The Finite Element Method

Linear Static and Dynamic Finite Element Analysis

Thomas J. R. Hughes

Professor of Mechanical Engineering
Chairman of the Division of Applied Mechanics
Stanford University

PRENTICE-HALL, INC., Englewood Cliffs, New Jersey 07632

Library of Congress Cataloging-in-Publication Data

HUGHES, THOMAS J. R.
The finite element method.

Bibliography.
Includes index.
1. Finite element method. 2. Boundary value problems. I. Title.
TA347.F5H84 1987 620'.0042 86-22558
ISBN 0-13-317025-X

Editorial/production supervision and interior design: Joan McCulley
Cover sketch: M. Igarashi and K. Shimomaki
Suzuki Motor Co., Ltd.
Cover design: 20/20 Services, Inc.
Manufacturing buyer: Rhett Conklin

Printed in the United States of America

10 9 8 7 6 5 4 3

ISBN 0-13-317025-X 025

PRENTICE-HALL INTERNATIONAL (UK) LIMITED, *London*
PRENTICE-HALL OF AUSTRALIA PTY. LIMITED, *Sydney*
PRENTICE-HALL CANADA INC., *Toronto*
PRENTICE-HALL HISPANOAMERICANA, S. A., *Mexico*
PRENTICE-HALL OF INDIA PRIVATE LIMITED, *New Delhi*
PRENTICE-HALL OF JAPAN, INC., *Tokyo*
PRENTICE-HALL OF SOUTHEAST ASIA PTE. LTD., *Singapore*
EDITORA PRENTICE-HALL DO BRASIL, LTDA., *Rio de Janeiro*

To my wife,
my mother,
and my children

Contents

Part Two Linear Dynamic Analysis

Preface

This book is based on courses that I have taught at the California Institute of Technology and Stanford University during the last 10 years. It deals with the finite element method in linear static and dynamic analysis. It is intended primarily for engineering and physical science students who wish to develop comprehensive skills in finite element methodology, from fundamental concepts to practical computer implementations.

Some sections of this book touch upon the frontiers of research and have been used as lecture notes in a number of more advanced short courses I have taught in Europe, Japan, and the United States. Consequently, I believe it will also be of interest to more experienced analysts and researchers working in the finite element field.

SUBJECTS COVERED

The first chapter of the book introduces the finite element method in the context of simple one-dimensional model problems. Chapter 2 deals with variational formulations of two- and three-dimensional boundary-value problems in heat conduction (in fact, all problems governed by Laplace/Poisson equations, such as electrostatics, potential flow, elastic membranes, and flow in porous media) and elasticity theory. These serve as the basis for finite element discretization and the techniques developed illustrate the relationship between "strong," or classical, statements of boundary-value problems and their "weak," or variational, counterparts. The Galerkin method of approximate solution is emphasized in Chapter 2 and throughout the book, rather than "variational principles" due to the significantly greater generality of the former approach. In Chapter 3 a variety of finite element interpolatory schemes are developed.

These apply to triangles, quadrilaterals, tetrahedra, hexahedra, wedges, etc. The isoparametric concept is emphasized, and special-purpose interpolatory strategies are also developed (e.g., "singular elements"). Programming techniques are introduced for numerically integrated finite elements. Chapter 4 deals with basic error estimates for standard "displacement" finite element methods and introduces mixed methods for constrained media applications such as incompressible elasticity and Stokes flow. A variety of "variational crimes" are described: for example, incompatible elements, reduced and selective integration, strain projection (i.e., $\bar{B}$-) methods, etc. (Most of these have been decriminalized in recent years.) The mathematical analysis of finite element methods for incompressible media is rather complex. David Malkus, an authority on this subject, explains some of the subtleties in an appendix to Chapter 4. Chapter 5 is concerned with finite element methods for Reissner-Mindlin plates and elastic frameworks composed of straight beam elements accounting for axial, torsional, bending, and transverse shear deformations (i.e., Timoshenko beams). Chapter 6 deals with three-dimensional curved shell elements and two-dimensional special cases such as rings, tubes, and shells of revolution. The problem classes discussed in Chapters 1–6 give rise to associated eigenvalue problems and initial-value problems. The formulations of problems of these types are the subjects of Chapter 7. Chapter 8 presents time-stepping algorithms for first-order ordinary differential equation systems such as those arising from unsteady heat conduction ("parabolic case"). Chapter 9 deals with algorithms for second-order ordinary differential equation systems such as those emanating from elastodynamics and structural dynamics ("hyperbolic case"). Chapter 10 presents basic algorithmic strategies for symmetric elliptic eigenvalue problems such as those encountered in free vibration and structural stability. A very efficient major software package for matrix eigenvalue and eigenvector calculations based on the Lanczos method is also presented in Chapter 10. The documentation of the Lanczos algorithm and software were written by Bahram Nour-Omid, an expert on procedures of this type. Chapter 11 presents a linear static and dynamic analysis computer program based on the methods developed in the book. My student, Robert Ferencz, and colleague, Arthur Raefsky, collaborated with me in the writing of the program. The program is named DLEARN and it contains a very complete library of finite element software tools. This program is suitable for homework assignments, projects (e.g., programming additional elements), and research studies. DLEARN is highly structured for readability, maintainability, and extendability and has been written specifically to complement and enhance the procedures described in the remainder of the book.

LEVEL AND BACKGROUND

This book is primarily intended for the graduate level, although advanced undergraduates will find much of it accessible. An undergraduate degree in engineering, mathematics, computer science, or any of the physical sciences constitutes essential background. Courses in applied linear algebra and elementary ordinary and partial differential equations are desirable prerequisites. A working knowledge of FORTRAN

(the dominant language used in finite element programming) is also necessary for understanding the software presented in Chapters 10 and 11.

The book emphasizes heat conduction and elasticity as primary vehicles for developing finite element methods because of the widespread interest in these theories in the applied sciences. By virtue of the fact that the partial differential equation of heat conduction, namely, the Laplace/Poisson equation, appears under different names in virtually all branches of engineering and physics, most students have had some familiarity with it. Some exposure to the theory of elasticity is also desirable. This can be obtained, for example, in a good advanced course on strength of materials. Students at Caltech and Stanford frequently have taken courses in elasticity simultaneously with finite elements. Background in structural mechanics (i.e., the theory of beams, plates, and shells) is certainly an asset when it comes to studying this book but is not essential. Only Chapters 5 and 6 deal exclusively with this subject, and these chapters may be ignored by students whose primary interests lie elsewhere. It is worth noting that students who have taken this finite element course at Stanford and Caltech have had very diverse backgrounds (e.g., geophysics, chemical engineering, planetary sciences, coastal engineering, electrical engineering, computer science, mathematics, bioengineering, aeronautics and astronautics, material science, physics, earth sciences, environmental engineering, biomechanics, thermosciences, engineering design, applied mechanics, earthquake engineering, and, of course, mechanical and civil engineering). In this spirit the book emphasizes fundamental finite element concepts and techniques applicable to a very broad range of problems and thus constitutes a suitable text for most students in the physical sciences.

It is my experience that most finite element methodology can and should be taught initially within the confines of linear behavior before introducing nonlinear effects. A solid understanding of finite elements in linear analysis is, of course, very useful in its own right. In addition, it makes the subject of nonlinear finite element analysis much more accessible.

WHAT IS UNIQUE ABOUT THIS BOOK

Although this book deals with many standard aspects of the finite element method, there are many unique features, some of which are enumerated below.

- A comprehensive presentation and analysis of algorithms for time-dependent phenomena.
- Beam, plate, and shell theories are derived directly from three-dimensional elasticity theory.
- An extensive static and dynamic finite element analysis program, DLEARN, was specially prepared based on the text material. It is written using structured programming concepts and contains many advanced procedures.
- Although written for students without serious mathematical training, the book contains introductory material on the mathematical theory of finite elements and

many important mathematical results. It thus serves as a primer for more advanced works on this subject.

- A systematic treatment of "weak," or variational, formulations for various classes of initial- and boundary-value problems.
- Many of the procedures described are presented in book form for the first time, for example, strain projection (i.e., $\overline{\boldsymbol{B}}$-) methods, implicit-explicit finite element mesh partitions in transient analysis, element-by-element iterative solvers, complete computer implementation of predictor/multicorrector implicit-explicit time-stepping algorithms based upon the Hilber-Hughes-Taylor α-method, etc.
- A Lanczos software package for eigenvalue/eigenvector extraction.

NOTATIONAL CONVENTIONS

As far as possible, we have adoped fairly standard notations. Vectors and matrices are denoted by boldface italic characters. Scalars and components of vectors, tensors, and matrices are denoted by lightface italic characters. Cartesian tensor notation is used in the presentation of heat conduction, elasticity, plates, etc. The symbol ■ is used to denote the end of a proof. Unavoidably, in a book of this size and scope there are some minor conflicts of notation. In instances when this occurs, it is hoped that surrounding explanation and context will make the intended meaning clear. All notations are defined when they are introduced. Principal notations are summarized in a brief glossary following the Preface.

NUMBERING CONVENTIONS

The book is divided into two parts. Chapters 1–6 deal with static analysis, and Chapters 7–11 deal with dynamic analysis. The chapters are subdivided into sections and several of the chapters include appendices. Some of the sections are further subdivided into subsections. Equations, tables, and figures are numbered consecutively within each section, with each number consisting of chapter, section, and item number. Exercises are numbered consecutively beginning with 1 in each section. References are numbered consecutively at the end of each chapter. Equation numbers are enclosed in parentheses, and reference numbers appearing in the text are enclosed in square brackets. If, for example, we wish to refer to Eq. 8 of Section 3, Chapter 4, then we would write this as (4.3.8).

If we wish to draw the reader's attention to a subsection, then the section number is followed by a decimal point and the subsection number. For example, subsection 2 of Section 3, Chapter 5, is written as Sec. 5.3.2.

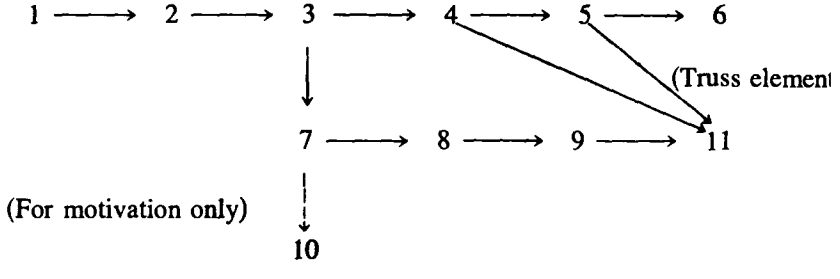
EXERCISES

Exercises appear throughout the text. In some cases answers or complete solutions a provided. Many of the exercises are placed in the proximity of the pertinent te material rather than at the ends of sections or chapters.

INTERDEPENDENCE OF CHAPTERS

Prior to consulting particular chapters it may be worthwhile to peruse the "readir paths," which indicate the interdependence of the chapters. Of course, if one h sufficient background, the suggested reading paths may be bypassed. Knowledge Chapters 1–3 is assumed for all subsequent chapters with the exception of Chapter 1(which is essentially self-contained.

Chapter Reading Paths



A NOTE FOR INSTRUCTORS REGARDING THE ORGANIZATION OF COURSES

Several courses can be developed based on the material in this book. Here are a fe examples.

Almost the entire book can be covered in a one-year course:

Title: Linear Static Finite Element Analysis (1 semester)
Syllabus: Chapters 1–3, selected topics from Chapters 4–6, and computing assign ments using DLEARN (Chapter 11).

Title: Linear Dynamic Finite Element Analysis (1 semester)
Syllabus: Chapters 7–10 and computing assignments using DLEARN (Chapter 11)

An abbreviated version of the above covering two quarters can be organized a follows:

Title: Linear Static Finite Element Analysis (1 quarter)

Syllabus: Chapters 1–3, selected topics from Chapter 4, and computing assignments using DLEARN (Chapter 11).

Title: Linear Dynamic Finite Element Analysis (1 quarter)
Syllabus: Sections 1–3 of Chapter 7; Sections 1, 2, and 5 of Chapter 8; Sections 1–4 of Chapter 9; Sections 1–5 of Chapter 10; and computing assignments using DLEARN (Chapter 11).

A course on finite element programming can be taught in one quarter or one semester:

Title: Finite Element Programming
Syllabus: Chapter 11 and selected sections from earlier chapters as background material. Lectures should describe overall program architecture (subroutines DLEARN and DRIVER), the structure of element routines (e.g., QUADC and subroutines called by QUADC), equation-solving (subroutines FACTOR and BACK), assembly routines (ADDLHS and ADDRHS), and important finite element utility routines (BC, COLHT, DCTNRY, DIAG, EQSET, FORMLM, LOCAL, and MPOINT). Programming assignments, such as adding new elements and extending capabilities, are essential.

ACKNOWLEDGMENTS

I would like to thank many former and current students, colleagues, and friends for their constructive comments and criticisms, especially Ted Belytschko, Alec Brooks, Marty Cohen, Leo Franca, Hans Hilber, Greg Hulbert, Itzhak Levit, Wing Kam Liu, Michel Mallet, Jerry Marsden, Isidoro Miranda, Arthur Muller, Juan Simo, Gary Stanley, Tayfun Tezduyar, Jim Winget, and Thomas Zimmermann.

I would also like to express my appreciation to those individuals who have made contributions to the writing of this book: Bob Ferencz, David Malkus, Bahram Nour-Omid, and Arthur Raefsky.

In addition, I would like to thank Alex Tessler who prepared notes for me on some recent developments in plate elements, which I used as the basis of comments included in Chapter 5. Wanda Gooden typed most of the manuscript. I wish to thank her for her patience, dedication, and good-naturedness in the face of a seemingly infinite number of corrections.

I would like to acknowledge the following organizations for permission to reprint various materials: the American Society of Mechanical Engineers for Figs. 5.3.21–5.3.31; John Wiley and Sons, Ltd., for Figs. 3.II.1, 3.II.2, 9.3.1–9.3.3, 9.3.5, and Table 3.I.1; Pergamon Journals, Ltd., for Figs. 5.3.11, 5.3.13–5.3.15, 5.3.18–5.3.20; and North-Holland Publishing Co. for portions of Chapters 8 and 9, which originally appeared in T. J. R. Hughes, "Analysis of Transient Algorithms with Particular Reference to Stability," pp. 67–155, *Computational Methods for Transient Analysis* (eds. T. Belytschko and T. J. R. Hughes), 1983.

Finally, I would like to thank M. Igarashi and K. Shimomaki of the Suzuk Motor Co., Ltd., for providing the mesh for the cover, and Dave Benson and Johr Hallquist of Lawrence Livermore National Laboratory for the accompanying colo graphics.

NOTE: The software in this book is available on diskette and magnetic tape for severa computers. Further information and an order form may be obtained by writing to the author: Professor Thomas J. R. Hughes, Division of Applied Mechanics, Durand Build ing, Stanford University, Stanford, California 94305.

A Brief Glossary of Notations

Sets

$\mathbb{R}$	real numbers
$\mathbb{C}$	complex numbers
$\cup$	set union
$\cap$	set intersection
$\varnothing$	empty set
$\in$	is a member of; belongs to
$\notin$	is not a member of; does not belong to
$\subset$	is a subset of
$\not\subset$	is not a subset of
$\forall$	for all; for each

Various Integers

n_{np}	number of nodal points
n_{en}	number of element nodes
n_{eq}	number of equations
n_{ee}	number of element equations
n_{dof}	number of degrees of freedom
n_{ed}	number of element degrees of freedom
n_{el}	number of elements
A, B	global node numbers $(1 \le A, B \le n_{np})$
a, b	element node numbers $(1 \le a, b \le n_{en})$
P, Q	global equation numbers $(1 \le P, Q \le n_{eq})$
p, q	element equation numbers $(1 \le p, q \le n_{ee})$
e	element number $(1 \le e \le n_{el})$

Element Arrays

m^e	mass matrix of the eth element
c^e	damping matrix of the eth element

k^e	stiffness matrix of the eth element
f^e	force vector of the eth element
a^e	acceleration vector of the eth element
v^e	velocity vector of the eth element
d^e	displacement vector of the eth element

Global Arrays

M	mass matrix
C	damping matrix
K	stiffness matrix
F	force vector
a	acceleration vector
v	velocity vector
d	displacement vector
$\mathbf{A}$	finite element assembly operator

$$\left(\text{e.g., } K = \mathop{\mathbf{A}}_{e=1}^{n_{el}} k^e\right)$$

Boundary-Value Problems

n_{sd}	number of space dimensions
i, j, k, l	spatial indices $(1 \le i, j, k, l \le n_{sd})$
$\mathbb{R}^{n_{sd}}$	Euclidean n_{sd}-space
Ω	domain in $\mathbb{R}^{n_{sd}}$
$\overline{\Omega}$	closure of Ω
$x_i, \boldsymbol{x}$	point in $\overline{\Omega}$
Γ	boundary of Ω
$n_i, \boldsymbol{n}$	unit outward normal vector to Γ
Γ_g, Γ_{g_i}	part of the boundary where Dirichlet conditions are specified
Γ_h, Γ_{h_i}	part of the boundary where Neumann conditions are specified
δ_{ij}	Kronecker delta
η	set of node numbers
η_g, η_{g_i}	set of node numbers at which Dirichlet conditions are specified

Variational Methods

h	mesh parameter
$\mathcal{S}, \mathcal{S}_i, \boldsymbol{\mathcal{S}}$	collections of trial solutions
$\mathcal{V}, \mathcal{V}_i, \boldsymbol{\mathcal{V}}$	collections of weighting functions
$\mathcal{S}^h, \mathcal{S}_i^h, \boldsymbol{\mathcal{S}}^h$	finite-dimensional collections of trial solutions
$\mathcal{V}^h, \mathcal{V}_i^h, \boldsymbol{\mathcal{V}}^h$	finite-dimensional collections of weighting functions
$u, u_i, \boldsymbol{u}$	trial solutions
$w, w_i, \boldsymbol{w}$	weighting functions
$u^h, u_i^h, \boldsymbol{u}^h$	finite-dimensional trial solutions
$w^h, w_i^h, \boldsymbol{w}^h$	finite-dimensional weighting functions
$\ell, \ell_i, \boldsymbol{\ell}$	source terms
$g, g_i, \boldsymbol{g}$	Dirichlet boundary data
$h, h_i, \boldsymbol{h}$	Neumann boundary data
$g^h, g_i^h, \boldsymbol{g}^h$	extensions of $g, g_i, \boldsymbol{g}$, respectively, to Ω
$e, e_i, \boldsymbol{e}$	errors
$L_2(\Omega)$	Hilbert space of square-integrable functions

$H^s(\Omega)$	Sobolev space of functions with s square-integrable generalized derivatives
$\|\cdot\|$	L_2 norm on Ω
$\|\cdot\|_s$	H^s norm on Ω
$a(\cdot, \cdot)$	symmetric bilinear form; strain-energy inner product
$(\cdot, \cdot)$	symmetric bilinear form; L_2 inner product on Ω
$(\cdot, \cdot)_\Gamma$	symmetric bilinear form; L_2 inner product on Γ

Data Processing Arrays

IEN (a, e)	element nodes matrix
ID (i, A)	destination matrix
LM (i, a, e)	location matrix

Heat Conduction

q_i	heat flux vector
κ_{ij}	conductivities
u	temperature
$u_{,i}$	temperature gradient
ℓ	prescribed heat supply
q	prescribed boundary temperature
h	prescribed boundary heat flux
d_A, d_a^e	nodal temperature
q_A, q_a^e	prescribed nodal temperature

Elasticity

ϵ_{ij}	infinitesimal strain tensor
σ_{ij}	Cauchy stress tensor
c_{ijkl}	elastic coefficients
u_i	displacement vector
$u_{(i,j)}$	symmetric part of the displacement gradients (equals ϵ_{ij})
ℓ_i	prescribed body force vector
q_i	prescribed boundary displacement vector
h_i	prescribed boundary traction vector
d_{iA}, d_{ia}^e	nodal displacements
q_{iA}, q_{ia}^e	prescribed nodal displacements
B	bulk modulus
E	Young's modulus
ν	Poisson's ratio
G, μ	shear modulus
λ	Lamé modulus
B, B_A, B_a	strain-displacement matrices
D, $\widetilde{D}$	material moduli matrices

Isoparametric Elements

Ω^e	element domain $\subset \Omega$
Γ^e	boundary of Ω^e
x_1, x_2, x_3, x, y, z	Cartesian coordinates
$\square$	parent domain of a quadrilateral or hexahedral element
$\boldsymbol{x}^e: \square \to \overline{\Omega}^e$	isoparametric mapping
ξ, η	element natural coordinates for quadrilaterals
ξ, η, ζ	element natural coordinates for hexahedra
$\boldsymbol{\xi}$	point in $\square$
j	Jacobian determinant of the mapping $\boldsymbol{x}^e(\boldsymbol{\xi})$

$C^0(\Omega)$	the class of continuous functions
$C^k(\Omega)$	the class of continuous functions possessing k continuous derivatives
N_A, N_a^e	shape functions
x_A, x_a^e	nodal points
r, s, t	triangular coordinates; area coordinates
r, s, t, u	tetrahedral coordinates; volume coordinates
$l_a(\xi)$	Lagrange polynomial associated with the ath element node
$\tilde{\xi}_l$	location of the lth integration point
W_l	weight assigned to lth integration point
$U1$, $U2$, $U3$	uniform integration elements
$S1$, $S2$, $S3$	selective integration elements

Incompressible and Nearly-Incompressible Elasticity

$\mathcal{P}$	pressure trial solution and weighting function space
$\mathcal{P}^h$	finite-dimensional pressure trial solution and weighting function space
$p, q \in \mathcal{P}$	trial pressure and weighting function
$p^h, q^h \in \mathcal{P}^h$	finite-dimensional trial pressure and weighting function
g^e, G	element and global gradient matrices
$(g^e)^T$, G^T	element and global divergence matrices
p	vector of nodal pressures
B_a^{dev}	deviatoric strain-displacement matr
B_a^{dil}	dilatational strain-displacement matr
$\overline{B}_a^{\text{dil}}$	improved dilatatio strain-displacemen matrix
$\overline{B}_a$	improved strain-displacement matri

Plates

α, β, γ, δ	tensor indices ($1 \leq \alpha, \beta, \gamma, \delta \leq$
$e_{\alpha\beta}$	alternator tensor
w	transverse displacement
θ_α	rotation vector
$\kappa_{\alpha\beta} = \theta_{(\alpha,\beta)}$	curvature tensor
γ_α	shear strain vector
$c_{\alpha\beta\gamma\delta}$, $c_{\alpha\beta}$	elastic coefficients
$m_{\alpha\beta}$	moment tensor
q_α	shear force vector
W	prescribed boundary displacement
Θ_α	prescribed boundary rotations
F	total applied transverse force
C_α	total applied couple
M_α	prescribed boundary moments
Q	prescribed boundary shear force
s	arc-length paramete of plate boundary
$\hat{\theta}_\alpha = e_{\alpha\beta}\theta_\beta$, $\hat{\theta}_s$, $\hat{\theta}_n$	right-hand-rule rotations
k_b^e, k_s^e	eth element bending and shear stiffness

D^b, D^s	matrices of bending and shear moduli
B^b, B^s	bending and shear strain-displacement matrices

Beams

w_i	displacements
θ_i	rotations
κ_α	curvatures
γ_α	shear strains
ϵ	axial strain
ψ	twist
m_i	moments
q_i	shear forces
W_i	prescribed end displacements
Θ_i	prescribed end rotations
F_i	total applied force per unit length
C_i	total applied couple per unit length
M_i	prescribed end moments
Q_i	prescribed end shears
A	cross-sectional area
A_α	shear areas
I_α	moments of inertia
J	polar moment of inertia
k_a^e, k_b^e, k_s^e, k_t^e	eth element axial, bending, shear, and torsional stiffness, respectively

Shells

x	position vector
$\bar{x}$, $\bar{x}_a$	position vectors to references surface
X, $\hat{X}_a$	fiber vectors
$z_a(\zeta)$	thickness function
e_1^l, e_2^l, e_3^l	lamina basis vectors
e_1^f, e_2^f, e_3^f	fiber basis vectors
u	displacement vector
$\bar{u}$, $\bar{u}_a$	displacement of reference surface
U, U_a, $\hat{U}_a$	fiber displacements
$\tilde{\sigma}^l = \tilde{D}^l \tilde{\epsilon}^l$	reduced constitutive equation in the lamina basis
$m_{\alpha\beta}$	moments
n_α	membrane forces
q_α	transverse shear forces

Dynamics

t	time
Δt	time step
n	time step number
a_n, v_n, d_n	approximations of $a(t_n)$, $v(t_n)$, and $d(t_n)$, respectively
$\tilde{a}_n$, $\tilde{v}_n$, $\tilde{d}_n$	predictor values of acceleration, velocity, and displacement, respectively
Δt_{crit}	critical time step
A	amplification factor
$\mathbf{A}$	amplification matrix
$\rho(\mathbf{A})$	spectral radius of amplification matrix
τ, $\boldsymbol{\tau}$	local truncation errors
λ, λ^h, $\bar{\lambda}^h$	eigenvalues
$\boldsymbol{\psi}$, $\boldsymbol{\phi}$	eigenvectors
ω, ω^h, $\bar{\omega}^h$	frequencies
$\bar{\xi}$	algorithmic damping ratio
$(\bar{T} - T)/T$	relative period error
Ω_{crit}	critical sampling frequency

$\boldsymbol{M}^I, \boldsymbol{C}^I, \boldsymbol{K}^I, \boldsymbol{F}^I$	implicit mesh partition of mass, damping, stiffness, and force
$\boldsymbol{M}^E, \boldsymbol{C}^E, \boldsymbol{K}^E, \boldsymbol{F}^E$	explicit mesh partition of mass, damping, stiffness, and force
$u_{0i}, \dot{u}_{0i}$	initial displacem and velocities
u_0	initial temperatu
ρ	density
c	capacity

1

Fundamental Concepts; a Simple One-Dimensional Boundary-Value Problem

1.1 INTRODUCTORY REMARKS AND PRELIMINARIES

The main constituents of a finite element method for the solution of a boundary-value problem are

i. The variational or weak statement of the problem; and
ii. The approximate solution of the variational equations through the use of "finite element functions."

To clarify concepts we shall begin with a simple example.
Suppose we want to solve the following differential equation for u:

$$u_{,xx} + \ell = 0 \tag{1.1.1}$$

where a comma stands for differentiation (i.e., $u_{,xx} = d^2u/dx^2$). We assume ℓ is a given smooth, scalar-valued function defined on the unit interval. We write

$$\ell : [0, 1] \to \mathbb{R} \tag{1.1.2}$$

where $[0, 1]$ stands for the unit interval (i.e., the set of points x such that $0 \leq x \leq 1$) and $\mathbb{R}$ stands for the real numbers. In words, (1.1.2) states that for a given x in $[0, 1]$, $\ell(x)$ is a real number. (Often we will use the notation $\in$ to mean "in" or "a member of." Thus for each $x \in [0, 1]$, $\ell(x) \in \mathbb{R}$.). Also, $[0, 1]$ is said to be the ***domain*** of ℓ, and $\mathbb{R}$ is its ***range***.

We have described the given function ℓ as being smooth. Intuitively, you probably know what this means. Roughly speaking, if we sketch the graph of the function ℓ, we want it to be a smooth curve without discontinuities or kinks. We do this to avoid technical difficulties. Right now we do not wish to elaborate further as

this would divert us from the main theme. At some point prior to moving on to the next chapter, the reader may wish to consult Appendix 1.I, "An Elementary Discussion of Continuity, Differentiability and Smoothness," for further remarks on this important aspect of finite element work. The exercise in Sec. 1.16 already uses a little of the language described in Appendix 1.I. The terminology may be somewhat unfamiliar to engineering and physical science students, but it is now widely used in the finite element literature and therefore it is worthwhile to become accustomed to it.

Equation (1.1.1) is known to govern the transverse displacement of a string in tension and also the longitudinal displacement of an elastic rod. In these cases, physical parameters, such as the magnitude of tension in the string, or elastic modulus in the case of the rod, appear in (1.1.1). We have omitted these parameters to simplify subsequent developments.

Before going on, we introduce a few additional notations and terminologies. Let $]0, 1[$ denote the unit interval without end points (i.e., the set of points x such that $0 < x < 1$). $]0, 1[$ and $[0, 1]$ are referred to as ***open and closed unit intervals,*** respectively. To simplify subsequent writing and tie in with notation employed later on in multidimensional situations, we shall adopt the definitions

$$\Omega =]0, 1[\quad \text{(open)} \tag{1.1.3}$$

$$\overline{\Omega} = [0, 1] \quad \text{(closed)} \tag{1.1.4}$$

See Fig. 1.1.1.

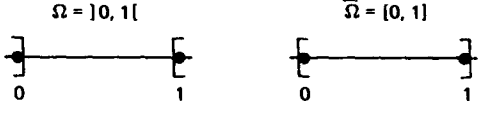


Figure 1.1.1

At this point, considerations such as these may seem pedantic. Our purpose, however, is to develop a language for the precise articulation of boundary-value problems, which is necessary for good finite element work.

1.2 STRONG, OR CLASSICAL, FORM OF THE PROBLEM

A boundary-value problem for (1.1.1) involves imposing ***boundary conditions*** on the function u. There are a variety of possibilities. We shall assume u is required to satisfy

$$u(1) = q \tag{1.2.1}$$

$$-u_{,x}(0) = h \tag{1.2.2}$$

where q and h are given constants. Equations (1.2.1) and (1.2.2) require that u take on the value q at $x = 1$ and the derivative of u (i.e., slope) take on the value $-h$ at $x = 0$, respectively. This set of boundary conditions will later enable us to illustrate certain key features of variational formulations. For obvious reasons, boundary conditions of the type (1.2.1) and (1.2.2) lead to so-called ***two-point boundary-value problems***.

The strong form of the boundary-value problem, (S), is stated as follows:

$$(S)\quad \begin{cases} \text{Given } \ell : \overline{\Omega} \to \mathbb{R} \text{ and constants } q \text{ and } h, \text{ find } u : \overline{\Omega} \to \mathbb{R}, \text{ such that} \\ u_{,xx} + \ell = 0 \qquad \text{on } \Omega \\ u(1) = q \\ -u_{,x}(0) = h \end{cases}$$

When we write $u_{,xx} + \ell = 0$ on Ω we mean $u_{,xx}(x) + \ell(x) = 0$ for all $x \in \Omega$. Of course, the exact solution of (S) is trivial to obtain, namely,

$$u(x) = q + (1 - x)h + \int_x^1 \left\{ \int_0^y \ell(z)\, dz \right\} dy \tag{1.2.3}$$

where y and z are used to denote dummy variables. However, this is not the main concern here. We are interested in developing schemes for obtaining approximate solutions to (S) that will be applicable to much more complex situations in which exact solutions are not possible.

Some methods of approximation begin directly with the strong statement of the problem. The most notable example is the finite difference method (e.g., see [1]). The finite element method requires a different formulation, which is treated in the next section.

1.3 WEAK, OR VARIATIONAL, FORM OF THE PROBLEM

To define the weak, or variational, counterpart of (S), we need to characterize two classes of functions. The first is to be composed of candidate, or trial, solutions. From the outset, we shall require these possible solutions to satisfy the boundary condition $u(1) = q$. The other boundary condition will not be required in the definition. Furthermore, so that certain expressions to be employed make sense, we shall require that the derivatives of the trial solutions be square-integrable. That is, if u is a trial solution, then

$$\int_0^1 (u_{,x})^2\, dx < \infty \tag{1.3.1}$$

Functions that satisfy (1.3.1) are called H^1-***functions***; we write $u \in H^1$. Sometimes the domain is explicitly included, i.e., $u \in H^1([0, 1])$.

Thus the collection of ***trial solutions***, denoted by $\mathcal{S}$, consists of all functions which have square-integrable derivatives and take on the value q at $x = 1$. This is written as follows:

$$\mathcal{S} = \{u \mid u \in H^1,\ u(1) = q\} \qquad \text{(trial solutions)} \tag{1.3.2}$$

The fact that $\mathcal{S}$ is a collection, or set, of objects is indicated by the curly brackets (called braces) in (1.3.2). The notation for the typical member of the set, in this case u, comes first inside the left-hand curly bracket. Following the vertical line ($|$) are the properties satisfied by members of the set.

The second collection of functions is called the ***weighting functions,*** or ***variations.*** This collection is very similar to the trial solutions except we require the homogeneous counterpart of the g-boundary condition. That is, we require weighting functions, w, to satisfy $w(1) = 0$. The collection is denoted by $\mathcal{V}$ and defined by

$$\mathcal{V} = \{w \mid w \in H^1,\ w(1) = 0\} \qquad \text{(weighting functions)} \qquad (1.3.3)$$

It simplifies matters somewhat to continue to think of $\ell : \Omega \to \mathbb{R}$ as being smooth. (However, what follows holds for a considerably larger class of ℓ's.)

In terms of the preceding definitions, we may now state a suitable weak form, (W), of the boundary-value problem.

$$(W)\begin{cases} \text{Given } \ell,\ g,\ \text{and } h, \text{ as before. Find } u \in \mathcal{S} \text{ such that for all } w \in \mathcal{V} \\ \boxed{\displaystyle\int_0^1 w_{,x} u_{,x}\, dx = \int_0^1 w\ell\, dx + w(0)h} \end{cases} \qquad (1.3.4)$$

Formulations of this type are often called ***virtual work,*** or ***virtual displacement, principles*** in mechanics. The w's are the ***virtual displacements.***

Equation (1.3.4) is called the ***variational equation,*** or (especially in mechanics) the ***equation of virtual work.***

The solution of (W) is called the ***weak,*** or ***generalized, solution.*** The definition given of a weak formulation is not the only one possible, but it is the most natural one for the problems we wish to consider.

1.4 EQUIVALENCE OF STRONG AND WEAK FORMS; NATURAL BOUNDARY CONDITIONS

Clearly, there must be some relationship between the strong and weak versions of the problem, or else there would be no point in introducing the weak form. It turns out that the weak and strong solutions are identical. We shall establish this assuming all functions are smooth. This will allow us to proceed expeditiously without invoking technical conditions with which the reader is assumed to be unfamiliar. "Proofs" of this kind are sometimes euphemistically referred to as "formal proofs." The intent is not to be completely rigorous but rather to make plausible the truth of the proposition. With this philosophy in mind, we shall "prove" the following.

Proposition

a. Let u be a solution of (S). Then u is also a solution of (W).
b. Let u be a solution of (W). Then u is also a solution of (S).

Another result, which we shall not bother to verify but is in fact easily established,

is that both (S) and (W) possess unique solutions. Thus, by (a) and (b), the strong and weak solutions are one and the same. Consequently, (W) is equivalent to (S).

Formal Proof

a. Since u is assumed to be a solution of (S), we may write

$$0 = -\int_0^1 w(u_{,xx} + \ell)dx \tag{1.4.1}$$

for any $w \in \mathcal{V}$. Integrating (1.4.1) by parts results in

$$0 = \int_0^1 w_{,x}u_{,x}\,dx - \int_0^1 w\ell\,dx - wu_{,x}\Big|_0^1 \tag{1.4.2}$$

Rearranging and making use of the fact that $-u_{,x}(0) = h$ and $w(1) = 0$ results in

$$\int_0^1 w_{,x}u_{,x}\,dx = \int_0^1 w\ell\,dx + w(0)h \tag{1.4.3}$$

Furthermore, since u is a solution of (S), it satisfies $u(1) = g$ and therefore is in $\mathcal{S}$. Finally, since u also satisfies (1.4.3) for all $w \in \mathcal{V}$, u satisfies the definition of a weak solution given by (W).

b. Now u is assumed to be a weak solution. Thus $u \in \mathcal{S}$; consequently $u(1) = g$, and

$$\int_0^1 w_{,x}u_{,x}\,dx = \int_0^1 w\ell\,dx + w(0)h$$

for all $w \in \mathcal{V}$. Integrating by parts and making use of the fact $w(1) = 0$ results in

$$0 = \int_0^1 w(u_{,xx} + \ell)\,dx + w(0)[u_{,x}(0) + h] \tag{1.4.4}$$

To prove u is a solution of (S) it suffices to show that (1.4.4) implies[1]

i. $u_{,xx} + \ell = 0$ on Ω; and
ii. $u_{,x}(0) + h = 0$

First we shall prove (i). Define w in (1.4.4) by

$$w = \phi(u_{,xx} + \ell) \tag{1.4.5}$$

where ϕ is smooth; $\phi(x) > 0$ for all $x \in \Omega =]0, 1[$; and $\phi(0) = \phi(1) = 0$. For example, we can take $\phi(x) = x(1 - x)$, which satisfies all the stipulated requirements (see Figure 1.4.1). It follows that $w(1) = 0$ and thus $w \in \mathcal{V}$, so (1.4.5) defines a

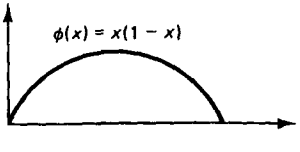


Figure 1.4.1

[1] These equations are sometimes called the ***Euler-Lagrange equations*** of the weak formulation.

legitimate member of $\mathcal{V}$. Substituting (1.4.5) into (1.4.4) results in

$$0 = \int_0^1 \underbrace{\phi(u_{,xx} + f)^2}_{\geq 0}\, dx + 0 \tag{1.4.6}$$

Since $\phi > 0$ on Ω, it follows from (1.4.6) that (i) must be satisfied.

Now that we have established (i), we may use it in (1.4.4) to prove (ii), namely,

$$0 = w(0)[u_{,x}(0) + h] \tag{1.4.7}$$

That $w \in \mathcal{V}$ puts no restriction whatsoever on its value at $x = 0$. Therefore, we may assume that the w in (1.4.7) is such that $w(0) \neq 0$. Thus (ii) is also shown to hold, which completes the proof of the proposition. ■

Remarks

1. The boundary condition $-u_{,x}(0) = h$ is not explicitly mentioned in the statement of (W). From the preceding proof, we see that this boundary condition is, however, implied by the satisfaction of the variational equation. Boundary conditions of this type are referred to as ***natural boundary conditions***. On the other hand, trial solutions are explicitly required to satisfy the boundary condition $u(1) = g$. Boundary conditions of this type are called ***essential boundary conditions***. The fact that solutions of the variational equation satisfy natural boundary conditions is extremely important in more complicated situations which we will consider later on.

2. The method used to prove part (b) of the proposition goes under the name of the ***fundamental lemma*** in the literature of the calculus of variations. In essence, it is the methodology that enables us to deduce the differential equations and boundary conditions implied by the weak formulation. To develop correct weak forms for complex, multidimensional problems, it is essential to have a thorough understanding of these procedures.

Now we see that to obtain approximate solutions to the original boundary-value problem we have alternative starting points, i.e., the strong or weak statements of the problem. *Finite element methods are based upon the latter*. Roughly speaking, the basic idea is to approximate $\mathcal{S}$ and $\mathcal{V}$ by convenient, finite-dimensional collections of functions. (Clearly, $\mathcal{S}$ and $\mathcal{V}$ contain infinitely many functions.) The variational equations are then solved in this finite-dimensional context. An explicit example of how to go about this is the subject of the next section. However, we first introduce some additional notations to simplify subsequent writing.

Let

$$a(w, u) = \int_0^1 w_{,x} u_{,x}\, dx \tag{1.4.8}$$

$$(w, f) = \int_0^1 wf\, dx \tag{1.4.9}$$

In terms of (1.4.8) and (1.4.9), the variational equation takes the form

$$\boxed{a(w, u) = (w, f) + w(0)h} \tag{1.4.10}$$

Here, $a(\cdot, \cdot)$ and $(\cdot, \cdot)$ are examples of ***symmetric, bilinear forms***. What this means is as follows: Let c_1 and c_2 be constants and let u, v, and w be functions. Then the ***symmetry*** property is

$$a(u, v) = a(v, u) \tag{1.4.11}$$

$$(u, v) = (v, u) \tag{1.4.12}$$

Bilinearity means linearity in each "slot"; for example,

$$a(c_1u + c_2v, w) = c_1a(u, w) + c_2a(v, w) \tag{1.4.13}$$

$$(c_1u + c_2v, w) = c_1(u, w) + c_2(v, w) \tag{1.4.14}$$

Exercise 1. Use the definitions of $a(\cdot, \cdot)$ and $(\cdot, \cdot)$ to verify the properties of symmetry and bilinearity.

The above notations are very concise; at the same time they capture essential mathematical features and thus are conducive to a mathematical understanding of variational and finite element methods. Diverse classes of physical problems can be written in essentially similar fashion to (1.4.10). Thus ideas developed and results obtained are seen at once to have very broad applicability.

1.5 GALERKIN'S APPROXIMATION METHOD

We shall now describe a method of obtaining approximate solutions to boundary-value problems based upon weak formulations. Our introduction to this subject is somewhat of an abstract treatment. However, the meaning should be significantly reinforced by the remaining sections of the chapter. It may be worthwhile for the reader to consult this section again after completing the rest of the chapter to make sure a full comprehension of the material is attained.

The first step in developing the method is to construct finite-dimensional approximations of $\mathcal{S}$ and $\mathcal{V}$. These collections of functions are denoted by $\mathcal{S}^h$ and $\mathcal{V}^h$, respectively. The superscript refers to the association of $\mathcal{S}^h$ and $\mathcal{V}^h$ with a ***mesh***, or ***discretization***, of the domain Ω, which is parameterized by a characteristic length scale h. We wish to think of $\mathcal{S}^h$ and $\mathcal{V}^h$ as being subsets of $\mathcal{S}$ and $\mathcal{V}$, respectively. This is written as

$$\mathcal{S}^h \subset \mathcal{S} \qquad \text{(i.e., if } u^h \in \mathcal{S}^h, \text{ then } u^h \in \mathcal{S}) \tag{1.5.1}$$

$$\mathcal{V}^h \subset \mathcal{V} \qquad \text{(i.e., if } w^h \in \mathcal{V}^h, \text{ then } w^h \in \mathcal{V}) \tag{1.5.2}$$

where the precise meaning is given in parentheses.[2] Consequences of (1.5.1) and (1.5.2) are (respectively) that if $u^h \in \mathcal{S}^h$ and $w^h \in \mathcal{V}^h$, then

$$u^h(1) = g \tag{1.5.3}$$

$$w^h(1) = 0 \tag{1.5.4}$$

The collections, $\mathcal{S}$, $\mathcal{V}$, $\mathcal{S}^h$, and $\mathcal{V}^h$, are often referred to as ***function spaces***. The terminology *space* in mathematics usually connotes a linear structure. This has the following meaning: If c_1 and c_2 are constants and v and w are in $\mathcal{V}$, then $c_1 v + c_2 w$ is also in $\mathcal{V}$. Both $\mathcal{V}$ and $\mathcal{V}^h$ are thus seen to possess the property of a linear space. However, this property is clearly not shared by $\mathcal{S}$ and $\mathcal{S}^h$ due to the inhomogeneous boundary condition. For example, if u_1 and u_2 are members of $\mathcal{S}$, then $u_1 + u_2 \notin \mathcal{S}$, since $u_1(1) + u_2(1) = g + g = 2g$ in violation of the definition of $\mathcal{S}$. Nevertheless, the terminology function space is still (loosely) applied to $\mathcal{S}$ and $\mathcal{S}^h$.

(Bubnov-) Galerkin Method

Assume the collection $\mathcal{V}^h$ is given. Then, to each member $v^h \in \mathcal{V}^h$, we construct a function $u^h \in \mathcal{S}^h$ by

$$u^h = v^h + g^h \tag{1.5.5}$$

where g^h is a *given* function satisfying the essential boundary condition, i.e.,

$$g^h(1) = g \tag{1.5.6}$$

Note that (1.5.5) satisfies the requisite boundary condition also:

$$\begin{aligned} u^h(1) &= v^h(1) + g^h(1) \\ &= 0 + g \end{aligned} \tag{1.5.7}$$

Thus (1.5.5) constitutes a definition of $\mathcal{S}^h$; that is, $\mathcal{S}^h$ is all functions of the form (1.5.5). The key point to observe is that, up to the function g^h, $\mathcal{S}^h$ and $\mathcal{V}^h$ are composed of *identical* collections of functions. This property will be shown later on to have significant consequences for certain classes of problems.

We now write a variational equation, of the form of (1.4.10), in terms of $w^h \in \mathcal{V}^h$ and $u^h \in \mathcal{S}^h$:

$$\boxed{a(w^h, u^h) = (w^h, \ell) + w^h(0)h} \tag{1.5.8}$$

This equation is to be thought of as defining an approximate (weak) solution, u^h.

[2] This condition may be considered standard. However, it is often violated in practice. Strang [2] coined the terminology "variational crimes" to apply to this, and other, situations in which the classical rules of variational methods are violated. Many "variational crimes" have been given a rigorous mathematical basis (e.g., see [2]). We shall have more to say about this subject in subsequent chapters.

Substitution of (1.5.5) into (1.5.8), and the bilinearity of $a(\cdot, \cdot)$ enables us to write

$$\boxed{a(w^h, v^h) = (w^h, \ell) + w^h(0)h - a(w^h, g^h)} \tag{1.5.9}$$

The right-hand side consists of the totality of terms associated with given data (i.e., ℓ, g, and h). Equation (1.5.9) is to be used to define v^h, the unknown part of u^h.

The (Bubnov-) Galerkin form of the problem, denoted by (G), is stated as follows:

$$(G)\begin{cases} \text{Given } \ell,\ g, \text{ and } h, \text{ as before, find } u^h = v^h + g^h, \text{ where } v^h \in \mathcal{V}^h, \\ \text{such that for all } w^h \in \mathcal{V}^h \\ \qquad a(w^h, v^h) = (w^h, \ell) + w^h(0)h - a(w^h, g^h) \end{cases}$$

Note that (G) is just a version of (W) posed in terms of a finite-dimensional collection of functions, namely, $\mathcal{V}^h$.

To make matters more specific, g^h and $\mathcal{V}^h$ have to be explicitly defined. Before doing this, it is worthwhile to mention a larger class of approximation methods, called ***Petrov-Galerkin methods,*** in which v^h is contained in a collection of functions other than $\mathcal{V}^h$. Recent attention has been paid to methods of this type, especially in the context of fluid mechanics. For the time being, we will be exclusively concerned with the Bubnov-Galerkin method. The Bubnov-Galerkin method is commonly referred to as simply the Galerkin method, terminology we shall adopt henceforth. Equation (1.5.9) is sometimes referred to as the ***Galerkin equation***.

Approximation methods of the type considered are examples of so-called ***weighted residual methods***. The standard reference dealing with this subject is Finlayson [3]. For a more succinct presentation containing an interesting historical account, see Finlayson and Scriven [4].

1.6 MATRIX EQUATIONS; STIFFNESS MATRIX K

The Galerkin method leads to a coupled system of linear algebraic equations. To see this we need to give further structure to the definition of $\mathcal{V}^h$. Let $\mathcal{V}^h$ consist of all linear combinations of given functions denoted by $N_A : \overline{\Omega} \to \mathbb{R}$, where $A = 1, 2, \ldots, n$. By this we mean that if $w^h \in \mathcal{V}^h$, then there exist constants c_A, $A = 1, 2, \ldots, n$, such that

$$\begin{aligned} w^h &= \sum_{A=1}^{n} c_A N_A \\ &= c_1 N_1 + c_2 N_2 + \cdots + c_n N_n \end{aligned} \tag{1.6.1}$$

The N_A's are referred to as ***shape, basis,*** or ***interpolation*** functions. We require that each N_A satisfies

$$N_A(1) = 0, \qquad A = 1, 2, \ldots, n \tag{1.6.2}$$

from which it follows by (1.6.1) that $w^h(1) = 0$, as is necessary. $\mathcal{V}^h$ is said to have dimension n, for obvious reasons.

To define members of $\mathcal{S}^h$ we need to specify q^h. To this end, we introduce another shape function, $N_{n+1} : \overline{\Omega} \to \mathbb{R}$, which has the property

$$N_{n+1}(1) = 1 \tag{1.6.3}$$

(Note $N_{n+1} \notin \mathcal{V}^h$.) Then q^h is given by

$$q^h = qN_{n+1} \tag{1.6.4}$$

and thus

$$q^h(1) = q \tag{1.6.5}$$

With these definitions, a typical $u^h \in \mathcal{S}^h$ may be written as

$$\begin{aligned} u^h &= v^h + q^h \\ &= \sum_{A=1}^{n} d_A N_A + qN_{n+1} \end{aligned} \tag{1.6.6}$$

where the d_A's are constants and from which it is apparent that $u^h(1) = q$.

Substitution of (1.6.1) and (1.6.6) into the Galerkin equation yields

$$\begin{aligned} a\left(\sum_{A=1}^{n} c_A N_A, \sum_{B=1}^{n} d_B N_B\right) = \left(\sum_{A=1}^{n} c_A N_A, \ell\right) + \left[\sum_{A=1}^{n} c_A N_A(0)\right] h \\ - a\left(\sum_{A=1}^{n} c_A N_A, qN_{n+1}\right) \end{aligned} \tag{1.6.7}$$

By using the bilinearity of $a(\cdot, \cdot)$ and $(\cdot, \cdot)$, (1.6.7) becomes

$$0 = \sum_{A=1}^{n} c_A G_A \tag{1.6.8}$$

where

$$G_A = \sum_{B=1}^{n} a(N_A, N_B) d_B - (N_A, \ell) - N_A(0)h + a(N_A, N_{n+1})q \tag{1.6.9}$$

Now the Galerkin equation is to hold for all $w^h \in \mathcal{V}^h$. By (1.6.1), this means for all c_A's, $A = 1, 2, \ldots, n$. Since the c_A's are arbitrary in (1.6.8), it necessarily follows that each G_A, $A = 1, 2, \ldots, n$, must be identically zero, i.e., from (1.6.9)

$$\boxed{\sum_{B=1}^{n} a(N_A, N_B) d_B = (N_A, \ell) + N_A(0)h - a(N_A, N_{n+1})q} \tag{1.6.10}$$

Note that everything is known in (1.6.10) except the d_B's. Thus (1.6.10) constitutes a system of n equations in n unknowns. This can be written in a more concise form as follows:

Let

$$K_{AB} = a(N_A, N_B) \tag{1.6.11}$$

$$F_A = (N_A, f) + N_A(0)t - a(N_A, N_{n+1})q \tag{1.6.12}$$

Then (1.6.10) becomes

$$\sum_{B=1}^{n} K_{AB} d_B = F_A, \qquad A = 1, 2, \ldots, n \tag{1.6.13}$$

Further simplicity is gained by adopting a matrix notation. Let

$$\boldsymbol{K} = [K_{AB}] = \begin{bmatrix} K_{11} & K_{12} & \cdots & K_{1n} \\ K_{21} & K_{22} & \cdots & K_{2n} \\ \vdots & \vdots & & \vdots \\ K_{n1} & K_{n2} & \cdots & K_{nn} \end{bmatrix} \tag{1.6.14}$$

$$\boldsymbol{F} = \{F_A\} = \begin{Bmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{Bmatrix} \tag{1.6.15}$$

and

$$\boldsymbol{d} = \{d_B\} = \begin{Bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{Bmatrix} \tag{1.6.16}$$

Now (1.6.13) may be written as

$$\boldsymbol{K}\boldsymbol{d} = \boldsymbol{F} \tag{1.6.17}$$

The following terminologies are frequently applied, especially when the problem under consideration pertains to a mechanical system:

$\boldsymbol{K}$ = stiffness matrix

$\boldsymbol{F}$ = force vector

$\boldsymbol{d}$ = displacement vector

A variety of physical interpretations are of course possible.

At this point, we may state the matrix equivalent, (M), of the Galerkin problem.

$$(M)\begin{cases} \text{Given the coefficient matrix } \boldsymbol{K} \text{ and vector } \boldsymbol{F}, \text{ find } \boldsymbol{d} \text{ such that} \\ \boldsymbol{K}\boldsymbol{d} = \boldsymbol{F} \end{cases}$$

The solution of (M) is, of course, just $\boldsymbol{d} = \boldsymbol{K}^{-1}\boldsymbol{F}$ (assuming the inverse of $\boldsymbol{K}$, $\boldsymbol{K}^{-1}$, exists). Once $\boldsymbol{d}$ is known, the solution of (G) may be obtained at any point $x \in \overline{\Omega}$ by employing (1.6.6), viz.,

$$u^h(x) = \sum_{A=1}^{n} d_A N_A(x) + gN_{n+1}(x) \tag{1.6.18}$$

Likewise, derivatives of u^h, if required, may be obtained by term-by-term differentiation. It should be emphasized, that the solution of (G) is an *approximate* solution of (W). Consequently, the differential equation and natural boundary condition are only approximately satisfied. The quality of the approximation depends upon the specific choice of N_A's and the number n.

Remarks

1. The matrix $\boldsymbol{K}$ is symmetric. This follows from the symmetry of $a(\cdot, \cdot)$ and use of Galerkin's method (i.e., the same shape functions are used for the variations and trial solutions):

$$\begin{aligned} K_{AB} &= a(N_A, N_B) \\ &= a(N_B, N_A) \\ &= K_{BA} \end{aligned} \tag{1.6.19}$$

In matrix notation

$$\boldsymbol{K} = \boldsymbol{K}^T \tag{1.6.20}$$

where the superscript T denotes transpose. The symmetry of $\boldsymbol{K}$ has important computational consequences.

2. Let us schematically retrace the steps leading to the matrix problem, as they are typical of the process one must go through in developing a finite element method for any given problem:

$$\boxed{(S) \Leftrightarrow (W) \approx (G) \Leftrightarrow (M)} \tag{1.6.21}$$

The only apparent approximation made thus far is in approximately solving (W) via (G). In more complicated situations, encountered in practice, the number of approximations increases. For example, the data ℓ, q, and h may be approximated, as well as the domain Ω, calculation of integrals, and so on. Convergence proofs and error analyses involve consideration of each approximation.

3. It is sometimes convenient to write

$$u^h(x) = \sum_{A=1}^{n+1} N_A(x) d_A \tag{1.6.22}$$

where $d_{n+1} = q$.

1.7 EXAMPLES: 1 AND 2 DEGREES OF FREEDOM

In this section we will carry out the detailed calculations involved in formulating and solving the Galerkin problem. The functions employed are extremely simple, thus expediting computations, but they are also primitive examples of typical finite element functions.

Example 1 (1 degree of freedom)

In this case $n = 1$. Thus $w^h = c_1 N_1$ and $u^h = v^h + g^h = d_1 N_1 + q N_2$. The only unknown is d_1. The shape functions must satisfy $N_1(1) = 0$ and $N_2(1) = 1$ (see (1.6.2) and (1.6.3)). Let us take $N_1(x) = 1 - x$ and $N_2(x) = x$. These are illustrated in Fig. 1.7.1 and clearly satisfy the required conditions. Since we are dealing with only 1 degree of freedom, the matrix paraphernalia collapses as follows:

$$\boldsymbol{K} = [K_{11}] = K_{11} \tag{1.7.1}$$

$$\boldsymbol{F} = \{F_1\} = F_1 \tag{1.7.2}$$

$$\boldsymbol{d} = \{d_1\} = d_1 \tag{1.7.3}$$

$$K_{11} = a(N_1, N_1) = \int_0^1 \underbrace{N_{1,x}}_{-1} \underbrace{N_{1,x}}_{-1} dx = 1 \tag{1.7.4}$$

$$\begin{aligned} F_1 &= (N_1, \ell) + N_1(0)h - a(N_1, N_2)q \\ &= \int_0^1 (1 - x)\ell(x)\, dx + h - \int_0^1 \underbrace{N_{1,x}}_{-1} \underbrace{N_{2,x}}_{+1} dx\, q \\ &= \int_0^1 (1 - x)\ell(x)\, dx + h + q \end{aligned} \tag{1.7.5}$$

$$d_1 = K_{11}^{-1} F_1 = F_1 \tag{1.7.6}$$

Consequently

$$u^h(x) = \underbrace{\left[\int_0^1 (1 - y)\ell(y)\, dy + h + q\right]}_{d_1} (1 - x) + qx \tag{1.7.7}$$

In (1.7.7), y plays the role of a dummy variable. An illustration of (1.7.7) appears in Fig. 1.7.2. To get a feel for the nature of the approximation, let us compare (1.7.7) with the exact solution (see (1.2.3)). It is helpful to consider specific forms for ℓ.

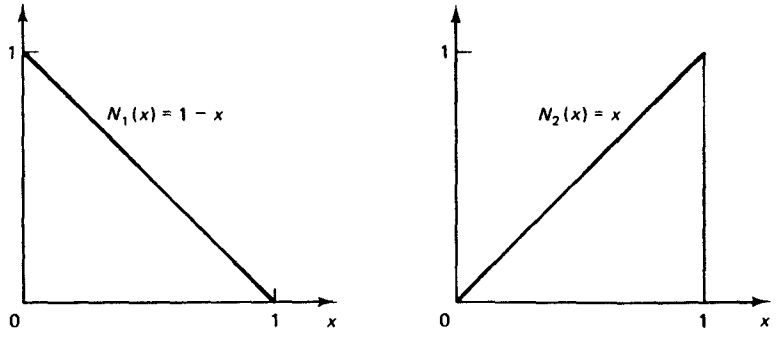


Figure 1.7.1 Functions for the 1 degree of freedom examples. (These functions are secretly the simplest finite element interpolation functions in a one-element context.)

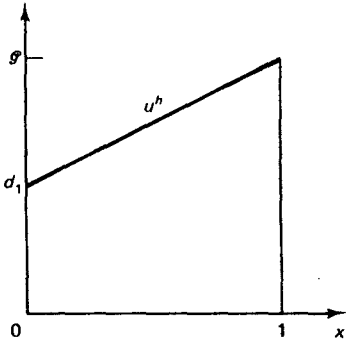


Figure 1.7.2 The Galerkin solution for the 1 degree of freedom example.

i. Let $f = 0$. Then

$$u^h(x) = u(x) = q + (1 - x)h \tag{1.7.8}$$

That is, the approximate solution is exact. In fact, it is clear by inspecting (1.7.7) and (1.2.3) that the homogeneous solution (i.e., the part of the solution corresponding to $f = 0$) is always exactly represented. The only approximation pertains to the particular solution (i.e., the part of the solution corresponding to $f \neq 0$).

ii. Now let us introduce a nonzero f. Assume $f(x) = p$, a constant. Then the particular solutions take the form

$$u_{\text{part}}(x) = \frac{p(1 - x^2)}{2} \tag{1.7.9}$$

and

$$u^h_{\text{part}}(x) = \frac{p(1 - x)}{2} \tag{1.7.10}$$

Equations (1.7.9) and (1.7.10) are compared in Fig. 1.7.3. Note that u^h_{part} is exact at $x = 0$ and $x = 1$ and that $u^h_{\text{part},x}$ is exact at $x = \frac{1}{2}$. (It should be clear that it is impossible for u^h_{part} to be exact at all x in the present circumstances. The exact solution, (1.7.9), contains a quadratic term in x, whereas the approximate solution is restricted to linear variation in x by the definitions of N_1 and N_2.)

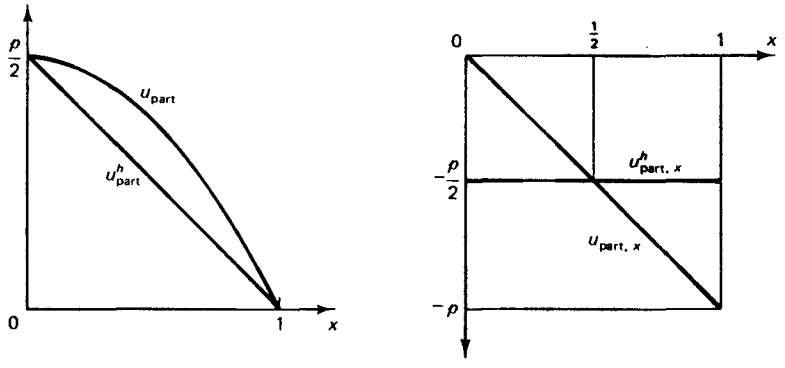


Figure 1.7.3 Comparison of exact and Galerkin particular solutions, Example 1, case (ii).

iii. This time let $f(x) = qx$, where q is a constant. This choice for f leads to

$$u_{\text{part}}(x) = \frac{q(1 - x^3)}{6} \tag{1.7.11}$$

and

$$u^h_{\text{part}}(x) = \frac{q(1 - x)}{6} \tag{1.7.12}$$

which are compared in Fig. 1.7.4. Again we note that the u^h_{part} is exact at $x = 0$ and $x = 1$. There is one point, $x = 1/\sqrt{3}$, at which $u^h_{\text{part},x}$ is exact.

Let us summarize what we have observed in this example:

a. The homogeneous part of u^h is exact in all cases.
b. In the presence of nonzero f, u^h is exact at $x = 0$ and $x = 1$.
c. For each case, there is at least one point at which $u^h_{,x}$ is exact.

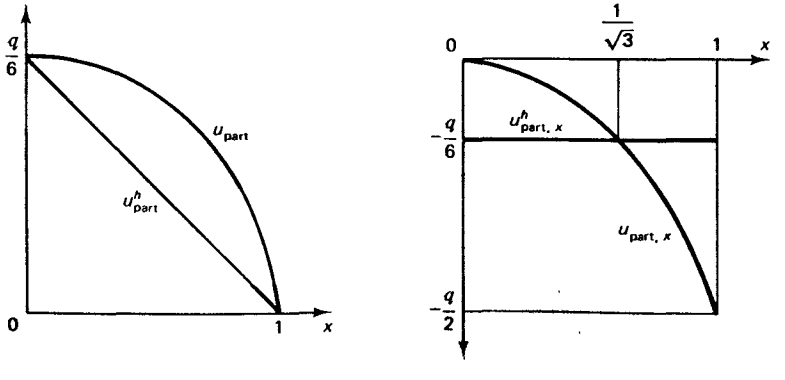


Figure 1.7.4 Comparison of exact and Galerkin particular solutions, Example 1, case (iii).

Example 2 (2 degrees of freedom)

In this case $n = 2$. Thus $w^h = c_1N_1 + c_2N_2$, where $N_1(1) = N_2(1) = 0$, and $u^h = d_1N_1 + d_2N_2 + gN_3$, where $N_3(1) = 1$. Let us define the N_A's as follows

$$N_1(x) = \begin{cases} 1 - 2x & 0 \le x \le \frac{1}{2} \\ 0 & \frac{1}{2} \le x \le 1 \end{cases} \tag{1.7.6}$$

$$N_2(x) = \begin{cases} 2x & 0 \le x \le \frac{1}{2} \\ 2(1 - x) & \frac{1}{2} \le x \le 1 \end{cases} \tag{1.7.7}$$

$$N_3(x) = \begin{cases} 0 & 0 \le x \le \frac{1}{2} \\ 2x - 1 & \frac{1}{2} \le x \le 1 \end{cases} \tag{1.7.8}$$

The shape functions are illustrated in Fig. 1.7.5. Typical $w^h \in \mathcal{V}^h$ and $u^h \in \mathcal{S}^h$ and their derivatives are shown in Fig. 1.7.6. Since $n = 2$, the matrix paraphernalia takes the following form:

$$\boldsymbol{K} = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \tag{1.7.9}$$

$$\boldsymbol{F} = \begin{Bmatrix} F_1 \\ F_2 \end{Bmatrix} \tag{1.7.10}$$

$$\boldsymbol{d} = \begin{Bmatrix} d_1 \\ d_2 \end{Bmatrix} \tag{1.7.11}$$

$$K_{AB} = a(N_A, N_B) = \int_0^1 N_{A,x}N_{B,x}\,dx = \int_0^{1/2} N_{A,x}N_{B,x}\,dx + \int_{1/2}^1 N_{A,x}N_{B,x}\,dx \tag{1.7.12}$$

$$K_{11} = 2, \qquad K_{12} = K_{21} = -2, \qquad K_{22} = 4 \tag{1.7.13}$$

$$\boldsymbol{K} = 2\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \tag{1.7.14}$$

$$\begin{aligned} F_A &= (N_A, \ell) + N_A(0)h - a(N_A, N_3)g \\ &= \int_0^1 N_A \ell\,dx + N_A(0)h - \int_{1/2}^1 N_{A,x}N_{3,x}\,dx\,g \end{aligned} \tag{1.7.15}$$

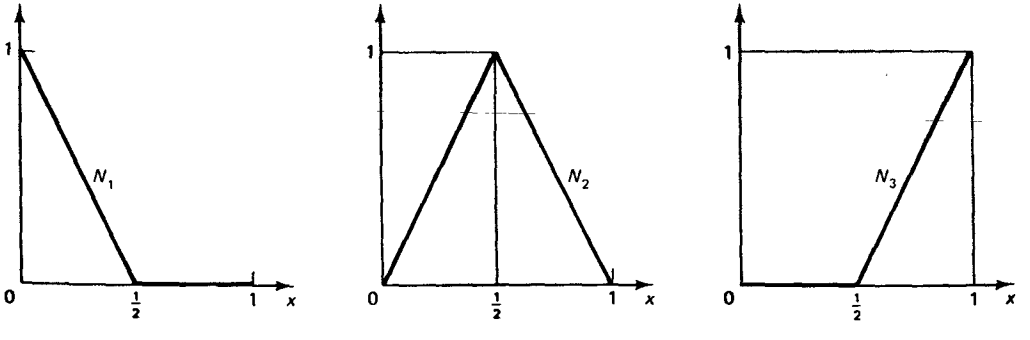


Figure 1.7.5 Functions for the 2 degree of freedom examples. (These functions are secretly the simplest finite element functions in a two-element context.)

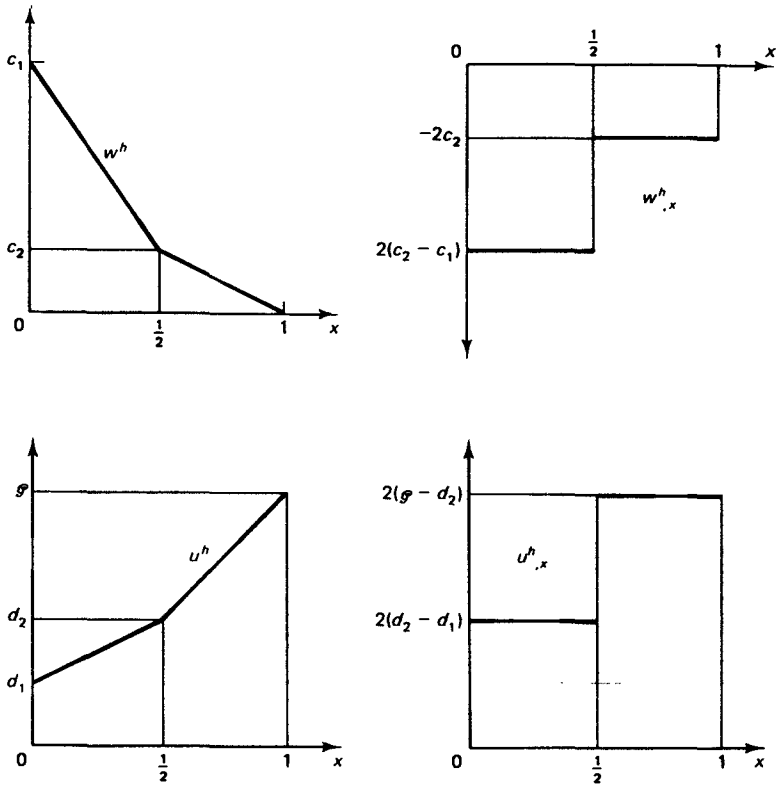


Figure 1.7.6 Typical weighting function and trial solution for the 2 degree of freedom example.

$$F_1 = \int_0^{1/2} (1 - 2x)\ell(x)\, dx + h \tag{1.7.16}$$

$$F_2 = 2\int_0^{1/2} x\ell(x)\, dx + 2\int_{1/2}^{1} (1 - x)\ell(x)\, dx + 2q \tag{1.7.17}$$

Note that due to the shape functions' discontinuities in slope at $x = \frac{1}{2}$, it is convenient to express integrals over the subintervals $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$ (e.g., see (1.7.12) and (1.7.15)). We need not worry about the value of the derivative of N_A at $x = \frac{1}{2}$ (it suffers a discontinuity there and thus is not well-defined classically) since it has no effect on the integrals in (1.7.12). This amounts to employing the notion of a ***generalized derivative.***

We shall again analyze the three cases considered in Example 1.

i. $\ell = 0$.

$$\boldsymbol{F} = \begin{Bmatrix} h \\ 2q \end{Bmatrix} \tag{1.7.18}$$

$$\boldsymbol{d} = \boldsymbol{K}^{-1}\boldsymbol{F}$$

$$= \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{Bmatrix} h \\ 2q \end{Bmatrix}$$

$$= \begin{Bmatrix} q + h \\ q + \frac{h}{2} \end{Bmatrix} \tag{1.7.19}$$

This results in

$$u^h = (q + h)N_1 + \left(q + \frac{h}{2}\right)N_2 + qN_3$$

$$= q(N_1 + N_2 + N_3) + h\left(N_1 + \frac{N_2}{2}\right) \tag{1.7.20}$$

$$u^h(x) = q + h(1 - x) \tag{1.7.21}$$

Again, the exact homogeneous solution is obtained. (The reason for this is that the exact solution is linear, and our trial solution is capable of exactly representing any linear function. Galerkin's method will give the exact answer whenever possible—that is, whenever the collection of trial solutions contains the exact solution among its members.)

ii. $f(x) = p$ = constant.

$$F_1 = \frac{p}{4} + h \tag{1.7.22}$$

$$F_2 = \frac{p}{2} + 2q \tag{1.7.23}$$

$$\boldsymbol{d} = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{Bmatrix} \frac{p}{4} + h \\ \frac{p}{2} + 2q \end{Bmatrix} = \begin{Bmatrix} \frac{p}{2} + q + h \\ \frac{3p}{8} + q + \frac{h}{2} \end{Bmatrix} \tag{1.7.24}$$

The solution takes the form

$$u^h(x) = q + h(1 - x) + u^h_{\text{part}}(x) \tag{1.7.25}$$

$$u^h_{\text{part}} = \frac{p}{2}N_1 + \frac{3p}{8}N_2 \tag{1.7.26}$$

The approximate particular solution is compared with the exact in Fig. 1.7.7, from which we see that agreement is achieved at $x = 0, \frac{1}{2}$ and 1, and derivatives coincide at $x = \frac{1}{4}$ and $\frac{3}{4}$.

iii. $f(x) = qx$, q = constant.

$$F_1 = \frac{q}{24} + h \tag{1.7.27}$$

$$F_2 = \frac{q}{4} + 2q \tag{1.7.28}$$

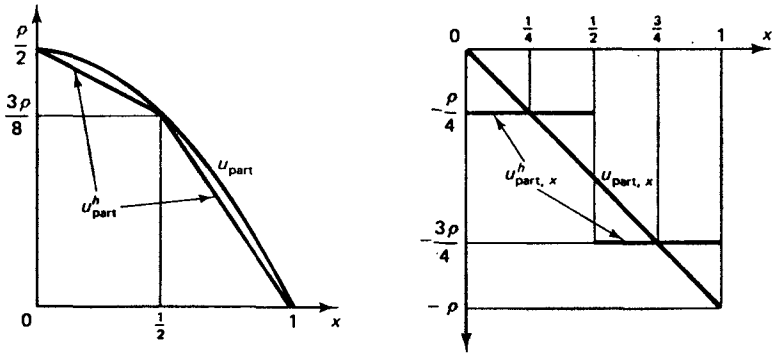


Figure 1.7.7 Comparison of exact and Galerkin particular solutions, Example 2, case (ii).

$$\boldsymbol{d} = \begin{Bmatrix} \frac{q}{6} + q + h \\ \frac{7q}{48} + q + \frac{h}{2} \end{Bmatrix}$$

Again u^h may be expressed in the form (1.7.25), where

$$u^h_{\text{part}} = \frac{q}{6}N_1 + \frac{7q}{48}N_2$$

A comparison is presented in Fig. 1.7.8. The Galerkin solution is seen to be ex again at $x = 0, \frac{1}{2}$, and 1, and the derivative is exact at two points.

Let us summarize the salient observations of Example 2:

a. The homogeneous part of u^h is exact in all cases, as in Example 1. (A 1 for this is given after Equation (1.7.21).)

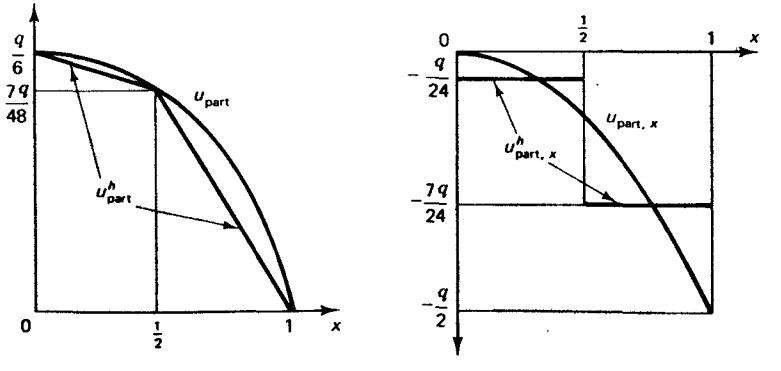


Figure 1.7.8 Comparison of exact and Galerkin particular solutions, Example 2, case (iii).

b. The Galerkin solution is exact at the endpoints of each subinterval for all cases.
c. In each case, there is at least one point in each subinterval at which u'_h is exact.

After generalizing to the case of n subintervals in the following section, we shall show in Sec. 1.10 that the above observations are not accidental.

Exercise 1. If the reader has not had experience with calculations of the type presented in this section, it would be worthwhile to reproduce all results, providing all omitted details.

1.8 PIECEWISE LINEAR FINITE ELEMENT SPACE

The examples of the preceding section employed definitions of $\mathcal{V}^h$ and $\mathcal{S}^h$ which were special cases of the so-called piecewise linear finite element space. To define the general case in which $\mathcal{V}^h$ is n-dimensional, we partition the domain $[0, 1]$ into n nonoverlapping subintervals. The typical subinterval is denoted by $[x_A, x_{A-1}]$, where $x_A < x_{A+1}$ and $A = 1, 2, \ldots, n$. We also require $x_1 = 0$ and $x_{n-1} = 1$. The x_A's are called ***nodal points,*** or simply ***nodes.*** (The terminologies ***joints*** and ***knots*** are also used.) The subintervals are sometimes referred to as the ***finite element domains,*** or simply ***elements.*** Notice that the lengths of the elements, $h_A = x_{A+1} - x_A$, are *not* required to be equal. The mesh parameter, h, is generally taken to be the length of the maximum subinterval (i.e., $h = \max h_A$, $A = 1, 2, \ldots, n$). The smaller h, the more "refined" is the partition, or mesh. If the subinterval lengths are equal, then $h = 1/n$.

The shape functions are defined as follows: Associated to a typical internal node (i.e., $2 \leq A \leq n$)

$$N_A(x) = \begin{cases} \dfrac{(x - x_{A-1})}{h_{A-1}}, & x_{A-1} \leq x \leq x_A \\ \dfrac{(x_{A+1} - x)}{h_A}, & x_A \leq x \leq x_{A-1} \\ 0, & \text{elsewhere} \end{cases} \tag{1.8.1}$$

whereas for the boundary nodes we have

$$N_1(x) = \frac{x_2 - x}{h_1}, \qquad x_1 \leq x \leq x_2 \tag{1.8.2}$$

$$N_{n+1}(x) = \frac{x - x_n}{h_n}, \qquad x_n \leq x \leq x_{n+1} \tag{1.8.3}$$

The shape functions are sketched in Fig. 1.8.1. For obvious reasons, they are referred to variously as "hat," "chapeau," and "roof" functions. Note that $N_A(x_B) = \delta_{AB}$, where

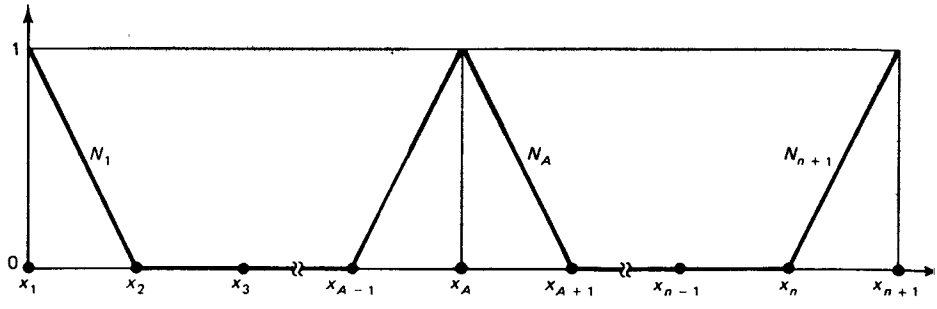


Figure 1.8.1 Basis functions for the piecewise linear finite element space.

δ_{AB} is the Kronecker delta (i.e., $\delta_{AB} = 1$ if $A = B$, whereas $\delta_{AB} = 0$ if $A \neq B$). In words, N_A takes on the value 1 at node A and is 0 at all other nodes. Furthermore, N_A is nonzero only in the subintervals that contain x_A.

A typical member $w^h \in \mathcal{V}^h$ has the form $\Sigma_{A=1}^{n} c_A N_A$ and appears as in Fig. 1.8.2. Note that w^h is continuous but has discontinuous slope across each element boundary. For this reason, $w^h_{,x}$, the generalized derivative of w^h, will be piecewise constant, experiencing discontinuities across element boundaries. (Such a function is sometimes called a ***generalized step function.***) Restricted to each element domain, w^h is a linear polynomial in x. In respect to the homogeneous essential boundary condition, $w^h(1) = 0$. Clearly, w^h is identically zero if and only if each $c_A = 0$, $A = 1, 2, \ldots, n$.

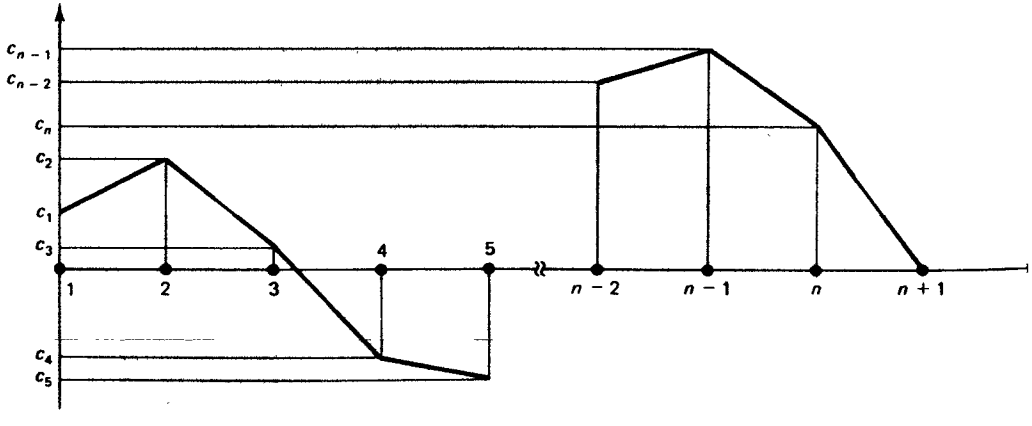


Figure 1.8.2 A typical member $w^h \in \mathcal{V}^h$.

Typical members of $\mathcal{S}^h$ are obtained by adding $g^h = gN_{n+1}$ to typical members of $\mathcal{V}^h$. This ensures that $u^h(1) = g$.

The piecewise linear finite element functions are the simplest and most widely used finite element functions for one-dimensional problems.

Exercise 1. Consider the weak formulation of the one-dimensional model problem:

$$\int_0^1 w_{,x}u_{,x}\,dx = \int_0^1 w\ell\,dx + w(0)h \tag{1.8.4}$$

where $w \in \mathcal{V}$ and $u \in \mathcal{S}$ are assumed to be smooth on element interiors (i.e., on $]x_A, x_{A+1}[$, $A = 1, 2, \ldots, n$), but may suffer slope discontinuities across element boundaries. (Functions of this class contain the piecewise linear finite element space described earlier.) From (1.8.4) and the assumed continuity of the functions, show that:

$$\boxed{\begin{aligned} 0 = \sum_{A=1}^{n} \int_{x_A}^{x_{A+1}} w(u_{,xx} + \ell)\,dx + w(0)[u_{,x}(0^+) + h] \\ + \sum_{A=2}^{n} w(x_A)[u_{,x}(x_A^+) - u_{,x}(x_A^-)] \end{aligned}} \tag{1.8.5}$$

Arguing as in Sec. 1.4, it may be concluded that the Euler-Lagrange conditions of (1.8.5) are

i. $u_{,xx}(x) + \ell(x) = 0$, where $x \in]x_A, x_{A+1}[$ and $A = 1, 2, \ldots, n$,
ii. $-u_{,x}(0^+) = h$; and
iii. $u_{,x}(x_A^-) = u_{,x}(x_A^+)$, where $A = 2, 3, \ldots, n$.

Observe that (i) is the differential equation *restricted to element interiors*, and (iii) is a continuity condition across element boundaries. This may be contrasted with the case in which the solution is assumed smooth. In this case the continuity condition is identically satisfied and the summation of integrals over element interiors may be replaced by an integral over the entire domain (see Sec. 1.4).

In the Galerkin finite element formulation, an *approximate* solution of (i)–(iii) is obtained.

1.9 PROPERTIES OF K

The shape functions N_A, $A = 1, 2, \ldots, n + 1$, are zero outside a neighborhood of node A. As a result, many of the entries of $\boldsymbol{K}$ are zero. This can be seen as follows. Let $B > A + 1$. Then (see Fig. 1.9.1)

$$K_{AB} = \int_0^1 \underbrace{N_{A,x}N_{B,x}}_{0}\,dx = 0 \tag{1.9.1}$$

The symmetry of $\boldsymbol{K}$ implies, in addition, that (1.9.1) holds for $A > B + 1$. One says that $\boldsymbol{K}$ is ***banded*** (i.e., its nonzero entries are located in a band about the main diagonal). Figure 1.9.2 depicts this property. Banded matrices have significant advantages in that the zero elements outside the band neither have to be stored nor operated

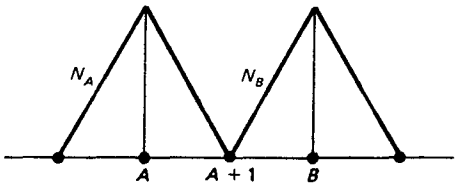


Figure 1.9.1 If $B > A + 1$, the n
zero portions of N_B and N_A do not o
lap.

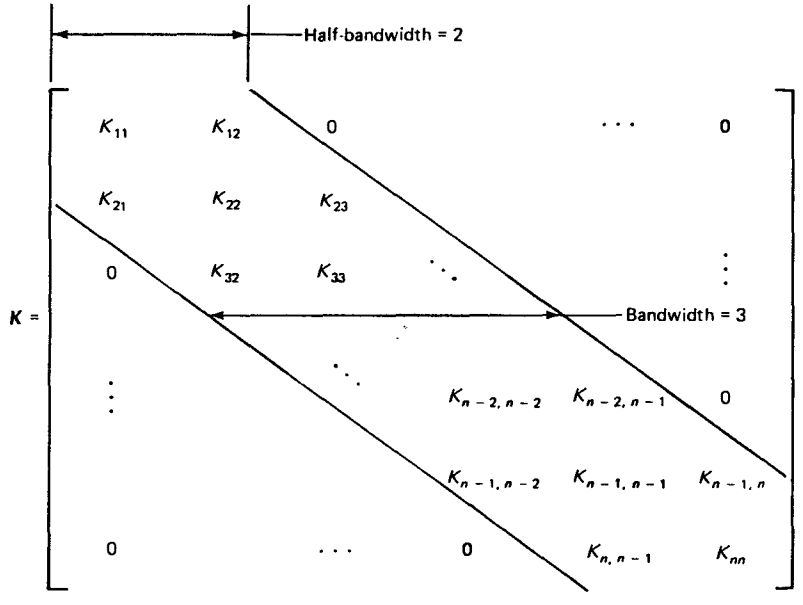


Figure 1.9.2 Band structure of K.

upon in the computer. The stiffness matrix arising in finite element analysis is, general, narrowly banded, lending itself to economical formation and solution.

Definition. An $n \times n$ matrix A is said to be ***positive definite*** if

i. $c^T Ac \geq 0$ for all n-vectors c; and
ii. $c^T Ac = 0$ implies $c = 0$.

Remarks

1. A symmetric positive-definite matrix posesses a unique inverse.
2. The eigenvalues of a positive-definite matrix are real and positive.

Theorem. The $n \times n$ matrix K defined by (1.6.11) is positive definite.

Proof

i. Let c_A, $A = 1, 2, \ldots, n$, be the components of c (i.e., $c = \{c_A\}$), arbitrary vector. Use these c_A's to construct a member of $\mathcal{U}^h$, $w^h = \sum_{A=1}^{n} c_A N_A$, wh

the N_A's are the basis functions for $\mathcal{V}^h$. Then

$$\begin{aligned} \boldsymbol{c}^T\boldsymbol{K}\boldsymbol{c} &= \sum_{A,B=1}^{n} c_A K_{AB} c_B \\ &= \sum_{A,B=1}^{n} c_A a(N_A, N_B) c_B && \text{(definition of } K_{AB}) \\ &= a\left(\sum_{A=1}^{n} c_A N_A, \sum_{B=1}^{n} c_B N_B\right) && \text{(bilinearity of } a(\cdot,\cdot)) \\ &= a(w^h, w^h) && \text{(definition of } w^h) \\ &= \int_0^1 \underbrace{(w^h_{,x})^2}_{\geq 0}\, dx && \text{(by (1.4.8)} \\ &\geq 0 \end{aligned}$$

ii. Assume $\boldsymbol{c}^T\boldsymbol{K}\boldsymbol{c} = 0$. By the proof of part (i),

$$\int_0^1 (w^h_{,x})^2\, dx = 0$$

and consequently w^h must be constant. Since $w^h \in \mathcal{V}^h$, $w^h(1) = 0$. Combining these facts, we conclude that $w^h(x) = 0$ for all $x \in [0, 1]$, which is possible only if each $c_A = 0$, $A = 1, 2, \ldots, n$. Thus $\boldsymbol{c} = \boldsymbol{0}$. ■

Note that part (ii) depended upon the definition of $\boldsymbol{K}$ and the zero essential boundary condition built into the definition of $\mathcal{V}^h$.

Summary. $\boldsymbol{K}$, defined by (1.6.11), is

i. Symmetric
ii. Banded
iii. Positive-definite

The practical consequence of the above properties is that a very efficient computer solution of $\boldsymbol{K}\boldsymbol{d} = \boldsymbol{F}$ may be performed.

1.10 MATHEMATICAL ANALYSIS

In this section we will show that the observations made with reference to the example problems of Sec. 1.7 are, in fact, general results. To establish these facts rigorously requires only elementary mathematical techniques.

Our first objective is to establish that the Galerkin finite element solution u^h is exact at the nodes. To do this we must introduce the notion of a Green's function.

Let $\delta_y(x) = \delta(x - y)$ denote the ***Dirac delta function.*** The Dirac function is not a function in the classical sense but rather an operator defined by its action on

(continuous) functions. Let w be continuous on [0, 1]; then we write

$$(w, \delta_y) = \int_0^1 w(x)\delta(x - y)\, dx$$
$$= w(y) \tag{1.10.1}$$

By (1.10.1), we see why attention is restricted to continuous functions—δ_y sifts o the value of w at y. If w were discontinuous at y, its value would be ambiguous. mechanics, we think of δ_y visually as representing a concentrated force of un amplitude located at point y.

The Green's function problem corresponding to (S) may be stated as follow Find a function g (i.e., the ***Green's function***) such that

$$g_{,xx} + \delta_y = 0 \qquad \text{on } \Omega \tag{1.10.}$$
$$g(1) = 0 \tag{1.10.}$$
$$g_{,x}(0) = 0 \tag{1.10.}$$

Note that (1.10.2)–(1.10.4) are simply (S) in which f is replaced by δ_y and q and are taken to be zero.

This problem may be solved by way of formal calculations with ***distribution*** or ***generalized functions***, such as δ_y. (The theory of distributions is dealt with Stakgold [5]. A good elementary account of formal calculations with distributions presented in Popov [9]. This latter reference is recommended to readers having h no previous experience with this topic.) To this end we note that the (formal) integr of δ_y is the ***Heaviside, or unit step, function:***

$$H_y(x) = H(x - y) = \begin{cases} 0, & x < y \\ 1, & x > y \end{cases} \tag{1.10.}$$

The integral of H_y is the ***Macaulay bracket:***

$$\langle x - y \rangle = \begin{cases} 0, & x \le y \\ x - y, & x > y \end{cases} \tag{1.10.}$$

The preceding functions are depicted in Fig. 1.10.1.

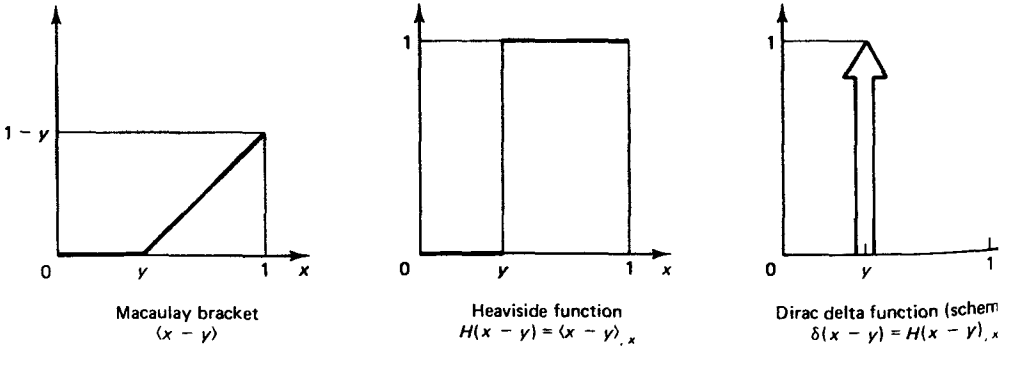


Figure 1.10.1 Elementary generalized functions (distributions).

To solve the Green's function problem, (1.10.2) is integrated, making use of (1.10.5), to obtain:

$$g_{,x} + H_y = c_1 \tag{1.10.7}$$

where c_1 is a constant of integration. A second integration and use of (1.10.6) yields

$$g(x) + \langle x - y \rangle = c_1 x + c_2 \tag{1.10.8}$$

where c_2 is another constant of integration. Evaluation of c_1 and c_2 is performed by requiring (1.10.7) and (1.10.8) to satisfy the boundary conditions. This results in (see Fig. 1.10.2)

$$g(x) = (1 - y) - \langle x - y \rangle \tag{1.10.9}$$

Observe that g is piecewise linear. Thus if $y = x_A$ (i.e., if y is a node), $g \in \mathcal{V}^h$.

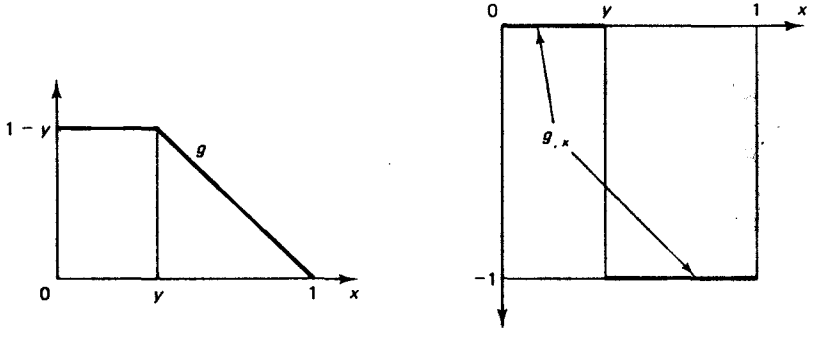


Figure 1.10.2 Green's function.

In the ensuing analysis we will need the variational equation corresponding to the Green's function problem. This can be deduced from (W) by replacing u by g, f by δ_y, and q and h by 0, viz.,

$$a(w, g) = (w, \delta_y) = w(y) \tag{1.10.10}$$

Equation (1.10.10) holds for all continuous $w \in \mathcal{V}$. The square-integrability of derivatives of functions $w \in \mathcal{V}$ actually implies the continuity of all $w \in \mathcal{V}$ by a well-known theorem in analysis due to Sobolev. (This result is true only in one dimension. The square-integrability of second derivatives is also required to ensure the continuity of functions defined on two- and three-dimensional domains.)

Theorem. $u^h(x_A) = u(x_A)$, $A = 1, 2, \ldots, n + 1$ (i.e., u^h is exact at the nodes). To prove the theorem, we need to establish two preliminary results.

Lemma 1. $a(u - u^h, w^h) = 0$ for all $w^h \in \mathcal{V}^h$.

Proof. We have observed previously that $\mathcal{V}^h \subset \mathcal{V}$, so we may replace w by w^h in the variational equation:

$$a(w^h, u) = (w^h, f) + w^h(0)h \tag{1.10.11}$$

Equation (1.10.11) holds for all $w^h \in \mathcal{V}^h$. Recall that the Galerkin equation is identical to (1.10.11) except that u^h appears instead of u. Subtracting the Galerkin equation

from (1.10.11) and using the bilinearity and symmetry of $a(\cdot, \cdot)$ yields the require result.

Lemma 2. $u(y) - u^h(y) = a(u - u^h, g)$, where g is the Green's functior

Proof

$$\begin{aligned} u(y) - u^h(y) &= (u - u^h, \delta_y) && \text{(definition of } \delta_y\text{)} \\ &= a(u - u^h, g) && \text{(by (1.10.10))} \end{aligned}$$

Note that line 2 is true since $u - u^h$ is in $\mathcal{U}$.

Proof of Theorem. As we have remarked previously, if $y = x_A$, a node $g \in \mathcal{U}^h$. Let us take this to be the case. Then

$$\begin{aligned} u(x_A) - u^h(x_A) &= a(u - u^h, g) && \text{(Lemma 2)} \\ &= 0 && \text{(Lemma 1)} \end{aligned}$$

The theorem is valid for $A = 1, 2, \ldots, n + 1$. Strang and Fix [6] attribute th argument to Douglas and Dupont. Results of this kind, embodying exceptional accu racy characteristics, are often referred to as ***superconvergence*** phenomena. Howeve the reader should appreciate that, in more complicated situations, we will not be able in practice, to guarantee nodal exactness. Nevertheless, as we shall see later or weighted residual procedures provide a framework within which optimal accurac properties of some sort may often be guaranteed.

Accuracy of the Derivatives

In considering the convergence properties of the derivatives, certain elementary nc tions of numerical analysis arise. The reader should make sure that he or she has complete understanding of these ideas as they subsequently arise in other contexts. W begin by introducing some preliminary mathematical results.

Taylor's Formula with Remainder

Let $f: [0, 1] \to \mathbb{R}$ possess k continuous derivatives and let y and z be two points i $[0, 1]$. Then there is a point c between y and z such that

$$\begin{aligned} f(z) = f(y) &+ (z - y)f_{,x}(y) + \frac{1}{2}(z - y)^2 f_{,xx}(y) \\ &+ \frac{1}{3!}(z - y)^3 f_{,xxx}(y) + \cdots + \\ &+ \frac{1}{k!}(z - y)^k f_{,\underbrace{x \ldots x}_{k \text{ times}}}(c) \end{aligned} \tag{1.10.12}$$

The proof of this formula may be found in [7]. Equation (1.10.12) is sometimes called a *finite Taylor expansion*.

Mean-Value Theorem

The mean-value theorem is a special case of (1.10.12) which is valid as long as $k \geq 1$ (i.e., f is continuously differentiable):

$$\boxed{f(z) = f(y) + (z - y)f_{,x}(c)} \tag{1.10.13}$$

Consider a typical subinterval $[x_A, x_{A+1}]$. We have already shown that u^h is exact at the endpoints (see Fig. 1.10.3). The derivative of u^h in $]x_A, x_{A+1}[$ is constant:

$$u^h_{,x}(x) = \frac{u^h(x_{A+1}) - u^h(x_A)}{h_A}, \qquad x \in]x_A, x_{A+1}[\tag{1.10.14}$$

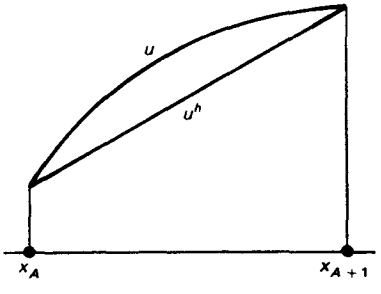


Figure 1.10.3

Theorem. Assume u is continuously differentiable. Then there exists at least one point in $]x_A, x_{A+1}[$ at which (1.10.14) is exact.

Proof. By the mean value theorem, there exists a point $c \in]x_A, x_{A+1}[$ such that

$$\frac{u(x_{A+1}) - u(x_A)}{h_A} = u_{,x}(c) \tag{1.10.15}$$

(We have used (1.10.13) with u, x_A, and x_{A+1}, in place of f, y, and z, respectively.) Since $u(x_A) = u^h(x_A)$ and $u(x_{A+1}) = u^h(x_{A+1})$, we may rewrite (1.10.15) as

$$\frac{u^h(x_{A+1}) - u^h(x_A)}{h_A} = u_{,x}(c) \tag{1.10.16}$$

Comparison of (1.10.16) with (1.10.14) yields the desired result. ■

Remarks

1. This result means that the constant value of $u^h_{,x}$ must coincide with $u_{,x}$ somewhere on $]x_A, x_{A+1}[$; see Fig. 1.10.4.

2. Without knowledge of u we have no way of determining the locations at which the derivatives are exact. The following results are more useful in that they tell us that the midpoints are, in a sense, optimally accurate, independent of u.

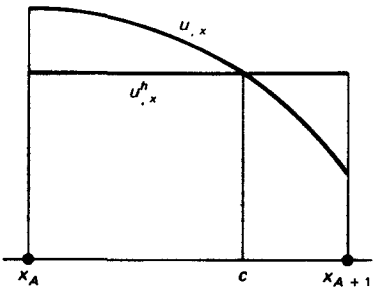


Figure 1.10.4

Let

$$e_{,x}(\alpha) \overset{\text{def.}}{=} u^h_{,x}(\alpha) - u_{,x}(\alpha) = \frac{u^h(x_{A+1}) - u^h(x_A)}{h_A} - u_{,x}(\alpha)$$

the ***error in the derivative*** at $\alpha \in [x_A, x_{A+1}]$. To establish the superiority of the midpoints in evaluating the derivatives, we need a preliminary result.

Lemma. Assume u is three times continuously differentiable. Then

$$e_{,x}(\alpha) = \left(\frac{x_{A+1} + x_A}{2} - \alpha\right) u_{,xx}(\alpha) + \frac{1}{3!\, h_A}[(x_{A+1} - \alpha)^3 u_{,xxx}(c_1) - (x_A - \alpha)^3 u_{,xxx}(c_2)] \tag{1.10.17}$$

where c_1 and c_2 are in $[x_A, x_{A+1}]$.

Proof. Expand $u(x_{A+1})$ and $u(x_A)$ in finite Taylor expansions about $\alpha \in [x_A, x_{A+1}]$, viz.,

$$u(x_{A+1}) = u(\alpha) + (x_{A+1} - \alpha)u_{,x}(\alpha) + \frac{1}{2}(x_{A+1} - \alpha)^2 u_{,xx}(\alpha) + \frac{1}{3!}(x_{A+1} - \alpha)^3 u_{,xxx}(c_1), \qquad c_1 \in [\alpha, x_{A+1}]$$

$$u(x_A) = u(\alpha) + (x_A - \alpha)u_{,x}(\alpha) + \frac{1}{2}(x_A - \alpha)^2 u_{,xx}(\alpha) + \frac{1}{3!}(x_A - \alpha)^3 u_{,xxx}(c_2), \qquad c_2 \in [x_A, \alpha]$$

Subtracting and dividing through by h_A yields

$$\frac{u(x_{A+1}) - u(x_A)}{h_A} = u_{,x}(\alpha) + \left(\frac{x_{A+1} + x_A}{2} - \alpha\right) u_{,xx}(\alpha) + \frac{1}{3!\, h_A}[(x_{A+1} - \alpha)^3 u_{,xxx}(c_1) - (x_A - \alpha)^3 u_{,xxx}(c_2)]$$

Replacing $u(x_{A+1})$ by $u^h(x_{A+1})$ and $u(x_A)$ by $u^h(x_A)$ in the left-hand side and rearranging terms completes the proof. ■

Discussion

To determine what (1.10.17) tells us about the accuracy of the derivatives, we wish to think of the situation in which the mesh is being systematically refined (i.e., we let h_A approach zero). In this case h_A^2 will be much smaller than h_A. Thus, for a given u, if the right-hand side of (1.10.17) *is* $O(h_A^2)$,[3] the error in the derivatives will be much smaller than if the right-hand side is only $O(h_A)$. The exponent of h_A is called the ***order of convergence*** or ***order of accuracy.*** In the former case we would have second-order convergence of the derivative, whereas in the latter case we would have only first-order convergence.

As an example, assume $\alpha \to x_A$. Then

$$e_{,x}(x_A) = \frac{h_A}{2}u_{,xx}(x_A) + \frac{h_A^2}{3!}u_{,xxx}(c_1) = O(h_A)$$

As $h_A \to 0$, the first term dominates. (We have seen from the example calculations in Sec. 1.8 that the endpoints of the subintervals are not very accurate for the derivatives.)

Clearly any point $\alpha \in [x_A, x_{A+1}]$ achieves first-order accuracy. We are thus naturally led to asking the question, are there any values of α at which higher-order accuracy is achieved?

Corollary. Let $x_{A+1/2} \equiv (x_A + x_{A+1})/2$ (i.e., the midpoint). Then

$$e_{,x}(x_{A+1/2}) = \frac{h_A^2}{24}u_{,xxx}(c), \qquad c \in [x_A, x_{A+1}]$$
$$= O(h_A^2)$$

Proof. By (1.10.17)

$$e_{,x}(x_{A+1/2}) = \frac{h_A^2}{48}[u_{,xxx}(c_1) + u_{,xxx}(c_2)]$$

By the continuity of $u_{,xxx}$, there is at least one point c between c_1 and c_2 such that

$$u_{,xxx}(c) = \frac{1}{2}[u_{,xxx}(c_1) + u_{,xxx}(c_2)]$$

Combining these facts completes the proof. ■

Remarks

1. From the corollary we see that the derivatives are second-order accurate at the midpoints.

[3]A function $f(x)$ is said to be $O(x^k)$ (i.e., order x^k) if $f(x)/x^k \to$ a constant as $x \to 0$. For example, $f(x) = x^k$ is $O(x^k)$, as is $f(x) = \sum_{j=k}^{k+l} x^j$, $l \geq 0$. But neither is $O(x^{k+1})$. (Verify.)

2. If the exact solution is quadratic (i.e., consists of a linear combination of the monomials $1, x, x^2$), then $u_{,xxx} = 0$ and—by (1.10.17)—the derivative is exact at the midpoints. This is the case when $f(x) = p = \text{constant}$.

3. In linear elastic rod theory, the derivatives are proportional to the stresses. The midpoints of linear "elements" are sometimes called the ***Barlow stress points,*** after Barlow [8], who first noted that points of optimal accuracy existed within elements.

Exercise 1. Assume the mesh length is constant (i.e., $h_A = h, A = 1, 2, \ldots, n$). Consider the standard finite difference "stencil" for $u_{,xx} + f = 0$ at a typical internal node, namely,

$$\frac{u_{A+1} - 2u_A + u_{A-1}}{h^2} + f_A = 0 \tag{1.10.18}$$

Assuming f varies in piecewise linear fashion and so can be expanded as

$$f = \sum_{A=1}^{n+1} f_A N_A \tag{1.10.19}$$

where the f_A's are the nodal values of f, set up the finite element equation associated with node A and contrast it with (1.10.18). Deduce when (1.10.18) will also be capable of exhibiting superconvergence phenomena. (That is, what is the restriction on f?) Set up the finite element equation associated with node 1, accounting for nonzero h. Discuss this equation from the point of view of finite differences. (For further comparisons along these lines, the interested reader is urged to consult [6], Chapter 1.)

Summary. The Galerkin finite element solution u^h, of the problem (S), possesses the following properties:

i. It is exact at the nodes.
ii. There exists at least one point in each element at which the derivative is exact.
iii. The derivative is second-order accurate at the midpoints of the elements.

1.11 INTERLUDE: GAUSS ELIMINATION; HAND-CALCULATION VERSION

It is important for anyone who wishes to do finite element analysis to become familiar with the efficient and sophisticated computer schemes that arise in the finite element method. It is felt that the best way to do this is to begin with the simplest scheme, perform some hand calculations, and gradually increase the sophistication as time goes on.

To do some of the problems we will need a fairly efficient method of solving matrix equations by hand. The following scheme is applicable to systems of equations

$\boldsymbol{Kd} = \boldsymbol{F}$ in which *no pivoting* (i.e., reordering) is necessary. For example, symmetric, positive-definite coefficient matrices never require pivoting. The procedure is as follows:

Gauss Elimination

- Solve the first equation for d_1 and elminate d_1 from the remaining $n - 1$ equations.
- Solve the second equation for d_2 and eliminate d_2 from the remaining $n - 2$ equations.

  .
  .
  .

- Solve the $n - 1$st equation for d_{n-1} and eliminate d_{n-1} from the nth equation.
- Solve the n-th equation for d_n.

The preceding steps are called ***forward reduction***. The original matrix is reduced to upper triangular form. For example, suppose we began with a system of four equations as follows:

$$\begin{bmatrix} K_{11} & K_{12} & K_{13} & K_{14} \\ K_{21} & K_{22} & K_{23} & K_{24} \\ K_{31} & K_{32} & K_{33} & K_{34} \\ K_{41} & K_{42} & K_{43} & K_{44} \end{bmatrix} \begin{Bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{Bmatrix} = \begin{Bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \end{Bmatrix}$$

The ***augmented matrix*** corresponding to this system is

$$\left[\begin{array}{cccc|c} K_{11} & K_{12} & K_{13} & K_{14} & F_1 \\ K_{21} & K_{22} & K_{23} & K_{24} & F_2 \\ K_{31} & K_{32} & K_{33} & K_{34} & F_3 \\ K_{41} & K_{42} & K_{43} & K_{44} & F_4 \end{array}\right] \quad \begin{matrix} \\ \\ \\ \\ \end{matrix} \quad \begin{matrix} \\ \\ \\ \\ \end{matrix}$$

$$\underbrace{\qquad\qquad\qquad}_{\boldsymbol{K}} \qquad \underbrace{\quad}_{\boldsymbol{F}}$$

After the forward reduction, the augmented matrix becomes

$$\left[\begin{array}{cccc|c} 1 & K'_{12} & K'_{13} & K'_{14} & F'_1 \\ 0 & 1 & K'_{23} & K'_{24} & F'_2 \\ 0 & 0 & 1 & K'_{34} & F'_3 \\ 0 & 0 & 0 & 1 & d_4 \end{array}\right] \quad \begin{matrix} \\ \\ \\ \\ \end{matrix} \quad \begin{matrix} \\ \\ \\ \\ \end{matrix} \qquad (1.11.1)$$

$$\underbrace{\qquad\qquad\qquad}_{\boldsymbol{U}} \qquad \underbrace{\quad}_{\boldsymbol{F}'}$$

corresponding to the upper triangular system $\boldsymbol{Ud} = \boldsymbol{F}'$.[4] It is a simply verified fact that if $\boldsymbol{K}$ is banded, then $\boldsymbol{U}$ will be also.

Employing the reduced augmented matrix, proceed as follows:

- Eliminate d_n from equations $n - 1, n - 2, \ldots, 1$.

[4]Primes will be used to denote intermediate quantities throughout this section.

- Eliminate d_{n-1} from equations $n-2, n-3, \ldots, 1$.

  .
  .
  .

- Eliminate d_2 from the first equation.

This procedure is called ***back substitution***. For example, in the example just given, after back substitution we obtain

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & d_1 \\ 0 & 1 & 0 & 0 & d_2 \\ 0 & 0 & 1 & 0 & d_3 \\ 0 & 0 & 0 & 1 & d_4 \end{array}\right] \qquad (1.11.2)$$

$\underbrace{\qquad\qquad\qquad\qquad}_{\boldsymbol{I}}$ $\underbrace{\quad}_{\boldsymbol{d}}$

corresponding to the identity $\boldsymbol{Id} = \boldsymbol{d}$. The solution winds up in the last column.

Hand-Calculation Algorithm

In a hand calculation, Gauss elimination can be performed on the augmented matrix as follows.

Forward reduction

- Divide row 1 by K_{11}.
- Subtract $K_{21} \times$ row 1 from row 2.
- Subtract $K_{31} \times$ row 1 from row 3.

  .
  .
  .

- Subtract $K_{n1} \times$ row 1 from row n.

Consider the example of four equations. The preceding steps reduce the first column to the form

$$\left[\begin{array}{cccc|c} 1 & K'_{12} & K'_{13} & K'_{14} & F'_1 \\ 0 & K''_{22} & K''_{23} & K''_{24} & F''_2 \\ 0 & K''_{32} & K''_{33} & K''_{34} & F''_3 \\ 0 & K''_{42} & K''_{43} & K''_{44} & F''_4 \end{array}\right]$$

Note that if $K_{A1} = 0$, then the computation for the Ath row can be ignored. Now reduce the second column

- Divide row 2 by K''_{22}.
- Subtract $K''_{32} \times$ row 2 from row 3.
- Subtract $K''_{42} \times$ row 2 from row 4.

.
.
.

- Subtract $K''_{n2} \times$ row 2 from row n.

The result for the example will look like

$$\left[\begin{array}{c|ccc|c} 1 & K'_{12} & K'_{13} & K'_{14} & F'_1 \\ \hline 0 & 1 & K'''_{23} & K'''_{24} & F''''_2 \\ 0 & 0 & K'''_{33} & K'''_{34} & F''''_3 \\ 0 & 0 & K'''_{43} & K'''_{44} & F''''_4 \end{array}\right]$$

Note that only the submatrix enclosed in dashed lines is affected in this procedure.

Repeat until columns 3 to n are reduced and the upper triangular form (1.11.1) is obtained.

Back substitution

- Subtract $K'_{n-1,n} \times$ row n from row $n-1$.
- Subtract $K'_{n-2,n} \times$ row n from row $n-2$.

.
.
.

- Subtract $K'_{1,n} \times$ row n from row 1.

After these steps the augmented matrix, for this example, will look like

$$\left[\begin{array}{ccc|c|c} 1 & K'_{12} & K'_{13} & 0 & F'''''_1 \\ 0 & 1 & K'_{23} & 0 & F'''''_2 \\ 0 & 0 & 1 & 0 & d_3 \\ 0 & 0 & 0 & 1 & d_4 \end{array}\right]$$

Note that the submatrix enclosed in dashed lines is unaffected by these steps, and, aside from zeroing the appropriate elements of the last column of the coefficient matrix, only the vector F' is altered.

Now clear the second-to-last column in the coefficient matrix:

- Subtract $K'_{n-2,n-1} \times$ row $n-1$ from row $n-2$.
- Subtract $K'_{n-3,n-1} \times$ row $n-1$ from row $n-3$.

.
.
.

- Subtract $K'_{1,n-1} \times$ row $n-1$ from row 1.

Again we mention that the only nontrivial calculations are being performed on the last column (i.e., on F).

Repeat as above until columns $n-2, n-3, \ldots, 2$ are cleared. The result is (1.11.2).

Remarks

1. In passing we note that the above procedure is *not* the same as the way one would implement Gauss elimination on a computer, which we shall treat later. In a computer program for Gauss elimination of symmetric matrices we would want all intermediate results to retain symmetry and thus save storage. This can be done by a small change in the procedure. However, it is felt that the given scheme is the clearest for hand calculations.

2. The numerical example with which we close this section illustrates the preceding elimination scheme. Note that the band is maintained (i.e., the zeros in the upper right-hand corner of the coefficient matrix remain zero throughout the calculations). The reader is urged to perform the calculations.

Example of Gauss elimination

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} \begin{Bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{Bmatrix} = \begin{Bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{Bmatrix}$$

Augmented matrix

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 1 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & 0 \end{array}\right]$$

Forward reduction

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & 0 \end{array}\right]$$

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & -1 & 2 & 0 \end{array}\right]$$

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

Back substitution

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

$$\begin{Bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{Bmatrix} = \begin{Bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{Bmatrix}$$

Exercise 1. Consider the boundary-value problem discussed in the previous sections:

$$u_{,xx}(x) + f(x) = 0 \qquad x \in]0, 1[$$
$$u(1) = q$$
$$-u_{,x}(0) = h$$

Assume $f = qx$, where q is constant, and $q = h = 0$.

a. Employing the linear finite element space with equally spaced nodes, set up and solve the Galerkin finite element equations for $n = 4 (h = \text{mesh parameter} = \frac{1}{4})$. Recall that in Sec. 1.7 this was carried out for $n = 1$ and $n = 2$ ($h = 1$ and $h = \frac{1}{2}$, respectively). Do *not* invert the stiffness matrix $\boldsymbol{K}$; use Gauss elimination to solve $\boldsymbol{Kd} = \boldsymbol{F}$ or a more sophisticated direct factorization scheme if you know one. You can check your answers since they must be exact at the nodes.

b. Let $re_{,x} = |u^h_{,x} - u_{,x}|/(q/2)$, the ***relative error*** in $u_{,x}$. Compute $re_{,x}$ at the midpoints of the four elements. They should all be equal. (This was also the case for $n = 2$.)

c. Employing the data for $h = 1, \frac{1}{2}$, and $\frac{1}{4}$, plot $\ln re_{,x}$ versus $\ln h$.

d. Using the error analysis for $re_{,x}$ at the midpoints presented in Sec. 1.10, answer the following questions:
   i. What is the significance of the slope of the graph in part (c)?
   ii. What is the significance of the y-intercept?

1.12 THE ELEMENT POINT OF VIEW

So far we have viewed the finite element method simply as a particular Galerkin approximation procedure applied to the weak statement of the problem in question. What makes what we have done a finite element procedure is the character of the selected basis functions; particularly their piecewise smoothness and "local support" (i.e., $N_A \equiv 0$ outside a neighborhood of node A). This is the mathematical point of view; it is a ***global*** point of view in that the basis functions are considered to be defined everywhere on the domain of the boundary-value problem. The global viewpoint is useful in establishing the mathematical properties of the finite element method. This can be seen in Sec. 1.10 and will be made more apparent later on.

Now we wish to discuss another point of view called the ***local***, or ***element***, point of view. This viewpoint is the traditional one in engineering and is useful in the computer implementation of the finite element method and in the development of finite elements.

We begin our treatment of the local point of view with a question: What is a finite element?

We shall attempt to give the answer in terms of the piecewise linear finite element space that we defined previously. An individual element consists of the following quantities.

Linear finite element (global description)

$(g1)$	Domain:	$[x_A, x_{A+1}]$
$(g2)$	Nodes:	$\{x_A, x_{A+1}\}$
$(g3)$	Degrees of freedom:	$\{d_A, d_{A+1}\}$
$(g4)$	Shape functions:[5]	$\{N_A, N_{A+1}\}$
$(g5)$	Interpolation function:	

$$u^h(x) = N_A(x)d_A + N_{A+1}(x)d_{A+1}, \qquad x \in [x_A, x_{A+1}]$$

(Recall $d_A = u^h(x_A)$.) In words, a linear finite element is just the totality of paraphernalia associated with the globally defined function u^h *restricted* to the *element* domain. The above quantities are in terms of *global* parameters—namely, the global coordinates, global shape functions, global node ordering, and so on. It is fruitful to introduce a *local* set of quantities, corresponding to the global ones, so that calculations for a typical element may be standardized. These are given as follows:

Linear finite element (local description)

$(l1)$	Domain:	$[\xi_1, \xi_2]$

[5] In weighted residual methods in which $\mathcal{S}^h$ and $\mathcal{V}^h$ are built up from different classes of functions (i.e., Petrov-Galerkin methods), we would also have to specify a set of ***weighting functions***, say $\{\tilde{N}_A, \tilde{N}_{A+1}\}$; the entire set of $\tilde{N}_A$'s would then constitute a basis for $\mathcal{V}^h$. In Galerkin's method $\tilde{N}_A = N_A$.

(*l*2) Nodes: $\{\xi_1, \xi_2\}$

(*l*3) Degrees of freedom: $\{d_1, d_2\}$

(*l*4) Shape functions: $\{N_1, N_2\}$

(*l*5) Interpolation function:

$$u^h(\xi) = N_1(\xi)d_1 + N_2(\xi)d_2$$

Note that in the local description, the nodal numbering begins with 1.

We shall relate the domains of the global and local descriptions by an "affine" transformation $\xi : [x_A, x_{A+1}] \to [\xi_1, \xi_2]$, such that $\xi(x_A) = \xi_1$ and $\xi(x_{A+1}) = \xi_2$. It is standard practice to take $\xi_1 = -1$ and $\xi_2 = +1$. Thus ξ may be represented by the expression

$$\xi(x) = c_1 + c_2 x \tag{1.12.1}$$

where c_1 and c_2 are constants which are determined by

$$\left.\begin{aligned} -1 &= c_1 + x_A c_2 \\ 1 &= c_1 + x_{A+1} c_2 \end{aligned}\right\} \tag{1.12.2}$$

Solving this system yields

$$\xi(x) = \frac{2x - x_A - x_{A+1}}{h_A} \tag{1.12.3}$$

(Recall $h_A = x_{A+1} - x_A$.) The inverse of ξ is obtained by solving for x:

$$x(\xi) = \frac{h_A \xi + x_A + x_{A+1}}{2} \tag{1.12.4}$$

In (1.12.1), ξ is a mapping and x is a point, whereas in (1.12.4), x is a mapping and ξ is a point.

In the sequel, we adopt the notational convention that subscripts $a, b, c, \ldots$ pertain to the local numbering system. The subscripts $A, B, C, \ldots$ will always pertain to the global numbering system. To control the proliferation of notations, we will frequently use the same notation for the local and global systems (e.g., d_a and d_A or N_a and N_A). This generally should not cause confusion as the context will make clear which point of view is being adopted. If there is danger of confusion, a superscript e will be introduced to denote a quantity in the local description associated with element number e (e.g., $d_a^e = d_A$, $N_a^e(\xi) = N_A(x^e(\xi))$, where $x^e : [\xi_1, \xi_2] \to [x_1^e, x_2^e] = [x_A, x_{A+1}]$, etc.).

In terms of ξ, the shape functions in the local description take on a standard form

$$N_a(\xi) = \tfrac{1}{2}(1 + \xi_a \xi), \qquad a = 1, 2 \tag{1.12.5}$$

Note also that (1.12.4) may be written in terms of (1.12.5):

$$\boxed{x^e(\xi) = \sum_{a=1}^{2} N_a(\xi) x_a^e.} \tag{1.12.6}$$

This has the same form as the interpolation function (cf. *15*).

For future reference, we note the following results:

$$N_{a,\xi} = \frac{\xi_a}{2} = \frac{(-1)^a}{2} \tag{1.12.7}$$

$$x^e_{,\xi} = \frac{h^e}{2} \tag{1.12.8}$$

where $h^e = x_2^e - x_1^e$ and

$$\xi^e_{,x} = (x^e_{,\xi})^{-1} = \frac{2}{h^e} \tag{1.12.9}$$

The local and global descriptions of the eth element are depicted in Fig. 1.12.1.

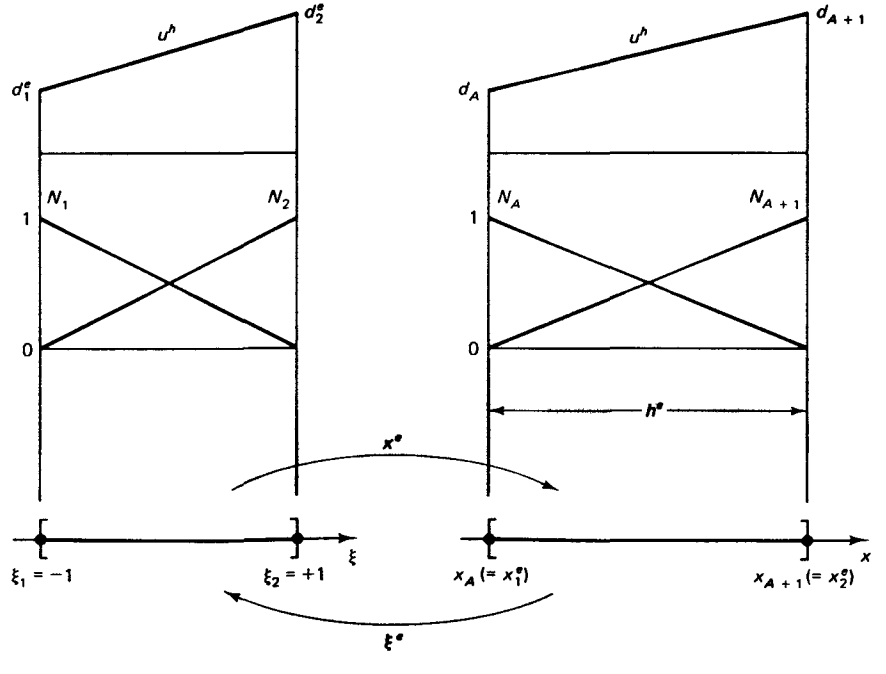


Figure 1.12.1 Local and global descriptions of the eth element.

1.13 ELEMENT STIFFNESS MATRIX AND FORCE VECTOR

To develop the element point of view further, let us assume that our model consists of n_{el} elements, numbered as shown in Figure 1.13.1. Clearly $n_{el} = n$ for this case. Let us take e to be the variable index for the elements; thus $1 \leq e \leq n_{el}$.

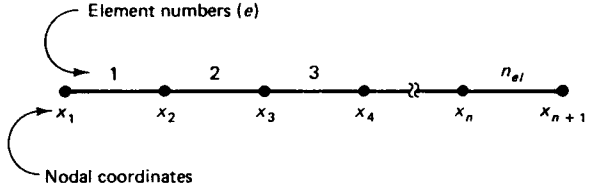


Figure 1.13.1

Now recall the definitions of the (global) stiffness matrix and force vector

$$K = \underbrace{[K_{AB}]}_{n \times n}, \qquad F = \underbrace{\{F_A\}}_{n \times 1} \tag{1.13.1}$$

where

$$K_{AB} = a(N_A, N_B) = \int_0^1 N_{A,x} N_{B,x}\, dx \tag{1.13.2}$$

$$\begin{aligned} F_A &= (N_A, \ell) + \delta_{A1} h - a(N_A, N_{n+1}) q \\ &= \int_0^1 N_A \ell\, dx + \delta_{A1} h - \int_0^1 N_{A,x} N_{n+1,x}\, dx\, q \end{aligned} \tag{1.13.3}$$

(In (1.13.3) we have assumed $N_A(x_1) = \delta_{A1}$, as for the piecewise linear finite element space.) The integrals over $[0, 1]$ may be written as sums of integrals over the element domains. Thus

$$K = \sum_{e=1}^{n_{el}} K^e, \qquad K^e = [K^e_{AB}] \tag{1.13.4}$$

$$F = \sum_{e=1}^{n_{el}} F^e, \qquad F^e = \{F^e_A\} \tag{1.13.5}$$

where

$$K^e_{AB} = a(N_A, N_B)^e = \int_{\Omega^e} N_{A,x} N_{B,x}\, dx \tag{1.13.6}$$

$$\begin{aligned} F^e_A &= (N_A, \ell)^e + \delta_{e1}\, \delta_{A1} h - a(N_A, N_{n+1})^e q \\ &= \int_{\Omega^e} N_A \ell\, dx + \delta_{e1}\, \delta_{A1} h - \int_{\Omega^e} N_{A,x} N_{n+1,x}\, dx\, q \end{aligned} \tag{1.13.7}$$

and $\Omega^e = [x_1^e, x_2^e]$, ***the domain of the eth element.***

The important observation to make is that K and F can be *constructed by*

summing the contributions of elemental matrices and vectors, respectively. In the literature, this procedure is sometimes called the ***direct stiffness method*** [10].

By the definitions of the N_A's, we have that

$$K^e_{AB} = 0, \qquad \text{if } A \neq e \text{ or } e+1 \text{ or } B \neq e \text{ or } e+1 \tag{1.13.8}$$

and

$$F^e_A = 0, \qquad \text{if } A \neq e \text{ or } e+1 \tag{1.13.9}$$

The situation for a typical element, e, is shown in Fig. 1.13.2. In practice we would not, of course, add in the zeros but merely add in the nonzero terms to the appropriate locations. For this purpose it is useful to define the ***eth element stiffness matrix*** $\boldsymbol{k}^e$ and ***element force vector*** $\boldsymbol{f}^e$ as follows:

$$\boldsymbol{k}^e = \underbrace{[k^e_{ab}]}_{2\times 2}, \qquad \boldsymbol{f}^e = \underbrace{\{f^e_a\}}_{2\times 1} \tag{1.13.10}$$

$$k^e_{ab} = a(N_a, N_b)^e = \int_{\Omega^e} N_{a,x} N_{b,x}\, dx \tag{1.13.11}$$

$$f^e_a = \int_{\Omega^e} N_a \ell\, dx + \begin{cases} \delta_{a1} h & e = 1 \\ 0 & e = 2, 3, \ldots, n_{el}-1 \\ -k^e_{a2} g & e = n_{el} \end{cases} \tag{1.13.12}$$

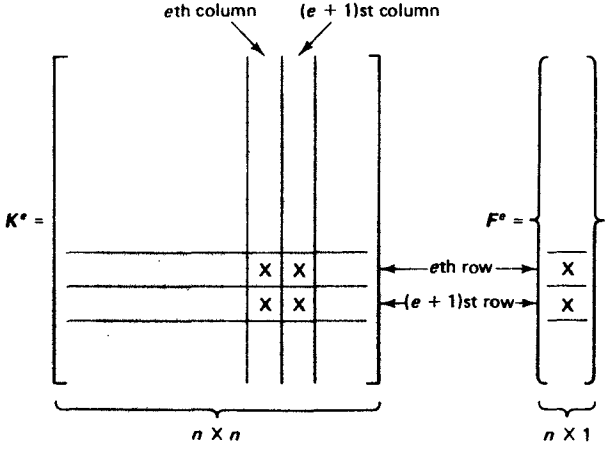


Figure 1.13.2 X's indicate nonzero terms; all other terms are zero.

Here $\boldsymbol{k}^e$ and $\boldsymbol{f}^e$ are defined with respect to the *local* ordering, whereas $\boldsymbol{K}^e$ and $\boldsymbol{F}^e$ are defined with respect to the global ordering. To determine where the components of $\boldsymbol{k}^e$ and $\boldsymbol{f}^e$ "go" in $\boldsymbol{K}$ and $\boldsymbol{F}$, respectively, requires keeping additional information. This is discussed in the following section.

1.14 ASSEMBLY OF GLOBAL STIFFNESS MATRIX AND FORCE VECTOR; LM ARRAY

In a finite element computer program, it is the task of a "finite element subroutine" to produce $\boldsymbol{k}^e$ and $\boldsymbol{f}^e$, $e = 1, 2, \ldots, n_{el}$, from given data and to provide an "assembly subroutine" enough information so that the terms in $\boldsymbol{k}^e$ and $\boldsymbol{f}^e$ can be added to the appropriate locations in $\boldsymbol{K}$ and $\boldsymbol{F}$, respectively. This assembly information is stored in an array named LM, the ***location matrix.***

Let us construct the LM array for the problem under consideration. The dimensions of LM are n_{en}, ***the number of element nodes,*** by the number of elements; in the present case, the numbers are 2 and n_{el}, respectively. Given a particular degree of freedom number and an element number (say a and e, respectively), the value returned by the LM array is the corresponding global equation number, A, viz.,

$$A = \mathrm{LM}(a, e) = \begin{cases} e & \text{if } a = 1 \\ e + 1 & \text{if } a = 2 \end{cases} \tag{1.14.1}$$

The complete LM array is depicted in Fig. 1.14.1. This is the way we envision it stored in the computer. Note that $\mathrm{LM}(2, n_{el}) = 0$. This indicates that degree of freedom 2 of element number n_{el} is prescribed and is not an unknown in the global matrix equation. Hence the terms $k_{12}^{n_{el}}$, $k_{21}^{n_{el}}$, $k_{22}^{n_{el}}$, and $f_2^{n_{el}}$ are *not* assembled into $\boldsymbol{K}$ and $\boldsymbol{F}$, respectively. (There are no places for them to go!)

Element numbers $1 \le e \le n_{el}$

Local node number	1	2	3	. . .	e	. . .	n_{el-1}	n_{el}
1	1	2	3	. . .	e	. . .	$n - 1$	n
2	2	3	4	. . .	$e + 1$	. . .	n	0

($n_{en} = 2$) $n_{en} \times n_{el}$

Figure 1.14.1 LM array for example problem.

As an example, assume we want to add the eth elemental contributions, where $1 \le e \le n_{el-1}$, to the partially assembled $\boldsymbol{K}$ and $\boldsymbol{F}$. From the LM array, we deduce the following assembly procedure:

$$K_{ee} \leftarrow K_{ee} + k_{11}^e \tag{1.14.2}$$

$$K_{e,e+1} \leftarrow K_{e,e+1} + k_{12}^e \tag{1.14.3}$$

$$K_{e+1,e} \leftarrow K_{e+1,e} + k_{21}^e \text{[6]} \tag{1.14.4}$$

$$K_{e+1,e+1} \leftarrow K_{e+1,e+1} + k_{22}^e \tag{1.14.5}$$

[6] Due to symmetry k_{21}^e would not actually be assembled in practice.

$$F_e \leftarrow F_e + f_1^e \tag{1.14.6}$$

$$F_{e+1} \leftarrow F_{e+1} + f_2^e \tag{1.14.7}$$

where the arrow (←) is read "is replaced by."

For element n_{el} we have only that

$$K_{nn} \leftarrow K_{nn} + k_{11}^{n_{el}} \tag{1.14.8}$$

$$F_n \leftarrow F_n + f_1^{n_{el}} \tag{1.14.9}$$

With these ideas, we may construct, in sketchy fashion, an algorithm for the assembly of $\boldsymbol{K}$ and $\boldsymbol{F}$; see Fig. 1.14.2.

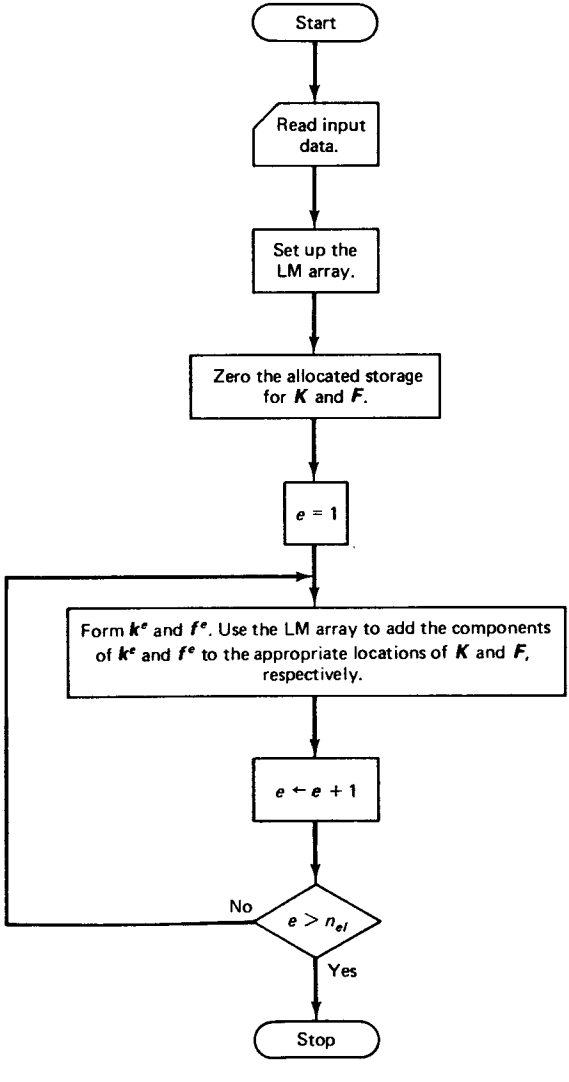


Figure 1.14.2 Flowchart of a finite element assembly algorithm.

The action of the assembly algorithm is denoted throughout by $\mathbf{A}$, the *assembly operator*, vis.,

$$\boxed{K = \mathop{\mathbf{A}}_{e=1}^{n_{el}} (k^e), \qquad F = \mathop{\mathbf{A}}_{e=1}^{n_{el}} (f^e)} \tag{1.14.10}$$

1.15 EXPLICIT COMPUTATION OF ELEMENT STIFFNESS MATRIX AND FORCE VECTOR

The explicit computation of k^e and f^e, for the problem under consideration, provides some preliminary insight into the type of calculations that must be performed in a finite element subroutine. Some preliminary results are required.

Change of Variables Formula (One-Dimensional Version)

Let $f:[x_1, x_2] \to \mathbb{R}$ be an integrable function and let $x:[\xi_1, \xi_2] \to [x_1, x_2]$ be continuously differentiable, with $x(\xi_1) = x_1$ and $x(\xi_2) = x_2$. Then

$$\boxed{\int_{x_1}^{x_2} f(x)\, dx = \int_{\xi_1}^{\xi_2} f(x(\xi)) x_{,\xi}(\xi)\, d\xi} \tag{1.15.1}$$

Chain Rule

Let f and x be as above, and, in addition, assume f is differentiable. Then

$$\boxed{\frac{\partial}{\partial \xi} f(x(\xi)) = f_{,x}(x(\xi)) x_{,\xi}(\xi)} \tag{1.15.2}$$

Proofs of these results may be found in [11].

The computation of k^e proceeds as follows:

$$\begin{aligned} k^e_{ab} &= \int_{\Omega^e} N_{a,x}(x) N_{b,x}(x)\, dx \qquad \text{(by definition)} \\ &= \int_{-1}^{+1} N_{a,x}(x(\xi)) N_{b,x}(x(\xi)) x_{,\xi}(\xi)\, d\xi \\ &\qquad \text{(Change of variables, where } x(\xi) \text{ is defined by (1.12.6))} \\ &= \int_{-1}^{+1} N_{a,\xi}(\xi) N_{b,\xi}(\xi) (x_{,\xi}(\xi))^{-1}\, d\xi \end{aligned}$$

(Chain rule; $N_{a,\xi}(\xi) = (\partial/\partial\xi)N_a(x(\xi)) = N_{a,x}(x(\xi))x_{,\xi}(\xi)$)

$= (-1)^{a+b}/h^e$ (by (1.12.7)–(1.12.9))

Thus

$$\boxed{k^e = \frac{1}{h^e}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}} \tag{1.15.3}$$

Observe that $N_{a,\xi}$ (see (1.12.7)) does not depend upon the particular element data, as $N_a = N_a(\xi)$. We shall see that this is generally true, and hence these computations may be done once and for all.

The derivatives $x_{,\xi}$ and $\xi_{,x}$ do depend on the particular element data (in the present case h^e), and subroutines will be necessary to compute the analogs of these quantities in more general cases.

Now we wish to compute f^e. However, this cannot be done without explicitly knowing what $\ell = \ell(x)$ is. In practice, it would be inconvenient to reprogram every time we wanted to solve a problem involving a different function ℓ. Generally a convenient approximation is made. For example, we might replace ℓ by its linear interpolate over each element, namely,

$$\ell^h = \sum_{a=1}^{2} \ell_a N_a \tag{1.15.4}$$

where $\ell_a = \ell(x(\xi_a))$; see Fig. 1.15.1. The notation ℓ^h is used to indicate that the approximation depends upon the mesh. This represents an approximation that is sufficient for most practical applications. (It is, of course, exact for constant or linear "loading" of the element.) Now standardization of input to the program may be facilitated; that is, the nodal values of ℓ are the required data. Let us employ this approximation in the explicit calculation of an element force vector:

$$\int_{\Omega^e} N_a(x)\ell^h(x)\,dx = \int_{-1}^{+1} N_a(x(\xi))\ell^h(x(\xi))x_{,\xi}(\xi)\,d\xi \qquad \text{(change of variables)}$$

$$= \frac{h^e}{2}\sum_{b=1}^{2}\int_{-1}^{+1} N_a(\xi)N_b(\xi)\,d\xi\,\ell_b \qquad \text{(by (1.12.8))} \tag{1.15.5}$$

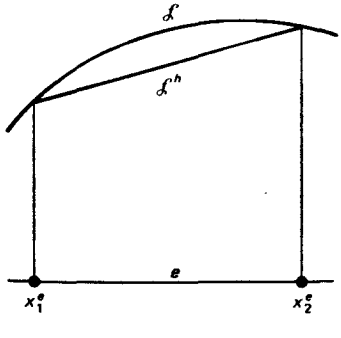


Figure 1.15.1 Approximation of ℓ by piecewise linear interpolation of nodal values.

Carrying out the integrations ($\int_{-1}^{+1} N_a N_b \, d\xi = (1 + \delta_{ab})/3$) yields

$$f^e = \frac{h^e}{6}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{Bmatrix} \ell_1 \\ \ell_2 \end{Bmatrix} \qquad \begin{matrix} (+ \text{ boundary terms,} \\ \text{cf. (1.13.12))} \end{matrix}$$

$$= \frac{h^e}{6}\begin{Bmatrix} 2\ell_1 + \ell_2 \\ \ell_1 + 2\ell_2 \end{Bmatrix} \qquad (+ \text{ boundary terms})$$

(1.15.6)

Remark. It can be shown that, under suitable hypotheses, piecewise linear nodal interpolation produces $O(h^2)$ errors in the data; in this case, ℓ. (See [12], pp. 56–57, for basic estimates of interpolation errors.) It can be shown that, in appropriate measures of the error, this produces at worst $O(h^2)$ errors in u^h and $u^h_{,x}$.

The following exercise indicates that there may be better ways to approximate given data.

Exercise 1. Suppose $\ell(x)$ is quadratic (i.e., consists of a linear combination of the monomials $1, x, x^2$). Determine a piecewise linear approximation—not necessarily continuous—to ℓ over each element which results in exact nodal values. *Hint:* The analysis may be performed with respect to one element.

Exercise 2. The equation of a string on an elastic foundation is given by:

$$u_{,xx} - \lambda u + \ell = 0 \qquad \text{on } \Omega =]0, 1[$$

where λ, a positive constant, is a measure of the foundation stiffness. Assuming the same boundary conditions as for the problem discussed previously in this chapter, it can be shown that an equivalent weak formulation is:

$$\int_\Omega (w_{,x} u_{,x} + w\lambda u) \, dx = \int_\Omega w\ell \, dx + w(0)h$$

where $u \in \mathcal{S}$, $w \in \mathcal{V}$, and so on. This can also be written as

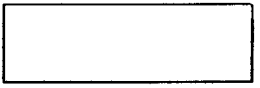
$$a(w, u) + (w, \lambda u) = (w, \ell) + w(0)h$$

i. Let $u^h = v^h + g^h$. Write the Galerkin counterpart of the weak formulation:

$$a(w^h, v^h) + \boxed{} =$$

$$(w^h, \ell) + w^h(0)h - a(w^h, g^h)$$

$$- \boxed{}$$

ii. Define $K_{AB} = a(N_A, N_B) +$ 

and

$$k^e_{ab} = a(N_a, N_b)^e +$$ 

iii. Determine k^e explicitly:

$$k^e = [k^e_{ab}] = \begin{bmatrix} & \\ & \end{bmatrix}$$

iv. Show that K is symmetric.

v. Show that K is positive definite. Is it necessary to employ the boundary condition $w^h(1) = 0$? Why?

vi. The Green's function for this problem satisfies

$$g_{,xx} - \lambda g + \delta_y = 0$$

and can be written as

$$g(x) = \begin{cases} c_1 e^{px} + c_2 e^{-px}, & 0 \le x \le y \\ c_3 e^{px} + c_4 e^{-px}, & y \le x \le 1 \end{cases}$$

where $p = \lambda^{1/2}$ and the c's are determined from the following four boundary and continuity conditions:

$$g(1) = 0$$
$$g_{,x}(0) = 0$$
$$g(y^+) = g(y^-)$$
$$g_{,x}(y^+) = g_{,x}(y^-) - 1$$

Why is the piecewise linear finite element space incapable of attaining nodally exact solutions in this case?

vii. Construct *exponential* element shape functions $N_1(x)$ and $N_2(x)$ such that

$$u^h(x) = d^e_1 N_1(x) + d^e_2 N_2(x), \qquad x \in \Omega^e$$

where

$$u^h(x) = c_1 e^{px} + c_2 e^{-px}$$

and the c's are determined from

$$d^e_a = u^h(x^e_a), \qquad a = 1, 2$$

What is the attribute which this choice of functions attains?

1.16 EXERCISE: BERNOULLI-EULER BEAM THEORY AND HERMITE CUBICS

This problem develops basic finite element results for Bernoulli-Euler beam theory. The strong form of a boundary-value problem for a thin beam (Bernoulli-Euler theory) fixed at one end and subjected to a shear force and moment at the other end, may be stated as follows:

Let the beam occupy the unit interval (i.e., $\Omega =]0, 1[$, $\overline{\Omega} = [0, 1]$).

(S) Given $\ell: \Omega \to \mathbb{R}$ and constants M and Q, find $u: \overline{\Omega} \to \mathbb{R}$ such that

$$EI\, u_{,xxxx} = \ell \text{ on } \Omega \qquad \text{(transverse equilibrium)}$$
$$u(1) = 0 \qquad \text{(zero transverse displacement)}$$
$$u_{,x}(1) = 0 \qquad \text{(zero slope)}$$
$$EI\, u_{,xx}(0) = M \qquad \text{(prescribed moment)}$$
$$EI\, u_{,xxx}(0) = Q \qquad \text{(prescribed shear)}$$

where E is Young's modulus and I is the moment of inertia, both of which are assumed to be constant.

The setup is shown in Fig. 1.16.1.

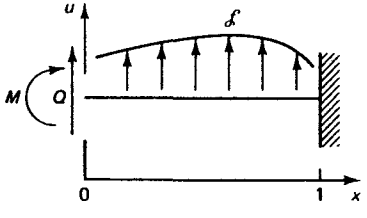


Figure 1.16.1

Let $\mathcal{S} = \mathcal{V} = \{w \mid w \in H^2(\Omega),\ w(1) = w_{,x}(1) = 0\}$[7]. Then a corresponding weak form of the problem is:

(W) Given ℓ, M, and Q, find $u \in \mathcal{S}$ such that for all $w \in \mathcal{V}$

$$\boxed{a(w, u) = (w, \ell) - w_{,x}(0)M + w(0)Q}$$

where

$$\boxed{\begin{aligned} a(w, u) &= \int_0^1 w_{,xx} EI\, u_{,xx}\, dx \\ (w, \ell) &= \int_0^1 w\ell\, dx \end{aligned}}$$

[7] $w \in H^2(\Omega)$ essentially means that $w_{,xx}$ is square-integrable (i.e., $\int_0^1 (w_{,xx})^2\, dx < \infty$).

The collection of functions, $\mathcal{V}$, may be thought of as the space of finite strain-energy configurations of the beam, satisfying the kinematic (essential) boundary conditions at $x = 1$. It is a consequence of Sobolev's theorem that each $w \in \mathcal{V}$ is continuously differentiable. For reasonable ℓ, these problems possess unique solutions.

Let $\mathcal{S}^h = \mathcal{V}^h$ be a finite-dimensional approximation of $\mathcal{S}$. In particular, we assume $w^h \in \mathcal{V}^h$ satisfies $w^h(1) = w^h_{,x}(1) = 0$.

The Galerkin statement of the problem goes as follows:

$$(G)\begin{cases}\text{Given } \ell,\ M, \text{ and } Q, \text{ find } u^h \in \mathcal{S}^h \text{ such that for all } w^h \in \mathcal{V}^h \\ \boxed{a(w^h, u^h) = (w^h, \ell) - w^h_{,x}(0)M + w^h(0)Q}\end{cases}$$

a. Assuming all functions are smooth and bounded, show that the solutions of (S) and (W) are identical. What are the natural boundary conditions?

b. Assume $0 = x_1 < x_2 < \cdots < x_{n+1} = 1$ and $\mathcal{V}^h = \{w^h \mid w^h \in C^1(\overline{\Omega}),\ w^h(1) = w^h_{,x}(1) = 0,$ and w^h restricted to $[x_A, x_{A+1}]$ is a cubic polynomial (i.e., consists of a linear combination of $1, x, x^2, x^3$)$\}$[8]. This is a space of ***piecewise cubic Hermite shape functions.*** Observe that $w^h \in \mathcal{V}^h$ need *not* have continuous second derivatives at the nodes.

On each subinterval, show that w^h may be written as

$$w^h(x) = N_1(x)w^h(x_1) + N_3(x)w^h(x_2) + N_2(x)w^h_{,x}(x_1) + N_4(x)w^h_{,x}(x_2)$$

where

$$N_1(x) = \frac{-(x - x_2)^2[-h + 2(x_1 - x)]}{h^3}$$

$$N_2(x) = \frac{(x - x_1)(x - x_2)^2}{h^2}$$

$$N_3(x) = \frac{(x - x_1)^2[h + 2(x_2 - x)]}{h^3}$$

$$N_4(x) = \frac{(x - x_1)^2(x - x_2)}{h^2}$$

Hint: Let $w^h(x) = c_1 + c_2x + c_3x^2 + c_4x^3$, where the c's are constants. Determine them by requiring the following four conditions hold:

$$w^h(x_1) = c_1 + c_2x_1 + c_3x_1^2 + c_4x_1^3$$

$$w^h(x_2) = c_1 + c_2x_2 + c_3x_2^2 + c_4x_2^3$$

$$w^h_{,x}(x_1) = c_2 + 2c_3x_1 + 3c_4x_1^2$$

$$w^h_{,x}(x_2) = c_2 + 2c_3x_2 + 3c_4x_2^2$$

[8] The notation $w^h \in C^1$ means w^h is continuously differentiable.

Sketch the element functions N_1, N_2, N_3, and N_4, and their typical global counterparts.

The finite element space described in part (b) results in *exact* nodal displacements and slopes (first derivatives), analogous to the case presented in Sec. 1.10. In part (g). you are asked to prove this. In problems of beam bending we are generally interested in curvatures (second derivatives) for bending moment calculations.

c. Locate the optimal curvature points in the sense of Barlow. *Warning:* The algebraic manipulations can be tiresome unless certain simplifications are observed. If we work in the ξ-element coordinate system introduced in Sec. 1.12 (recall $\xi = (2x - x_A - x_{A+1})/h_A$). the location of the Barlow curvature points may be expressed as $\xi = \pm 1/\sqrt{3}$. That is, there are two symmetrically spaced optimal locations to compute curvature.

d. What is the rate of convergence of curvature at these points? (*Ans.* $O(h^3)$).

e. If the segment of the beam $[x_A, x_{A+1}]$ is unloaded (i.e., $u_{,xxxx} = 0$, where u is the exact solution), which points are optimal?

f. Assume $n_{el} = 1$ (i.e.. one element) and $\ell(x) = c$ = constant. Set up and solve the Galerkin-finite element equations. Plot u^h and u; $u^h_{,x}$ and $u_{,x}$; and $u^h_{,xx}$ and $u_{,xx}$. Indicate the locations of the Barlow curvature points.

g. Prove that

$$u^h(x_A) = u(x_A)$$

$$u^h_{,x}(x_A) = u_{,x}(x_A)$$

where x_A is a typical node (i.e.. prove ***the displacements and slopes are exact at the nodes***). To do the second part you will have to be familiar with the ***dipole,*** $\delta_{,x}(x - x_A)$, which is the generalized derivative of the delta function.

h. Show that the Barlow curvature points are exact when $\ell(x) = c$ = constant.

i. Why do we require that the functions in $\mathcal{V}^h$ have continuous first derivatives?

j. Calculate the 4×4 element stiffness matrix,

$$k^e_{pq} = \int_{x_1^e}^{x_2^e} N_{p,xx} EI\, N_{q,xx}\, dx \qquad 1 \le p, q \le 4$$

where $h^e = x_2^e - x_1^e$.

k. (See the exercise in Sec. 1.8.) Consider the weak formulation. Assume $w \in \mathcal{V}$ and $u \in \mathcal{S}$ are smooth on element interiors (i.e., on $]x_A, x_{A+1}[$) but may exhibit discontinuities in second, and higher, derivatives across element boundaries. (Functions of this type contain the piecewise-cubic Hermite functions.) Show that

$$
\begin{aligned}
0 = \sum_{A=1}^{n} \int_{x_A}^{x_{A+1}} & w(EI\, u_{,xxxx} - \ell)\, dx \\
& - w_{,x}(0)(EI\, u_{,xx}(0^+) - M) \\
& + w(0)(EI\, u_{,xxx}(0^+) - Q) \\
& - \sum_{A=2}^{n} w_{,x}(x_A)\, EI\, (u_{,xx}(x_A^+) - u_{,xx}(x_A^-)) \\
& + \sum_{A=2}^{n} w(x_A)\, EI\, (u_{,xxx}(x_A^+) - u_{,xxx}(x_A^-))
\end{aligned}
$$

from which it may be concluded that the Euler-Lagrange conditions are

i. $EI\, u_{,xxxx}(x) = \ell(x)$, where $x \in]x_A, x_{A+1}[$ and $A = 1, 2, \ldots, n$
ii. $EI\, u_{,xx}(0^+) = M$
iii. $EI\, u_{,xxx}(0^+) = Q$
iv. $EI\, u_{,xx}(x_A^+) = EI\, u_{,xx}(x_A^-)$, where $A = 2, 3, \ldots, n$
v. $EI\, u_{,xxx}(x_A^+) = EI\, u_{,xxx}(x_A^-)$, where $A = 2, 3, \ldots, n$

Note that (i) is the equilibrium equation *restricted to the element interiors*, and (iv) and (v) are continuity conditions across element boundaries of moment and shear, respectively. Contrast these results with those obtained for functions w and u, which are *globally* smooth.

The Galerkin finite element formulation yields a solution that *approximates* (i) through (v).

Appendix 1.1

An Elementary Discussion of Continuity, Differentiability, and Smoothness

Throughout Chapter 1 we have introduced mathematical terminologies and ideas in a gradual, as-needed format. Many of these ideas had to do with the continuity and differentiability of functions. The presentation was, admittedly, somewhat vague on these points in order that the main ideas would not be overencumbered. Careful characterization of the properties of functions is an essential ingredient in the development and analysis of finite element methods. However, to pursue this subject deeply would take us into the realm of serious mathematical analysis, which is outside the scope of this book. Nevertheless, we feel compelled to say a few additional words on the subject to round out the presentation in Chapter 1 and to expose the reader to notations and ideas that will probably be encountered if he or she attempts to read published papers on finite elements.

The discussion here will be restricted to one dimension. In Chapter 1 we spoke of continuously differentiable functions. If we have a grasp of the notion of a continuous function, then continuously differentiable functions pose no problem.

Definition: A function f: $\Omega \to \mathbb{R}$ (recall $\Omega =]0, 1[$) is said to be ***k-times continuously differentiable,*** or ***of class*** $C^k = C^k(\Omega)$, if its derivatives of order j, where $0 \leq j \leq k$, exist and are continuous functions.

A C^0 function is simply a continuous function. A C^∞ function is one that possesses a continuous derivative of any order (i.e., $j = 0, 1, \ldots, \infty$).

Definition: A function f is said to be ***of class*** C_b^k if it is C^k and bounded (i.e., $|f(x)| < c$, where c is a constant, for all $x \in \Omega$).

Example 1

The functions defined by monomials (i.e., $f(x) = 1, x, x^2$, etc.) are C_t^∞

Example 2

The function $f(x) = 1/x$ is continuous on Ω, as are all its derivatives; hence it is C^∞, but it is not bounded (i.e., there does not exist a constant c such that $|1/x| < c$ for all $x \in \Omega$; see Fig. 1.I.1). Consequently this function is not of class C_b^k for any $k \geq 0$.

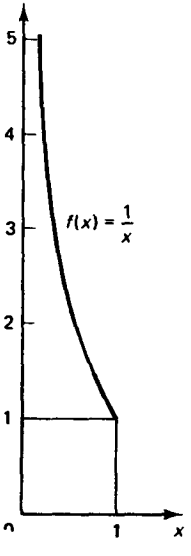


Figure 1.I.1 A continuous function that is not bounded.

Example 3

The function

$$f(x) = \begin{cases} x, & x \leq \frac{1}{2} \\ 1/2, & x > \frac{1}{2} \end{cases} \tag{1.I.1}$$

is continuous but not continuously differentiable (i.e., it is C_b^0 but not C_b^1). Functions in C_b^k, $k \geq 1$, but not in C_b^{k+1} may be constructed by integrating (1.I.1) k times. For example,

$$f(x) = \begin{cases} \dfrac{x^2}{2}, & x \leq \frac{1}{2} \\ \dfrac{(x - \frac{1}{4})}{2}, & x > \frac{1}{2} \end{cases} \tag{1.I.2}$$

is in C_b^1 but not C_t^2. (The reader may wish to verify this.)

There is no universally accepted definition of what is meant by a "smooth" function. However, it is generally taken to mean that at least one derivative exists and is continuous (i.e., either C^1 or C_b^1) and sometimes means $k > 1$, even ∞.

The C^k and C_b^k functions employ the classical notion of a derivative in their definitions. If we employ the closed unit interval, $\overline{\Omega} = [0, 1]$, instead of $\Omega =]0, 1[$, the difference between C^k and C_b^k disappears. This is because if f is $C^k([0, 1])$, $f(0)$ and $f(1)$ are real numbers and are not allowed to be ∞. Thus unboundedness, as in the example above, is precluded. Very often, we think of C^k functions in this light. However, in some situations the differences between $C^k(\Omega)$ and $C_b^k(\Omega)$ must be kept in mind.

Generally, finite element functions are smooth on element interiors (there are exceptions, however) but possess only low-order continuity across element boundaries. One might be tempted to characterize them as locally smooth but globally "rough." The piecewise linear finite element functions discussed in Sec. 1.8 are of class C_b^0. The Hermite cubics employed in Sec. 1.16 are C_b^1. To calculate derivatives of such functions we need to employ the notion of a "generalized derivative," as was used in solving the Green's function problem of Sec. 1.10. For example, the first derivative of a piecewise linear finite element function is a generalized step function; the second derivative is a generalized Dirac delta function (i.e., delta functions, of various amplitudes, acting at the nodes). In the case of the Hermite cubics, the first derivative is continuous, the second a generalized step function, and so on. We have seen in Sec. 1.16 that other generalized functions also arise in the analysis of finite element behavior (namely, the dipole). The useful examples of generalized functions are by no means exhausted by what we have seen thus far. However, the ones we have introduced are perhaps the most basic.

In the mathematical analysis of boundary-value problems, and consequently in finite element analysis, we need to introduce classes of functions that possess generalized derivatives and, in addition, certain integrability properties. We have encountered such functions in the statements of weak formulations in Sec. 1.3 and 1.16. These are particular examples of ***Sobolev spaces of functions*** defined as follows:

$$H^k = H^k(\Omega) = \{w \mid w \in L_2;\ w_{,x} \in L_2;\ \ldots;\ w_{,\underbrace{x \ldots x}_{k \text{ times}}} \in L_2\} \tag{1.I.3}$$

where

$$L_2 = L_2(\Omega) = \{w \mid \int_0^1 w^2\, dx < \infty\} \tag{1.I.4}$$

In words, the Sobolev space of degree k, denoted by H^k, consists of functions that possess square-integrable generalized derivatives through order k. A square-integrable function is called an L_2-function, by virtue of (1.I.4). From (1.I.3), we see that $H^0 = L_2$ and that $H^{k+1} \subset H^k$. The Sobolev spaces are the most important for studying elliptic boundary-value problems.

The question naturally arises as to the relation between Sobolev spaces and the classical spaces of differentiable functions introduced previously. In particular, when is an H^k-function smooth in the classical sense? The answer is provided by ***Sobolev's theorem,*** which states that, in one dimension, $H^{k+1} \subset C_b^k$. That is, if a function is of class H^{k+1}, then it is actually a C_b^k function. For example, in Sec. 1.3 we required H^1 functions. By Sobolev's theorem, such functions are, additionally, continuous and bounded. In Sec 1.16, we employed H^2 functions. These are C_b^1 by Sobolev's theorem and thus possess bounded, continuous, classical derivatives.

Certain "singularities" are precluded by square-integrability. For example, $x^{-1/4}$ is in L_2, but $x^{-1/2}$ is not. (Verify!) Such considerations become important in many physical circumstances (e.g., in fracture mechanics).

The number of other types of function spaces that arise in mathematical analysis is large, and many are difficult to comprehend without serious training in "functional analysis." These topics are outside the scope of this book. The reader who wishes to delve further may consult [13, 14, 15] and references therein.

REFERENCES

Section 1.2

1. A. R. Mitchell and D. F. Griffiths, *The Finite Difference Method in Partial Differential Equations*. New York: John Wiley, 1980.

Section 1.5

2. G. Strang and G. J. Fix, *An Analysis of the Finite Element Method*. Englewood Cliffs, N. J.: Prentice-Hall, 1973.
3. B. A. Finlayson, *The Method of Weighted Residuals and Variational Principles*. New York: Academic Press, 1972.
4. B. A. Finlayson and L. E. Scriven, "The Method of Weighted Residuals—A Review," *Applied Mechanics Reviews*, 19, (1966), 735–738.

Section 1.10

5. I. Stakgold, *Boundary-Value Problems of Mathematical Physics*. Vols. I and II, New York: Macmillan, 1968.
6. G. Strang and G. J. Fix, *An Analysis of the Finite Element Method*. Englewood Cliffs, N. J.: Prentice-Hall, 1973.
7. J. E. Marsden, *Elementary Classical Analysis*. San Francisco: W. H. Freeman, 1974.
8. J. Barlow, "Optimal Stress Locations in Finite Element Models," *International Journal for Numerical Methods in Engineering*, 10 (1976), 243–251.
9. E. Popov, *Introduction to Mechanics of Solids*. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

Section 1.13

10. M. J. Turner, R. W. Clough, H. C. Martin, and L. J. Topp, "Stiffness and deflection analysis of complex structures," *Journal of Aeronautical Sciences*, 23 (1956), 805–823.

Section 1.15

11. J. E. Marsden, *Elementary Classical Analysis*. San Francisco: W. H. Freeman, 1974.
12. P. J. Davis, *Interpolation and Approximation*. New York: Blaisdell, 1963.

Appendix 1.I

13. P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*. New York: North-Holland, 1978.
14. J. T. Oden and J. N. Reddy, *An Introduction to the Mathematical Theory of Finite Elements*. New York: Academic Press, 1978.
15. J. T. Oden, *Applied Functional Analysis*. Englewood Cliffs, N. J.: Prentice-Hall, 1979.