this information, the formula for evaluating the integral is

$$\int_{\Omega_{ecg}} \phi_3 d\omega = \frac{\det [P]_{ecg} / 2}{3} \quad (9.78)$$

where $\det [P]_{ecg}$ is the determinant of the matrix $[P]$ defined for triangle ecg .

Next subtract the integral

$$\int_{\Omega_{eba}} \phi_3 d\omega = \frac{\det [P]_{eba} / 2}{6} = \frac{1}{24} \quad (9.79)$$

and the integral

$$\int_{\Omega_{adg}} \phi_3 d\omega = \frac{\det [P]_{adg} / 2}{6} = \frac{1}{24} \quad (9.80)$$

from Eq. (9.78). The result is the portion of the integral in Eq. (9.78) under the weighting function associated with the finite volume located at 3. This calculation provides the element coefficient value at $i = 3, j = 3$. More specifically the coefficient value is

$$\begin{aligned} a_{33} &= \int_{\Omega_3^1} \phi_3(x, y) d\omega = \frac{(\det \Delta ecg) / 2}{3} - \frac{(\det \Delta eba) / 2}{6} - \frac{(\det \Delta adg) / 2}{6} \quad (9.81) \\ &= \frac{1}{3} - \frac{0.25}{6} - \frac{0.25}{6} = \frac{1}{4}. \end{aligned}$$

To obtain the value for coefficient at $i = 3, j = 4$, one uses the value obtained in Eq. (9.79) and for coefficient at $i = 3, j = 1$ the value is given by Eq. (9.80). Evaluation of all of the integrals of the form $\int_{\Omega_i} \phi_j(x, y) d\omega$ for element 1 gives the coefficient matrix:

$$\begin{bmatrix} 1/6 & 1/24 & 1/24 \\ 1/24 & 1/4 & 1/24 \\ 1/24 & 1/24 & 1/6 \end{bmatrix}. \quad (9.82)$$

Because of the geometry, the coefficient matrix for each element, in terms of the local coordinates $I, I + 1$, and $I + 2$, is the same in this example.

Numerical Integration Approach

While the above strategy works for the topology of the special case we are considering, it is awkward to use in a more general case. An alternative approach uses the isoparametric transformation concept which we introduced in Section 8.3. In this approach we transform our coordinate system from (x, y) to (ξ, η) where $-1 \leq \xi, \eta \leq 1$.

The first step is to formulate the transformation equations. As earlier we proceed as follows:

$$x = \sum_{i=1}^{i=4} x_i \psi_i (\xi, \eta) \equiv P (\xi, \eta) \quad (9.83)$$

and

$$y = \sum_{i=1}^{i=4} y_i \psi_i (\xi, \eta) \equiv Q (\xi, \eta) \quad (9.84)$$

where, referring to Fig. 8.17 we have

$$\psi_1 (\xi, \eta) = \frac{1}{4} (1 - \xi) (1 - \eta) \quad (9.85)$$

$$\psi_2 (\xi, \eta) = \frac{1}{4} (1 + \xi) (1 - \eta) \quad (9.86)$$

$$\psi_3 (\xi, \eta) = \frac{1}{4} (1 + \xi) (1 + \eta) \quad (9.87)$$

$$\psi_4 (\xi, \eta) = \frac{1}{4} (1 - \xi) (1 + \eta) . \quad (9.88)$$

For the problem at hand we need to transform our integral from the (x, y) system to the (ξ, η) system. The new integral takes the form

$$\int_{\Omega_i^e} \phi_i (x, y) dx = \int_{\xi=-1}^{\xi=1} \int_{\eta=-1}^{\eta=1} \phi_i [P (\xi, \eta), Q (\xi, \eta)] |J (\xi, \eta)| d\xi d\eta \quad (9.89)$$

where $J(\xi, \eta)$ is the Jacobian of the transformation and defined as

$$\det [J (\xi, \eta)] = \det \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \end{bmatrix} . \quad (9.90)$$

In our example, the square area of interest is such that the (x, y) coordinates are colinear with (ξ, η) . This simplifies the problem since some of the terms in the transformation vanish and we get for $\frac{\partial x}{\partial \xi}$

$$\begin{aligned}
\frac{\partial x}{\partial \xi} &= x_1 \frac{\partial}{\partial \xi} \psi_1(\xi, \eta) + x_2 \frac{\partial}{\partial \xi} \psi_2(\xi, \eta) + x_3 \frac{\partial}{\partial \xi} \psi_3(\xi, \eta) + x_4 \frac{\partial}{\partial \xi} \psi_4(\xi, \eta) \quad (9.91) \\
&= \frac{1}{4} \{(0)[-(1-\eta)] + (0.35)(1-\eta) + (0.35)(1+\eta) + (0)[-(1+\eta)]\} \\
&= \frac{1}{4} [(0.35)(1-\eta) + (0.35)(1+\eta)] \\
&= \frac{0.7}{2} = 0.35
\end{aligned}$$

and for $\frac{\partial y}{\partial \xi}$

$$\begin{aligned}
\frac{\partial y}{\partial \xi} &= y_1 \frac{\partial}{\partial \xi} \psi_1(\xi, \eta) + y_2 \frac{\partial}{\partial \xi} \psi_2(\xi, \eta) + y_3 \frac{\partial}{\partial \xi} \psi_3(\xi, \eta) + y_4 \frac{\partial}{\partial \xi} \psi_4(\xi, \eta) \quad (9.92) \\
&= \frac{1}{4} \{(0)[-(1-\eta)] + (0)(1-\eta) + (0.35)(1+\eta) + (0.35)[-(1+\eta)]\} \\
&= \frac{1}{4} (0.7)(1+\eta) - (0.7)(1+\eta) \\
&= 0.
\end{aligned}$$

Using a similar strategy for the other elements of the Jacobian we get

$$\det [J] = \det \begin{bmatrix} 0.35 & 0 \\ 0 & 0.35 \end{bmatrix} = 0.125. \quad (9.93)$$

The integral appearing in [Eq. \(9.89\)](#) now takes the form

$$\begin{aligned}
\int_{\Omega_i^c} \phi_i(x, y) dx &= \int_{\xi=-1}^{\xi=1} \int_{\eta=-1}^{\eta=1} \phi_i[P(\xi, \eta), Q(\xi, \eta)] |J(\xi, \eta)| d\xi d\eta \quad (9.94) \\
&= \int_{\xi=-1}^{\xi=1} \int_{\eta=-1}^{\eta=1} \phi_i[P(\xi, \eta), Q(\xi, \eta)] (0.12) d\xi d\eta.
\end{aligned}$$

So, for the case of $\phi_i(x, y) = \phi_3(x, y)$ we have

$$\int_{\Omega_i^c} \phi_3(x, y) dx = \int_{\xi=-1}^{\xi=1} \int_{\eta=-1}^{\eta=1} \phi_3[P(\xi, \eta), Q(\xi, \eta)] (0.12) d\xi d\eta. \quad (9.95)$$

The next step is to determine $\phi_3[P(\xi, \eta), Q(\xi, \eta)]$. To do this we need to write out the definition of $\phi_3(x, y)$ which is obtained from the definition we have used earlier, namely

$$\phi_I(x, y) = \frac{x(y_{I+1} - y_{I+2}) + y(x_{I+2} - x_{I+1}) + (x_{I+1}y_{I+2} - x_{I+2}y_{I+1})}{\det[P]} \quad (9.96)$$

Substitution for the values of $I, I + 2$, and $I + 3$ from [Fig. 9.8](#) into this expression we get

$$\phi_3 [P(\xi, \eta), Q(\xi, \eta)] = \frac{(P)(y_4 - y_1) + (Q)(x_1 - x_4) + (x_4y_1 - x_1y_4)}{\det[P]} \quad (9.97)$$

which, after providing the location values, gives

$$\phi_3(\xi, \eta) = \frac{(P)(\sqrt{2} - 0) + (Q)(0 - \sqrt{2}) + ((\sqrt{2})(0) - (0)(\sqrt{2}))}{\det[P]} \quad (9.98)$$

or

$$\phi_3(\xi, \eta) = \frac{(P)(\sqrt{2}) - (Q)(\sqrt{2})}{2} \quad (9.99)$$

$$\int_{\Omega_3^c} \phi_3(x, y) dx = \int_{\xi=-1}^{\xi=1} \int_{\eta=-1}^{\eta=1} \phi_3(P(\xi, \eta), Q(\xi, \eta)) (0.12) d\xi d\eta \quad (9.100)$$

$$\begin{aligned} &= \int_{\xi=-1}^{\xi=1} \int_{\eta=-1}^{\eta=1} \frac{(P)(\sqrt{2}) - (Q)(\sqrt{2})}{2} (0.125) d\xi d\eta \\ &= \int_{\xi=-1}^{\xi=1} \int_{\eta=-1}^{\eta=1} \left(\frac{\sqrt{2}}{2} \right) \left(\sum_{i=1}^{i=4} x_i \psi_i(\xi, \eta) - \sum_{i=1}^{i=4} y_i \psi_i(\xi, \eta) \right) (0.125) \\ &\quad d\xi d\eta. \end{aligned}$$

This looks pretty messy, but there is an easy solution to the evaluation of this expression. Recall that we introduced Gaussian quadrature in Section 3.1.6 Let us try one Gauss point to see how well this procedure approximates this integral. The value of the single Gauss point is (0, 0) and the weighting is 2. Using this information, [Eq. \(9.100\)](#) becomes

$$\begin{aligned}
\int_{\Omega_3^c} \phi_3(x, y) dx &\approx \int_{\xi=-1}^{\xi=1} \int_{\eta=-1}^{\eta=1} \frac{(\sqrt{2})}{2} \left(\sum_{i=1}^{i=4} x_i \psi_i(\xi, \eta) - \sum_{i=1}^{i=4} y_i \psi_i(\xi, \eta) \right) (0.12) d\xi d\eta \\
&= 2\sqrt{2} [x_1 \psi_1(0, 0) + x_2 \psi_2(0, 0) + x_3 \psi_3(0, 0) + x_4 \psi_4(0, 0) \\
&\quad - y_1 \psi_1(0, 0) - y_2 \psi_2(0, 0) - y_3 \psi_3(0, 0) - y_4 \psi_4(0, 0)] (0.12) \\
&= 2\sqrt{2} \left[\left(\frac{\sqrt{2}}{2} \right) (0.25) + (\sqrt{2}) (0.25) + (\sqrt{2}) (0.25) + \left(\frac{\sqrt{2}}{2} \right) 0.25 \right. \\
&\quad \left. - (0) (0.25) - (0) (0.25) - \left(\frac{\sqrt{2}}{2} \right) (0.25) - \left(\frac{\sqrt{2}}{2} \right) (0.25) \right] (0.12) \\
&= 0.25.
\end{aligned}$$

This is the same value we obtained from geometric consideration in [Eq. \(9.81\)](#). Note that this approach is general and holds for any finite volume mesh.

9.2.6 Global Matrix Assembly

We are now in a position to evaluate our global coefficients a_{31} , a_{33} and a_{34} . We do this using the element matrices presented earlier and [Table 9.1](#) relating the global and local matrix coefficients. For example, the line integral global values for a_{31} , a_{33} and a_{34} we obtain are as shown in the following table:

Table 9.1 Table relating global and local coordinates.

		Global					
		1	2	3	4	5	Element
		I		I+1	I+2		1
Local	I+2	I	I+1				2
			I+2	I+1		I	3
				I+1	I	I+2	4

		element				
		1	2	3	4	
		a_{31}	1.0	0.5		
coefficient	a_{33}	-2	-1	-2	-1	
		a_{34}	1.0			0.5

For the area integral we get

		element			
		1	2	3	4

		1	2	3	4
	a_{31}	0.125	0.125		
coefficient	a_{33}	0.25	0.25	0.25	0.25
	a_{34}	0.125			0.125

The complete line-integral global matrix is

$$\begin{bmatrix} 0 & 0.75 & -1.5 & 0.75 & 0 \\ 0.75 & 0 & -1.5 & 0 & 0.75 \\ 1.5 & 1.5 & -6.0 & 1.5 & 1.5 \\ 0.75 & 0 & -1.5 & 0 & 0.75 \\ 0 & 0.75 & -1.5 & 0.75 & 0 \end{bmatrix} \quad (9.102)$$

and the complete area integral global matrix is

$$\begin{bmatrix} 0.33 & .042 & 0.084 & 0.042 & 0 \\ 0.042 & 0.33 & 0.084 & 0 & 0.042 \\ 0.25 & 0.25 & 1.0 & 0.25 & 0.25 \\ 0.042 & 0 & 0.084 & 0.33 & 0.042 \\ 0 & 0.042 & 0.084 & 0.042 & 0.33 \end{bmatrix}. \quad (9.103)$$

The combined final coefficient matrix is

$$\begin{bmatrix} 0.33 & 0.79 & -1.4 & 0.79 & 0 \\ 0.79 & 0.33 & -1.4 & 0 & 0.79 \\ 1.7 & 1.7 & -5.0 & 1.7 & 1.7 \\ 0.79 & 0 & -1.42 & 0.33 & 0.79 \\ 0 & 0.79 & -1.42 & 0.79 & 0.33 \end{bmatrix}. \quad (9.104)$$

Collecting the information from the above table and assuming $\kappa = 1$, we obtain

$$a_{31} = 1.7 \quad (9.105)$$

$$a_{33} = -5 \quad (9.106)$$

$$a_{34} = 1.75. \quad (9.107)$$

Returning to our equation we have

$$\begin{aligned} (a_{33})(u_3) &= f_3 - (a_{31})(1) - (a_{34})(1) \\ -5u_3 &= 0 - (1.7)(1) - (1.7)(1) \\ u_3 &= 0.68. \end{aligned} \quad (9.108)$$

If κ is taken as 0, the result becomes $u_3 = 0.5$.

9.3 Finite Volume Example Problem Two

9.3.1 Problem Definition

Example problem 2 is provided to illustrate the application of the finite volume approach when a *cross derivative is involved*. The problem is generated by rotation of the problem presented as example 1 and is shown in [Fig. 9.9](#). The equations in the rotated system exhibit a cross derivative. The solution to both the first and second problems should be the same for the case of $\kappa = 0$:

(9.109)

$$1.5 \frac{\partial^2 u(x, y)}{\partial x^2} - 0.5 \frac{\partial^2 u(x, y)}{\partial x \partial y} + 1.5 \frac{\partial^2 u(x, y)}{\partial y^2} - Q = 0 \quad x, y \in \Omega$$

$$u(x, y) = 1 \quad \text{line 1 - 4} \quad (9.110)$$

$$u(x, y) = 0 \quad \text{line 2 - 5} \quad (9.111)$$

$$\frac{\partial u(x, y)}{\partial n} = 0 \quad \text{line 1 - 2} \quad (9.112)$$

$$\frac{\partial u(x, y)}{\partial n} = 0 \quad \text{line 4 - 5.} \quad (9.113)$$

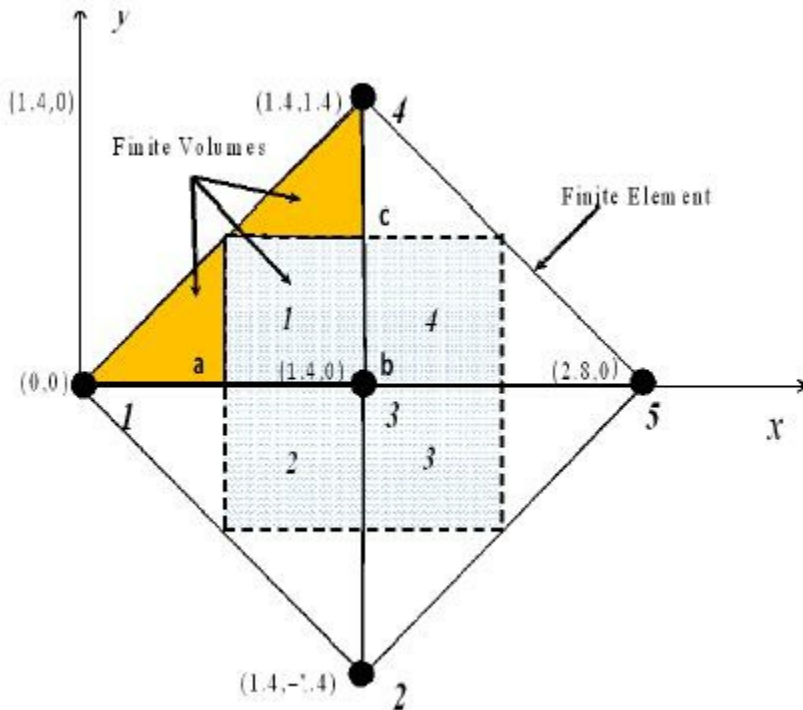


Figure 9.9: Definition diagram for example two of the finite volume methods considered in Section 9.3.

The domain of interest is presented in [Fig. 9.9](#), and as in example problem 1, has four triangular elements and five finite volumes. The finite volumes consist of the square in the middle and the four triangles in the corners (note that each corner volume is actually made up of two smaller elements).

9.3.2 Weighted Residual Formulation

Beginning with [Eq. \(9.109\)](#) we have, after substituting our approximation $\hat{u}(x, y)$ for $u(x, y)$, applying the method of weighted residuals using $w_i(x, y)$ as the weighting function, and finally, employing Green's theorem to the second-order term

(9.114)

$$\sum_{e=1}^{E_i} \sum_{j=1}^5 u_j \int_{\Omega_i^e} \left(-1.5 \frac{\partial \phi_j}{\partial x} \frac{\partial w_i}{\partial x} + 0.5 \frac{\partial \phi_j}{\partial y} \frac{\partial w_i}{\partial x} + 0.5 \frac{\partial \phi_j}{\partial x} \frac{\partial w_i}{\partial y} - 1.5 \frac{\partial \phi_j}{\partial y} \frac{\partial w_i}{\partial y} \right) d\omega - \int_{\Omega} Q w_i d\Omega + \int_{\partial \Omega} \frac{\partial \hat{u}}{\partial n} w_i(x, y) ds = 0 \quad i = 1, 2, \dots, 5$$

where, as earlier in [Section 9.2](#) on page 246, the number of elements identified with (attached to) each node is given by E_i .

Let us now return to our example problem. We can write [Eq. \(9.114\)](#) in matrix form for the problem presented in [Fig. 9.9](#) as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & 0 \\ a_{21} & a_{22} & a_{23} & 0 & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & 0 & a_{43} & a_{44} & a_{45} \\ 0 & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \end{bmatrix}. \quad (9.115)$$

where the coefficients a_{ij} are given by

(9.116)

$$a_{ij} = \sum_{e=1}^{E_i} \int_{\Omega_i^e} \left(-1.5 \frac{\partial \phi_j}{\partial x} \frac{\partial w_i}{\partial x} + 0.5 \frac{\partial \phi_j}{\partial y} \frac{\partial w_i}{\partial x} + 0.5 \frac{\partial \phi_j}{\partial x} \frac{\partial w_i}{\partial y} - 1.5 \frac{\partial \phi_j}{\partial y} \frac{\partial w_i}{\partial y} \right) d\omega.$$

Note that each row of the matrix corresponds to a finite volume and each column contains information from all finite element nodes connected to that finite volume. Thus, for example, a_{15} is zero because the volume centered at node 1 does not have a connection to the finite element node 5.

At this point we realize, as in the first example, that we have only one unknown, namely node 3, because Dirichlet conditions are specified at all the other nodes and thus the values there are known. The resulting system of equations reduces to

$$\begin{bmatrix} a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ u_3 \\ 1 \\ 0 \end{bmatrix} = [f_3] \quad (9.117)$$

or

$$(a_{33}) (u_3) = f_3 - (a_{31}) (1) - (a_{34}) (1). \quad (9.118)$$

The coefficient a_{31} , is of the form (where, as earlier, the first index, in this case 3, is the element number and the second index 1 is the connecting node).

$$a_{31} = \sum_{e=1}^{E_3} \int_{\partial\Omega_3^e} \left(\left\{ 1.5 \frac{\partial\phi_1}{\partial x} - 0.5 \frac{\partial\phi_1}{\partial y} \right\} n_x + \left\{ 1.5 \frac{\partial\phi_1}{\partial y} - 0.5 \frac{\partial\phi_1}{\partial x} \right\} n_y \right) ds. \quad (9.119)$$

9.3.3 Element Coefficient Matrices

Let us evaluate this coefficient for element 1. As in example problem 1, we need to consider two sides of the finite volume centered at node 3, namely the side \overline{ab} and \overline{bc} . From

$$\frac{\partial\phi_J}{\partial x} = \frac{y_{J+1} - y_{J+2}}{\det[P]} \quad (9.120)$$

we obtain for $\frac{\partial\phi_1}{\partial x}$

$$\begin{aligned} \frac{\partial\phi_1}{\partial x} &= \frac{y_3 - y_4}{\det[P]} & (9.121) \\ &= \frac{0 - 1.4}{2} \\ &= \frac{-1}{\sqrt{2}} \end{aligned}$$

Similarly,

$$\frac{\partial\phi_J}{\partial y} = \frac{x_{J+2} - x_{J+1}}{\det[P]} \quad (9.122)$$

which gives for $\frac{\partial\phi_1}{\partial y}$

$$\begin{aligned}\frac{\partial\phi_1}{\partial y} &= \frac{x_4 - x_3}{\det[P]} & (9.123) \\ &= \frac{1 - 1}{2} \\ &= 0\end{aligned}$$

where $\det[P]$ is given by [Eq. \(9.56\)](#).

Using [Eqs. \(9.22\)](#) and [9.23](#) on page 243 we obtain

$$n_x|_{\overline{ab}} = \frac{x_1 - x_3}{\sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2}} = \frac{-\sqrt{2}}{\sqrt{2}} = -1 \quad (9.124)$$

$$n_x|_{\overline{bc}} = \frac{x_4 - x_3}{\sqrt{(x_3 - x_4)^2 + (y_3 - y_4)^2}} = \frac{0}{\sqrt{2}} = 0 \quad (9.125)$$

$$n_y|_{\overline{ab}} = \frac{y_1 - y_3}{\sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2}} = \frac{0}{\sqrt{2}} = 0 \quad (9.126)$$

$$n_y|_{\overline{bc}} = \frac{y_4 - y_3}{\sqrt{(x_3 - x_4)^2 + (y_3 - y_4)^2}} = \frac{\sqrt{2}}{\sqrt{2}} = 1. \quad (9.127)$$

We now substitute [Eqs. 9.121–9.127](#) into [Eq. \(9.119\)](#) and obtain for the coefficient a_{31} of element 1

$$\begin{aligned}a_{31} &= \int_{\partial\Omega_3^1} \left(\left\{ 1.5 \frac{\partial\phi_1}{\partial x} - 0.5 \frac{\partial\phi_1}{\partial y} \right\} n_x|_{ab} + \left\{ 1.5 \frac{\partial\phi_1}{\partial y} - 0.5 \frac{\partial\phi_1}{\partial x} \right\} n_y|_{ab} \right) ds \\ &\quad + \int_{\partial\Omega_3^1} \left(\left\{ 1.5 \frac{\partial\phi_1}{\partial x} - 0.5 \frac{\partial\phi_1}{\partial y} \right\} n_x|_{bc} + \left\{ 1.5 \frac{\partial\phi_1}{\partial y} - 0.5 \frac{\partial\phi_1}{\partial x} \right\} n_y|_{bc} \right) ds \\ &= \left[\left\{ (1.5) \left(\frac{-1}{\sqrt{2}} \right) - (0.5)(0) \right\} (-1) \left(\frac{\sqrt{2}}{2} \right) + \left\{ (1.5)(0) - (0.5) \left(\frac{-1}{\sqrt{2}} \right) \right\} (0) \left(\frac{\sqrt{2}}{2} \right) \right] \\ &\quad + \left\{ (1.5) \left(\frac{-1}{\sqrt{2}} \right) - (0.5)(0) \right\} (0) \left(\frac{\sqrt{2}}{2} \right) + \left\{ (1.5)(0) - (0.5) \left(\frac{-1}{\sqrt{2}} \right) \right\} (1) \left(\frac{\sqrt{2}}{2} \right) \\ &= \frac{1.5}{2} + \frac{0.5}{2} = 1.\end{aligned} \quad (9.128)$$

Notice that this coefficient, namely that for a_{31} for element 1, is equivalent to the element coefficient at $I + 1, J$ in element 1. Returning to element coefficient matrix 1 in example

problem 1, (the matrix noted as [Eq. \(9.74\)](#) on page 256) you will observe that the coefficient at $I + 1, J$ has the value 1.0. In fact all the coefficient values in all the element matrices are the same for the line-integral part of problem 1 and Problem 2. In fact, since [Table 9.1](#) on page 249 that relates the global coordinates to the local coordinates is the same for both example problems, the global matrix for the line-integral part is also the same for both example problems. This is a result of the peculiar nature of the specific example problems considered here and would not generally be the case. As a result, it is not necessary to detail the coefficient evaluation of the matrix equation for this problem.

9.3.4 Evaluation of the Source Term

The remaining term has the following form

$$f_i = \int_{\Omega_i^e} Q w_i(x, y) d\Omega \quad (9.129)$$

which for constant Q over each finite volume becomes

$$f_i = Q_i \int_{\Omega_i^e} w_i(x, y) d\Omega. \quad (9.130)$$

As an example, for element 1 and finite volume 3 we obtain

$$\begin{aligned} f_3 &= Q_3 \int_{\Omega_3^1} w_3(x, y) d\Omega \\ &= Q_3 \det[P] \end{aligned} \quad (9.131)$$

where $[P]$ is identified with element 1.

If we let $Q = 0$. we have for our equation

$$\begin{aligned} (a_{33}) (u_3) &= f_3 - a_{31} (1) - a_{34} (1) \\ (-6) (u_3) &= 0 - (1.5) (1) - (1.5) (1) \\ u_3 &= 0.50 \end{aligned} \quad (9.132)$$

which is the same answer we obtained in the original system before rotation. If we let $Q_3 = 1$ we obtain

$$\begin{aligned} f_3 &= \sum_{e=1}^4 Q_3 \int_{\Omega_3^e} w_3(x, y) d\Omega \\ &= 2 \end{aligned} \quad (9.133)$$

and

$$\begin{aligned}(a_{33}) (u_3) &= f_3 - a_{31} (1) - a_{34} (1) && (9.134) \\ -6u_3 &= 2 - (1.5) (1) - (1.5) (1) \\ u_3 &= 0.17.\end{aligned}$$

The above provides the protocol for solving time-independent problems using the finite-volume method for problems with variable coefficients and cross-derivatives. Extension to time dependent problems follows the strategy to be provided in Chapter 10.

9.4 Chapter Summary

The finite volume method is presented. A weighted residual strategy is used that is consistent with the finite element method. The formulation of the element coefficient matrices and their collection into the global coefficient matrix is analogous to the procedure used earlier with the finite element method. Two examples are presented in detail. The second is the equivalent of the first but for the rotation of the coordinate system so that a cross-derivative appears in the equation.

9.5 Problems

1. [Fig. 9.10](#) illustrates five finite volumes and four finite elements. The finite volume associated with node 5 is square and shaded. The other finite volumes are triangular and centered at the nodes 1–4. Assume the element coefficient matrices are as follows (e is the number of the element)

$$e = 1 \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad (9.135)$$

$$e = 2 \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \quad (9.136)$$

$$e = 3 \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \quad (9.137)$$

$$e = 4 \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} \quad (9.138)$$

The relationship between the global and local element node numbers is given below. The top row of numbers are the global node numbers. The second row of numbers contains the corresponding element node numbers for the first element, element 1 with element entries a_{ij} .

	Global				Element
	1	2	3	4	5
	1			2	3
Local	1	3			2
		3	2		1
				1	3

		2	1	5	4 (d)
--	--	---	---	---	-------

The global coefficient matrix will have 5 rows and 5 columns as shown below. Your task is to populate this matrix with the values from the four element matrices shown in [Eqs. \(9.135\)–\(9.140\)](#). One coefficient is completed as an example:

				$a_{13} + b_{12}$

2. Please fill in the spaces in the columns with an X where appropriate. For example in the first row we ask what methods generate a 4×4 element coefficient matrix. Finite volume may do so and isoparametric always does, so a X is placed in the first and third columns. A blank is not counted one way or the other.

Attribute	FV	FE	Iso
method generates a 4×4 element coefficient matrix	X		X
uses prism with polygonal cross section for weighting function			
uses basis function as weighting function			
employs basis functions defined on triangular subspaces			
uses basis functions defined on a natural coordinate system			
approximation uses Green's theorem (integration by parts)			
elements are defined by straight lines connecting nodes			
uses Jacobian matrix in formulation			
conserves extensive property (for example mass)			
type two (specified flux) boundary conditions are			
an integral part of the approximate equation			
employs the Dirac delta function in the equation formulation			
approximation of a derivative is constant over an element			
elements are, in general, four-sided shapes with			
straight sides and no sides parallel			
can easily provide small elements in the			
neighborhood of singularities, for example, a well			
creates a 4×4 element coefficient matrix			
uses method of weighted residuals in formulation			

FV = is finite Volume,
 FE = is finite Element,

iso = is isoparametric

1. Consider the finite volume mesh presented in [Fig. 9.11](#) the following figure. the letters are the finite volumes, the italic numbers are the elements and the rest are nodes. Please answer the following:
 - a. In the global matrix, how many non-zero coefficients will be found in the first line of the matrix _____?
 - b. How many zero coefficients will be found in the matrix of the first line of the global matrix equation _____?
 - c. What equations (row numbers) will contain information from node 1 _____?
 - d. Given the equations

(9.139)

$$\begin{aligned}
 & \sum_{e=1}^{E_i} \sum_{j=1}^n u_j \int_{\partial\Omega_i^e} \left(2 \frac{\partial \phi_j(x, y)}{\partial x} n_x + \frac{\partial \phi_j(x, y)}{\partial y} n_y \right) ds \\
 & + \sum_{e=1}^{E_i} \sum_{j=1}^5 u_j \int_{\Omega_e} (\kappa \phi_j^e(x, y)) w_i(x, y) d\Omega - \int_{\Omega} Q w_i(x, y) d\Omega + \\
 & \int_{\partial\Omega} \left[2 \frac{\partial \hat{u}(x, y)}{\partial x} n_x + \frac{\partial \hat{u}(x, y)}{\partial y} n_y \right] w_i(x, y) ds \\
 & = 0 \quad i = 1, 2, \dots, 5
 \end{aligned}$$

(9.139)

what is the value of E_8 _____?

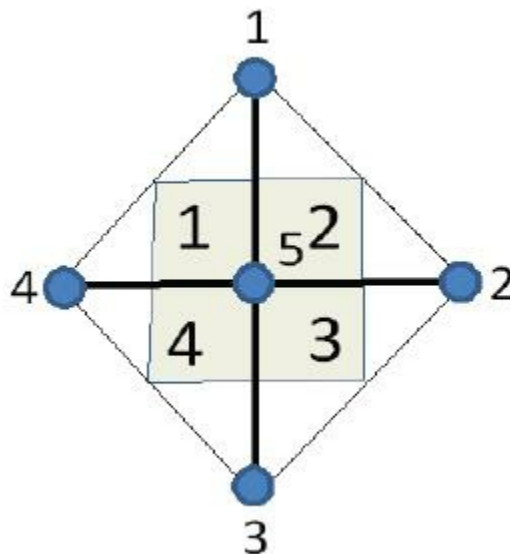


Figure 9.10: The figure illustrates four triangular elements and five finite volumes. The finite volume associated with node 5 is shaded. The finite volume numbers correspond

to the nodal numbers and the bold numbers are associated with the finite elements.

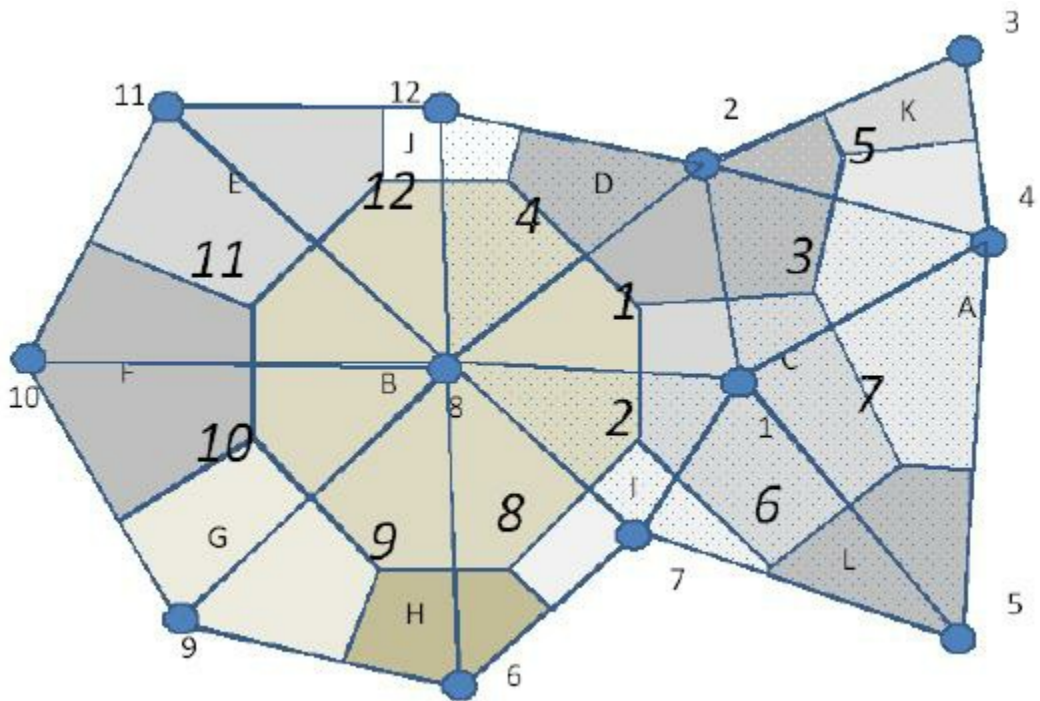


Figure 9.11: Definition sketch for question 9.5.

Bibliography

Cai, Z. (1991) "On the finite volume element method," *Numer. Math.* 58, 713-735.

Eymard, R., T. Gallouët and R. Herbin (2000) "Finite volume methods, in Techniques of Scientific Computing," Part III, *Handbook of Numerical Analysis*, P.G Ciarlet and J.L. Lions, editors, North-Holland, Amsterdam.

Li, R., Z. Chen, and W. Wu (2000) "Generalized difference methods for differential equations," *Pure and Applied Mathematics*, Marcel Dekker, New York.

Chapter 10

Initial Boundary-Value Problems

In this section we will extend our two-space dimensional analysis to consider time-dependent problems. The equation we will consider is the convective-diffusion equations, that is

$$a \frac{\partial u}{\partial t} + b \nabla^2 u + c \nabla u = 0 \quad (10.1)$$

with boundary and initial conditions

$$u(x, y, 0) = u_0(x, y) \quad x, y \in \Omega \quad (10.2)$$

$$u(x, y, t) = u_b(x, y) \quad x, y \in \Omega_D \quad (10.3)$$

$$\frac{\partial u(x, y, t)}{\partial n} = \frac{\partial u(x, y, t)}{\partial n_b} \quad x, y \in \Omega_N. \quad (10.4)$$

We now proceed in a fashion analogous to that presented in Section 6.5 on page 158. The first step is to define the approximation to $u(x, y)$, namely $\hat{u}(x, y)$

$$\hat{u}(x, y) = \sum_{j=1}^N u_j(t) \phi_j(x, y). \quad (10.5)$$

Note that the unknown coefficients $u_j(t)$ are now a function of time and the basis functions $\phi_i(x, y)$ are a function only of space.

We now define a residual as

$$a \frac{\partial \hat{u}}{\partial t} + b \nabla^2 \hat{u} + c \nabla \hat{u} = R(x, y, t). \quad (10.6)$$

From Galerkin's method we have

$$\int_{\Omega} R(x, y, t) \phi_i(x, y) d\Omega = 0 \quad i = 1, \dots, N. \quad (10.7)$$

Substituting [Eq. \(10.6\)](#) into [Eq. \(10.7\)](#) we have

$$\int_{\Omega} \left(a \frac{\partial \hat{u}}{\partial t} + b \nabla^2 \hat{u} + c \nabla \hat{u} \right) \phi_i(x, y) d\Omega = 0 \quad i = 1, \dots, N. \quad (10.8)$$

If we now substitute [Eq. \(10.5\)](#) into [Eq. \(10.8\)](#) we obtain

(10.9)

$$\sum_{j=1}^N \int_{\Omega} \left(a \frac{du_j(t)}{dt} \phi_j(x, y) + bu_j \nabla^2 \phi_j(x, y) + cu_j \nabla \phi_j(x, y) \right) \phi_i(x, y) d\Omega = 0 \quad i = 1, \dots, N.$$

$$\sum_{j=1}^N \int_{\Omega} \left[a \frac{du_j(t)}{dt} \phi_j(x, y) \phi_i(x, y) - bu_j \nabla \phi_j(x, y) \nabla \phi_i(x, y) + cu_j \nabla \phi_j(x, y) \phi_i(x, y) \right] d\Omega \quad (10.10)$$

$$+ \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_i(x, y) dl = 0 \quad i = 1, \dots, N. \quad (10.11)$$

Following our earlier protocol we now apply Green's theorem to the second-order term (it is also possible to apply Green's theorem to the first order in space term). You obtain

We can now write this expression in matrix form to obtain

$$[M] \{u\} + [N] \left\{ \frac{du}{dt} \right\} + \{f\} = 0 \quad (10.12)$$

where

(10.13)

$$m_{ij} = - \int_{\Omega} (b \nabla \phi_j(x, y) \nabla \phi_i(x, y) - c \nabla \phi_j(x, y) \phi_i(x, y)) d\Omega$$

(10.14)

$$n_{ij} = \int_{\Omega} a \phi_j(x, y) \phi_i(x, y) d\Omega$$

(10.15)

$$f_i = \int_{\partial\Omega} \frac{\partial \hat{u}}{\partial n} \phi_i(x, y) dl.$$

Of the above equations, only [Eq. \(10.14\)](#) is new; the rest we have encountered earlier. Let us consider the specific case of triangular elements. We have the following for one element:

$$\int_{\Omega_e} a \phi_j(x, y) \phi_i(x, y) d\Omega. \quad (10.16)$$

Note that, assuming the coefficient a to be a constant over the element we have

$$a \int_{\Omega_e} \phi_j(x, y) \phi_i(x, y) d\Omega. \quad (10.17)$$

We can integrate the term appearing in [Eq. \(10.17\)](#) using the relationship given earlier for integrating terms of this kind; the relationship is

$$\int_{\Omega_e} \phi_I^{m_I}(x, y) \phi_{I+1}^{m_{I+1}}(x, y) \phi_{I+2}^{m_{I+2}}(x, y) d\Omega_e = 2A \frac{m_I! m_{I+1}! m_{I+2}!}{(m_I + m_{I+1} + m_{I+2} + 2)!}. \quad (10.18)$$

The coefficients m_I , m_{I+1} , and m_{I+2} are for [Eq. \(10.17\)](#) 1, 1, and 0. The zero value exists because we have only two linear basis functions represented in [Eq. \(10.17\)](#). We can now evaluate the integral as follows.

$$\begin{aligned} & \int_{\Omega_e} \phi_j(x, y) \phi_i(x, y) d\Omega && (10.19) \\ &= \int_{\Omega_e} \phi_I^1(x, y) \phi_{I+1}^1(x, y) d\Omega_e \\ &= 2A_e \frac{1!1!0!}{(1 + 1 + 0 + 2)!} \\ &= 2A_e \frac{1}{4!} \\ &= \frac{A_e}{12}. \end{aligned}$$

For the case of $i = j$ we have

$$\begin{aligned}
& \int_{\Omega_e} \phi_j(x, y) \phi_j(x, y) d\Omega & (10.20) \\
& = \int_{\Omega_e} \phi_I^1 \phi_I^1 d\Omega_e \\
& = \int_{\Omega_e} \phi_I^2 d\Omega_e \\
& = 2A_e \frac{2!}{(2+0+0+2)!} \\
& = 2A_e \frac{2!}{4!} \\
& = \frac{A_e}{6}.
\end{aligned}$$

The matrix equation [Eq. \(10.12\)](#) represents a set of N ordinary differential equations in time. We now approximate the time derivative using finite difference methods to obtain

$$[M] \{u\}_{t+\Delta t} + [N] \left(\frac{\{u\}_{t+\Delta t} - \{u\}_t}{\Delta t} \right) + \{f\} = 0. \quad (10.21)$$

Grouping terms we obtain

$$\left([M] + \frac{[N]}{\Delta t} \right) \{u\}_{t+\Delta t} = \frac{[N]}{\Delta t} \{u\}_t - \{f\} \quad (10.22)$$

which can be simplified to read

$$[C] \{u\}_{t+\Delta t} = \{g\} \quad (10.23)$$

where

$$[C] = \left([M] + \frac{[N]}{\Delta t} \right) \quad (10.24)$$

and

$$\{g\} = \frac{[N]}{\Delta t} \{u\}_t - \{f\}. \quad (10.25)$$

Thus we can solve [Eq. \(10.23\)](#) symbolically as

$$\{u\}_{t+\Delta t} = [C]^{-1} \{g\}. \quad (10.26)$$

10.1 Mass Lumping

While the form of the $[N]$ matrix as presented can be, and is used in practice, there are alternative strategies. One approach is to lump the matrix $[N]$. Lumping is defined as follows

$$n_{ii} = \sum_{j=1}^N \int_{\Omega} a \phi_j(x, y) \phi_i(x, y) d\Omega \quad (10.27)$$

$$n_{ij} = 0 \quad i \neq j. \quad (10.28)$$

Under this strategy the diagonal elements of $[N]$ are the only non-zero elements.

10.2 Chapter Summary

In this chapter we extended the finite difference and finite element methods introduced earlier to tackle the transient case.

10.3 Problems

1. Assume that you want to solve the equation, New

$$a \frac{\partial u}{\partial t} + b \nabla^2 u = 0 \quad (10.29)$$

using linear the approximation

$$u(x, y, t) \approx \hat{u}(x, y, t) = \sum_{j=1}^n u_j \phi_j(x, y, t) \quad (10.30)$$

- a. assuming you want to employ a linear approximation, that is one where there are no second degree terms in $\phi_j(x, y, t)$, what two geometric forms can $\phi_j(x, y, t)$ take?
- b. sketch one three-dimensional finite element of each form with views showing the $x - y$ and $x - t$ planes.
- c. explain why the bandwidth of the coefficient matrix increases when the formulation based upon $\hat{u}(x, y, t) = \sum_{j=1}^n u_j \phi_j(x, y, t)$ is used rather than $\hat{u}(x, y, t) = \sum_{j=1}^n u_j(t) \phi_j(x, y)$.

Bibliography

Allen, M.B. III and E.L. Isaacson (1988) *Numerical Analysis for Applied Science*, John Wiley and Sons, New York.

Bathe, K.J. (1976) *Numerical Methods in Finite Element Analysis*, Prentice-Hall.

Becker, E.B., G.F. Carey, and J.T. Oden (1981) *Finite Elements: An Introduction*, Vol. 1., Prentice Hall, New Jersey New York.

Celia, M.A. and W.G. Gray (1992) *Numerical Methods for Differential Equations: Fundamental Concepts for Scientific and Engineering Applications*, Prentice-Hall, Englewood Cliffs.

Collatz, L. (1966) *The Numerical Treatment of Differential Equations*, Springer Verlag, New York.

Finlayson, B.A. (1972) *The Method of Weighted Residuals and Variational*

Principles, Academic Press, New York.

Hildebrand, F.B. (1968) *Finite-Difference Equations and Simulations*, Prentice Hall, New York.

Lapidus, L. and G.F. Pinder (1982) *Numerical Solution of Partial Differential Equations in Science and Engineering*, Wiley-Interscience, John Wiley and Sons, MitNew York.

Mitchell, A.R. and D.F. Griffiths (1980) *The Finite Difference Method in Partial Differential Equations*, John Wiley and Sons.

Morton, K.W. and D.F. Mayers (2005) *Numerical Solution of Partial Differential Equations, An Introduction*, Cambridge University Press.

Reddy, J.N. (2006) *An Introduction to the Finite Element Method*, McGraw-Hill, New York.

Richtmeyer, R.F. (1957) *Difference Methods for Initial-Value Problems*, John Wiley and Sons, New York.

Smith, G.D. (1985) *Numerical Solution of Partial Differential Equations: Finite Difference Methods* 3rd ed., Oxford University Press, Oxford.

Strang, G. and G. Fix (1973) *Analysis of the Finite Element Method*, Prentice-Hall, Englewood Cliffs.

Zienkiewicz, O.C., R.L. Taylor and J.Z. Zhu (2005) *The Finite Element Method: Its Basis and Fundamentals*, 6th ed. Butterworth-Heinemann.

Chapter 11

Boundary-Value Problems in Three Space

In this chapter we will extend the concepts introduced earlier for the solution of problems in two space dimensions to problems in three space dimensions. We begin with the finite difference approximations.

11.1 Finite Difference Approximations

The point of departure for the approximation of three-dimensional problems using finite difference methods is the material found in Section 7 on page 167. [Eq. \(7.5\)](#) which represented a function in the (x, y) coordinate system using Lagrange polynomials is now expanded to consider three space dimensions and is written as

$$u(x, y, z) = \sum_{i=0}^n \sum_{j=0}^m \sum_{k=0}^n \ell_i^n(x) \ell_j^m(y) \ell_k^p(z) u_{i,j} + E_{nmp} \quad (11.1)$$

where k is the index and p is the order of the Lagrange polynomial in the z direction.

We can now obtain an approximation of the derivatives we normally encounter in the solution of the spatially-dependent partial-differential equations of interest to us. For the first order derivatives we have

$$\left. \frac{\partial u}{\partial x} \right|_{i,j,k} = \frac{1}{2\Delta x} [u_{i+1,j,k} - u_{i-1,j,k}] + \frac{\partial}{\partial x} E_{i,j,k} \quad (11.2)$$

$$\left. \frac{\partial u}{\partial y} \right|_{i,j,k} = \frac{1}{2\Delta y} [u_{i,j+1,k} - u_{i,j-1,k}] + \frac{\partial}{\partial y} E_{i,j,k} \quad (11.3)$$

$$\left. \frac{\partial u}{\partial z} \right|_{i,j,k} = \frac{1}{2\Delta z} [u_{i,j,k+1} - u_{i,j,k-1}] + \frac{\partial}{\partial z} E_{i,j,k} \quad (11.4)$$

and for the second derivatives we obtain

$$\left. \frac{\partial^2 u}{\partial x^2} \right|_{i,j,k} = \frac{1}{\Delta x^2} [u_{i-1,j,k} - 2u_{i,j,k} + u_{i+1,j,k}] + \frac{\partial^2}{\partial x^2} E_{i,j,k} \quad (11.5)$$

$$\left. \frac{\partial^2 u}{\partial y^2} \right|_{i,j,k} = \frac{1}{\Delta y^2} [u_{i,j-1,k} - 2u_{i,j,k} + u_{i,j+1,k}] + \frac{\partial^2}{\partial y^2} E_{i,j,k} \quad (11.6)$$

$$\left. \frac{\partial^2 u}{\partial z^2} \right|_{i,j,k} = \frac{1}{\Delta z^2} [u_{i,j,k-1} - 2u_{i,j,k} + u_{i,j,k+1}] + \frac{\partial^2}{\partial z^2} E_{i,j,k}. \quad (11.7)$$

Now, let us use these approximations in the equation

$$\begin{aligned} & \alpha_{xx}(x, y, z) \frac{\partial^2 u(x, y, z)}{\partial x^2} + \alpha_{yy}(x, y, z) \frac{\partial^2 u(x, y, z)}{\partial y^2} + \alpha_{zz}(x, y, z) \frac{\partial^2 u(x, y, z)}{\partial z^2} \\ & - \beta_x(x, y, z) \frac{\partial u(x, y, z)}{\partial x} - \beta_y(x, y, z) \frac{\partial u}{\partial y} - \beta_z(x, y, z) \frac{\partial u}{\partial z} + u(x, y, z) = 0. \end{aligned} \quad (11.8)$$

where α_{xx} , α_{yy} , α_{zz} , β_x , β_y , and β_z are known coefficients which may be a function of space. Substitution of the approximations for the node $(i\Delta x, j\Delta y, k\Delta z)$ in [Eq. \(11.8\)](#) yields

$$\begin{aligned} & \frac{\alpha_{xx} \ i, j, k}{\Delta x^2} [u_{i-1, j, k} - 2u_{i, j, k} + u_{i+1, j, k}] \\ & + \frac{\alpha_{yy} \ i, j, k}{\Delta y^2} [u_{i, j-1, k} - 2u_{i, j, k} + u_{i, j+1, k}] \\ & + \frac{\alpha_{zz} \ i, j, k}{\Delta z^2} [u_{i, j, k-1} - 2u_{i, j, k} + u_{i, j, k+1}] \\ & - \frac{\beta_x \ i, j, k}{2\Delta x} [u_{i+1, j, k} - u_{i-1, j, k}] \\ & - \frac{\beta_y \ i, j, k}{2\Delta y} [u_{i, j+1, k} - u_{i, j-1, k}] \\ & - \frac{\beta_z \ i, j, k}{2\Delta z} [u_{i, j, k+1} - u_{i, j, k-1}] + u_{i, j, k} \\ & = 0. \end{aligned} \quad (11.9)$$

The computational molecule for this finite difference approximation is seen in [Fig. 11.1](#). The reference point is at $(i\Delta x, j\Delta y, k\Delta z)$. It is a straight-forward extension of the two dimensional finite difference in space molecule; the only change is the number of subscripts and the addition of the vertical space dimension. In contrast to the two dimensional case, each equation (row) of the matrix equation will now have seven non-zero coefficients organized into seven diagonal bands. The matrix is generated by using the template of [Eq. \(11.9\)](#) for each node in the finite difference mesh. Boundary conditions are handled in the same manner as for the two-dimensional problem.

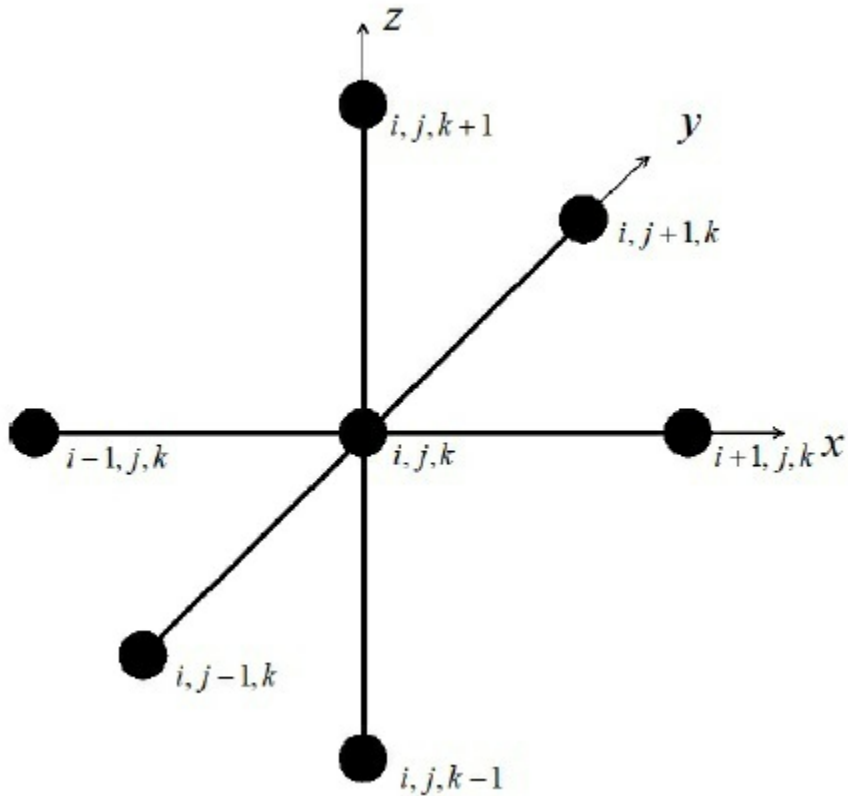


Figure 11.1: Computational molecule for three space dimension finite difference approximation.

11.2 Finite Element Approximations

The extension of the two-dimensional finite element approximation method can be approached in two ways, the choice depending upon the nature of physical system being considered. We will illustrate the methodology using the example equation presented above, that is

$$\begin{aligned}
 & \alpha_{xx}(x, y, z) \frac{\partial^2 u(x, y, z)}{\partial x^2} + \alpha_{yy}(x, y, z) \frac{\partial^2 u(x, y, z)}{\partial y^2} + \alpha_{zz}(x, y, z) \frac{\partial^2 u(x, y, z)}{\partial z^2} \\
 & - \beta_x(x, y, z) \frac{\partial u(x, y, z)}{\partial x} - \beta_y(x, y, z) \frac{\partial u}{\partial y} - \beta_z(x, y, z) \frac{\partial u}{\partial z} + u(x, y, z) = 0.
 \end{aligned} \tag{11.10}$$

The first step is to employ the Galerkin method of approximation to obtain the weighted residual formulation. To do this we first define a residual by substituting the sought-after function $u(x, y, z)$ with an approximate value $\hat{u}(x, y, z)$. The result is

$$\begin{aligned}
 & \alpha_{xx}(x, y, z) \frac{\partial^2 \hat{u}(x, y, z)}{\partial x^2} + \alpha_{yy}(x, y, z) \frac{\partial^2 \hat{u}(x, y, z)}{\partial y^2} + \alpha_{zz}(x, y, z) \frac{\partial^2 \hat{u}(x, y, z)}{\partial z^2} \\
 & - \beta_x(x, y, z) \frac{\partial \hat{u}(x, y, z)}{\partial x} - \beta_y(x, y, z) \frac{\partial \hat{u}(x, y, z)}{\partial y} - \beta_z(x, y, z) \frac{\partial \hat{u}(x, y, z)}{\partial z} + \hat{u}(x, y, z) \equiv R(x, y, z).
 \end{aligned} \tag{11.11}$$

The second step is to weight the residual defined in [Eq. \(11.11\)](#) by a test function $\phi_i(x, y, z)$, the nature of which will be explained shortly, and integrate over the domain of interest. The region $\Omega(x, y, z)$ is now three-dimensional (see [Fig. 11.2](#)), rather than the two-dimensional form shown in [Fig. 8.9](#) on page 8.9. This process yields

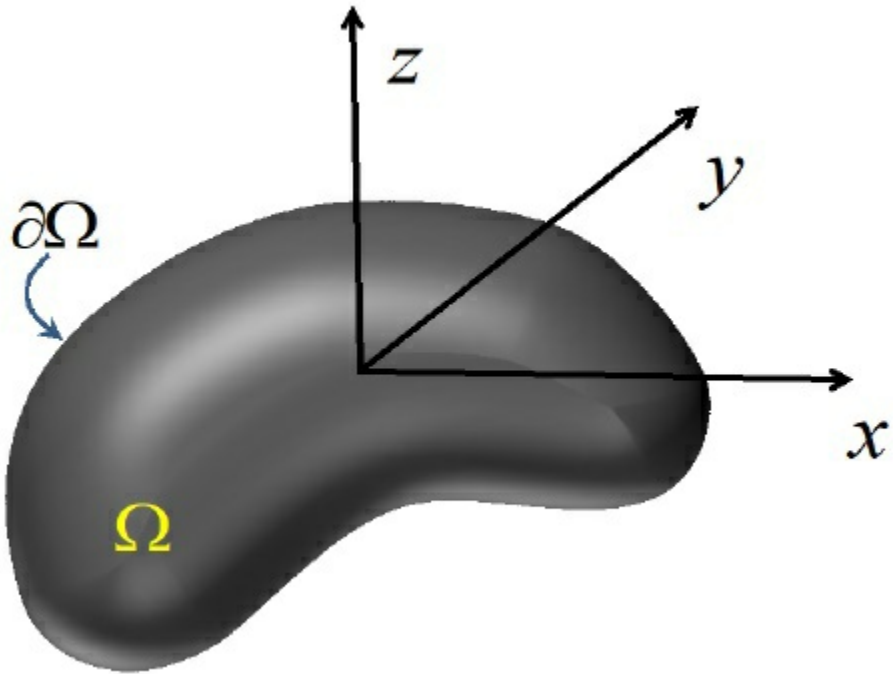


Figure 11.2: Three-dimensional region as encountered in three-dimensional finite element analysis.

$$\int_{\Omega} R(x, y, z) \phi_i(x, y, z) = 0 \quad i = 1, \dots, N. \tag{11.12}$$

The approximating function $\hat{u}(x, y, z)$ is now approximated as

$$\hat{u}(x, y, z) = \sum_{j=1}^N u_j \phi_j(x, y, z) \tag{11.13}$$

where the functions $\phi_j(x, y, z)$ are the same as those used as the weighting functions in [Eq. \(11.12\)](#). Substitution of [Eqs. \(11.13\)](#) and [\(11.11\)](#) into [Eq. \(11.12\)](#) yields

$$\begin{aligned}
& \sum_{j=1}^N u_j \int_{\Omega} \left[\alpha_{xx}(x, y, z) \frac{\partial^2}{\partial x^2} \phi_j(x, y, z) \right. \\
& + \alpha_{yy}(x, y, z) \frac{\partial^2}{\partial y^2} \phi_j(x, y, z) + \alpha_{zz}(x, y, z) \frac{\partial^2}{\partial z^2} \phi_j(x, y, z) \\
& - \beta_x(x, y, z) \frac{\partial}{\partial x} \phi_j(x, y, z) - \beta_y(x, y, z) \frac{\partial}{\partial y} \phi_j(x, y, z) \\
& \left. - \beta_z(x, y, z) \frac{\partial \hat{u}}{\partial z} + \hat{u}(x, y, z) \phi_i(x, y, z) \right] d\Omega \\
& = 0 \quad i = 1, 2, \dots, N.
\end{aligned} \tag{11.14}$$

The next step is to introduce the finite element notation which replaces the domain Ω with a series of E finite elements. in [Eq. \(11.14\)](#). Application of Green's theorem changes the form of the second order terms in this equation. In the process, a line integral along the boundary $\partial\Omega$ is created. The result is

$$\begin{aligned}
& \sum_{j=1}^N u_j \int_{\Omega} \left[\alpha_{xx}(x, y, z) \frac{\partial}{\partial x} \phi_j(x, y, z) \frac{\partial}{\partial x} \phi_i(x, y, z) \right. \\
& + \alpha_{yy}(x, y, z) \frac{\partial}{\partial y} \phi_j(x, y, z) \frac{\partial}{\partial y} \phi_i(x, y, z) \\
& + \alpha_{zz}(x, y, z) \frac{\partial}{\partial z} \phi_j(x, y, z) \frac{\partial}{\partial z} \phi_i(x, y, z) + \\
& \beta_x(x, y, z) \frac{\partial}{\partial x} \phi_j(x, y, z) + \beta_y(x, y, z) \frac{\partial}{\partial y} \phi_j(x, y, z) \\
& \left. + \beta_z(x, y, z) \frac{\partial \hat{u}}{\partial z} + \hat{u}(x, y, z) \phi_i(x, y, z) \right] d\Omega_e - \\
& \int_{\partial\Omega} a_n \frac{\partial \hat{u}(x, y, z)}{\partial n} \phi_i(x, y, z) d\omega = 0 \quad i = 1, 2, \dots, N
\end{aligned} \tag{11.15}$$

We now introduce the concept of integration over finite elements where e denotes the element number. This yields

$$\begin{aligned}
& \sum_{\epsilon=1}^E \sum_{j=1}^N u_j \int_{\Omega_{\epsilon}} \left[\alpha_{xx_{\epsilon}}(x, y, z) \frac{\partial}{\partial x} \phi_j(x, y, z) \frac{\partial}{\partial x} \phi_i(x, y, z) \right. \\
& + \alpha_{yy_{\epsilon}}(x, y, z) \frac{\partial}{\partial y} \phi_j(x, y, z) \frac{\partial}{\partial y} \phi_i(x, y, z) \\
& + \alpha_{zz_{\epsilon}}(x, y, z) \frac{\partial}{\partial z} \phi_j(x, y, z) \frac{\partial}{\partial z} \phi_i(x, y, z) \\
& + \beta_{x_{\epsilon}}(x, y, z) \frac{\partial}{\partial x} \phi_j(x, y, z) + \beta_{y_{\epsilon}}(x, y, z) \frac{\partial}{\partial y} \phi_j(x, y, z) \\
& \left. + \beta_{z_{\epsilon}}(x, y, z) \frac{\partial \hat{u}}{\partial z} + \hat{u}(x, y, z) \phi_i(x, y, z) \right] d\Omega_{\epsilon} - \\
& \int_{\partial\Omega_{\epsilon}} a_{n_{\epsilon}} \frac{\partial \hat{u}(x, y, z)}{\partial n} \Big|_{\epsilon} \phi_i(x, y, z) d\omega = 0 \quad i = 1, 2, \dots, N
\end{aligned} \tag{11.16}$$

where the subscript e on the coefficients a indicate that they are considered as element-wise constant values. Imbedded in [Eq. \(11.16\)](#) are the $\phi_i(x, y, z)$ and $\phi_j(x, y, z)$ basis and weighting functions respectively.

The remaining task is to define the functions $\phi_i(x, y, z)$ and $\phi_j(x, y, z)$ which are defined element-wise as was done in Section 8.2.1 on page 196. Their natural choice for three dimensional simulation is the three-dimensional extension of the triangle, the tetrahedron. A typical tetrahedral element is seen in [Fig. 11.3](#).

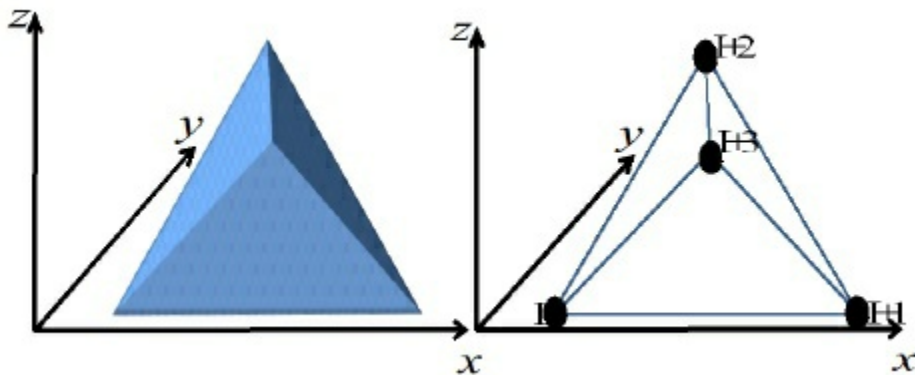


Figure 11.3: Tetrahedral finite elements to be used in three-dimensional finite element analysis.

To formulate the appropriate tetrahedral basis functions we follow the same procedure we did for triangles (see Section 8.2.1). We begin by expressing the functions $\phi_i(x, y, z)$ using the expression definition of a the triangular-based pyramid, that is,

$$\phi_i(x, y, z) = a_i + b_i x + c_i y + d_i z$$

and impose the requirements of a basis function, namely that $\phi_i(x, y, z)$ must be unity at the node for which it is defined (here node i) and zero at all other nodes. Thus we can write the following set of equations for $\phi_i(x, y, z)$, where I is the local element index as seen in [Fig. 11.3](#)

$$\begin{bmatrix} \phi_I(x_I, y_I, z_I) \\ \phi_I(x_{I+1}, y_{I+1}, z_{I+1}) \\ \phi_I(x_{I+2}, y_{I+2}, z_{I+2}) \\ \phi_I(x_{I+3}, y_{I+3}, z_{I+3}) \end{bmatrix} = \begin{bmatrix} 1 & x_I & y_I & z_I \\ 1 & x_{I+1} & y_{I+1} & z_{I+1} \\ 1 & x_{I+2} & y_{I+2} & z_{I+2} \\ 1 & x_{I+3} & y_{I+3} & z_{I+3} \end{bmatrix} \begin{bmatrix} a_I \\ b_I \\ c_I \\ d_I \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (11.17)$$

The solution to [Eq. \(11.17\)](#) is given as

$$\phi_i(x, y, z) = a_i + b_i x + c_i y + d_i z \quad (11.18)$$

where

$$a_I = \det \begin{bmatrix} x_{I+1} & y_{I+1} & z_{I+1} \\ x_{I+2} & y_{I+2} & z_{I+2} \\ x_{I+3} & y_{I+3} & z_{I+3} \end{bmatrix} \quad (11.19)$$

$$b_I = \det \begin{bmatrix} 1 & y_{I+1} & z_{I+1} \\ 1 & y_{I+2} & z_{I+2} \\ 1 & y_{I+3} & z_{I+3} \end{bmatrix} \quad (11.20)$$

$$c_I = \det \begin{bmatrix} x_{I+1} & 1 & z_{I+1} \\ x_{I+2} & 1 & z_{I+2} \\ x_{I+3} & 1 & z_{I+3} \end{bmatrix} \quad (11.21)$$

$$d_i = \det \begin{bmatrix} x_{I+1} & y_{I+1} & 1 \\ x_{I+2} & y_{I+2} & 1 \\ x_{I+3} & y_{I+3} & 1 \end{bmatrix} \quad (11.22)$$

and

$$6V = \det \begin{bmatrix} 1 & x_I & y_I & z_I \\ 1 & x_{I+1} & y_{I+1} & z_{I+1} \\ 1 & x_{I+2} & y_{I+2} & z_{I+2} \\ 1 & x_{I+3} & y_{I+3} & z_{I+3} \end{bmatrix}. \quad (11.23)$$

The equations for x_{I+2} , y_{I+3} , and z_{I+3} can be obtained by rotating through the remaining nodes and populating the various determinants using the right-hand rule. As an example,

$$a_{I+1} = \begin{bmatrix} x_{I+2} & y_{I+2} & z_{I+2} \\ x_{I+3} & y_{I+3} & z_{I+3} \\ x_I & y_I & z_I \end{bmatrix} \quad (11.24)$$

Recall that the area of an element is related to its determinant. (see p. 199)

Remember also that with triangles we had to number the nodes (corners) in a counterclockwise manner to assure that the determinant was positive. Similarly in the case of tetrahedral elements the numbering matters. One scheme that assures that the determinant in [Eq. \(11.23\)](#) will be positive is the following: take any node location and call it I . Now select a side and number the nodes increasing counterclockwise. For example in [Fig. 11.3](#) we number the side as $I, I+1, I+2$ and the fourth node, off the plane of the others, on the opposite side of the tetrahedral element, is $I+3$. In the evaluation process, the values of $I, I+1, I+2$ and $I+3$ would be replaced by global nodal values with their associated locations when populating the determinants noted above.

The rest of the process of formulating and solving the discrete equations follows the process outlined in Section 8.2.1. The extension to the transient case is also achieved using the same procedure as outlined for the triangular element case in Section 10 on page 273.

There remains to be considered the term

$$\int_{\partial\Omega_\epsilon} a_{n_\epsilon} \frac{\partial \hat{u}(x, y, z)}{\partial n} \Big|_e \phi_i(x, y, z) d\omega. \quad (11.25)$$

In the three dimensional formulation this is a integral over the boundary of the area of interest, considered on element at a time. Consequently this integration is the same as that considered for triangular elements. In this case, the $\phi_i(x, y, z)$ term is a triangle on the surface of the domain. The integrated values are then allocated to each of the three nodes per element.

Integrations in three dimensions can be achieved using a formula analogous to that used for triangles. The formula is

$$\int_{\Omega_e} \phi_I^{m_I} \phi_{I+1}^{m_{I+1}} \phi_{I+2}^{m_{I+2}} \phi_{I+3}^{m_{I+3}} d\Omega_e = 6V \frac{m_I! m_{I+1}! m_{I+2}! m_{I+3}!}{m_I + m_{I+1} + m_{I+2} + m_{I+3} + 3} \quad (11.26)$$

where, as earlier, m_I is the power to which the polynomial is raised

11.3 Chapter Summary

In this chapter we extended our earlier analysis of finite difference and finite element from two space dimensions to three. The fundamental formulation in three dimensions for finite difference is very similar to that for two dimensions. In the case of finite elements, it is necessary to introduce a new finite element based upon the tetrahedron. Again, the basic theory is analogous to that developed for two dimensions. Although we did not consider collocation or finite volume extensions to three dimensions, such extensions introduce no new concepts.

Bibliography

Felippa, C.A. (2015) "Advanced finite element methods,"
<http://www.colorado.edu/engineering/cas/courses.d/AFEM.d/Home.html>.

Gallagher, R.H. (1975) *Finite Element Analysis, Fundamentals* Prentice-Hall, New Jersey.

Moaveni, S. (2008) *Finite Element Analysis, Theory and Application with ANSYS*, Prentice-Hall, New Jersey.

Chapter 12

Nomenclature

Latin alphabet	
$\overline{2G}$	line segment connection node 2 to point G
a_i	known constant
a_{ij}	matrix coefficient of matrix $[A]$
$a_{xx}, a_{xy}, a_{yx}, a_{yy}$	components of coefficient tensor
\overline{AF}	line segment connecting point A to point F
b	constant
b_j	known constant
b_{ij}	coefficient in matrix $[B]$
c	constant
$c(x, t)$	concentration
d	a constant
e_i	value in vector $[e]$
e	element number index when used as superscript
D	dispersion coefficient
$f(x)$	function of coordinate x
$\hat{f}(x)$	approximation of function $f(x)$
$f(x_i)$	function $f(x)$ evaluated at location x_i

\overline{EA}	line segment connecting point E to point A
\overline{FB}	line segment connecting point F to point B
$g(x)$	function of coordinate x
\overline{GB}	line segment connecting point G to point B
\overline{GC}	line segment connecting point G to point C
h_i^k	Hermite cubic basis function of type 0 at node i
h_i^1	Hermite cubic basis function of type 1 at node i
\mathbf{i}	unit vector in x coordinate direction
\mathbf{j}	unit vector in y coordinate direction
l_j^n	Lagrange polynomial of degree n defined at node j
L	characteristic length and also natural coordinate system coordinate
m_I	power to which the function φ_I is raised e.g. $\varphi_I^{m_I}$
m	degree of Lagrange polynomial in time
n	normal coordinate in (n, τ) coordinate system
$O(x)$	order of x
P_e	Peclet number
$P_n(x)$	polynomial of degree n defined in the x coordinate direction
Q	flow rate entering a pipe
$u(x)$	a function of coordinate x

$\hat{u}(x)$	approximation of $u(x)$
u_i	value of variable u at node location i
v	velocity
$u_{i,j}$	value of u at location $x + i\Delta x, j + j\Delta y$
n	degree of Lagrange polynomial in space
x_I	coordinate location for local element index I
y_I	coordinate location for local element index I
k_i	values associated with Runge-Kutta formulations
T	temperature and also indicates transpose of a matrix when used as a superscript
$w_i(x)$	weighting function defined in x coordinate system
$w(x, y)$	weighting function defined in $x - y$ coordinate system
$R(x)$	residual defined in x coordinate direction
$R(x, y)$	residual defined in x coordinate direction
x_i	nodal location i in along the $x - y$ coordinate system
Greek alphabet	
β	real coefficient
β_i	constant
δ_{ij}	Kronecker delta
Γ	diffusivity
$\delta(x - x_i)$	Dirac delta function defined for location \bar{x}_i

.	measure of error
ξ	natural coordinate
η	natural coordinate
θ	trigonometric angle
κ	coefficient
λ_i	eigenvalue
$\cdot P_i^j$	multiplication operator
$\cdot \rho$	density
$\cdot \sigma$	coordinate direction for integration along a two-dimensional element boundary
$\cdot \tau$	unit vector in τ coordinate direction
$\cdot \tau$	tangential coordinate in (n, τ) coordinate system
$\cdot \varphi_j(x, y)$	triangular basis function defined at node j
$\cdot \chi$	local coordinate
$\cdot \psi_j(\xi, \eta)_i$	basis function defined at node j
Matrices	
$\cdot [A]$	space matrix
$\cdot \{ \}$	error vector
$\cdot [J]$	Jacobian matrix
$\cdot [P]$	matrix of nodal location information used in defining element matrix coefficients

Expressions	
$\tilde{e}_i \equiv u_i - \tilde{u}_i$	error between function u and its approximation \tilde{u}
$E(\theta) \equiv f(\theta) - P_2(\theta)$	error as a function of coordinate θ
$h \equiv x_i - x_{i-1}$	incremental distance
$L(\cdot) \equiv \frac{d^2(\cdot)}{dx^2} + \frac{d(\cdot)}{dx} + 0.5(\cdot)$	mathematical operator as defined
$p_n(x) = \prod_{i=0}^{n-1} (x - x_i)$	
$\Delta x \equiv x_{i+1} - x_i$	spatial increment in coordinate x
$\Delta y \equiv y_{j+1} - y_j$	spatial increment in coordinate y
$\Delta t \equiv t - t_{i-1}$	increment in time t
$\partial\Omega$	line or surface of Ω when Ω is an area or volume respectively
$\partial\Omega_I^e$	line emanating from node I in element e
$\partial\Omega_D$	segment of boundary of Ω designated as a Dirichlet boundary condition
$\partial\Omega_N$	segment of boundary of Ω designated as a Neumann boundary condition

Index

- numerical integration common formulas, 56
- approximation of a function, 4
- backward difference approximation, 36, 161
- backward in time approximation, 147
- basis functions, 85, 95, 184, 217, 273
 - Lagrange, 85
 - triangular, 196
- boundary conditions, 107, 120, 144, 163, 273
 - Dirichlet, 149
 - Neumann, 206
- boundary-value problems in three space coordinates, 279
- centered in space time approximation, 147
- chapter summary
 - chapter eight, 230
 - chapter eleven, 285
 - chapter five, 133
 - chapter four, 78
 - chapter nine, 266
 - chapter one, 24
 - chapter seven, 175
 - chapter six, 162
 - chapter ten, 276
 - chapter three, 62
 - chapter two, 46
- collocation method, 123, 124
 - first-order equations, 123
 - second-order equations, 126
- collocation points, 124
- computational molecule, 143
- conditional stability, 162

- consistency
 - definition, 72
- convergence
 - definition, 69
- coordinate transformation function, 216, 219
- coordinates
 - global, 196
 - local, 211
 - natural, 211
- cross derivative, 262
- finite volume formulation, 241

- difference approximation
 - first derivative, 102
 - second derivative, 279
- difference equation, 73
- Dirac delta function, 113, 124, 243
- Dirichlet boundary condition, 150, 263
- discretization, 2

- eigenvalue, 151
- eigenvector, 151
- element coefficient matrix, 92, 105
 - for triangular elements, 202
 - linear elements, 120
- error of the approximation, 4
 - finite difference formulae, 36
 - first-order derivative, 42, 43
 - Hermite, 23
 - quadratic Lagrange polynomial, 9
- Euler forward integration formula, 68
- Euler forward integration method example, 68

- finite difference approximation, 35, 169
 - in three space variables, 279
 - two space dimensions, 141

finite difference interpretation

of the first-order Galerkin approximation, 102

of the second-order Galerkin approximation, 111

finite element

approximation in three space variables, 280

approximation over triangles, 195

example problem, 199

approximations over rectangles, 181

Galerkin approximation in time, 158, 161

forward difference approximation, 160

isoparametric, 211

basis functions, 217

tetrahedral, 283

triangular, 196

finite volume

application of Green's theorem, 241

approximation, 239

approximation to second-order equations, 112

example problem, 262

example problem 1, 246

element coefficient matrices, 248

evaluation of the area integral, 256

evaluation of the line integral, 249

global matrix assembly, 260

weighted residual formulation, 246

example problem 2

element coefficient matrices, 263

evaluation of the source term, 265

problem definition, 262

weighted residual formulation, 262

finite difference interpretation, 122

finite difference interpretation for second-order equations, 122

finite element and polygon domain, 240

interpretation in terms of finite difference method, 93

mesh, 240

- method, 133
- normal-tangential boundary condition, 242
- vector normalization, 243
- weighted residual formulation, 240
- weighting function, 84
- weighting function in two space dimensions, 240
- first derivative difference approximation, 41, 102
- flux type boundary condition, 106
- forward difference, 35, 160

- Galerkin method, 94, 159, 200
 - finite element approximation in time, 158
 - for first-order equations, 94
 - for second-order equations, 102
 - time approximation
 - backward difference approximation, 161
- Gauss integration, 61
- Gauss points, 124
- Gauss quadrature, 59, 124
- global coefficient matrix, 93, 120, 204, 265
- global coordinates, 196
- global system, 89
- Green's theorem, 184, 201, 241, 274

- Hermite polynomials, 20, 127
 - development of expression for, 20

- implicit approximation, 148
- incremental distance h , 35
- initial boundary value problems, 273
 - mass lumping, 276
- initial conditions, 67, 144, 163, 273
- initial value problems, 67
- inner product, 84
- integration
 - finite volume method, 251

- for isoparametric finite elements, 227
- formula for triangular element, 198
- three dimensional finite elements (tetrahedral elements), 285
- integration by parts, 103, 117, 159, 184
- interpolant, 7
- interpolation, 1
- interpretation of finite volume method in terms of the finite difference method, 93
- isoparametric finite element, 211
 - basis functions, 217, 218
 - coordinate transformation functions, 216
 - example problem, 223
 - Gauss-quadrature integration, 227
 - Jacobian calculation, 219
 - matrix assembly, 228
 - natural coordinate system, 211
 - relationship between natural and global coordinates, 225
 - two-dimensional natural coordinate system, 212
- Jacobian, 219
 - calculation of, 219
- Jacobian matrix, 221
- Kronecker delta, 5
- Lagrange basis functions, 86
- Lagrange interpolation, 4, 8, 55
- Lagrange polynomials, 35, 55, 96, 141, 162, 170
 - linear, 5, 85
 - local coordinate system, 107
 - quadratic, 8
- linear Lagrange polynomial, 207
- local coefficient matrix, 187
- local coordinates, 88, 98, 104, 191
- local Lagrange polynomials, 120
- local system, 90
- local – global coefficient transformation matrix, 189
- locally defined Lagrange polynomials, 96

- mass lumping, 276
- matrix assembly isoparametric elements, 228
- matrix method of stability analysis, 149
- method of weighted residuals, 83, 112, 123, 159, 183, 240
- multiple elements, 14

- natural coordinate systems
 - one dimensional natural coordinate systems, 211
 - two dimensional natural coordinate systems, 212
 - use in isoparametric integration, 222
- Neumann boundary condition, 106, 206
- Newton–Cotes quadrature formulae, 55
- node, 4
- numerical differentiation, 33
 - general theory, 33
- numerical integration, 55
- numerical solution of boundary value problems in two space dimensions
 - finite difference approximations, 169
 - finite element approximations over rectangles, 181
 - finite element approximations over triangles, 195
 - finite volume approximation, 239
- numerical solution of boundary value problems in two space dimensions and time, 169
- numerical solution of initial boundary-value problems, 139
- numerical solution of ordinary differential equations, 67

- II, definition of, 6

- quadratic Lagrange polynomial, 8

- residual error, 83
- round-off error, 150
- Runge-Kutta
 - geometric interpretation, 77
 - second-order form, 75
- Runge-Kutta
 - type formulae, 75

- second type or Neumann boundary value problem, 206
- second-order derivatives, 43
- σ coordinate system, 207
- Simpsons rule, 57
- stability, 73, 151
 - example of, 153
 - example simulation, 156
- stability of finite difference approximations to boundary value problems, 149
- steady state behavior, 169
- subdomain method, 84

- tetrahedral element, 284
- three space dimensions, 279
 - finite element approximations, 280
 - finite-difference approximations, 279
- three-point difference formulae, 40
- trapezoidal rule, 56
- triangular basis functions
 - formulation, 196
- tridiagonal matrix, 161
- two dimensional natural coordinate systems, 212
- two-dimensional polynomial approximation, 139, 169
- two-point difference formulae, 34
- two-point formulae from Taylor Series, 37

- weighting function, 83, 106, 124, 184
 - finite volume method, 84
 - Galerkin's method, 95
- Weirstraus Approximation Theorem, 3

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley's ebook EULA.