

Universität Stuttgart

INSTITUT FÜR  
MECHANIK

# Discretization Methods

An Introduction

---

K. Willner

A. Schmidt

# Chapter 1

## Introduction

### What are discretization methods?

A way to solve a differential equation or a system of differential equations.

### Why do we need them?

Nearly every physical problem in engineering can be modeled by a single differential equation or a system thereof. Some simple examples are<sup>1</sup>

- one-dof oscillator

$$m u(t)_{,tt} + c u(t) = f(t) \quad (1.1)$$

- 1-D steady-state heat conduction

$$-\kappa \Theta(x)_{,xx} = \varrho r(x) \quad (1.2)$$

- 2-D steady-state heat conduction

$$-\kappa_x \Theta(x, y)_{,xx} - \kappa_y \Theta(x, y)_{,yy} = \varrho r(x, y) \quad (1.3)$$

- 1-D transient heat conduction

$$\varrho c_v \Theta(x, t)_{,t} - \kappa \Theta(x, t)_{,xx} = \varrho r(x, t) \quad (1.4)$$

- equation for axial motion of a rod

$$\varrho u(x, t)_{,tt} - E u(x, t)_{,xx} = \varrho f(x, t) \quad (1.5)$$

- equation for bending motion of a beam

$$\varrho A w(x, t)_{,tt} - \varrho I w(x, t)_{,tttxx} + EI w(x, t)_{,xxxx} = q(x, t) \quad (1.6)$$

### How do they work?

Discretization methods transform the differential equations into algebraic equations, which can be easily solved by a computer.

---

<sup>1</sup>The notation  $a_{,x}$  denotes partial derivative of the argument  $a$  with respect to  $x$ :

$$a_{,x} = \frac{\partial a}{\partial x}$$

The discrete thing about them is that the solution is represented by data at discrete points in space and/or time and an interpolation between these data points. Now, instead of seeking an unknown function in a function space, only discrete values of the data points are sought. By this method the solution space is reduced from the infinite dimensional space of the exact analytical solution to a finite dimensional space of the approximate discrete solution.

### Pros

- well suited to computer implementation
- virtually any differential equation can be solved

### Cons

- only approximate solution
- may be very costly in terms of computation time if we use a lot of free parameters, i.e., if we choose a very fine discretization.

### What methods will be presented?

In this course we will treat the *Finite Difference Method* (FDM), the *Method of Weighted Residuals*, and the *Finite Element Method* (FEM). The FDM, the FEM and an additional method known as the *Boundary Element Method* (BEM) are the three basic approaches for discretizing differential equations. Since this is an introductory course, only the basic idea of the methods mentioned above will be presented and applied to some simple examples. The finite element and boundary element methods are nowadays the most often used computational tools in engineering. Students who wish to deepen their understanding of these methods are encouraged to take further courses offered in these subjects.

### Are there other techniques?

Of course, there are other techniques, as well analytical solution techniques, e.g., integral transform, as well as other approximation techniques, e.g., global Ritz and Galerkin methods. These methods are generally applied on simple domains (rectangular, circular, infinite, semi-infinite) having simple boundary conditions. However, an analytical solution on a simple domain can serve as a test case to verify a numerical technique.

### Is there any teaching aid besides this manuscript available?

To the best knowledge of the authors, there is no textbook available that completely covers the contents of this lecture. This manuscript presents an overview of discretization methods which is detailed enough so as to equip students with the necessary tools for engaging in more advanced treatments of specific discretization methods. For in-depth treatment of the FEM and the FDM, Bathe [2, 3] and Özisik [14] are recommended, respectively. Further literature for interested readers is given throughout the text whenever necessary or meaningful. The theory is illustrated using various numerical examples that can be downloaded as MATLAB<sup>TM</sup> files from the homepage [http://www.mecha.uni-stuttgart.de/Lehre/lehre\\_downl\\_discmeth\\_en.htm](http://www.mecha.uni-stuttgart.de/Lehre/lehre_downl_discmeth_en.htm). For those who have no access to the MATLAB<sup>TM</sup> software, we recommend installing the freeware program Octave which will also run the download files (see <http://www.gnu.org/software/octave/download.html>). A recommended textbook

for an introduction into the use of MATLAB<sup>TM</sup> and Octave in scientific computing is given in [16].

# Chapter 3

## The Finite Difference Method

The most direct way to solve a differential equation numerically is the finite difference method (FDM).

The basic idea is to replace the derivatives in the differential equation by finite difference approximations. The problem domain is *discretized* so that only values of the unknown function at discrete points, the *nodal points* are considered. The problem of solving a differential equation is therefore reduced to the problem of solving a set of algebraic equations.

### 3.1 Discrete approximation of derivatives

We recall the definition of the derivative of a function  $f(x)$  at  $x = x_0$

$$\frac{df}{dx} = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}. \quad (3.1)$$

For a sufficiently continuous function  $f$  the right-hand side of (3.1) is a reasonable approximation of the derivative, if a small but finite  $h$  is chosen.

#### 3.1.1 Approximation of the first derivative

A formal way to obtain finite difference approximations is the Taylor series expansion. Consider the Taylor series expansion of the function  $f(x)$  in the *forward* direction about a point  $x_0$ ,

$$f(x_0 + h) = f(x_0) + \left. \frac{df}{dx} \right|_{x=x_0} h + \left. \frac{d^2 f}{dx^2} \right|_{x=x_0} \frac{(h)^2}{2!} + \left. \frac{d^3 f}{dx^3} \right|_{x=x_0} \frac{(h)^3}{3!} + \dots \quad (3.2)$$

and also the Taylor series expansion of the function  $f(x)$  in the *backward* direction about a point  $x_0$ ,

$$f(x_0 - h) = f(x_0) - \left. \frac{df}{dx} \right|_{x=x_0} h + \left. \frac{d^2 f}{dx^2} \right|_{x=x_0} \frac{(h)^2}{2!} - \left. \frac{d^3 f}{dx^3} \right|_{x=x_0} \frac{(h)^3}{3!} + \dots \quad (3.3)$$

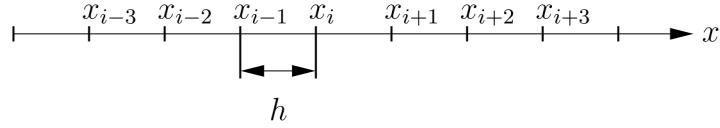


Figure 3.1: Notation of discretized variables

Rearranging these two expansions gives the *forward* and *backward* finite difference approximations of the first derivative

$$\frac{df}{dx}\Big|_{x=x_0} = \frac{f(x_0 + h) - f(x_0)}{h} + O(h) \quad (\text{forward}) , \quad (3.4)$$

$$\frac{df}{dx}\Big|_{x=x_0} = \frac{f(x_0) - f(x_0 - h)}{h} + O(h) \quad (\text{backward}) . \quad (3.5)$$

The notation  $O(h)$  characterizes the order of the *truncation error* in the finite difference approximation. While the forward and backward schemes are of first order in  $h$ , an higher order scheme can be obtained by subtracting the Taylor series expansions (3.2) and (3.3) to obtain

$$\frac{df}{dx}\Big|_{x=x_0} = \frac{f(x_0 + h) - f(x_0 - h)}{2h} + O(h^2) . \quad (3.6)$$

This scheme is called a *central difference formula*. Again only two points are involved, but now the truncation error is of second order in  $h$  resulting in a better approximation. Higher order formulae can be obtained by considering more points. Table 3.1 summarizes some common schemes. The finite difference formula are also called *stencils* and are sometimes represented in a graphical form as shown in Fig. 3.2. The involved nodal points are called the *molecule* of the formula.

The following notation is adopted: Let  $x_i$  be the point under consideration, then

$$x_{i-1} = x_i - h , \quad x_{i+1} = x_i + h , \quad x_{i+2} = x_i + 2h , \dots \quad (3.7)$$

as shown in Fig. 3.1, and also

$$f_{i-1} = f(x_{i-1}) , \quad f_{i+1} = f(x_{i+1}) , \quad f_{i+2} = f(x_{i+2}) , \dots \quad (3.8)$$

For the derivatives the following shorthand notation is introduced

$$\frac{\partial f}{\partial x}\Big|_{x_i} = f_{i,x} , \quad \frac{\partial^2 f}{\partial x^2}\Big|_{x_i} = f_{i,xx} . \quad (3.9)$$

### 3.1.2 Approximation of the second derivative

By the same technique of Taylor series expansion one can also obtain formulae for higher derivatives.

From adding (3.2) and (3.3) one obtains

$$f_{i,xx} = \frac{f_{i-1} - 2f_i + f_{i+1}}{h^2} + O(h^2) \quad (3.10)$$

which is a *central difference formula*. To obtain forward and backward finite difference approximations of the second derivative, the functions  $f(x_0 + 2h)$  and  $f(x_0 - 2h)$  have to be expanded into Taylor series. Again table 3.2 summarizes some common formula for the second derivative.

Two-point formulae	
forward	$f_{i,x} = \frac{f_{i+1} - f_i}{h} + O(h)$
backward	$f_{i,x} = \frac{f_i - f_{i-1}}{h} + O(h)$
central	$f_{i,x} = \frac{f_{i+1} - f_{i-1}}{2h} + O(h^2)$
Three-point formulae	
forward	$f_{i,x} = \frac{-3f_i + 4f_{i+1} - f_{i+2}}{2h} + O(h^2)$
backward	$f_{i,x} = \frac{f_{i-2} - 4f_{i-1} + 3f_i}{2h} + O(h^2)$
Four-point formulae	
forward	$f_{i,x} = \frac{-11f_i + 18f_{i+1} - 9f_{i+2} + 2f_{i+3}}{6h} + O(h^3)$
	$f_{i,x} = \frac{-2f_{i-1} - 3f_i + 6f_{i+1} - f_{i+2}}{6h} + O(h^3)$
	$f_{i,x} = \frac{f_{i-2} - 6f_{i-1} + 3f_i + 2f_{i+1}}{6h} + O(h^3)$

Table 3.1: Finite difference approximations of first derivative on an equally spaced grid

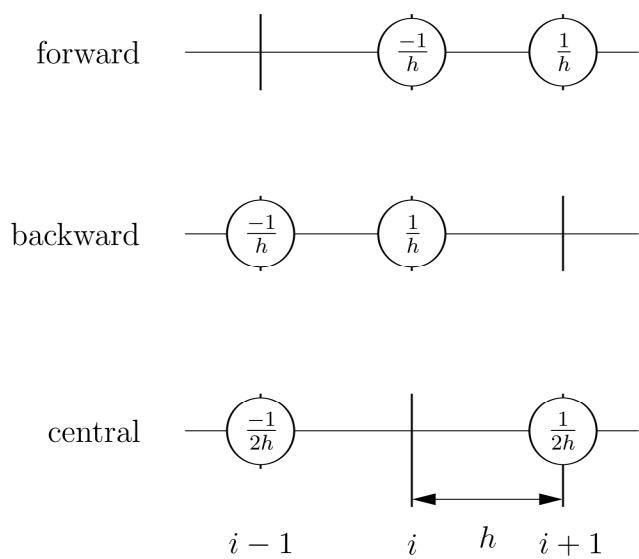


Figure 3.2: Two-point stencils

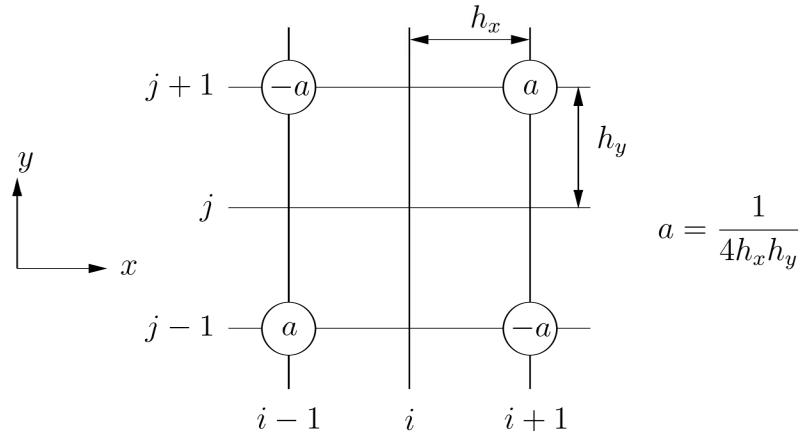


Figure 3.3: Four-point stencil for a central/central partial derivative

### 3.1.3 Approximation of mixed partial derivatives

For partial differential equations the approximation of mixed partial derivatives will be necessary. These can be obtained by applying the one-dimensional formulae successively.

Lets consider the mixed partial derivative  $\partial^2 f / \partial x \partial y$  and use central difference approximation in both directions. We have

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) = \frac{1}{2h_x} \left( \frac{\partial f}{\partial y} \Big|_{i+1,j} - \frac{\partial f}{\partial y} \Big|_{i-1,j} \right) + O(h_x^2) \quad (3.11)$$

where the subscripts  $i$  and  $j$  refer to the discretization in the  $x$  and  $y$  direction, respectively.

Applying the central difference formula once more to discretize the partial derivatives with respect to  $y$  we obtain

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{1}{2h_x} \left( \frac{f_{i+1,j+1} - f_{i+1,j-1}}{2h_y} - \frac{f_{i-1,j+1} - f_{i-1,j-1}}{2h_y} \right) + O(h_x^2, h_y^2) \quad (3.12)$$

$$= \frac{1}{4h_xh_y} (f_{i+1,j+1} - f_{i+1,j-1} - f_{i-1,j+1} + f_{i-1,j-1}) + O(h_x^2, h_y^2). \quad (3.13)$$

Figure 3.3 illustrates the points and factors involved. The order of the approximations is of no importance as long as

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}, \quad (3.14)$$

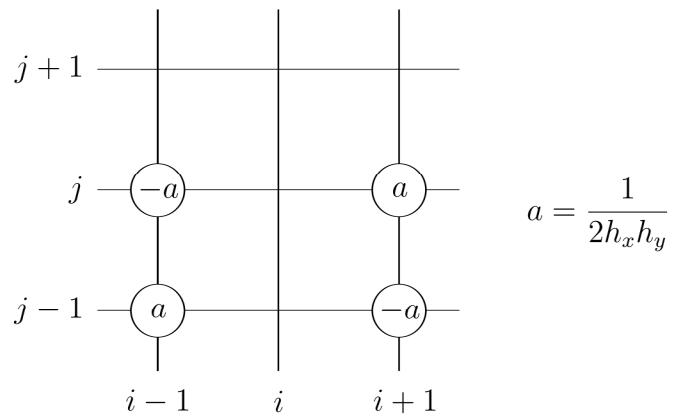
which is assured if the derivatives are continuous.

Of course any combination of difference formulae can be used for approximation of a mixed partial derivative.

If we restrict ourselves to two-point formulae, which are the forward (3.4), the backward (3.5) and the central difference formula (3.6), a total of nine combinations is possible, which are given in table 3.3.

#### Example 3.1

The stencil of the mixed central/backward formula is given by:



Three-point formulae	
forward	$f_{i,xx} = \frac{f_i - 2f_{i+1} + f_{i+2}}{h^2} + O(h)$
backward	$f_{i,xx} = \frac{f_{i-2} - 2f_{i-1} + f_i}{h^2} + O(h)$
central	$f_{i,xx} = \frac{f_{i-1} - 2f_i + f_{i+1}}{h^2} + O(h^2)$
Four-point formulae	
forward	$f_{i,xx} = \frac{2f_i - 5f_{i+1} + 4f_{i+2} - f_{i+3}}{h^2} + O(h^2)$
backward	$f_{i,xx} = \frac{-f_{i-3} + 4f_{i-2} - 5f_{i-1} + 2f_i}{h^2} + O(h^2)$

Table 3.2: Finite difference approximations of second derivative on an equally spaced grid

x	y	finite difference formula	order of error
F F		$\frac{1}{h_x} \left( \frac{f_{i+1,j+1} - f_{i+1,j}}{h_y} - \frac{f_{i,j+1} - f_{i,j}}{h_y} \right)$	$O(h_x, h_y)$
	B	$\frac{1}{h_x} \left( \frac{f_{i+1,j} - f_{i+1,j-1}}{h_y} - \frac{f_{i,j} - f_{i,j-1}}{h_y} \right)$	$O(h_x, h_y)$
	C	$\frac{1}{h_x} \left( \frac{f_{i+1,j+1} - f_{i+1,j-1}}{2h_y} - \frac{f_{i,j+1} - f_{i,j-1}}{2h_y} \right)$	$O(h_x, h_y^2)$
B F		$\frac{1}{h_x} \left( \frac{f_{i,j+1} - f_{i,j}}{h_y} - \frac{f_{i-1,j+1} - f_{i-1,j}}{h_y} \right)$	$O(h_x, h_y)$
	B	$\frac{1}{h_x} \left( \frac{f_{i,j} - f_{i,j-1}}{h_y} - \frac{f_{i-1,j} - f_{i-1,j-1}}{h_y} \right)$	$O(h_x, h_y)$
	C	$\frac{1}{h_x} \left( \frac{f_{i,j+1} - f_{i,j-1}}{2h_y} - \frac{f_{i-1,j+1} - f_{i-1,j-1}}{2h_y} \right)$	$O(h_x, h_y^2)$
C F		$\frac{1}{2h_x} \left( \frac{f_{i+1,j+1} - f_{i+1,j}}{h_y} - \frac{f_{i-1,j+1} - f_{i-1,j}}{h_y} \right)$	$O(h_x^2, h_y)$
	B	$\frac{1}{2h_x} \left( \frac{f_{i+1,j} - f_{i+1,j-1}}{h_y} - \frac{f_{i-1,j} - f_{i-1,j-1}}{h_y} \right)$	$O(h_x^2, h_y)$
	C	$\frac{1}{2h_x} \left( \frac{f_{i+1,j+1} - f_{i+1,j-1}}{2h_y} - \frac{f_{i-1,j+1} - f_{i-1,j-1}}{2h_y} \right)$	$O(h_x^2, h_y^2)$

F=forward difference, B=backward difference, C=central difference

Table 3.3: Finite difference approximations of mixed derivative  $f_{i,xy}$  on a regular grid

## 3.2 Simple boundary value problems

Here we will consider simple boundary value problems based on ordinary differential equations.

### 3.2.1 One-dimensional steady-state heat flow

Lets consider the 1-D steady-state heat flow problem governed by

$$\kappa \Theta(x)_{,xx} = -\varrho r(x) \quad (3.15)$$

on a domain  $0 < x < \ell$ .

To solve this problem, we discretize this domain using  $n$  nodes with a regular spacing  $h$ , as shown in Fig. 3.4. Obviously the spacing is connected to the number of nodes by

$$h = \frac{\ell}{n-1}. \quad (3.16)$$

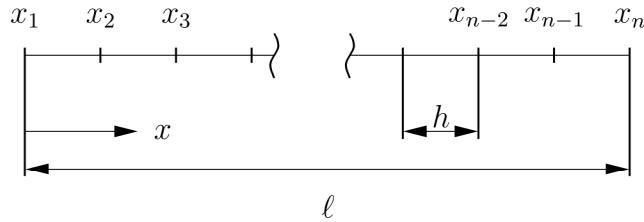


Figure 3.4: Discretization of one dimensional steady-state heat flow problem

Using the central difference formula (3.10) we can now approximate the differential equation as

$$\kappa \frac{\Theta_{i-1} - 2\Theta_i + \Theta_{i+1}}{h^2} = -\varrho r_i. \quad (3.17)$$

This approximation holds on all interior points  $i = 2, 3, \dots, n-1$ . It can be written in matrix form as

$$\begin{bmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} \Theta_1 \\ \Theta_2 \\ \Theta_3 \\ \vdots \\ \Theta_{n-2} \\ \Theta_{n-1} \\ \Theta_n \end{bmatrix} = -\frac{h^2}{\kappa} \begin{bmatrix} \varrho r_2 \\ \varrho r_3 \\ \vdots \\ \varrho r_{n-2} \\ \varrho r_{n-1} \end{bmatrix}. \quad (3.18)$$

This system provides us with  $n-2$  equations for the  $n$  unknown nodal temperatures. So two more equations are necessary, which are given by the boundary conditions.

### 3.2.2 Dirichlet boundary conditions

Lets first have a look at the Dirichlet problem with boundary conditions

$$\Theta(x = 0) = \bar{\Theta}_1 \quad \text{and} \quad \Theta(x = \ell) = \bar{\Theta}_n . \quad (3.19)$$

Now the temperatures at the boundary nodes are given by the boundary conditions and are therefore known.

Rearranging the system (3.18) by applying the boundary conditions results in a square symmetric tridiagonal system matrix,

$$\begin{bmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & \ddots & \ddots & \ddots \\ & & 1 & -2 & 1 \\ & & & 1 & -2 \end{bmatrix} \begin{bmatrix} \Theta_2 \\ \Theta_3 \\ \vdots \\ \Theta_{n-2} \\ \Theta_{n-1} \end{bmatrix} = -\frac{h^2}{\kappa} \begin{bmatrix} \varrho r_2 \\ \varrho r_3 \\ \vdots \\ \varrho r_{n-2} \\ \varrho r_{n-1} \end{bmatrix} - \begin{bmatrix} \bar{\Theta}_1 \\ 0 \\ \vdots \\ 0 \\ \bar{\Theta}_n \end{bmatrix} . \quad (3.20)$$

### 3.2.3 Neumann boundary conditions

While the treatment of Dirichlet conditions poses non problems, the introduction of Neumann conditions requires some additional considerations. In the following we will still assume a Dirichlet condition at  $x = \ell$ , but we now consider a Neumann condition which prescribes the heat flux at  $x = 0$

$$q(x = 0) = \bar{q} . \quad (3.21)$$

Recalling that the heat flux is governed by Fourier's law  $q = -\kappa \Theta_{,x}$ , we can write the Neumann boundary condition as

$$\Theta(0)_{,x} = -\frac{\bar{q}}{\kappa} , \quad (3.22)$$

which means that actually the first derivative of the unknown function  $\Theta$  is prescribed at  $x = 0$ .

An obvious way to introduce this condition in the finite difference scheme is to recast it in a finite difference form. The easiest way would be to use the forward difference formula (3.4) to obtain

$$-\Theta_1 + \Theta_2 = -\frac{h}{\kappa} \bar{q} . \quad (3.23)$$

This equation provides the additional equation which is needed to determine the now  $n - 1$  unknown temperatures  $\Theta_1, \dots, \Theta_{n-1}$ . Together with the  $n - 2$  equations (3.18) the complete set of equations can be written as

$$\begin{bmatrix} -1 & 1 & & \\ 1 & -2 & 1 & \\ & 1 & -2 & 1 \\ & & \ddots & \ddots & \ddots \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{bmatrix} \begin{bmatrix} \Theta_1 \\ \Theta_2 \\ \Theta_3 \\ \vdots \\ \Theta_{n-2} \\ \Theta_{n-1} \end{bmatrix} = -\frac{h^2}{\kappa} \begin{bmatrix} 0 \\ \varrho r_2 \\ \varrho r_3 \\ \vdots \\ \varrho r_{n-2} \\ \varrho r_{n-1} \end{bmatrix} - \frac{h}{\kappa} \begin{bmatrix} \bar{q} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \bar{\Theta}_n \end{bmatrix} . \quad (3.24)$$

This is a working scheme, however not a very good one. If you recall the truncation errors of the different formulae, you see that the central difference formula used for the approximation of the differential equation has an error of order  $O(h^2)$ . The simple forward formula used for the approximation of the Neumann boundary condition is just of the order  $O(h)$ . This degrades the whole system to an error order of  $O(h)$ . So whatever efforts you make, the approximation error is only as good as the lowest order of any approximation you have made. The second problem is that a possible heat source at  $x = 0$  is not accounted for.

There are two ways to improve the solution. One is to use a higher order forward formula, like the three-point forward formula of table 3.1. With this scheme the boundary condition becomes

$$-3\Theta_1 + 4\Theta_2 - \Theta_3 = -\frac{2h}{\kappa}\bar{q}. \quad (3.25)$$

The complete system then looks like

$$\begin{bmatrix} -3 & 4 & -1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \\ \dots & \dots & \dots \\ 1 & -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} \Theta_1 \\ \Theta_2 \\ \Theta_3 \\ \vdots \\ \Theta_{n-2} \\ \Theta_{n-1} \end{bmatrix} = -\frac{h^2}{\kappa} \begin{bmatrix} 0 \\ \varrho r_2 \\ \varrho r_3 \\ \vdots \\ \varrho r_{n-2} \\ \varrho r_{n-1} \end{bmatrix} - \frac{2h}{\kappa} \begin{bmatrix} \bar{q} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \bar{\Theta}_n \end{bmatrix}. \quad (3.26)$$

This is a set of equations with an overall approximation error of order  $O(h^2)$ . Unfortunately the desirable property of a tridiagonal system matrix is lost and still no heat source at  $x = 0$  is taken into consideration.

An alternative way to produce a set of equations with an error of order  $O(h^2)$  is the introduction of *fictitious nodes* or *ghost points* outside the real domain, as shown in Fig. 3.5

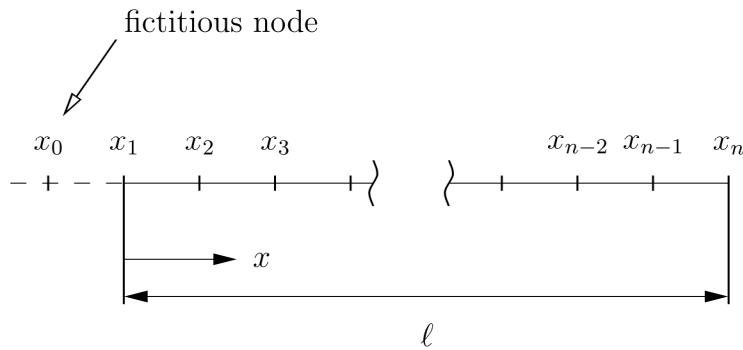


Figure 3.5: Use of fictitious node for Neumann boundary conditions

If we introduce an additional node on the left side, we could approximate the Neumann condition using the central difference formula

$$-\Theta_0 + \Theta_2 = -\frac{2h}{\kappa}\bar{q}. \quad (3.27)$$

So far nothing has been gained, we have an additional equation but we have also introduced an additional unknown  $\Theta_0$ . The solution is to treat the point  $x_1$  as an interior point.

The difference equation at  $x = 0$  is

$$\Theta_0 - 2\Theta_1 + \Theta_2 = -\frac{h^2}{\kappa} \varrho r_1 . \quad (3.28)$$

We can now eliminate  $\Theta_0$  using (3.27) to obtain

$$-2\Theta_1 + 2\Theta_2 = -\frac{h^2}{\kappa} \varrho r_1 - \frac{2h}{\kappa} \bar{q} . \quad (3.29)$$

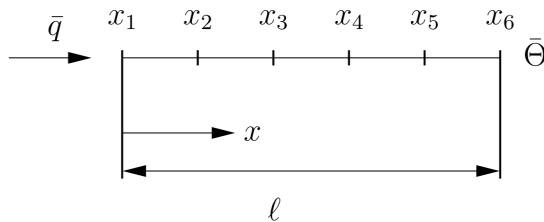
The complete system now has the tridiagonal, however non-symmetric form

$$\begin{bmatrix} -2 & 2 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{bmatrix} \begin{bmatrix} \Theta_1 \\ \Theta_2 \\ \Theta_3 \\ \vdots \\ \Theta_{n-2} \\ \Theta_{n-1} \end{bmatrix} = -\frac{h^2}{\kappa} \begin{bmatrix} \varrho r_1 \\ \varrho r_2 \\ \varrho r_3 \\ \vdots \\ \varrho r_{n-2} \\ \varrho r_{n-1} \end{bmatrix} - \frac{2h}{\kappa} \begin{bmatrix} \bar{q} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \bar{\Theta}_n \end{bmatrix} . \quad (3.30)$$

This is a system with an error order of  $O(h^2)$  and a possible heat source at  $x = 0$  is present. However if you compare this result carefully with the system of equations (3.24) you will notice that they are the same for a vanishing source  $r_1 = 0$ . For this special problem a first order approximation of the Neumann condition does not degrade the system.

### Example 3.2

We are now able to solve the problem posed in example 2.2, using the formulation (3.30). A more abstract view of this problem is presented below, already introducing a 6 point discretization.



With vanishing source term the system becomes

$$\begin{bmatrix} -2 & 2 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & 1 & -2 & 1 & \\ & & & 1 & -2 & \end{bmatrix} \begin{bmatrix} \Theta_1 \\ \Theta_2 \\ \Theta_3 \\ \Theta_4 \\ \Theta_5 \end{bmatrix} = -\frac{2h}{\kappa} \begin{bmatrix} \bar{q} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \bar{\Theta} \end{bmatrix} . \quad (3.31)$$

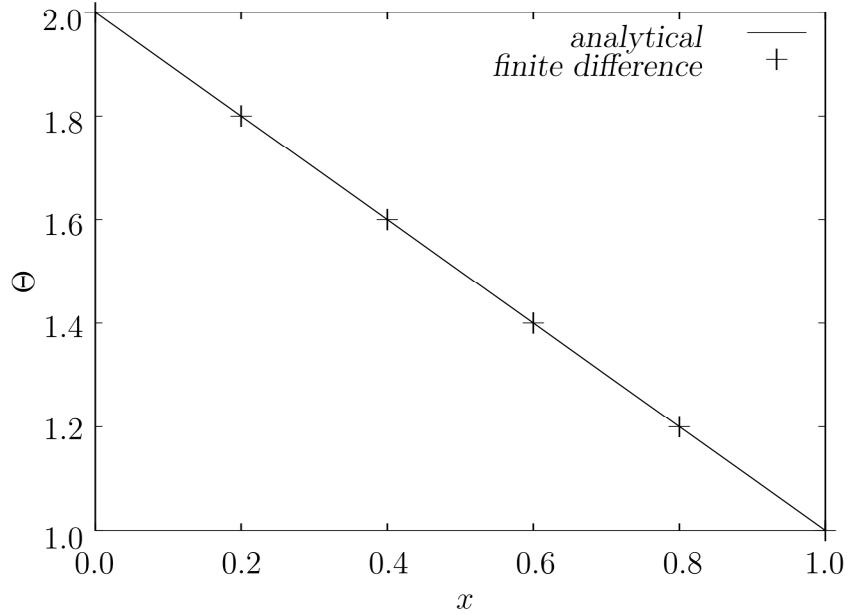
The analytical solution of this problem is very easy to obtain. Since no sources are present the temperature distribution has to be linear,

$$\kappa \Theta_{,xx} = 0 \quad \longrightarrow \quad \Theta(x) = C_1 x + C_2 . \quad (3.32)$$

Applying the boundary conditions results in

$$\Theta(x) = \bar{\Theta} + \frac{\bar{q}}{\kappa}(\ell - x) . \quad (3.33)$$

The following figure shows the analytical solution in comparison to the finite difference approximation for  $\ell = 1.0$ ,  $\kappa = 1.0$ ,  $\bar{q} = 1.0$  and  $\bar{\Theta} = 1.0$ . Obviously the finite difference approximation captures the exact solution.



### Example 3.3

We consider a heat conducting rod with Dirichlet conditions at both ends

$$\Theta(0) = \Theta(\ell) = 0 , \quad (3.34)$$

and a sinusoidal source distribution

$$\varrho r(x) = \sin\left(\frac{\pi x}{\ell}\right) . \quad (3.35)$$

The analytical solution to this problem is

$$\kappa\Theta_{,xx} = -\sin\left(\frac{\pi x}{\ell}\right) \quad \longrightarrow \quad \Theta(x) = \frac{\ell^2}{\pi^2\kappa} \sin\left(\frac{\pi x}{\ell}\right) + C_1x + C_2 . \quad (3.36)$$

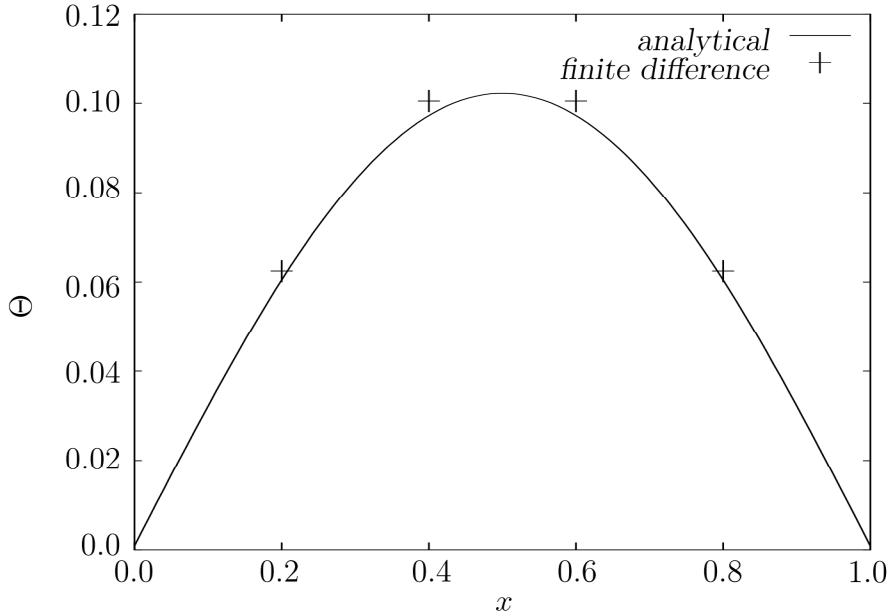
Applying the boundary conditions leads to

$$\Theta(x) = \frac{\ell^2}{\pi^2\kappa} \sin\left(\frac{\pi x}{\ell}\right) . \quad (3.37)$$

The equivalent finite difference system is

$$\begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & 1 & -2 & \end{bmatrix} \begin{bmatrix} \Theta_2 \\ \Theta_3 \\ \Theta_4 \\ \Theta_5 \end{bmatrix} = -\frac{h^2}{\kappa} \begin{bmatrix} \varrho r_2 \\ \varrho r_3 \\ \varrho r_4 \\ \varrho r_5 \end{bmatrix} . \quad (3.38)$$

Even with this rather coarse mesh of only 4 interior points the exact solution is quite well reproduced by the numerical method, as can be seen in the figure below.



### 3.2.4 Diffusive-convective system

Let us consider the one-dimensional steady-state heat flow through a fin, where heat is transported away from the surface of the fin by convection. This is an example of a diffusive-convective system. The diffusion part is the heat conduction within the fin, while the convective part is given by the convective heat transport at the surface. Typical applications include heat radiators, electronic equipment cooling, electrical transformers, automobile radiators etc.

The governing differential equation can be obtained by considering the balance equation of energy for a heat conductor without internal sources

$$\int_V \varrho c_v \dot{\Theta} dv = - \int_A \vec{q} \cdot \vec{n} da , \quad (3.39)$$

where  $(\cdot)$  denotes total differentiation with respect to time

$$\dot{\Theta} = \frac{d\Theta}{dt} . \quad (3.40)$$

We restrict ourselves to a one-dimensional problem with varying cross-section and write down the energy balance for a small section of length  $dx$ , as shown in Fig. 3.6. The surface can be split into the cross sections to the left and right,  $A^-$  and  $A^+$ , and the lateral surface  $dA_l$ . Accordingly the heat flow through the cross-sections is given by

$$q^- = \vec{q} \cdot \vec{n}^- = \kappa \Theta_{,x}^- \quad (3.41)$$

$$q^+ = \vec{q} \cdot \vec{n}^+ = -\kappa \Theta_{,x}^+ . \quad (3.42)$$

The heat flow through the lateral area is given by

$$q_l = \vec{q} \cdot \vec{n}_l = h(\Theta - \Theta_\infty) , \quad (3.43)$$

with the *convection coefficient*  $h$  and the *ambient temperature*  $\Theta_\infty$ . This results in

$$\varrho c_v \dot{\Theta} Adx = \kappa(A^+ \Theta_{,x}^+ - A^- \Theta_{,x}^-) - h(\Theta - \Theta_\infty) dA_l , \quad (3.44)$$

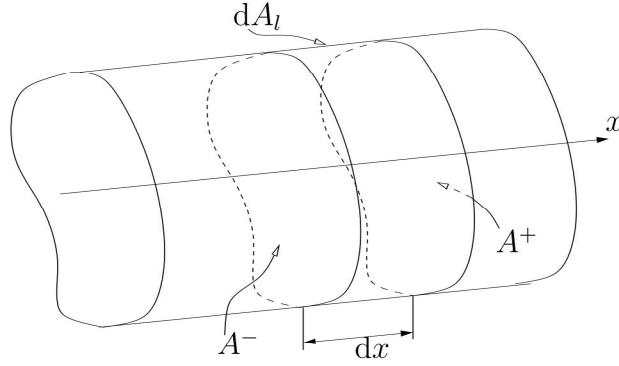


Figure 3.6: Infinitesimal section of one-dimensional heat conductor

which can be recast to

$$\varrho c_v A \dot{\Theta} = \kappa \frac{d}{dx} (A \Theta_{,x}) - h \frac{dA_l}{dx} (\Theta - \Theta_\infty) . \quad (3.45)$$

Since we are dealing with a steady-state rigid body problem, we have

$$\dot{\Theta} = 0 , \quad (3.46)$$

which simplifies equation (3.45) to

$$0 = \frac{d}{dx} (A \Theta_{,x}) - \frac{h}{\kappa} \frac{dA_l}{dx} (\Theta - \Theta_\infty) . \quad (3.47)$$

For a fin with constant cross-section  $A$  this can be further reduced to

$$0 = \Theta_{,xx} - \frac{h}{A\kappa} \frac{dA_l}{dx} (\Theta - \Theta_\infty) . \quad (3.48)$$

If we consider for example a fin with a circular cross-section with radius  $R$ , the cross-section area and the lateral area are given by

$$A = \pi R^2 \quad \text{and} \quad dA_l = 2\pi R dx . \quad (3.49)$$

The governing differential equation is then

$$\Theta_{,xx} - m^2 (\Theta - \Theta_\infty) = 0 , \quad (3.50)$$

with

$$m^2 = \frac{2h}{R\kappa} . \quad (3.51)$$

To simplify this further we change variables by

$$\theta = \Theta - \Theta_\infty \quad (3.52)$$

to obtain finally

$$\theta_{,xx} - m^2 \theta = 0 . \quad (3.53)$$

If the fin base is kept at a constant temperature  $\bar{\Theta}$  at the base  $x = 0$  and the heat loss at the tip cross-section  $x = \ell$  is negligible compared to the lateral convection loss, the boundary conditions are given by

$$\theta(0) = \bar{\theta} = \bar{\Theta} - \Theta_\infty \quad (3.54)$$

$$\theta(\ell),_x = 0 . \quad (3.55)$$

The exact analytical solution for this problem is (ÖZİŞİK [14])

$$\frac{\theta(x)}{\bar{\theta}} = \frac{\cosh[m(\ell - x)]}{\cosh(m\ell)} . \quad (3.56)$$

From this solution we get for example the total heat flow through the fin as

$$qA = -\kappa A \theta_x|_{x=0} = \kappa A \bar{\theta} m \tanh(m\ell) . \quad (3.57)$$

We now use the central difference approximation (3.10) to obtain the second order accurate finite difference approximation

$$\theta_{i-1} - (2 + m^2 \Delta x^2) \theta_i + \theta_{i+1} = 0 \quad (3.58)$$

for the interior points  $i = 2, \dots, n - 1$ . At the fin tip a ghost point has to be introduced to obtain a second order accurate approximation. Using

$$\theta_{n+1} - \theta_{n-1} = 0 \quad (3.59)$$

and

$$\theta_{n-1} - (2 + m^2 \Delta x^2) \theta_n + \theta_{n+1} = 0 \quad (3.60)$$

results in

$$2\theta_{n-1} - (2 + m^2 \Delta x^2) \theta_n = 0 . \quad (3.61)$$

The heat flow can be discretized either by the first order accurate forward difference scheme (3.4) as

$$q = -\frac{\kappa}{\Delta x} (\theta_2 - \theta_1) , \quad (3.62)$$

or by the second order accurate central difference formula (3.6) using a ghost point to the left

$$q = -\kappa \frac{\theta_2 - \theta_0}{2\Delta x} \quad (3.63)$$

and

$$\theta_0 - (2 + m^2 \Delta x^2) \theta_1 + \theta_2 = 0 . \quad (3.64)$$

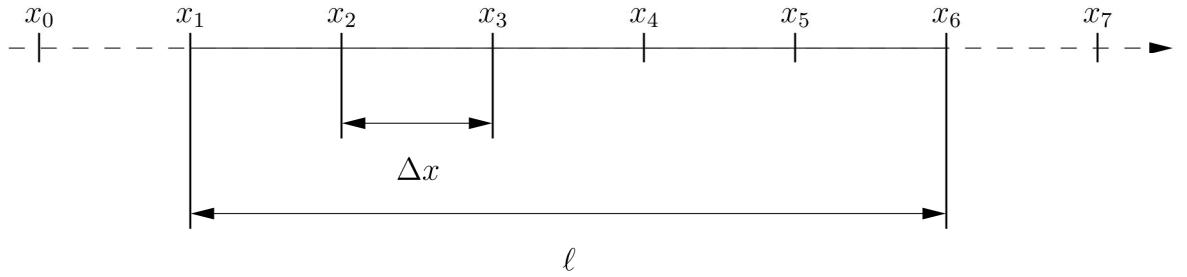
Eliminating  $\theta_0$  results in

$$q = -\frac{\kappa}{\Delta x} (\theta_2 - \theta_1) + \frac{\kappa m^2 \Delta x}{2} \bar{\theta} . \quad (3.65)$$

### Example 3.4

To obtain some numerical values, we consider a steel rod of length  $\ell = 0.05$  m and radius  $R = 0.01$  m. This rod is protruding from a wall into an ambient temperature of  $\Theta_\infty = 30^\circ$  C. The base of the rod is kept at a constant temperature of  $\bar{\Theta} = 330^\circ$  C. The thermal conductivity of the rod is  $\kappa = 50$  W/mK and the convection coefficient is  $h = 100$  W/m<sup>2</sup>K.

We discretize the problem with 5 equal sections and 6 nodes with two additional ghost points  $x_0$  and  $x_7$ .



With

$$\Delta x = 0.01 \text{ m} , \quad m^2 = 400 \text{ } 1/\text{m}^2 , \quad m^2 \Delta x^2 = 0.04 , \quad m^2 \Delta x^2 \Theta_\infty = 1.2 \text{ K} \quad (3.66)$$

we obtain the system of equations

$$\begin{bmatrix} -2.04 & 1.00 & 0.00 & 0.00 & 0.00 \\ 1.00 & -2.04 & 1.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & -2.04 & 1.00 & 0.00 \\ 0.00 & 0.00 & 1.00 & -2.04 & 1.00 \\ 0.00 & 0.00 & 0.00 & 2.00 & -2.04 \end{bmatrix} \begin{bmatrix} \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \\ \theta_6 \end{bmatrix} = \begin{bmatrix} -300.0 \\ 0.0 \\ 0.0 \\ 0.0 \\ 0.0 \end{bmatrix} \quad (3.67)$$

The results are in excellent agreement with the analytical solution:

$x/\ell$	$\theta_i$	Exact	FDM
0.2	$\theta_2$	260.0	260.1
0.4	$\theta_3$	230.5	230.6
0.6	$\theta_4$	210.2	210.4
0.8	$\theta_5$	198.3	198.6
1.0	$\theta_6$	194.4	194.7

The total heat flow through the fin is found to be:

Exact	Forward	Central
71.8	62.6	72.1

The result for the central difference approximation compares favorably with the analytical solution, while the neglect of the convective term in the forward difference approximation leads to a significant error.

### 3.2.5 Diffusive-convective system with flow

The problem presented in the last section can be expanded to a diffusive-convective system with flow.

Instead of a rod radiating into an ambient, the heat conductor can be considered as a fluid filled duct, where the fluid can transfer heat by convection to the duct wall. If the fluid is not moving, this is exactly the same problem as the fin problem examined before. Again the governing differential equation is (3.48)

$$\Theta_{,xx} - m^2(\Theta - \Theta_W) = 0 , \quad (3.68)$$

with the wall temperature  $\Theta_W$  now replacing the ambient temperature  $\Theta_\infty$ .

However, if we assume a steady-state flow of the fluid with velocity  $v$  through the duct, the energy balance has to be modified to include the energy transport by the flowing fluid. Starting with (3.45)

$$\varrho c_v A \dot{\Theta} = \kappa \frac{d}{dx} (A \Theta_{,x}) - h \frac{dA_l}{dx} (\Theta - \Theta_W) , \quad (3.69)$$

we have now to consider the total time derivative  $\dot{\Theta}$  by

$$\dot{\Theta} = \frac{d\Theta}{dt} = \Theta_{,x} v + \Theta_{,t} . \quad (3.70)$$

Since we still have a steady-state system  $\Theta_{,t}$  vanishes. But the energy equation now becomes

$$\varrho c_v A \Theta_{,x} v = \kappa \frac{d}{dx} (A \Theta_{,x}) - h \frac{dA_l}{dx} (\Theta - \Theta_W) . \quad (3.71)$$

If we again assume a constant circular cross-section, rearrangement leads to

$$\Theta_{,xx} - \frac{v}{\alpha} \Theta_{,x} - m^2 (\Theta - \Theta_W) = 0 , \quad (3.72)$$

with the *diffusivity*  $\alpha$  of the fluid given by

$$\alpha = \frac{\kappa}{\varrho c_v} . \quad (3.73)$$

It is clear that Eq. (3.72) reduces to (3.68) for vanishing flow velocity  $v$ . Again change of variables with  $\theta = \Theta - \Theta_W$  simplifies (3.72) to

$$\theta_{,xx} - \frac{v}{\alpha} \theta_{,x} - m^2 \theta = 0 . \quad (3.74)$$

While it may look attractive to use the central difference formula (3.6) to discretize the first derivative to obtain a fully second order accurate scheme, it will be shown that this scheme gives rise to numerical instabilities under certain conditions. Therefore an alternative scheme called an *upwind-difference* scheme, will be presented, which is unconditionally stable.

If we use the central difference scheme for both derivatives, we obtain

$$\frac{\theta_{i+1} - 2\theta_i + \theta_{i-1}}{\Delta x^2} - \frac{v}{\alpha} \frac{\theta_{i+1} - \theta_{i-1}}{2\Delta x} - m^2 \theta_i = 0 , \quad (3.75)$$

which can be rearranged to

$$\left(1 + \frac{v\Delta x}{2\alpha}\right) \theta_{i-1} - (2 + m^2 \Delta x^2) \theta_i + \left(1 - \frac{v\Delta x}{2\alpha}\right) \theta_{i+1} = 0 . \quad (3.76)$$

Introducing the *Peclet number*  $Pe$  by

$$Pe = \frac{v\Delta x}{\alpha} \quad (3.77)$$

one obtains

$$\left(1 + \frac{Pe}{2}\right) \theta_{i-1} - (2 + m^2 \Delta x^2) \theta_i + \left(1 - \frac{Pe}{2}\right) \theta_{i+1} = 0 . \quad (3.78)$$

If we use the backward difference scheme for the first derivative, we obtain

$$\frac{\theta_{i+1} - 2\theta_i + \theta_{i-1}}{\Delta x^2} - \frac{v}{\alpha} \frac{\theta_i - \theta_{i-1}}{\Delta x} - m^2 \theta_i = 0, \quad (3.79)$$

which can be rearranged to

$$(1 + \text{Pe}) \theta_{i-1} - (2 + m^2 \Delta x^2 + \text{Pe}) \theta_i + \theta_{i+1} = 0. \quad (3.80)$$

Both formulations are of the form

$$A\theta_{i-1} + B\theta_i + C\theta_{i+1} = 0. \quad (3.81)$$

The stability of such formulations can be assessed by an argument given by SHIH [23]:

### Stability criterion

Let us arrange terms in (3.81) such that  $B < 0$ . Physical meaningful solutions will result only if  $A > 0$  and  $C > 0$ .

For a flow in positive  $x$ -direction the Peclet number is positive. Then, for the central difference formulation (3.78) we always have  $A > 0$ . But  $C > 0$  will only hold, if

$$\text{Pe} < 2 \quad \longrightarrow \quad \Delta x < \frac{2\alpha}{v}. \quad (3.82)$$

The central difference scheme is stable only for certain step sizes which relate to the flow velocity. Such a scheme is called *conditionally stable*.

For the backward difference formulation (3.80) the requirements  $A > 0$  and  $C > 0$  for the stability are always met. The scheme is therefore *unconditionally stable*. However it is less accurate than the central difference scheme. Since the difference formula for the first derivative was in the upstream or upwind direction, this scheme is called an *upwind difference scheme*. Notice that a forward difference formula would have been necessary if the flow had been in the negative  $x$ -direction.

### Example 3.5

We consider the above flow problem in a duct with length  $\ell = 0.5$ . The problem is discretized by 51 nodes, giving an equal spacing of  $\Delta x = 0.01$ . The parameters are  $m^2 = 6 \cdot 10^4$  and  $v/\alpha = 2 \cdot 10^4$ , which give a Peclet number of  $\text{Pe} = 200$ .

As boundary conditions the temperature is prescribed at both ends to  $\bar{\theta}(0) = 100$  and  $\bar{\theta}(\ell) = 0$ , respectively.

The analytical solution of the governing differential equation (3.74) is given by

$$\theta(x) = C_1 \exp[\gamma_1 x] + C_2 \exp[\gamma_2 x], \quad (3.83)$$

with

$$\gamma_{1/2} = \frac{v}{2\alpha} \pm \sqrt{m^2 + \left(\frac{v}{2\alpha}\right)^2}. \quad (3.84)$$

For the given parameters we find

$$\gamma_1 = 20003 \quad \text{and} \quad \gamma_2 = -3. \quad (3.85)$$

Using the boundary conditions to determine the unknown constants  $C_1$  and  $C_2$  gives

$$C_2 = \frac{\bar{\theta}(\ell) - \bar{\theta}(0) \exp [\gamma_1 \ell]}{\exp [\gamma_2 \ell] - \exp [\gamma_1 \ell]} \quad (3.86)$$

$$C_1 = \bar{\theta}(0) - C_2 . \quad (3.87)$$

For the given numerical values this can be safely set to

$$C_2 \approx \bar{\theta}(0) \quad \text{and} \quad C_1 \approx 0 . \quad (3.88)$$

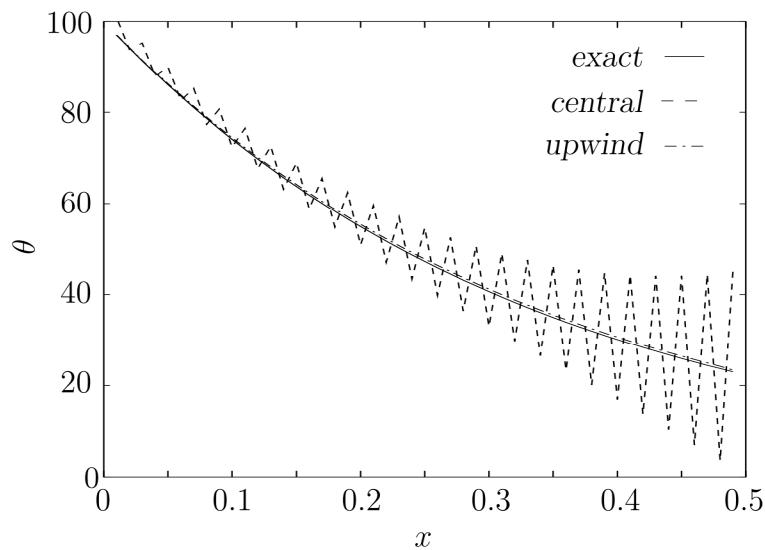
The analytical solution for the given values is therefore

$$\theta(x) = \bar{\theta}(0) \exp [\gamma_2 x] . \quad (3.89)$$

The exact solution is very good captured by the upwind scheme, but the central difference scheme gives a wildly oscillating solution which is completely useless, since here

$$2 \frac{\alpha}{v} = 10^{-4} \quad \text{and} \quad \Delta x = 0.01 \quad (3.90)$$

and therefore the stability requirement (3.82) is not satisfied.



# Chapter 4

## The Finite Element Method

### 4.1 Weighted residuals

A totally different approach to solving a differential equation is the method of weighted residuals, which forms a starting point for the finite element method as well as the boundary element method.

The basic idea is to use a suitable trial function as an inspired guess of the solution of the differential equation and then tune this trial function to obtain the best approximation.

#### 4.1.1 Residuum and weak form

Lets assume a boundary value problem with a differential equation

$$F(u) = 0 \quad (4.1)$$

and given boundary conditions. We now choose a *trial function*  $\tilde{u}$  given by a set of *shape functions*  $\tilde{u}_i$  and some free parameters  $c_i$ . A typical trial function may have the linear form

$$\tilde{u} = \sum_{i=1}^N c_i \tilde{u}_i , \quad (4.2)$$

but other forms are possible. In general this trial function will neither fulfill the differential equation nor the boundary conditions. For example, inserting this trial function in the differential equation will lead to an error or a *residual*  $R$ ,

$$F(\tilde{u}) = R(c_i) \neq 0 . \quad (4.3)$$

The same holds for the boundary conditions.

Obviously the residual depends on the free parameters of the trial function. We now seek parameters which will minimize the residual in a certain sense. In the best case we will find parameters for which the residual vanishes completely and everywhere, meaning we have found an exact solution of the differential equation.

Since Eq. (4.3) provides us with only one equation for  $N$  unknown parameters  $c_i$ , we have to generate some more equations. This is achieved by the *method of weighted residuals*.

DE satisfied	BC satisfied	resulting method
no	yes	interior method
yes	no	boundary method
no	no	mixed method

Table 4.1: Choice of trial functions

We introduce a set of  $N$  distinct *weight functions*  $w_j$ ,  $j = 1, \dots, N$  and require

$$\int_{\Omega} w_j R \, dv = 0 . \quad (4.4)$$

Equation (4.4) is called a *weak form* of the differential equation: Instead of requiring the residuum  $R$  to vanish everywhere, we now only require the integral of the residuum weighted with some function to vanish. Since this is a much weaker requirement as the exact satisfaction of the differential equation, one says the differential equation is satisfied only in a *weak sense*.

The crucial point is now to choose appropriate trial and weight functions.

### 4.1.2 Trial functions

Three basic alternatives for the choice of trial functions are possible:

1. The first one is to choose trial functions which satisfy the boundary conditions but not the differential equation. This is usually relatively easy and will lead to an *interior method*, since we now only seek parameters to satisfy the differential equation in the weak sense in the interior of the domain.
2. The second alternative is to choose trial functions which satisfy the differential equation and then seek parameters for which the solution satisfies the boundary conditions, again in a weak sense. Such a method is called a *boundary method* and is usually harder to obtain since at least some solutions to the differential equation have to be known.
3. The third way is to choose an arbitrary set of trial functions which do neither satisfy the differential equation nor the boundary conditions, thus leading to a *mixed method*. This is very easy, but the results will be usually worse than for the other methods, where some information has already been put into the trial functions.

The methods are listed in Tab. 4.1.

### 4.1.3 Weight functions

A lot of different choices of weight functions have been used over the time. We will discuss here the more popular ones which have been proven to give good results.

### Method of subdomains

In the method of subdomains we split the domain in  $N$  subdomains  $\Omega_j$  and choose the weight function as

$$w_j = \begin{cases} 1 & \text{if } x \in \Omega_j \\ 0 & \text{if } x \notin \Omega_j \end{cases} \quad (4.5)$$

This gives us  $N$  equations

$$\int_{\Omega_j} R(c_i) \, dv = 0 \quad j = 1, \dots, N \quad (4.6)$$

for  $N$  unknown parameters  $c_i$ . For the most simple case of just one single free parameter we simply require that the integral of the residuum vanishes,

$$\int_{\Omega} R(c) \, dv = 0. \quad (4.7)$$

Usually the subdomains are chosen equally but this is not necessary.

### Collocation method

The collocation method uses the *Dirac distribution* as a weight function.

We now take a little detour to discuss this special distribution<sup>1</sup> which will also be used excessively in the development of the boundary element method. The Dirac distribution has the following definition:

$$\delta(x - x_j) = \begin{cases} 0 & \text{if } x \neq x_j \\ \infty & \text{if } x = x_j \end{cases} \quad \text{and} \quad \int_{\Omega} \delta(x - x_j) \, dv = 1 \quad x_j \in \Omega. \quad (4.8)$$

The most useful property of the Dirac distribution is its filtering property,

$$\int_{\Omega} \delta(x - x_j) f(x) \, dv = f(x_j). \quad (4.9)$$

If we use this distribution as a weight function, we obtain

$$\int_{\Omega} \delta(x - x_j) R(c_i) \, dv = R(c_i)|_{x=x_j} = 0 \quad j = 1, \dots, N. \quad (4.10)$$

This means, we require the residuum to vanish at a set of  $N$  preassigned *collocation points*  $x_j$ , giving us  $N$  equations for  $N$  unknown parameters  $c_i$ . The method strongly depends on a sensible choice of the collocation points, which requires some physical insight by the user.

### Least squares method

The least squares method requires the minimization of

$$\int_{\Omega} R^2(c_i) \, dv = \min. \quad (4.11)$$

---

<sup>1</sup>The Dirac distribution lacks some properties which qualifies it as a true function in a strict mathematical sense, therefore it is only called a distribution. However the distinction does not matter here.

resulting in

$$\int_{\Omega} \frac{\partial R(c_i)}{\partial c_j} R(c_i) \, dv = 0 \quad j = 1, \dots, N. \quad (4.12)$$

Again this results in  $N$  equations. The weight functions are

$$w_j = \frac{\partial R(c_i)}{\partial c_j}. \quad (4.13)$$

### Galerkin's method

Galerkin's method assumes that we deal with an interior method where the trial functions satisfy the boundary conditions. We then use as weight functions also a set of functions which satisfy the boundary conditions. In the standard Galerkin procedure the trial functions are in the linear form of (4.2) and we use the shape functions  $\tilde{u}_i$  themselves as weight functions,

$$w_j = \tilde{u}_j, \quad (4.14)$$

to obtain

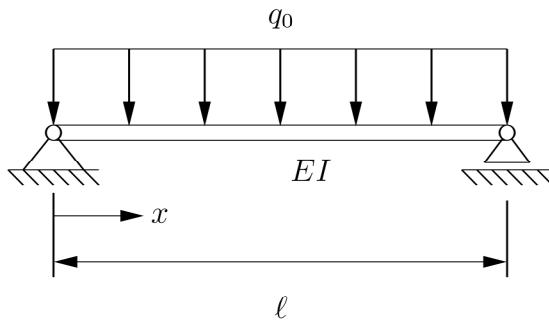
$$\int \tilde{u}_j R(c_i) \, dv = 0 \quad j = 1, \dots, N. \quad (4.15)$$

This is also called the *Bubnov-Galerkin method*. If we use other functions than the shape functions, but still satisfying the boundary conditions we have a *Petrov-Galerkin method*.

The requirement for the weight functions to satisfy the boundary conditions can be lifted, if one considers additional residuals on the boundaries using a mixed method. We will learn about this later, when dealing with the finite element method.

### Example 4.1

We consider a simply supported Euler-Bernoulli beam of length  $\ell$  under a constant distributed load  $q_0$ .



The differential equation is

$$w_{xxxx} = \frac{q_0}{EI} \quad (4.16)$$

and the boundary conditions are

$$w(x = 0) = 0 \quad w_{xx}(x = 0) = 0 \quad (4.17)$$

$$w(x = \ell) = 0 \quad w_{xx}(x = \ell) = 0. \quad (4.18)$$

From this we obtain the exact solution as

$$w(x) = \frac{q_0}{EI} \frac{\ell^4}{24} \left\{ \left(\frac{x}{\ell}\right)^4 - 2\left(\frac{x}{\ell}\right)^3 + \left(\frac{x}{\ell}\right)^2 \right\}. \quad (4.19)$$

The maximum displacement at  $x = \ell/2$  is

$$w_{\max} = \frac{5}{384} \frac{q_0 \ell^4}{EI} \approx 0.01302 \frac{q_0 \ell^4}{EI}. \quad (4.20)$$

We will now solve this problem by the methods of weighted residuals. As a trial function we choose

$$\tilde{w} = c \sin \frac{\pi x}{\ell}, \quad (4.21)$$

$$\tilde{w}_{,xx} = -c \left(\frac{\pi}{\ell}\right)^2 \sin \frac{\pi x}{\ell}, \quad (4.22)$$

$$\tilde{w}_{,xxxx} = c \left(\frac{\pi}{\ell}\right)^4 \sin \frac{\pi x}{\ell}, \quad (4.23)$$

which has just one unknown parameter  $c$  and fulfills all boundary conditions, as it is easy to check. The parameter has the physical interpretation of the midlength displacement and can therefore directly compared to the exact solution at  $x = \ell/2$ .

The residual is

$$R = \tilde{w}_{,xxxx} - \frac{q_0}{EI} = c \left(\frac{\pi}{\ell}\right)^4 \sin \frac{\pi x}{\ell} - \frac{q_0}{EI}. \quad (4.24)$$

We now apply all of the above methods to this problem:

### Method of subdomains

With the methods of subdomains we find according to (4.7)

$$\int_0^\ell \left( c \left(\frac{\pi}{\ell}\right)^4 \sin \frac{\pi x}{\ell} - \frac{q_0}{EI} \right) \cdot 1 \, dx = 0. \quad (4.25)$$

Resulting in

$$c \left(\frac{\pi}{\ell}\right)^4 \left[ -\cos \frac{\pi x}{\ell} \right]_0^\ell \frac{\ell}{\pi} = \frac{q_0 \ell}{EI} \quad (4.26)$$

$$c 2 \left(\frac{\pi}{\ell}\right)^3 = \frac{q_0 \ell}{EI} \quad (4.27)$$

and finally

$$c = \frac{1}{2\pi^3} \frac{q_0 \ell^4}{EI} \approx 0.01613 \frac{q_0 \ell^4}{EI}. \quad (4.28)$$

### Collocation method

We use as collocation point  $x = \ell/2$  and obtain

$$c \left(\frac{\pi}{\ell}\right)^4 \sin \frac{\pi}{2} = \frac{q_0}{EI}. \quad (4.29)$$

This gives

$$c = \frac{1}{\pi^4} \frac{q_0 \ell^4}{EI} \approx 0.01027 \frac{q_0 \ell^4}{EI}. \quad (4.30)$$

### Least squares method

With

$$\frac{\partial R}{\partial c} = \left(\frac{\pi}{\ell}\right)^4 \sin \frac{\pi x}{\ell} \quad (4.31)$$

we obtain

$$\int_0^\ell \left( c \left(\frac{\pi}{\ell}\right)^4 \sin \frac{\pi x}{\ell} - \frac{q_0}{EI} \right) \left( \left(\frac{\pi}{\ell}\right)^4 \sin \frac{\pi x}{\ell} \right) dx = 0. \quad (4.32)$$

From this we find

$$c \left(\frac{\pi}{\ell}\right)^8 \int_0^\ell \sin^2 \frac{\pi x}{\ell} dx = \frac{q_0}{EI} \left(\frac{\pi}{\ell}\right)^4 \int_0^\ell \sin \frac{\pi x}{\ell} dx \quad (4.33)$$

$$c \left(\frac{\pi}{\ell}\right)^4 \left[ -\frac{1}{4} \sin \frac{2\pi x}{\ell} + \frac{\pi x}{2\ell} \right]_0^\ell \frac{\ell}{\pi} = \frac{q_0}{EI} \left[ -\cos \frac{\pi x}{\ell} \right]_0^\ell \frac{\ell}{\pi} \quad (4.34)$$

$$c \left(\frac{\pi}{\ell}\right)^4 \frac{\ell}{2} = 2 \frac{\ell}{\pi} \frac{q_0}{EI} \quad (4.35)$$

and finally

$$c = \frac{4}{\pi^5} \frac{q_0 \ell^4}{EI} \approx 0.01307 \frac{q_0 \ell^4}{EI}. \quad (4.36)$$

### Galerkin method

The Galerkin methods uses

$$\int_0^\ell \left( c \left(\frac{\pi}{\ell}\right)^4 \sin \frac{\pi x}{\ell} - \frac{q_0}{EI} \right) \left( \sin \frac{\pi x}{\ell} \right) dx = 0. \quad (4.37)$$

This result in

$$c \left(\frac{\pi}{\ell}\right)^4 \int_0^\ell \sin^2 \frac{\pi x}{\ell} dx = \frac{q_0}{EI} \int_0^\ell \sin \frac{\pi x}{\ell} dx, \quad (4.38)$$

which is the same equation as for the least squares method. Here, least squares and the Galerkin method give the same value, but this is not a general result.

The collocation method is the easiest to use, since no integrals have to be solved to set up the equations. However, as Tab. 4.2 shows, the least squares and the Galerkin method give the best results with only one trial function. Since in general the least square method results in more complicated equations, the Galerkin method is usually preferred.

The Galerkin method is also closely related to variational methods. In fact it was originally developed to solve variational problems, and we will discuss this relationship in the following.

#### 4.1.4 Relation to variational methods

For some differential equations an equivalent variational form can be found, which means the differential equation can be obtained by variation of a functional. We will discuss this here only briefly to complete the picture. For a detailed discussion of variational formulations in mechanics the reader is directed to COURANT and HILBERT [4], GURTIN [8],

method	$c \frac{EI}{q_0 \ell^4}$	error
exact	0.01302	—
subdomains	0.01613	24%
collocation	0.01027	21%
least squares	0.01307	0.4%
Galerkin	0.01307	0.4%

Table 4.2: Results for different approximation methods for beam example

LANCZOS [11], ODEN and REDDY [13], REDDY [17], REISSNER [18], and WASHIZU [27]. General textbooks on the calculus of variations are for example EWING [5], FOX [7], SAGAN [19], and WEINSTOCK [28].

As an example we will use *Hamilton's functional* in its most simple form for a one-dof oscillator:

$$I(x) = \int_0^T \left( \frac{m}{2} \dot{x}^2 - \frac{c}{2} x^2 \right) dt = \min. \quad (4.39)$$

We assume given boundary conditions at  $t = 0$  and  $t = T$ . This is somewhat artificial, since for a problem in time domain we usually have initial values, but we want the example as simple as possible.

To find the minimum of this functional we use the calculus of variations, which will be outlined now.

We introduce a trial function

$$x(t, \varepsilon) = x_0(t) + \varepsilon \eta(t), \quad (4.40)$$

where  $x_0$  is assumed to be the desired solution and  $\eta(t)$  is a fixed but arbitrary function. For a vanishing *variational parameter*  $\varepsilon$  the exact solution is recovered. One says the exact solution is *embedded* in the trial function. Since we want the trial function to satisfy the boundary conditions, we must require

$$\eta(t=0) = \eta(t=T) = 0. \quad (4.41)$$

We develop the trial function in a Taylor series about  $\varepsilon = 0$  to obtain

$$x(t, \varepsilon) = x(t, \varepsilon)|_{\varepsilon=0} + \frac{\partial x(t, \varepsilon)}{\partial \varepsilon} \Big|_{\varepsilon=0} \varepsilon + \dots \quad (4.42)$$

$$= x_0(t) + \delta x(t, \varepsilon) + \dots \quad (4.43)$$

The expression<sup>2</sup>

$$\delta x(t, \varepsilon) = \frac{\partial x(t, \varepsilon)}{\partial \varepsilon} \Big|_{\varepsilon=0} \varepsilon = \varepsilon \eta(t) \quad (4.44)$$

---

<sup>2</sup>Here the symbol  $\delta$  is used to denote a variation. It has not to been mistaken for the Dirac distribution. Unfortunately no alphabet contains enough symbols to denote all the different functions and operators, even when using Latin and Greek letters.

is called the *first variation* of  $x$ . It is easy to show that

$$\delta \frac{\partial x}{\partial t} = \frac{\partial \delta x}{\partial t} . \quad (4.45)$$

Using this variation in Hamilton's principle, we obtain

$$\delta I = \left. \frac{\partial I}{\partial \varepsilon} \right|_{\varepsilon=0} \varepsilon = \int_0^T (m \dot{x} \delta \dot{x} - c x \delta x) dt = 0 . \quad (4.46)$$

Partial integration of the first term leads to

$$\int_0^T (m \ddot{x} + c x) \delta x dt - [m \dot{x} \delta x]_0^T = 0 . \quad (4.47)$$

Since the variation  $\delta x$  vanishes at the boundary due to (4.41), the boundary term vanishes, and since  $\delta x$  is an arbitrary function, Eq. (4.47) can only be satisfied, if the term in parentheses vanishes

$$m \ddot{x} + c x = 0 . \quad (4.48)$$

This is the equation of motion (1.1) for a one-dof oscillator.

We now consider a method to obtain an approximative solution on the basis of the variational formulation:

### Ritz method

The Ritz method introduces a trial function of the type (4.2) into the functional,

$$\tilde{x}(t) = \sum_{i=1}^N \tilde{x}_i(t) c_i . \quad (4.49)$$

The variation is now carried out with respect to the unknown parameters  $c_i$  to obtain

$$\delta \tilde{x}(t) = \sum_{i=1}^N \tilde{x}_i(t) \delta c_i . \quad (4.50)$$

Using this in (4.47) gives

$$\delta \tilde{I} = \int_0^T \left( m \sum_{i=1}^N \ddot{\tilde{x}}_i c_i + c \sum_{i=1}^N \tilde{x}_i c_i \right) \sum_{j=1}^N \tilde{x}_j \delta c_j dt = 0 . \quad (4.51)$$

Now, since the variations  $\delta c_j$  are arbitrary, we find the  $N$  independent equations

$$\int_0^T \left( m \sum_{i=1}^N \ddot{\tilde{x}}_i c_i + c \sum_{i=1}^N \tilde{x}_i c_i \right) \tilde{x}_j dt = 0 \quad j = 1, \dots, N , \quad (4.52)$$

or in a more compact form

$$\int_0^T R(c_i) \tilde{x}_j dt = 0 \quad j = 1, \dots, N . \quad (4.53)$$

If we compare this to (4.15), we see that this is exactly the form of the Bubnov-Galerkin method.

Besides Ritz's method some other methods can be used to obtain approximative solutions from a variational functional, like the *Trefftz method*. An account of these methods as well as a detailed discussion of Ritz and Galerkin methods can be found in LEIPHOLZ [12].

Variational methods are usually advantageous, since they provide certain additional information on the problem, like error bounds. However, it has to be pointed out that not for every differential equation an equivalent variational form can be found. A discussion of this problem is given by FINLAYSON [6].

We will not go into further details, but stick to the method of weighted residuals in the form of Galerkin's method to develop the *finite element method*.

# Chapter 6

## Comparison of Methods

The following table shows a short overview of the abilities of each of the presented methods:

	FDM	FEM	BEM
type of differential equation	any	any	only linear
implementation	easy	moderate	difficult
access of error	easy	difficult	difficult
arbitrary domains	difficult	yes	yes
infinite domains	no	no	yes
nonlinear problems	yes	yes	difficult
transient problems	yes	yes	difficult
commercial programs available	few	many	few

By this short table the application range of each method should be clear:

**Finite Difference Methods** should be used, if the solution on a very simple domain like a rectangle is wanted and no commercial FEM package is available to handle this differential equation. Irregular domains can be treated by special techniques, however FEM is usually easier to handle and general-purpose packages are readily available. An exemption to this rule is Computational Fluid Dynamics (CFD) where FDM and related methods are wide spread and commercial software is available.

**Finite Element Methods** are the most versatile tool nowadays available to the engineer. Their ability to cope with almost any type of differential equation on almost any domain (infinite domains are a problem) has led to their widespread use and the availability of an abundant number of commercial FEM packages either for general purpose or for very special applications.

**Boundary Element Methods** handle infinite or semi-infinite domains properly and are therefore specially suited for problems like acoustics or foundation-soil problems. Nonlinear and transient problems can be handled only by resorting to special techniques, which substantially subtract from the elegance of the formulation.