THE FINITE ELEMENT METHOD AND ITS APPLICATIONS

MASATAKE MORI

THE FINITE ELEMENT METHOD AND ITS APPLICATIONS

MASATAKE MORI

MACMILLAN PUBLISHING COMPANY
NEW YORK

COLLIER MACMILLAN CANADA, INC.
TORONTO

COLLIER MACMILLAN PUBLISHERS
LONDON

Originally published in Japanese by Iwanami Shoten, Publishers, Tokyo.
This English edition published in 1986 by Macmillan Publishing Company, New York, by arrangement with Iwanami Shoten, Publishers, Tokyo.

Macmillan Publishing Company,
866 Third Avenue, New York, NY 10022

Collier Macmillan Canada, Inc.

Printed in the United States of America

Printing: 1 2 3 4 5 6 7 8 9 10   Year: 6 7 8 9 0 1 2 3 4 5

Library of Congress Cataloging-in-Publication Data

Mori, Masatake.
The finite element method and its applications.

Translation of: Yūgen yōsohō to sono ōyō.
Bibliography: p.
Includes index.
1. Finite element method. I. Title.
TA347.F5M6713 1986 620'.0015153'53 86-87
ISBN 0-02-948621-1

Preface to English Edition

This is a translation of *Yuugen-youso-hou To Sono Ouyou,* written in Japanese. The translation was done by the author of the original edition. Since the original edition was intended to present the fundamentals of the finite element method, and to avoid including material that might be out of date in several years, the original edition has been translated faithfully into English without changing the contents, except that books [7], [8], and [10] on the reference list, which were written in Japanese, have been replaced by equivalent books written in English.

I have attempted to correct errors in the original edition and wish to thank those who have pointed out such errors, in particular, Shinsuke Suga. Thanks are also due to Atsuko Yamamoto and Tami Nakata for typing the manuscript.

Masatake Mori

Preface

The purpose of this book is to introduce the finite element method from the standpoint of applied mathematics. While this technique is indispensable in the field of structural mechanics, it is also a powerful tool for the numerical solution of partial differential equations related to various natural phenomena. Actually it has been a long time since the finite element method established itself in the natural sciences and in engineering. Accordingly, a variety of books about this method have been published. There are two conventional ways to approach the finite element method—from structural mechanics and from the solution of partial differential equations. Consequently, there are also two standpoints when writing a book on this subject. I intended to write this book from the standpoint of solving partial differential equations. Furthermore, books on the finite element method may be classified into two groups—books about techniques and books about mathematics. This book belongs in the latter group.

Although in practice the finite element method is applied to problems in two or three space dimensions, this book starts with applications to problems in one dimension in order to provide an easily understood description of the basic idea of the method, and then presents the mathematical background and error analysis. The main part of the book describes the application of the finite element method to partial differential equations in two dimensions, including time-dependent equations such as the heat equation and the wave equation. Chapters 3, 6, 7, and 14 are devoted mainly to error analysis, and the description is more mathematical, so that these chapters may be omitted on a first reading if the reader is more interested in the techniques of the finite element method.

In practice, in the final stage of the finite element method a system of linear equations with a large, sparse coefficient matrix of order, say, several tens of thousands must be olved, or an eigenvalue problem with respect to such a matrix must be solved. Therefore, in order to master thoroughly the techniques of the finite element method numerical analysis of such large-scale matrices should be learned. However, the numerical analysis of linear algebra is itself a large problem, and to describe it is beyond the scope of this book, hence only brief comments on this subject are included.

The amount of work published in the field of the finite element method is enormous, and it is impossible to describe all of it in a single book. Therefore the material presented is not exhaustive but is limited to what is necessary for readers studying this procedure for the first time. In particular, in Chap. 13, methods but not theories for solving nonlinear problems are presented. I hope that those who become interested in this method after reading the book will expand their knowledge by consulting the references listed at the end.

I intended to write this book so that it could be read without referring to other books or papers. In order to learn about the finite element method knowledge of the variational principle is necessary. And in order to understand it mathematically a knowledge of functional analysis is required. However, this book does not assume this specialized knowledge. On the contrary, the reader will be able to learn about the variational principle and functional analysis in a practical way using the finite element method. Although this book belongs to the field of mathematics, I have avoided descriptions that are too abstract and tried to write so that the reader can understand the mathematical theorems intuitively. In this way, I hoped to bridge the gap between the theory and the practice of the finite element method. Furthermore, in the theoretical discussions I have tried to explain modern ideas using classical terms so that physicists and engineers who are not familiar with the technical terms of modern mathematics can read the book without difficulty. I hope that this book will be useful not only as an introductory text for students studying the finite element method for the first time but also in satisfying the mathematical interests of scientists and engineers who have used the finite element method as a tool for solving problems.

Special thanks are due to Professor Hiroshi Hujita who aroused my interest in the mathematical aspect of the finite element method. I also thank my many colleagues, especially Makoto Natori, Masaaki Naka-

mura, and Masaaki Sugihara, who read the manuscript carefully and provided many suggestions. Finally, thanks are due to Mr. Hisao Miyauchi of Iwanami Shoten.

Masatake Mori

Contents

1

The Basic Idea of the Finite Element Method

1.1 *A Two-Point Boundary Value Problem*

In order to sketch the basic idea of the *finite element method* (FEM) we first consider a two-point boundary value problem in one dimension:

$$-\frac{d}{dx}\left(p\frac{du}{dx}\right) + qu = f(x) \qquad 0 < x < 1 . \tag{1.1.1}$$

$$u(0) = u(1) = 0 \tag{1.1.2}$$

where p and q are given positive constants and $f(x)$ is a given function. The boundary condition (1.1.2), which prescribes the values at both end points to vanish, is a typical example of a homogeneous Dirichlet condition.

Since (1.1.1) is a linear differential equation with constant coefficients having an inhomogeneous term, it is easy to solve for u in a closed form. However, we consider here an approximate solution in terms of a Fourier series with N terms as follows in order to explain the basic idea of the FEM:

$$u_N(x) = \sum_{j=1}^{N} a_j \sin j\pi x \tag{1.1.3}$$

The terms of $\cos j\pi x$ that may appear in a general Fourier series are discarded at the beginning in the present series so that it satisfies the boundary condition (1.1.2). In order to obtain the values of the coefficients a_j we use a conventional procedure; that is, we substitute u_N in (1.1.3) for u in (1.1.1), multiply both sides by $\sin k\pi x$, and integrate over (0,1). Then we have

$$\sum_{j=1}^{N} a_j p(j\pi)^2 \int_0^1 \sin k\pi x \sin j\pi x \, dx + \sum_{j=1}^{N} a_j q \int_0^1 \sin k\pi x \sin j\pi x \, dx$$

$$= \int_0^1 f(x) \sin k\pi x \, dx \tag{1.1.4}$$

From the orthogonality of $\{\sin j\pi x\}$ over (0,1),

$$\int_0^1 \sin k\pi x \sin j\pi x \, dx = \begin{cases} \frac{1}{2} & k = j \\ 0 & k \neq j \end{cases} \tag{1.1.5}$$

we have

$$a_k = \frac{2}{(k\pi)^2 p + q} \int_0^1 f(x) \sin k\pi x \, dx \tag{1.1.6}$$

Therefore if we substitute this expression into (1.1.3) we eventually obtain an approximate solution to the problem (1.1.1) and (1.1.2).

1.2 *A Solution in Terms of a Generalized Fourier Series*

We can generalize the procedure stated above in the following way. First we choose a set of linearly independent functions

$$\varphi_j(x) \qquad j = 1, 2, \ldots, N \tag{1.2.1}$$

and construct an approximate solution to the boundary value problem (1.1.1) and (1.1.2) in terms of a linear combination of these functions:

$$u_N(x) = \sum_{j=1}^{N} a_j \varphi_j(x) \tag{1.2.2}$$

We assume here that each $\varphi_j(x)$ satisfies

$$\varphi_j(0) = \varphi_j(1) = 0 \tag{1.2.3}$$

This assumption forces $u_N(x)$ to satisfy the boundary condition

$$u_N(0) = u_N(1) = 0 \tag{1.2.4}$$

corresponding to (1.1.2) from the beginning. Substituting $u_N(x)$ for u in (1.1.1), multiplying both sides by $\phi_k(x)$, and integrating over (0,1), we have

$$-\sum_{j=1}^{N} a_j p \int_0^1 \varphi_k(x) \frac{d^2\varphi_j(x)}{dx^2} \, dx + \sum_{j=1}^{N} a_j q \int_0^1 \varphi_k(x) \varphi_j(x) \, dx$$

$$= \int_0^1 f(x)\varphi_k(x)\, dx \tag{1.2.5}$$

Integrating the first term by parts using the boundary condition (1.2.3) results in

$$\sum_{j=1}^{N} a_j\left\{p\int_0^1 \frac{d\varphi_k}{dx}\frac{d\varphi_j}{dx}\,dx + q\int_0^1 \varphi_k\varphi_j\,dx\right\} = \int_0^1 f(x)\varphi_k(x)\,dx$$

$$k = 1, 2, \ldots, N \tag{1.2.6}$$

The set of functions

$$\varphi_j(x) = \sin j\pi x \qquad j = 1, 2, \ldots, N \tag{1.2.7}$$

mentioned above has an orthogonality in the sense of (1.1.5) over (0,1). In addition, their derivatives

$$\frac{d\varphi_j(x)}{dx} = j\pi \cos j\pi x \qquad j = 1, 2, \ldots, N \tag{1.2.8}$$

also have the same type of orthogonality. Therefore only the term with $j = k$ on the left side of (1.2.6) remains without vanishing, and the coefficients $\{a_j\}$ can be obtained by simple algebraic division. An arbitrarily given set of functions $\{\phi_j(x)\}$, on the other hand, usually has no orthogonality. Even if these functions $\{\phi_j(x)\}$ satisfy an orthogonality requirement by themselves, their derivatives $\{d\phi_j(x)/dx\}$ in general will not. Also, if p or q is a function of x, the orthogonality of the functions $\{\phi_j(x)\}$ will gain nothing at all.

The system of equations (1.2.6), which is obtained on the basis of $\{\phi_j(x)\}$, is a system of simultaneous linear equations with N unknowns $\{a_j\}$. If the coefficient matrix of (1.2.6) is not singular, we can obtain a solution by solving this system of equations for a_j, $j = 1, 2, \ldots, N$. Each member of the set of functions given by (1.2.1) is called a *basis function,* and the method for obtaining a solution in the form of a Fourier series in a wider sense as shown above is called *Galerkin's method.* Typical basis functions that have been conventionally used in applied analysis are trigonometric functions, simple monomials, and orthogonal polynomials.

1.3 *Piecewise Linear Basis Functions*

Consider a pyramid-shaped set of basis functions $\phi_k(x)$ as shown in Fig. 1.1. In order to define this set of functions we first divide the interval [0,1] into n subintervals with an equal mesh size

$$h = \frac{1}{n} \tag{1.3.1}$$

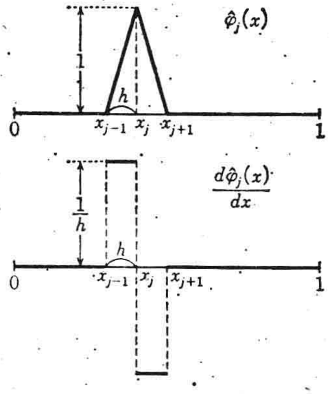

Fig. 1.1 Piecewise linear basis function and its derivative.

Then, in each subinterval bounded by the *nodes*

$$x_k = kh \qquad k = 0, 1, 2, \ldots, n \tag{1.3.2}$$

we define

$$\varphi_k(x) = \begin{cases} 0 & 0 \le x < x_{k-1} \\ \dfrac{x - x_{k-1}}{h} & x_{k-1} \le x < x_k \\ \dfrac{x_{k+1} - x}{h} & x_k \le x < x_{k+1} \\ 0 & x_{k+1} \le x \le 1 \end{cases} \tag{1.3.3}$$

where $\hat{\phi}_0(x)$ and $\hat{\phi}_n(x)$ are defined by the right half and left half of (1.3.3), respectively. This type of basis function is called a *piecewise linear basis function*. One of the remarkable characteristics of this function is that its support is extremely localized to a small domain. Such a function is sometimes called a *local basis function*. The fact that the support of the basis function is localized is, as will be shown later, significant from the standpoint of numerical computation.

The derivatives of the basis function (1.3.3) are

$$\frac{d\varphi_k}{dx} = \begin{cases} 0 & 0 \le x < x_{k-1} \\ \dfrac{1}{h} & x_{k-1} \le x < x_k \\ -\dfrac{1}{h} & x_k \le x < x_{k+1} \\ 0 & x_{k+1} \le x \le 1 \end{cases} \tag{1.3.4}$$

as shown in Fig. 1.1.

1.4 *Construction of an Approximate Equation*

We write here an approximate solution of (1.1.1) and (1.1.2) in terms of a linear combination of $\{\phi_j(x)\}$:

$$\hat{u}_n(x) = \sum_{j=1}^{n-1} a_j\hat{\varphi}_j(x) \tag{1.4.1}$$

Because of the boundary condition (1.1.2) we have omitted the terms corresponding to $j=0$ and $j=n$ from the beginning. It is evident that the function given by (1.4.1) consists of polygonal lines as shown in Fig. 1.2. A function of this shape is called a *piecewise linear polynomial.* The expansion (1.4.1) satisfies

$$\hat{u}_n(x_j) = a_j \tag{1.4.2}$$

at the node $x = x_j$; that is, the coefficient a_j of the expansion is equal to the value of $\hat{u}_n(x)$ at the node, which is quite convenient for practical purposes.

Now when we try to apply the procedure stated in the previous section starting with the expansion (1.4.1), we encounter an unfavorable situation. That is, although $\hat{u}_n(x)$ can be differentiated once, it cannot be differentiated twice. The first derivative of $\hat{u}_n(x)$ is discontinuous as seen from (1.3.4), and when we try to differentiate further, the *Dirac δ-function* appears at every node in the second derivative of $\phi_j(x)$. It is evident that this is not consistent with (1.1.1).

In order to avoid this inconsistency we consider an equation of the following form instead of (1.1.1):

$$\int_0^1 \left(p\,\frac{d\hat{u}_n}{dx}\,\frac{d\hat{\varphi}_j}{dx} + q\hat{u}_n\hat{\varphi}_j\right) dx = \int_0^1 f\hat{\varphi}_j\, dx \qquad j = 1, 2, \ldots, n-1 \tag{1.4.3}$$

That is, we start with (1.2.6) which is obtained by multiplying (1.1.1) by $\hat{\phi}_j$ followed by integration by parts.

We substitute (1.4.1) for $\hat{u}_n$ on the left-hand side of (1.4.3) and carry out termwise integration using (1.3.3) or (1.3.4). Then, corresponding to each integral we have

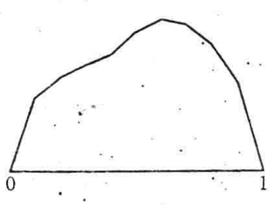

Fig. 1.2 Piecewise linear polynomial $\hat{u}_n(x)$.

$$\int_0^1 \hat{\varphi}_k(x)\hat{\varphi}_j(x)\,dx = \begin{cases} 0 & j < k-1 \\ \frac{1}{6}h & j = k-1 \\ \frac{2}{3}h & j = k \\ \frac{1}{6}h & j = k+1 \\ 0 & j > k+1 \end{cases} \tag{1.4.4}$$

$$\int_0^1 \frac{d\varphi_k}{dx}\frac{d\hat{\varphi}_j}{dx}\,dx = \begin{cases} 0 & j < k-1 \\ -\frac{1}{h} & j = k-1 \\ \frac{2}{h} & j = k \\ -\frac{1}{h} & j = k+1 \\ 0 & j > k+1 \end{cases} \tag{1.4.5}$$

Substitution of these values into (1.2.6) leads to the following system of linear equations with respect to $\{a_j\}$:

$$(K + M)\boldsymbol{a} = \boldsymbol{f} \tag{1.4.6}$$

where

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n-1} \end{pmatrix} \tag{1.4.7}$$

and, if we define

$$f_j = \int_0^1 f(x)\hat{\varphi}_j(x)\,dx \tag{1.4.8}$$

then

$$\boldsymbol{f} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \end{pmatrix} \tag{1.4.9}$$

K and M are $(n-1) \times (n-1)$ matrices defined as

$$K = \frac{p}{h}\begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & 0 & \\ & -1 & 2 & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & \ddots & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix} \tag{1.4.10}$$

$$M = \frac{qh}{6}\begin{pmatrix} 4 & 1 & & & & \\ 1 & 4 & 1 & & 0 & \\ & 1 & 4 & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & \ddots & 4 & 1 \\ & & & & 1 & 4 \end{pmatrix} \tag{1.4.11}$$

In accordance with the traditional terminology of structural mechanics K and M are called the *stiffness matrix* and the *mass matrix*, respectively.

1.5 *Properties of Matrices and the Finite Element Solution*

One of the most remarkable features of the matrices mentioned above is that they are *tridiagonal.* A tridiagonal matrix is a matrix whose entries vanish except on the diagonal and on the subdiagonal. In the first example (1.1.3), in which trigonometric functions were used as basis functions that were orthogonal to each other along with their derivatives, the matrices K and M were both diagonal. On the other hand, in the present example in which the basis functions (1.3.3) are used, the matrices are not completely diagonal but *nearly diagonal,* that is, tridiagonal. It is evident that the reason why they are nearly diagonal is that the support of the basis functions (1.3.3) is localized in a very small domain. Generally speaking, if functions whose support is localized in a small domain are used as basis functions, the nonzero entries of the coefficient matrix of the system of linear equations will be concentrated close to the diagonal, although not completely on the diagonal. As will be seen later, this matrix pattern leads to highly efficient numerical computation and substantial saving of computer memory.

The coefficient matrices K and M of the system of linear equations (1.4.6) are *symmetric*. In addition, K is *positive definite* in the present case, because for any vector

$$\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \end{pmatrix} \neq \boldsymbol{0} \tag{1.5.1}$$

we have

$$b^T K b = \int_0^1 p\left(\sum_{j=1}^{n-1} b_j \frac{d\hat{\varphi}_j}{dx}\right)^2 dx > 0 \qquad (1.5.2)$$

since p is assumed to be positive. Here b^T is the transposed vector of b. The matrix M is also shown to be positive definite in the same way. When p or q is a function of x, K and M also become positive definite provided that $p(x) > 0$ and $q(x) > 0$. If K and M are positive definite, then $K + M$ is also positive definite, so that (1.4.6) can be solved for $\{a_j\}$. Then, if we substitute $\{a_j\}$ into (1.4.1), we have an approximate solution $\hat{u}_n(x)$. This solution is just the *finite element solution*, which is the main topic of this book. It will become clear in Chap. 3 in what sense it is an approximation to an exact solution of the problem.

Although the example stated above is a simple model problem in one dimension, the same idea applies directly to problems in two or three dimensions. For example, in a two-dimensional problem, we divide the given domain into small triangles and choose piecewise polynomial basis functions each of which vanishes except in a small number of triangles adjacent to each other. Then we construct an approximate solution in terms of a linear combination of these basis functions and apply Galerkin's method.

The method stated above in which we divide the whole domain into subdomains, for example, triangles with (not an infinitely small but) a finite area, choose basis functions such that each of them does not vanish only in the near neighborhood of a particular node, construct an approximate solution to the given problem in terms of a linear combination of piecewise polynomials, and apply the Galerkin method, is generically called the finite element method (FEM).

In the FEM we usually must solve a large system of linear equations such as (1.4.6). For this reason no one thought of applying the FEM as a powerful tool for solving practical problems until high-speed computers with large storage capacities became available.