HIGHER ORDER DISCRETIZATION OF ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS

A thesis submitted to

Lakehead University

in partial fulfillment of the requirements

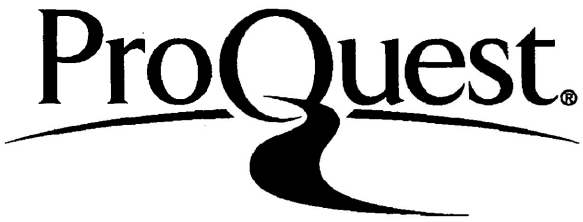
for the degree of

Master of Science

by

Golam Mosthafa Pathan

1979



ACKNOWLEDGMENT

I wish to thank my supervisor, Dr. D. J. Walton, for his advice and encouragement during the preparation of this thesis.

ABSTRACT

Higher order finite difference methods are discussed with respect to speed and accuracy when used in the solution of elliptic partial differential equations.

Although fast direct methods for solving elliptic partial differential equations are currently often discussed in the literature, the methods usually lean towards using the conventional five-point differencing on a uniform rectangular mesh which gives rise to block tridiagonal and tridiagonal matrices of Toeplitz form. For the solution of large linear systems which result from the use of a finite difference formula involving more mesh-points, the matrix equation

$$XA + AY = F$$

is used instead of the usual composite matrix approach. Although the matrices involved become less sparse, the operation count remains $O(n^3)$ when using an $n \times n$ mesh. However, for a comparable accuracy, n is much smaller for a higher order finite difference formula than that required for a standard five-point formula.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Partial Differential Equations are of interest since these arise in the mathematical formulation of many physical problems, for example, in equilibrium or steady-state problems the equilibrium configuration ϕ in a domain D is to be determined by solving the differential equation

$$L[\phi] = F$$

within D, *subject to certain conditions*

$$B_i[\phi] = g_i$$

on the boundary, ∂D, of D. Usually the integration domain D is closed and bounded. Such problems are known as boundary value problems. Steady viscous flow, steady temperature distribution and equilibrium stress in elasticity can be mentioned as examples of steady-state problems. The governing equations for such problems are elliptic.

1.1 NOTATION AND CONVENTIONS.

Unless otherwise mentioned, the following notation and conventions are assumed.

Scalar variables are denoted by lowercase letters, e.g. a, b, x, y, α, β .

Column vectors are denoted by underscored lowercase letters, e.g. $\underline{v}$, $\underline{w}$, $\underline{z}$.

Matrices are denoted by capital letters, e.g. A, B, C .

The elements of a column vector $\underline{v}$ are usually indicated as

$$\underline{v} = [v_1, v_2, v_3, \cdots, v_n]^T .$$

The elements of a row vector $\underline{v}^T$ are denoted by

$$\underline{v}^T = [v_1, v_2, v_3, \cdots, v_n] \quad .$$

The elements of a matrix, A , are usually indicated as

$$A = [a_{i,j}] \ .$$

The value of a function $f(x, y)$, evaluated at a point (x_j, y_i) is denoted as $f_{i,j}$.

It is also understood that

$$u^{(p,q)}(x, y) \equiv \frac{\partial^{p+q}}{\partial x^p \partial y^q} u(x, y) \ .$$

The usual notation $O(h^m)$ is used to indicate a truncation error of order h^m .

1.2 TYPES OF PROBLEMS TO BE SOLVED.

In the ensuing work the numerical solution of the elliptic partial differential equation

$$\alpha(x)\frac{\partial^2}{\partial x^2}u(x, y) + \beta(x)\frac{\partial}{\partial x}u(x, y) + \gamma(x)u(x, y) + \theta(y)\frac{\partial^2}{\partial y^2}u(x, y)$$

$$+ \Psi(y)\frac{\partial}{\partial y}u(x, y) + \xi(y)u(x, y) = f(x, y) ,$$

$$\alpha(x) , \theta(y) > 0 \qquad (1.2.1)$$

is considered on the rectangular region

$$R: x_0 \le x \le x_n , y_0 \le y \le y_m ,$$

with Dirichlet boundary conditions. The solution $u(x, y)$ of the equation (1.2.1) is required to take on prescribed values on the boundary ∂R of the region R where $u(x, y)$ is assumed to be sufficiently smooth on and within the region R. For the existence of a unique solution to (1.2.1), it is further assumed that the coefficients α, β, γ, θ, Ψ, ξ and the function f satisfy the required conditions (Courant and Hilbert [8], page 334).

Well known examples of elliptic partial differential equations are Poisson's equation

$$\frac{\partial^2}{\partial x^2}u(x, y) + \frac{\partial^2}{\partial y^2}u(x, y) = f(x, y) \qquad (1.2.2)$$

and Laplace's equation,

$$\frac{\partial^2}{\partial x^2}u(x, y) + \frac{\partial^2}{\partial y^2}u(x, y) = 0 \qquad (1.2.3)$$

which can be obtained from equation (1.2.1) by setting

$$\alpha(x) \equiv \Theta(y) \equiv 1 \ ,$$

$$\beta(x) \equiv \gamma(x) \equiv \Psi(y) \equiv \xi(y) \equiv 0$$

and for Laplace's equation, $f(x, y) = 0$.

For convenience, only equations of the form (1.2.2) and (1.2.3) are considered. The generalization for the equation (1.2.1) is straight forward but involves tedious manipulation.

1.3 DISCRETIZATION AND MATRIX EQUATION.

The method of finite differencing deals with the discretization of an arbitrary problem involving partial differential equations (Forsythe and Wasow [10], page 178) and, in particular, in this work, with the problem outlined in section 1.2. To apply this method, a network of mesh-points is first established through out the region of interest. These mesh-points are the points of intersection of mesh-lines drawn parallel to the axes covering the region. The terms 'grid-point', 'pivotal-point', 'nodal-point' and 'lattice-point' also refer to a mesh-point. After this point the term 'mesh' will be used to denote the network thus obtained, and synonymous to lattice or grid and 'point' will be referred to as mesh-point.

Along the x-axis the mesh-lines are drawn through,

$$x_0 < x_1 < \ldots < x_j < \ldots < x_{n+1} \ ,$$

and those along the y-axis are drawn through

$$y_0 < y_1 < \ldots < y_i < \ldots < y_{m+1} \ ,$$

and the respective mesh-spacings are defined as follows:

$$h_j = x_{j+1} - x_j \ , \ j = 0(1)n \ ,$$
$$k_i = y_{i+1} - y_i \ , \ i = 0(1)m \ . \tag{1.3.1}$$

It is shown in fig. (1.1) how the basic approximation involves the replacement of a continuous region by a mesh of discrete points within R .

Let

$$U = [u_{i,j}]$$

denote the true solution to the equation (1.2.2) at the internal points of

$$\{(x_j, y_i) : i = 0(1)m + 1 \ , \ j = 0(1)n + 1\} \tag{1.3.2}$$

where

$$u_{i,j} = u(x_j, y_i)$$

is the exact value of the solution at (x_j, y_i) .

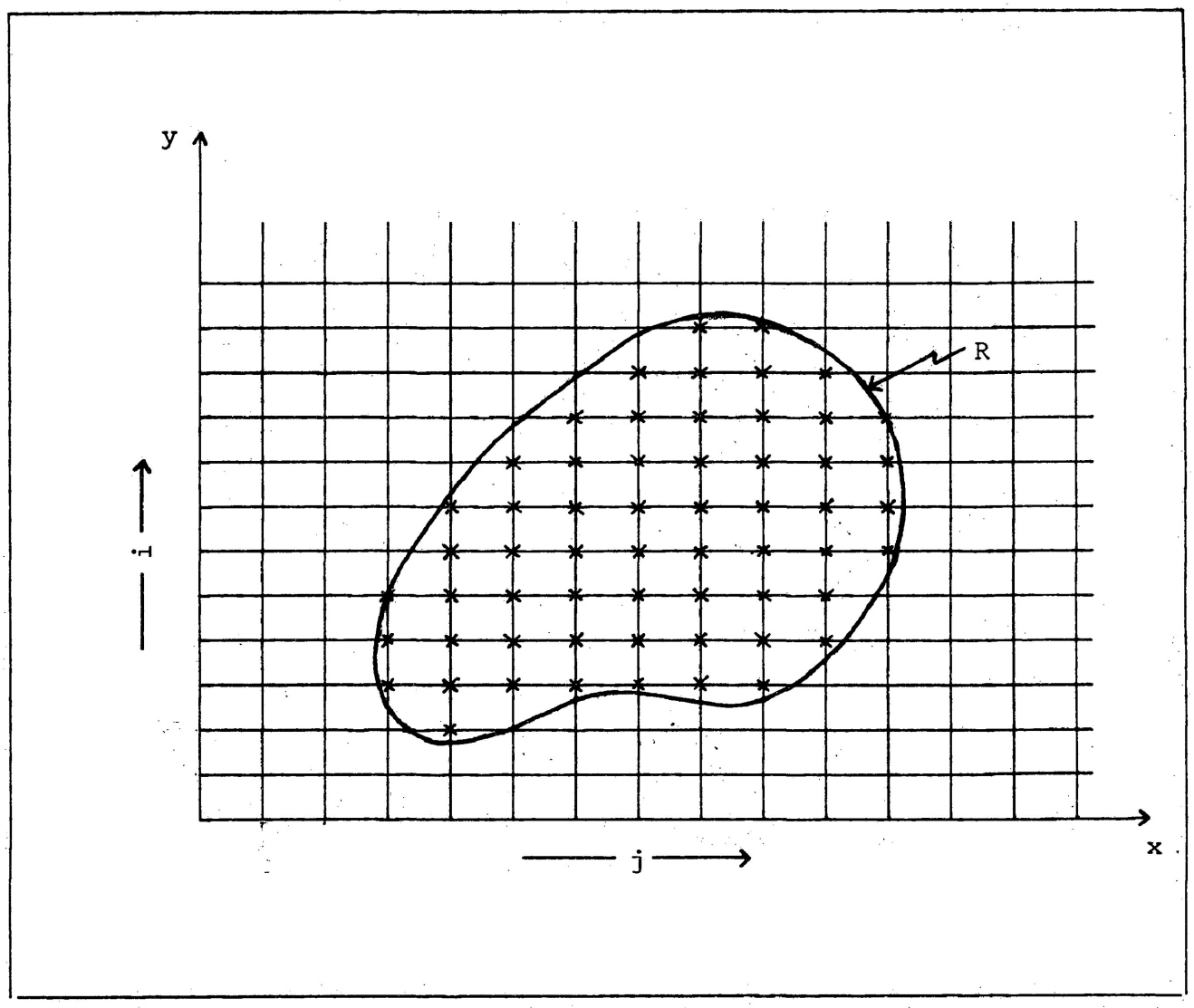


Fig. 1.1 DISCRETE APPROXIMATION OF A
CONTINUOUS TWO DIMENSIONAL REGION.

The transition from the entire continuous region to a finite set of points destroy the possibility of an exact calculation of the derivatives (Kantorovich and Krylov [15], page 199). The derivatives in equation (1.2.2) are approximated in terms of h, k, and their finite difference expressions involving central, forward or backward

differencing at each point within the region of interest. The boundary conditions are also approximated as such. This is a method of reducing the problem of differential equations to a linear algebraic system by using a mesh (Kantorovich and Krylov [15], page 199, Forsythe and Wasow [10], page 175-176). This process is called discretization. Let

$$A = [a_{ij}] \quad \begin{matrix} i = 1(1)m \\ j = 1(1)n \end{matrix},$$

denote the solution of the system of finite difference equations thus arrived at. In general the true solution U differs from A at a particular point. The difference revealing the discrepancy between the solution of the differential equation and the solution of the system of approximating difference equations on a mesh of particular size is called the discretization error. Taylor's series expansions may be used to investigate this error of discretization for each replacement.

It can be noted that approximate values at non-mesh-points may be evaluated from the discrete solution by interpolatory techniques (Ames [2], page 15).

The discretization of equation (1.2.2) by finite difference technique leads to a matrix equation (Bickley and McNamee [5]) of the form

$$AV + WA = G \qquad (1.3.3)$$

where V, W, G are known matrices and A is the solution matrix. In discretizing equation (1.2.2), the axes are so chosen that the x variation is indicated by the column suffix j and the y variation by the row suffix i , that is, i increases vertically upward and j horizontally to the right as in fig. 1.2.

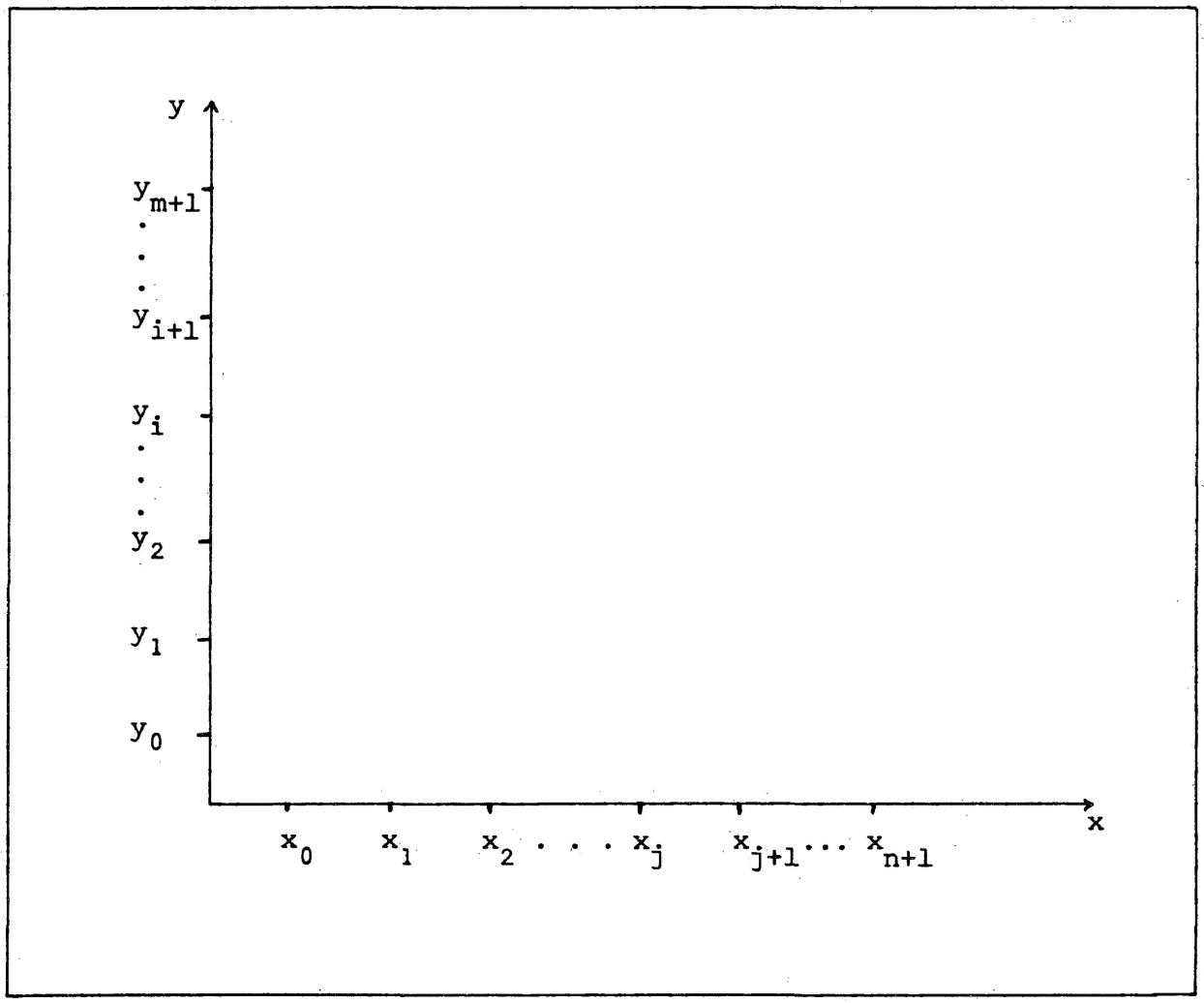


Fig. 1.2 SYSTEM OF AXES SHOWING PROJECTIONS OF DISCRETIZATION POINTS

With this setup, the matrix

$$V = [v_{i,j}] \quad \begin{matrix} i = 1(1)n \ , \\ j = 1(1)n \ , \end{matrix} \tag{1.3.4}$$

arises when a finite difference expression is substituted in equation (1.2.2) for the x-derivative at the internal points of (1.3.2).

Likewise, the matrix

$$W = [w_{i,j}] \quad \begin{matrix} i = 1(1)m \ , \\ j = 1(1)m \ , \end{matrix} \tag{1.3.5}$$

is obtained when y-derivatives in (1.2.2) are replaced by the finite difference expression at the same internal points of (1.3.2).

The values of the function $f(x, y)$ evaluated at the internal points of (1.3.2) in the process of discretization is denoted as

$$F = [f_{i,j}] \quad \begin{matrix} i = 1(1)m \ , \\ j = 1(1)n \ , \end{matrix} \tag{1.3.6}$$

where

$$f_{i,j} = f(x_j, y_i) \ .$$

The prescribed values of the function at the boundary points are denoted as

$$B = [b_{i,j}] \quad \begin{matrix} i = 1(1)m \ , \\ j = 1(1)n \ . \end{matrix} \tag{1.3.7}$$

Equations (1.3.6) and (1.3.7) can be combined to give

$$G = [f_{i,j} + b_{i,j}] \quad \begin{matrix} i = 1(1)m \ , \\ j = 1(1)n \ . \end{matrix}$$

The matrices A and G are of the same order.

It can be observed that if the x-increment is indicated by row suffix i and the y-increment by column suffix j then with the altered notation

$$\{(x_i, y_j): i = 0(1)m + 1, j = 0(1)n + 1\} \ ,$$

the matrix equation (1.4.3) takes the form

$$VA + AW = G$$

where

$$V = [v_{i,j}] \quad \begin{matrix} i = 1(1)m \ , \\ j = 1(1)m \ , \end{matrix}$$

$$W = [w_{i,j}] \quad \begin{matrix} i = 1(1)n \ , \\ j = 1(1)n \ , \end{matrix}$$

and the matrices A and G remain of the same order.

1.4 HIGHER ORDER DISCRETIZATION.

The recent literature (Hockney [13], Buzbee, Golub and Nielson [6], Swarztrauber [21], Bank and Rose [4], Sweet [20]) seem to concentrate on the use of five-point difference formula on a

uniform rectangular mesh for the solution of equations of the form (1.2.1). The conventional composite matrix formulation using tridiagonal and block tridiagonal matrices of Toeplitz form appears to be used frequently.

In the ensuing work the numerical solution of elliptic partial differential equations is investigated using higher order discretization formulae on a uniform rectangular mesh. The solution of the corresponding system of difference equations is obtained by solving a matrix equation of the form (1.3.3) rather than using a linear system which involves a composite matrix.

The solution matrix, A, of the finite difference system (1.3.3) at each internal point of (1.3.2) is found by using the algorithm SOLVEXAAY (Hoskins, Meek and Walton [14]).

CHAPTER 2

DISCRETIZATION FORMULAE

2.1 INTRODUCTION.

Finite difference schemes can be used for the solution of a variety of problems in physics and engineering. The region on which the solution is desired is replaced by a finite set of points and the governing partial differential equation of the problem is approximated by finite difference formulae at each of these points. Finite difference formulae for discretization of some partial differential equations can be found in Abramowitz and Stegun [1], and Collatz [7]. Such discretizations may lead to a matrix equation of the form (1.3.3).

2.2 FINITE DIFFERENCE OPERATORS AND THEIR RELATION TO DERIVATIVES.

Consider a mesh defined in section 1.3 where mesh-spacings h_j and k_i are given in (1.3.1). The following notation for various differences and related operators is used. They are applied to a function

$$u = u(x_j, y_i) \ , \qquad \begin{matrix} i = 0(1)m + 1 \ , \\ j = 0(1)n + 1 \ , \end{matrix}$$

over a constant mesh-spacing

$$h_j = h \ , \ j = 0(1)n \quad ,$$

and (2.2.1)

$$k_i = k \ , \ i = 0(1)m$$

The following are standard definitions for difference operators (S. Goldberg [12]).

Central difference:

$$\delta_x u_{i,j} = u_{i,j+\frac{1}{2}} - u_{i,j-\frac{1}{2}} \tag{2.2.2}$$

Forward difference:

$$\Delta_x u_{i,j} = u_{i,j+1} - u_{i,j} \tag{2.2.3}$$

Backward difference:

$$\nabla_x u_{i,j} = u_{i,j} - u_{i,j-1} \tag{2.2.4}$$

Differential Operator:

$$D_x u_{i,j} = \left.\frac{\partial u}{\partial x}\right|_{x=x_j} \tag{2.2.5}$$

Shift Operator:

$$E_x u_{i,j} = u_{i,j+1} \tag{2.2.6}$$

with similar expressions for the y-direction. In subsequent developments x-directional expressions are derived and y-directional expressions are taken to be analogous.

The following operational identities are immediate from (2.2.2) to (2.2.4) and (2.2.6):

$$\delta_x = E_x^{\frac{1}{2}} - E_x^{-\frac{1}{2}} , \tag{2.2.7}$$

$$\Delta_x = E_x - 1 , \tag{2.2.8}$$

$$\nabla_x = 1 - E_x^{-1} . \tag{2.2.9}$$

The finite difference approximation to derivatives can be obtained by relating the operator D_x with others in (2.2.2) to (2.2.6). In deriving relations between operators the Taylor's series expansion

$$u_{i,j+1} = u_{i,j} + \frac{h}{1!}\frac{\partial u_{i,j}}{\partial x} + \frac{h^2}{2!}\frac{\partial^2 u_{i,j}}{\partial x^2} + \frac{h^3}{3!}\frac{\partial^3 u_{i,j}}{\partial x^3} + \ldots$$

can be re-written as

$$\begin{aligned} E_x u_{i,j} &= \left(1 + \frac{h}{1!}D_x + \frac{h^2}{2!}D_x^2 + \frac{h^3}{3!}D_x^3 + \ldots\right) u_{i,j} \\ &= e^{hD_x} u_{i,j} . \end{aligned}$$

The relation

$$E_x = e^{hD_x} \tag{2.2.10}$$

is useful since the equality between operators as in (2.2.10) means that E and $\sum_{\ell}^{n} \frac{h^\ell D_x^\ell}{\ell!}$ give identical results when used for any polynomial of degree n for any n (Fox, L. [11], page 4). It is

known that all finite difference formulae are based upon polynomial approximation, that is, they give exact results when operating upon a polynomial of proper degree. In all other cases the formulae are approximations and are usually expressed in series form. Since only a finite number of terms can be used, the truncation error is of concern. The presence of such errors are indicated by using the O notation.

The relations (2.2.7) - (2.2.9) and (2.2.10) give rise to the following:

$$hD_x = \log_e E_x$$

$$= \log_e(1 + \Delta_x) \tag{2.2.11}$$

$$= -\log_e(1 - \nabla_x) \tag{2.2.12}$$

$$= 2 \sinh^{-1} \frac{\delta_x}{2} \tag{2.2.13}$$

The first derivative of u with respect to x at $x = x_j$ can be expressed in terms of forward differences as follows:

$$\left.\frac{\partial u}{\partial x}\right|_{x=x_j} = \frac{1}{h}\left(\Delta_x - \frac{1}{2}\Delta_x^2 + \frac{1}{3}\Delta_x^3 - \frac{1}{4}\Delta_x^4 + \ldots\right)u_{i,j} \tag{2.2.14}$$

from which an expression for $\frac{\partial^2}{\partial x^2} u(x, y)$ at $x = x_j$ can be formed, viz:

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_{x=x_j} = \frac{1}{h^2}\left[\log(1+\Delta_x)\right]^2 u_{i,j}$$

$$= \frac{1}{h^2}\left(\Delta_x^2 - \Delta_x^3 + \frac{11}{12}\Delta_x^4 - \frac{5}{6}\Delta_x^5 + \frac{137}{180}\Delta_x^6 - \ldots\right)u_{i,j} \qquad (2.2.15)$$

A similar formula can also be derived at the point x_{j-1} by setting

$$u_{i,j} = E_x u_{i,j-1}$$

in equation (2.2.14) (Fox, L. [11], page 7), i.e.

$$\left.\frac{\partial u}{\partial x}\right|_{x=x_j} = \frac{1}{h}\left[\log_e(1+\Delta_x)\right]u_{i,j}$$

$$= \frac{1}{h}\left[\log_e(1+\Delta_x)\right](1+\Delta_x)u_{i,j-1}$$

$$= \frac{1}{h}\left(\Delta_x + \frac{1}{2}\Delta_x^2 - \frac{1}{6}\Delta_x^3 + \frac{1}{12}\Delta_x^4 - \frac{1}{20}\Delta_x^5 + \ldots\right)u_{i,j-1}$$

and for the second derivative one can obtain

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_{x=x_j} = \frac{1}{h^2}\left(\Delta_x^2 - \frac{1}{12}\Delta_x^4 + \frac{1}{12}\Delta_x^5 - \frac{13}{180}\Delta_x^6 + \frac{11}{180}\Delta_x^7 \ldots\right)u_{i,j-1} \qquad (2.2.16)$$

This process can be repeated for the point (j-2) and (j-3) to yield

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_{x=x_j} = \frac{1}{h^2}\left(\Delta_x^2 + \Delta_x^3 - \frac{1}{12}\Delta_x^4 + \frac{1}{90}\Delta_x^6 - \frac{1}{90}\Delta_x^7 + \ldots\right)u_{i,j-2} \qquad (2.2.17)$$

and

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_{x=x_j} = \frac{1}{h^2}\left(\Delta_x^2 + 2\Delta_x^3 + \frac{11}{12}\Delta_x^4 - \frac{1}{12}\Delta_x^5 + \frac{1}{90}\Delta_x^6 \ldots\right)u_{i,j-3} \qquad (2.2.18)$$

The forward differences in the expressions for $u_{xx}(x_j, y_i)$ can be evaluated using the following convenient form:

$$\Delta^p u_{i,j} = \sum_{\ell=0}^{n} (-1)^{\ell}\binom{p}{\ell} u_{i,j+p-\ell}$$

where $\binom{p}{\ell}$ are binomial coefficients.

Similar formulae involving backward differences can be established using equation (2.2.12). The central difference expression for second derivatives is obtained from equation (2.2.13):

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_{x=x_j} = \left(\frac{2}{h}\sinh^{-1}\frac{\delta_x}{2}\right)^2 u_{i,j}$$

$$= \frac{1}{h^2}\left(\delta_x - \frac{1^2}{2^2\cdot 3!}\delta_x^3 + \frac{1^2\cdot 3^2}{2^4\cdot 5!}\delta_x^5 - \frac{1^2\cdot 3^2\cdot 5^2}{2^6\cdot 7!}\delta_x^7 + \ldots\right)^2 u_{i,j}$$

$$= \frac{1}{h^2}\left(\delta_x^2 - \frac{1}{12}\delta_x^4 + \frac{1}{90}\delta_x^6 - \frac{1}{560}\delta_x^8 + \frac{1}{3150}\delta_x^{10} - \ldots\right)u_{i,j}$$

(2.2.19)

2.3 FINITE DIFFERENCE FORMULAE.

2.3.1 FIVE-POINT FORMULA.

Consider the points (x_j, y_i), (x_j, y_{i+1}), (x_j, y_{i-1}), (x_{j-1}, y_i) and (x_{j+1}, y_i) on the rectangular mesh (1.3.2), as illustrated in fig. (2.1). An approximate expression for Laplace's operator

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

at a point (x_j, y_i) can be obtained by forming the differences of the values of u at the point

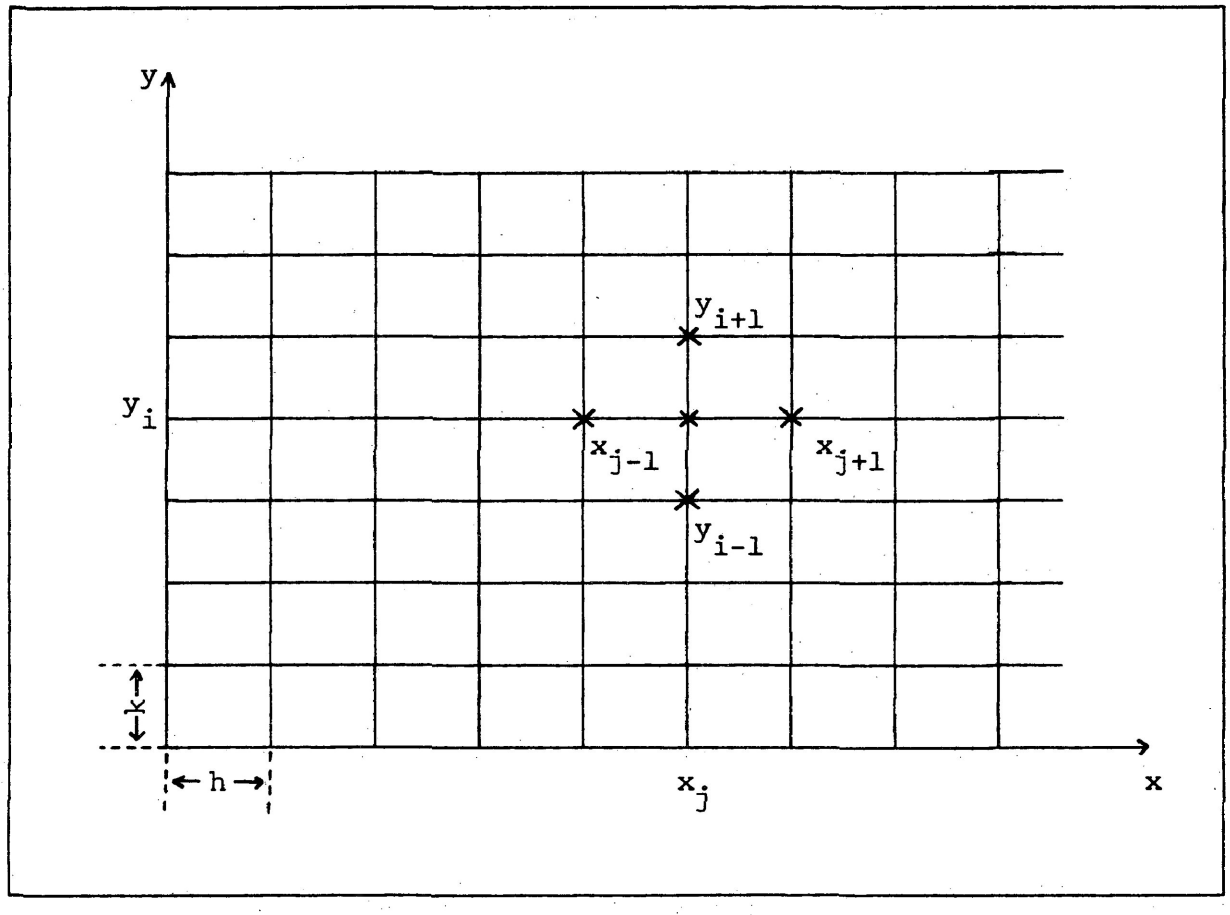


Fig. 2.1 FIVE-POINT MODE

(x_j, y_i) and the four points closest to it (Kantorovich and Krylov [15], page 181). The following expressions are obtained by using Taylor's series expansion for a uniform mesh:

$$u_{i,j+1} - u_{i,j} = hu^{(1,)}_{i,j} + \frac{h^2}{2!} u^{(2,)}_{i,j} + \frac{h^3}{3!} u^{(3,)}_{i,j} + \frac{h^4}{4!} u^{(4,)}_{i,j} + \ldots$$

$$u_{i,j-1} - u_{i,j} = -hu^{(1,)}_{i,j} + \frac{h^2}{2!} u^{(2,)}_{i,j} - \frac{h^3}{3!} u^{(3,)}_{i,j} + \frac{h^4}{4!} u^{(4,)}_{i,j} - \ldots$$

(2.3.1)

$$u_{i+1,j} - u_{i,j} = ku^{(,1)}_{i,j} + \frac{k^2}{2!} u^{(,2)}_{i,j} + \frac{k^3}{3!} u^{(,3)}_{i,j} + \frac{k^4}{4!} u^{(,4)}_{i,j} + \ldots$$

$$u_{i-1,j} - u_{i,j} = -ku^{(,1)}_{i,j} + \frac{k^2}{2!} u^{(,2)}_{i,j} + \frac{k^3}{3!} u^{(,3)}_{i,j} + \frac{k^4}{4!} u^{(,4)}_{i,j} - \ldots$$

The replacement for Laplace's operator is then obtained by adding all the equations in (2.3.1) term by term (Kantorovich and Krylov [15], page 181):

$$\frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h^2} + \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{k^2} = u^{(2,)}_{i,j} + u^{(,2)}_{i,j} + \frac{h^2}{12} u^{(4,)}_{i,j} + \frac{k^2}{12} u^{(,4)}_{i,j} + \ldots \qquad (2.3.2)$$

Hence, the approximation

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_{x=x_j} + \left(\frac{\partial^2 u}{\partial y^2}\right)_{y=y_i} \simeq \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h^2} + \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{k^2} \qquad (2.3.3)$$

with a truncation error of $O(h^2) + O(k^2)$.

The five-point approximation (2.3.3) for Laplace's operator can also be obtained from equation (2.2.19) and its analogue for the y-derivative using term by term addition.

Thus the approximation is

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_{x=x_j} + \left.\frac{\partial^2 u}{\partial y^2}\right|_{y=y_i} \simeq \frac{1}{h^2}\delta_x^2 u_{i,j} + \frac{1}{k^2}\delta_y^2 u_{i,j} + \cdots$$

$$= \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h^2} + \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{k^2} \qquad (2.3.4)$$

which has the error of the same order as in (2.3.3).

Discretization of Poisson's equation (1.2.2) with Dirichlet boundary conditions over a rectangular region by the five-point formula (2.2.4) leads to the matrix equation (1.3.3) (Bickley and McNamee [5]) where

$$V = \frac{1}{h^2}\begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & & -1 & 2 \end{bmatrix}_{n\times n},$$

$$W = \frac{1}{k^2}\begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & & -1 & 2 \end{bmatrix}_{m\times m},$$

and $G = -F + B_1 + B_2$

with $F = (f_{i,j}) = \{f(x_j, y_i): i = 1(1)m \ , \ j = 1(1)n\}$,

$A = \{(a_{ij}): i = 1(1)m \ , \ j = 1(1)n\}$

$$B_1 = \frac{1}{h^2} \begin{bmatrix} u_{1,0} & & u_{1,n+1} \\ u_{2,0} & & u_{2,n+1} \\ u_{3,0} & & u_{3,n+1} \\ \cdot & 0 & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ u_{m,0} & & u_{m,n+1} \end{bmatrix} ,$$

$$B_2 = \frac{1}{k^2} \begin{bmatrix} u_{0,1} & u_{0,2} & u_{0,3} & \cdots & u_{0,n} \\ & & 0 & & \\ u_{m+1,1} & u_{m+1,2} & u_{m+1,3} & \cdots & u_{m+1,n} \end{bmatrix}$$

It may be noted that the central difference operator $-\delta^2$ is used rather than δ^2 in order that V and W may have positive eigenvalues.

It is understood that the unspecified elements in the matrices are zero.

2.3.2 A NINE-POINT FORMULA.

Consider Poisson's equation (1.2.2). The solutions of the exact equation (2.3.2) and the approximate equation (2.3.3) in finite

differences for the same boundary conditions do not in general agree exactly. (Kantorovich and Krylov [15], page 182). The measure of the discrepancy between them is indicated by the truncation error $O(h^2) + O(k^2)$. The accuracy may be improved by a higher order finite difference formula designed to reduce the truncation error.

Such techniques are discussed in Kantorovich and Krylov [15], page 182-199, Collatz [7], Fox [11], page 260. In addition to the values of the function u at the points used in section (2.3.1), consider the values at (x_{j-2}, y_i), (x_{j+2}, y_i), (x_j, y_{i-2}) and (x_j, y_{i+2}) as well. An analysis similar to that in section (2.3.1) may be carried out to produce an approximation of the form (2.3.3) This can also be accomplished by taking the first two terms from equation (2.2.19) and its analogue for the y-derivative and then adding them together. Hence

$$\frac{1}{h^2}\left[\delta_x^2 - \frac{1}{12}\delta_x^4\right]u_{i,j} + \frac{1}{k^2}\left[\delta_y^2 - \frac{1}{12}\delta_y^4\right]u_{i,j}$$

$$= \frac{-u_{i,j-2} + 16u_{i,j-1} - 30u_{i,j} + 16u_{i,j+1} - u_{i,j+2}}{h^2}$$

$$+ \frac{-u_{i-2,j} + 16u_{i-1,j} - 30u_{i,j} + 16u_{i+1,j} - u_{i+2,j}}{k^2} \qquad (2.3.5)$$

By Taylor's series expansion it can be shown that

$$\frac{1}{h^2}\left[\delta_x^2 - \frac{1}{12}\delta_x^4\right]u_{i,j} + \frac{1}{k^2}\left[\delta_y^2 - \frac{1}{12}\delta_y^4\right]u_{i,j} = u_{i,j}^{(2,)} + u_{i,j}^{(,2)}$$

$$- \frac{1}{90}h^4 u_{i,j}^{(6,)} - \frac{1}{90}k^4 u_{i,j}^{(,6)} + \ldots$$

which indicates a local truncation error of $O(h^4) + O(k^4)$. The stencil in (2.3.5) may be referred to as a 9-point cross of mesh-points.

2.3.3 COMPLEXITY NEAR BOUNDARIES.

Consider the discretization of Poisson's equation (1.2.2) with Dirichlet boundary conditions using the stencil (2.3.5) on a rectangular region assuming that h and k can be chosen so that the boundaries are mesh-lines. When the stencil (2.3.5) is applied to points near the boundaries, values of the function u are required at some points outside the solution region as indicated in fig. (2.2).

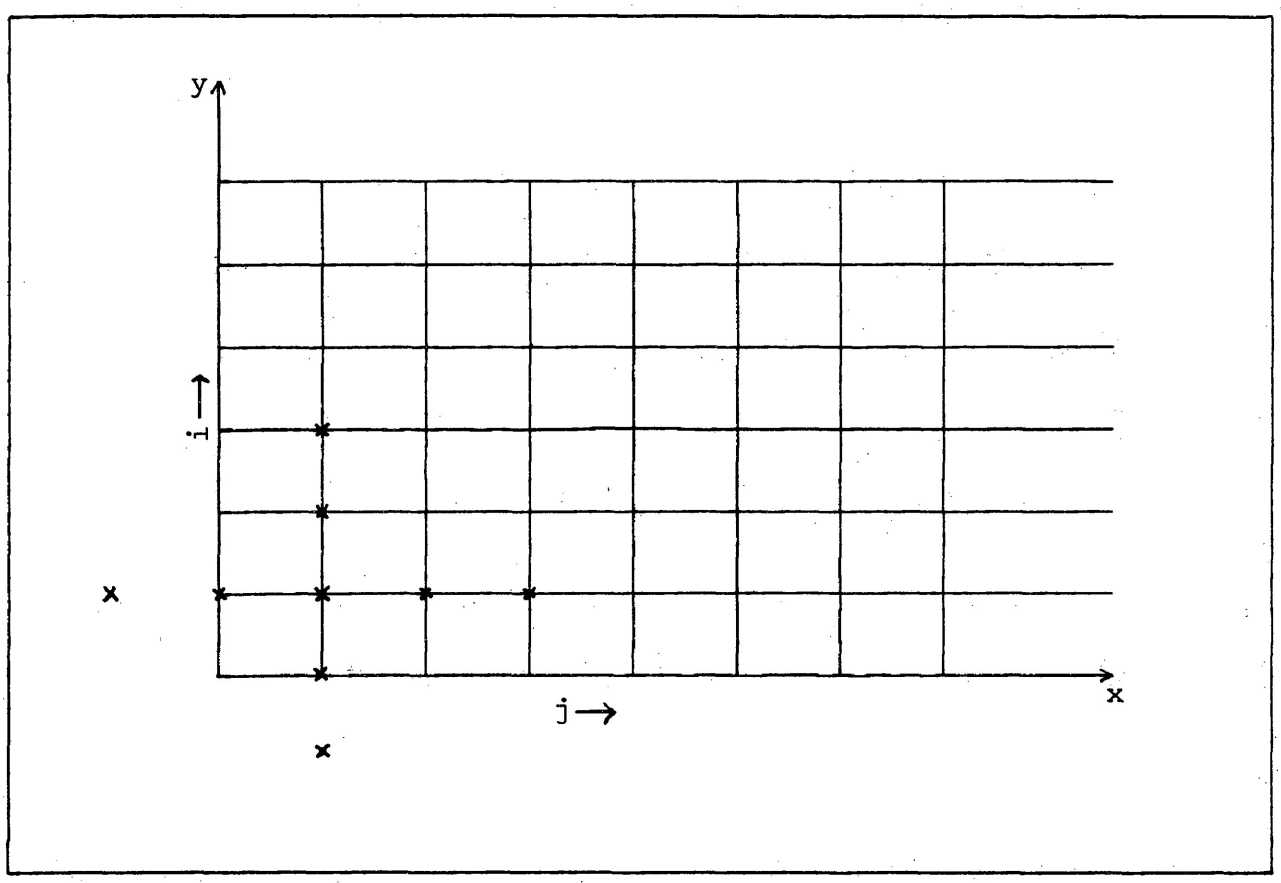


Fig. 2.2 9-POINT CROSS OF MESH-POINT

To avoid such situations forward-difference or a backward-difference formulae involving the same number of points as in (2.3.5) can be used for points near the boundaries, but the truncation error is likely to be of lower order in h or k, being reduced, for example from $O(h^4)$ to $O(h^3)$.

2.3.4 BOUNDARY ADJUSTMENT FOR A NINE-POINT FORMULA.

Consider the first two terms in equation (2.2.16) for the replacement of the second derivative, u_{xx} at points on the mesh-line

$x = x_0 + h$, thus

$$u_{xx} \simeq \frac{1}{h^2}\left(\Delta_x^2 - \frac{1}{12}\Delta_x^4\right)u_{i,j-1} \qquad (2.3.6)$$

and the corresponding replacement for u_{yy} at points on the mesh-line $y = y_0 + k$ is

$$u_{yy} \simeq \frac{1}{k^2}\left(\Delta_y^2 - \frac{1}{12}\Delta_y^4\right)u_{i-1,j} \qquad (2.3.7)$$

A similar replacement can be done using backward differencing for the points on mesh-line $x = x_{n+1} - h$ and $y = y_{m+1} - k$.

Addition of equations (2.3.6) and (2.3.7) and subsequent Taylor's series expansion produce

$$\frac{1}{h^2}\left(\Delta_x^2 - \frac{1}{12}\Delta_x^4\right)u_{i,j-1} + \frac{1}{k^2}\left(\Delta_y^2 - \frac{1}{12}\Delta_y^4\right)u_{i-1,j} = u_{i,j}^{(2,)}$$

$$+ u_{i,j}^{(,2)} - \frac{h^3}{12}u_{i,j}^{(5,)} - \frac{k^3}{12}u_{i,j}^{(,5)} + \ldots \qquad (2.3.8)$$

which indicates a local truncation error of $O(h^3) + O(k^3)$.

It may be noted that the discretization of equation (1.2.2) using equation (2.3.6) and equation (2.3.5) at the points

$$y_i = y_0 + ik \ , \ i = 2(1)m - 1$$

along $x = x_0 + h$ gives a local truncation error of $O(h^3) + O(k^3)$ and that the error due to the discretization by (2.3.7) and (2.3.5) at the points

$$x_j = x_0 + jh \ , \ j = 2(1)n - 1$$

along $y = y_0 + k$ is of order $O(h^4) + O(k^3)$. A similar analysis may be carried out for the points near the other boundaries.

Discretization of equation (1.2.2) with Dirichlet boundary condition by (2.3.5) and (2.3.8) leads to a matrix equation (1.3.3) where

$$V = \frac{1}{12h^2}\begin{bmatrix} 20 & -16 & 1 & & & & & \\ -6 & 30 & -16 & \cdot & & & & \\ -4 & -16 & 30 & \cdot & \cdot & & & \\ 1 & 1 & -16 & \cdot & \cdot & & & \\ & & 1 & \cdot & \cdot & \cdot & 1 & \\ & & & \cdot & \cdot & \cdot & -4 & \\ & & & & \cdot & \cdot & -6 & \\ & & & & \cdot & \cdot & 20 & \end{bmatrix}_{n\times n},$$

$$W = \frac{1}{12k^2}\begin{bmatrix} 20 & -6 & -4 & 1 & & & & & \\ -16 & 30 & -16 & 1 & & & & & \\ 1 & -16 & 30 & -16 & 1 & & & & \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot & \cdot & \cdot & \\ & & & & & \cdot & \cdot & \cdot & \cdot \\ & & & & & 1 & -4 & -6 & 20 \end{bmatrix}_{m\times m},$$

$$g = -F + B_1 + B_2$$

with,

$$B_1 = \frac{1}{12h^2}\begin{pmatrix} llu_{1,0} & -u_{1,0} & & -u_{1,n+1} & llu_{1,n+1} \\ llu_{2,0} & -u_{2,0} & & -u_{2,n+1} & llu_{2,n+1} \\ \vdots & \vdots & 0 & \vdots & \vdots \\ llu_{m,0} & -u_{m,0} & & -u_{m,n+1} & llu_{m,n+1} \end{pmatrix},$$

$$B_2 = \frac{1}{12k^2}\begin{pmatrix} llu_{0,1} & llu_{0,2} & llu_{0,3} & llu_{0,n-1} & llu_{0,n} \\ -u_{0,1} & -u_{0,2} & -u_{0,3} & -u_{0,n-1} & -u_{0,n} \\ & & 0 & & \\ -u_{m+1,1} & -u_{m+1,2} & -u_{m+1,3} & -u_{m+1,n-1} & -u_{m+1,n} \\ llu_{m+1,1} & llu_{m+1,2} & llu_{m+1,3} & llu_{m+1,n-1} & llu_{m+1,n} \end{pmatrix}$$

and matrices A and F are as in section 2.3.1.

2.3.5 A MODIFIED NINE-POINT FORMULA.

It is clear from equation (2.3.5) and (2.3.8) that the discretization error varies from $O(h^3) + O(k^3)$ to $O(h^4) + O(k^4)$. To increase the accuracy of (2.3.8) to that of (2.3.5), consider the value of the function at an additional internal point while approximating the second derivative by forward or backward differencing at a

relevant point (Kantorovich and Krylov [15], page 196) to obtain

$$u_{xx} + u_{yy} \approx \frac{1}{h^2}\left(\Delta_x^2 - \frac{1}{12}\Delta_x^4 + \frac{1}{12}\Delta_x^5\right)u_{i,j-1}$$

$$+ \frac{1}{k^2}\left(\Delta_y^2 - \frac{1}{12}\Delta_y^4 + \frac{1}{12}\Delta_y^5\right)u_{i-1,j} \qquad (2.3.9)$$

The Taylor's series expansion gives

$$\frac{1}{h^2}\left(\Delta_x^2 - \frac{1}{12}\Delta_x^4 + \frac{1}{12}\Delta_x^5\right)u_{i,j-1} + \frac{1}{k^2}\left(\Delta_y^2 - \frac{1}{12}\Delta_y^4 + \frac{1}{12}\Delta_y^5\right)u_{i-1,j}$$

$$= u_{i,j}^{(2,)} + u_{i,j}^{(,2)} + \frac{26}{360}h^4 u_{i,j}^{(6,)} + \frac{26}{360}k^4 u_{i,j}^{(,6)} + \ldots \qquad (2.3.10)$$

which indicates a local truncation error of the same order as in (2.3.5).

Discretization of equation (1.2.2) with Dirichlet boundary condition by (2.3.5) and (2.3.9) leads to the matrix equation (1.3.3) where

$$V = \frac{1}{12h^2}\begin{bmatrix} 15 & -16 & 1 & & & & \\ 4 & 30 & -16 & \cdot & & & \\ -14 & -16 & 30 & \cdot & \cdot & & -1 \\ 6 & 1 & -16 & \cdot & \cdot & \cdot & 6 \\ -1 & & 1 & \cdot & \cdot & \cdot & -14 \\ & & & \cdot & \cdot & \cdot & 4 \\ & & & & \cdot & \cdot & 15 \end{bmatrix}_{n\times n},$$

$$W = \frac{1}{12k^2}\begin{bmatrix} 15 & 4 & -14 & 6 & -1 & & & & \\ -16 & 30 & -16 & 1 & & & & & \\ 1 & -16 & 30 & -16 & 1 & & & & \\ & \cdot & \cdot & \cdot & \cdot & \cdot & & & \\ & & \cdot & \cdot & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \cdot & \cdot & \\ & & & & \cdot & \cdot & \cdot & \cdot & \\ & & & -1 & 6 & -14 & 4 & 15 & \end{bmatrix}_{m\times m},$$

and the matrices A, F are as before. The matrices B_1 and B_2 are also as in section (2.3.4) except that the coefficients of $u_{i,0}$ and $u_{i,n+1}$, $(i = 1(1)m)$ in the first and last columns of B_1, as well as the coefficients of $u_{0,j}$ and $u_{m+1,j}$ $(j = 1(1)n)$ in the first and last rows of B_2 are now 10 rather than 11.

2.3.6 AN ALTERNATIVE NINE-POINT FORMULA.

In section (2.3.4) a less accurate nine-point formula is used at points close to the boundaries and a nine-point formula of section (2.3.2) is used at the remaining interior points. Later in section (2.3.5) a modification is done by taking the value of the function at an additional internal point while using the one-sided (forward or backward) differencing. It may be desirable to use a nine-point formula at all points in the solution region. Such a nine-point

formula, which may be referred to as a 9-point square of mesh-points, is as follows.

From equation (1.2.2) it follows that

$$\frac{\partial^4}{\partial x^4} u(x, y) + \frac{\partial^4}{\partial x^2 \partial y^2} u(x, y) = \frac{\partial^2}{\partial x^2} f(x, y)$$

$$\frac{\partial^4}{\partial x^2 \partial y^2} u(x, y) + \frac{\partial^4}{\partial y^4} u(x, y) = \frac{\partial^2}{\partial y^2} f(x, y) \ . \tag{2.3.11}$$

Taylor's series expansion of equation (2.3.4) are

$$\frac{1}{h^2} \delta_x^2 u_{i,j} = u_{i,j}^{(2,)} + \frac{h^2}{12} u_{i,j}^{(4,)} + \frac{h^4}{360} u_{i,j}^{(6,)} + \frac{h^6}{20160} u_{i,j}^{(8,)} + \ldots, \tag{2.3.12}$$

$$\frac{1}{k^2} \delta_y^2 u_{i,j} = u_{i,j}^{(,2)} + \frac{k^2}{12} u_{i,j}^{(,4)} + \frac{k^4}{360} u_{i,j}^{(,6)} + \frac{k^6}{20160} u_{i,j}^{(,8)} + \ldots, \tag{2.3.13}$$

$$i = 1(1)m, \ j = 1(1)n \ .$$

The difference between $u_{i,j}$ and values of u at additional points (circled in fig. 2.3) are tabulated and Taylor's series expansion gives:

$$\frac{1}{12h^2} \delta_x^2 \delta_y^2 u_{i,j} = \frac{k^2}{12} u_{i,j}^{(2,2)} + \frac{h^2 k^2}{144} u_{i,j}^{(4,2)} + \frac{h^4 k^2}{4320} u_{i,j}^{(6,2)} + \ldots$$

$$+ \frac{k^4}{144} u_{i,j}^{(2,4)} + \frac{h^2 k^4}{1728} u_{i,j}^{(4,4)} + \ldots + \frac{k^6}{4320} u_{i,j}^{(6,2)} + \ldots \ , \tag{2.3.14}$$

$$\frac{1}{12k^2}\delta_x^2\delta_y^2 u_{i,j} = \frac{h^2}{12}u_{i,j}^{(2,2)} + \frac{h^2k^2}{144}u_{i,j}^{(2,4)} + \frac{h^2k^2}{4320}u_{i,j}^{(2,6)} + \ldots$$

$$+ \frac{h^4}{144}u_{i,j}^{(4,2)} + \frac{h^4k^2}{1728}u_{i,j}^{(4,4)} + \ldots + \frac{h^6}{4320}u_{i,j}^{(2,6)} + \ldots ,$$

(2.3.15)

$$i = 1(1)m \ , \ j = 1(1)n \ .$$

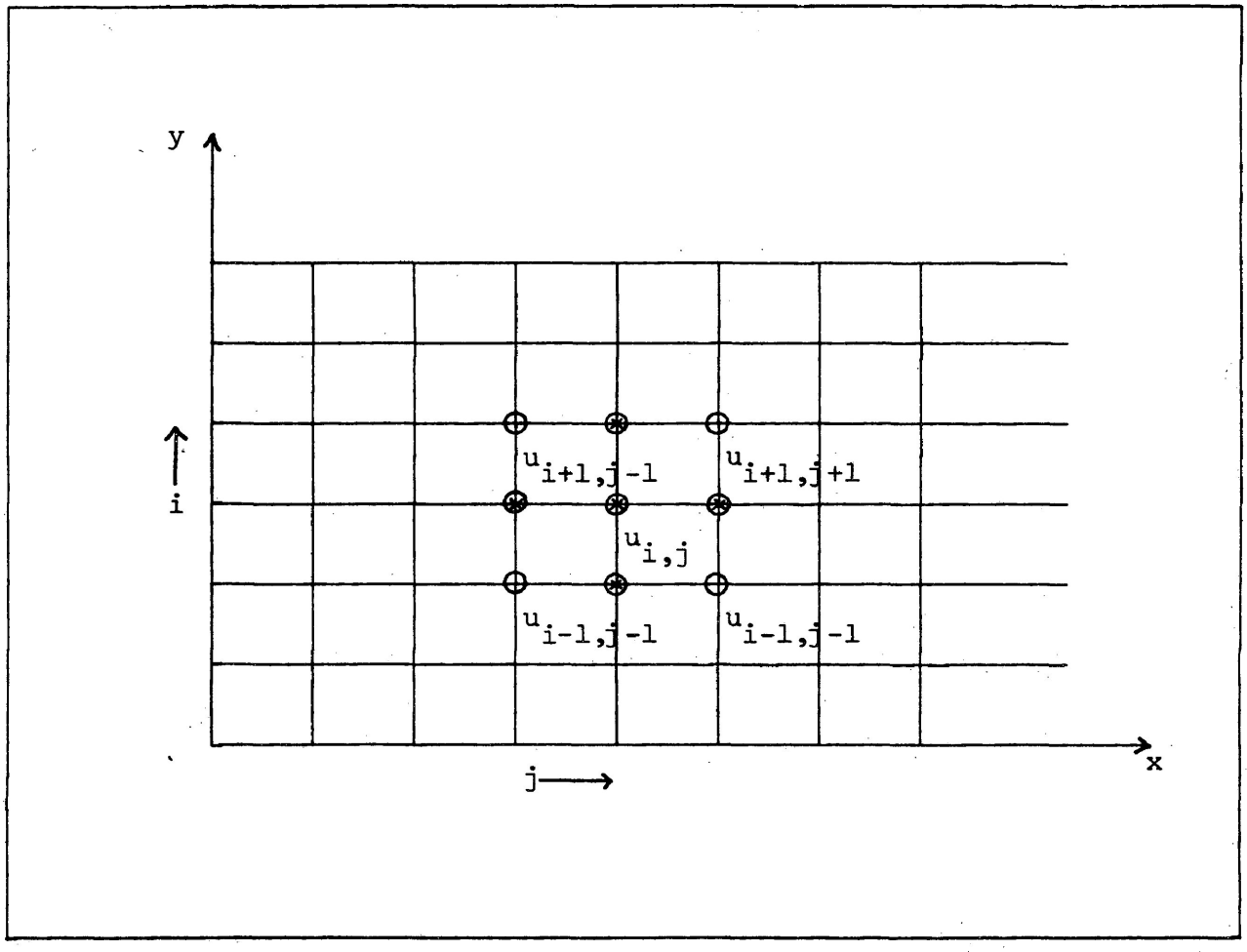


Fig. 2.3 9-POINT SQUARE OF MESH-POINT

Addition of equations from (2.3.12) to (2.3.15) and use of equation (2.3.11) gives (Kantorovich and Krylov [15], page 185, Forsythe and Wasow [10], page 193)

$$\left\{ \frac{1}{h^2}\delta_x^2 + \frac{1}{k^2}\delta_y^2 + \frac{1}{12}\left(\frac{1}{h^2} + \frac{1}{k^2}\right)\delta_x^2\delta_y^2 \right\} u_{i,j} = f_{i,j} + \frac{h^2}{12} f^{(2,)}_{i,j}$$

$$+ \frac{k^2}{12} f^{(,2)}_{i,j} + O(h^4) + O(k^4) ,$$

$$i = 1(1)m \ , \ j = 1(1)n \ . \qquad (2.3.16)$$

If $h = k$ the equations (2.3.11) through (2.3.16) yield (Kantorovich and Krylov [15], page 210, Forsythe and Wasow [10], page 194-195, Smith [19] page 143)

$$\frac{1}{h^2}\left(\delta_x^2 + \delta_y^2 + \frac{1}{6}\delta_x^2\delta_y^2\right) u_{i,j} = f_{i,j} + \frac{h^2}{12}\left(f^{(2,)}_{i,j} + f^{(,2)}_{i,j}\right)$$

$$+ \frac{h^4}{360}\left(f^{(4,)}_{i,j} + f^{(2,2)}_{i,j} + f^{(,4)}_{i,j}\right) + O(h^6)$$

$$i = 1(1)m \ , \ j = 1(1)n \qquad (2.3.17)$$

The system of difference equations (2.3.16) can be represented by a matrix equation as follows (Walton [25], page 139)

$$AV + WA - \frac{1}{12}\left(h^2 + k^2\right) WAV = G \qquad (2.3.18)$$

where

$$G = -F - \frac{h^2}{12} R - \frac{k^2}{12} S + B_1 + B_2 + \frac{1}{12}\left(\frac{1}{h^2} + \frac{1}{k^2}\right) C$$

with matrices F, A, B_1, B_2, V and W as in section (2.3.1) and

$$\left.\begin{aligned} R &= \left(f^{(2,)}_{i,j}\right) , \\ S &= \left(f^{(,2)}_{i,j}\right) , \\ C &= (c_{i,j}) , \end{aligned}\right\} \quad i = 1(1)m \ , \ j = 1(1)n \ .$$

$$c_{11} = u_{00} - 2u_{0,1} + u_{0,2} - 2u_{1,0} + u_{2,0}$$

$$c_{1,n} = u_{0,n-1} - 2u_{0,n} + u_{0,n+1} - 2u_{1,n+1} + u_{2,n+1}$$

$$c_{m,1} = u_{m-1,0} - 2u_{m,0} + u_{m+1,0} - 2u_{m+1,1} + u_{m+1,2}$$

$$c_{m,n} = u_{m-1,n} - 2u_{m,n} + u_{m+1,n-1} - 2u_{m+1,n} + u_{m+1,n+1}$$

$$c_{1,j} = u_{0,j-1} - 2u_{0,j} + u_{0,j+1} \ , \ j = 2(1)n - 1$$

$$c_{m,j} = u_{m+1,j-1} - 2u_{m+1,j} + u_{m+1,j+1} \ , \ j = 2(1)n - 1$$

$$c_{i,1} = u_{i-1,0} - 2u_{i,0} + u_{i+1,0} \ , \ i = 2(1)m - 1$$

$$c_{i,n} = u_{i-1,n+1} - 2u_{i,n+1} + u_{i+1,n+1} \ , \ i = 2(1)m - 1$$

$$c_{i,j} = 0 \quad \text{for} \quad i = 2(1)m - 1 \quad \text{and} \quad j = 2(1)n - 1 \ .$$

For Laplace's equation the coefficients of h^2 and k^4 in equation (2.3.17) vanish. Hence this nine-point formula is a more accurate finite difference approximation of Laplace's equation for $h = k$ (Fox [11], page 261, Smith [19], page 143). But for $h \neq k$,

the finite difference approximation of Laplace's equation is given by (2.3.16) which indicates a local truncation error of $O(h^4) + O(k^4)$ since cancellation does not occur as in (2.3.17).

2.3.7 A THIRTEEN-POINT FORMULA.

Following the analysis given in section (2.3.2) a thirteen-point formula using central differencing can be obtained by taking the first three terms from equation (2.2.19) and its y-analogue.

Addition and Taylor's series expansion give:

$$\frac{1}{h^2}\left(\delta_x^2 - \frac{1}{12}\delta_x^4 + \frac{1}{90}\delta_x^6\right)u_{i,j} + \frac{1}{k^2}\left(\delta_y^2 - \frac{1}{12}\delta_y^4 + \frac{1}{190}\delta_y^6\right)u_{i,j}$$

$$= u_{i,j}^{(2,)} + u_{i,j}^{(,2)} + \frac{1}{560}h^6 u_{i,j}^{(8,)} + \frac{1}{560}k^6 u_{i,j}^{(,8)} + \ldots,$$

$$i = 1(1)m \ , \ j = 1(1)n \ . \qquad (2.3.18)$$

The difficulties mentioned in section (2.3.3) also arise in applying this formula at points on mesh-lines $x_0 + h$, $x_0 + 2h$, $x_{n+1} - h$, $x_{n+1} - 2h, y_0 + k$, $y_0 + 2k$, $y_{m+1} - k$, $y_{m+1} - 2k$. Proceeding as in section (2.3.4), the replacement for second derivatives at points on mesh-lines $x_0 + h$ and $y_0 + k$ may be obtained by taking the first four terms in equation (2.2.16) and its y-analogue; thus

$$u_{xx} \simeq \frac{1}{h^2}\left(\Delta_x^2 - \frac{1}{12}\Delta_x^4 + \frac{1}{12}\Delta_x^5 - \frac{13}{180}\Delta_x^6\right)u_{i,j-1} \ \ldots \qquad (2.3.19)$$

and

$$u_{yy} \simeq \frac{1}{k^2}\left(\Delta_y^2 - \frac{1}{12}\Delta_y^4 + \frac{1}{12}\Delta_y^5 - \frac{13}{180}\Delta_y^6\right)u_{i-1,j} \ \ldots \qquad (2.3.20)$$

Addition of equations (2.3.19) and (2.3.20) and subsequent Taylor's expansion yield,

$$\frac{1}{h^2}\left(\Delta_x^2 - \frac{1}{12}\Delta_x^4 + \frac{1}{12}\Delta_x^5 - \frac{13}{180}\Delta_x^6\right)u_{i,j-1} + \frac{1}{k^2}\left(\Delta_y^2 - \frac{1}{12}\Delta_y^4 + \frac{1}{12}\Delta_y^5 - \frac{13}{180}\Delta_y^6\right)u_{i-1,j} = u_{i,j}^{(2,)} + u_{i,j}^{(,2)} - \frac{11}{180}h^5 u_{i,j}^{(7,)} - \frac{11}{18}k^5 u_{i,j}^{(,7)} + \ldots,$$

$$i = 1(1)m \ , \ j = 1(1)n \ . \qquad (2.3.21)$$

A similar replacement for u_{xx} and u_{yy} for the points on mesh-lines $x_0 + 2h$ and $y_0 + 2k$ can be made by using the first four terms in equation (2.2.17) and its y-analogue respectively to obtain:

$$\frac{1}{h^2}\left(\Delta_x^2 + \Delta_x^3 - \frac{1}{12}\Delta_x^4 + \frac{1}{90}\Delta_x^6\right)u_{i,j-2} + \frac{1}{k^2}\left(\Delta_y^2 + \Delta_y^3 - \frac{1}{12}\Delta_y^4 + \frac{1}{90}\Delta_y^6\right)u_{i-2,j} = u_{i,j}^{(2,)} + u_{i,j}^{(,2)} + \frac{1}{90}h^5 u_{i,j}^{(7,)} + \frac{1}{90}k^5 u_{i,j}^{(,7)} + \ldots$$

$$i = 1(1)m \ , \ j = 1(1)n \qquad (2.3.22)$$

Approximations similar to (2.3.21) and (2.3.22) involving backward differencing for the points near the other boundaries can also be constructed as above.

The discretization of equation (1.2.2) using the formula (2.3.18), (2.3.21) and (2.3.22) for the respective internal points will produce a matrix equation of the form (1.3.3) where

$$V = \frac{1}{180h^2}\begin{pmatrix} 147 & -228 & 27 & -2 & & & & & \\ 225 & 420 & -270 & 27 & . & & & & \\ -470 & -200 & 490 & -270 & . & . & & & \\ 285 & -15 & -270 & 490 & . & . & . & & \\ -93 & 12 & 27 & -270 & . & . & . & -2 & 13 \\ 13 & -2 & -2 & 27 & . & . & . & 12 & -93 \\ & & & -2 & . & . & . & -15 & 285 \\ & & & & . & . & . & -200 & -470 \\ & & & & & . & . & 420 & 225 \\ & & & & & & . & -228 & 147 \end{pmatrix}_{n\times n},$$

$$W = \frac{1}{180k^2}\begin{pmatrix} 147 & 225 & -470 & 285 & -93 & 13 & & & & & \\ -228 & 420 & -200 & -15 & 12 & -2 & & & & & \\ 27 & -270 & 490 & -270 & 27 & -2 & & & & & \\ -2 & 27 & -270 & 490 & -270 & 27 & -2 & & & & \\ & . & . & . & . & . & . & . & & & \\ & & . & . & . & . & . & . & . & & \\ & & & . & . & . & . & . & . & . \\ & & & & -2 & 12 & -15 & -200 & 420 & -228 \\ & & & & 13 & -93 & 285 & -470 & 225 & 147 \end{pmatrix}_{m\times m},$$

$$G = -F + B_1 + B_2$$

with

$$B_1 = \frac{1}{180h^2}\begin{bmatrix} 137u_{1,0} & -13u_{1,0} & 2u_{1,0} & 0 & 2u_{1,m+1} & -13u_{1,m+1} & 137u_{1,m+1} \\ 137u_{2,0} & -13u_{2,0} & 2u_{2,0} & & 2u_{2,n+1} & -13u_{2,m+1} & 137u_{2,m+1} \\ \vdots & \vdots & \vdots & 0 & \vdots & \vdots & \vdots \\ 137u_{n,0} & -13u_{n,0} & 2u_{n,0} & 0 & 2u_{n,m+1} & -13u_{n,m+1} & 137u_{n,m+1} \end{bmatrix},$$

$$B_2 = \frac{1}{180k^2}\begin{bmatrix} 137u_{0,1} & 137u_{0,2} & 137u_{0,3} & \cdots & 137u_{0,m} \\ -13u_{0,1} & -13u_{0,2} & -13u_{0,3} & \cdots & -13u_{0,m} \\ 2u_{0,1} & 2u_{0,2} & 2u_{0,3} & \cdots & 2u_{0,m} \\ 0 & & 0 & & 0 \\ 2u_{n+1,1} & 2u_{n+1,2} & 2u_{n+1,3} & \cdots & 2u_{n+1,m} \\ -13u_{n+1,1} & -13u_{n+1,2} & -13u_{n+1,3} & \cdots & -13u_{n+1,m} \\ 137u_{n+1,1} & 137u_{n+1,2} & 137u_{n+1,3} & \cdots & 137u_{n+1,m} \end{bmatrix}$$

The only non-zero elements of $B_1 + B_2$ occur in the first, second, third, last, second to last, and third to last rows and columns.

It may be noted that the matrices V and W become less sparse as more points enter into the difference formulae.

2.3.8 A MODIFIED THIRTEEN-POINT FORMULA.

It appears from formula (2.3.21) and (2.3.22) that the accuracy near the boundary is of order $O(h^5) + O(k^5)$ due to the use of one-sided (forward or backward) differencing involving the same number of points as in the central differencing of the function at the internal points. In order to have the same order for all the truncation error terms, a technique similar to that discussed in section (2.3.5) can be applied using equations (2.2.16) and (2.2.17). The stencil thus formulated for discretizing equation (1.2.2) leads to the matrix equation (1.3.3) where

$$180h^2V = Z$$

and

$$Z = \begin{pmatrix}
70 & -214 & 27 & -2 & & & & & & \\
486 & 378 & -270 & 27 & . & & & & & \\
-855 & -130 & 490 & -270 & . & . & & & & \\
670 & -85 & -270 & 490 & . & . & . & & & \\
-324 & 54 & 27 & -270 & . & . & . & 2 & -11 & \\
90 & -16 & -2 & 27 & . & . & . & -16 & 90 & \\
-11 & 2 & & -2 & . & . & . & 54 & -324 & \\
 & & & & . & . & . & -85 & 670 & \\
 & & & & & & & -130 & -855 & \\
 & & & & & & . & 378 & 486 & \\
 & & & & & & . & -214 & 70 &
\end{pmatrix}_{n\times n},$$

$$180k^2W = Z_2 \quad \text{and}$$

$$Z_2 = \begin{bmatrix} 70 & 486 & -855 & 670 & -324 & 90 & -11 & & & & \\ -214 & 378 & -130 & -85 & 54 & -16 & 2 & & & & \\ 27 & -270 & 490 & -270 & 27 & -2 & & & & & \\ -2 & 27 & -270 & 490 & -270 & 27 & -2 & & & & \\ & . & . & . & . & . & . & . & & & \\ & & . & . & . & . & . & . & . & & \\ & & & . & . & . & . & . & . & . & \\ & & & 2 & -16 & 54 & -85 & -130 & 378 & -214 \\ & & & -11 & 90 & -324 & 670 & -855 & 486 & 70 \end{bmatrix}_{m\times m},$$

and matrices A and F are as before. The matrices B_1 and B_2 are also as in section (2.3.7) except that the coefficient of $u_{i,0}$ and $u_{i,n+1}$ $(i = 1(1)m)$ in the first and last columns of B_1 as well as the coefficients of $u_{0,j}$ and $u_{m+1,j}$ $(j = 1(1)n)$ in the first and last rows of B_1 are now 126 rather than 137 and the coefficients of a $u_{i,0}$ and $u_{i,n+1}$ $(i = 1(1)m)$ in the second and one but last columns of B_1 as well as the coefficients of $u_{0,j}$ and $u_{m+1,j}$ $(j = 1(1)n)$ in the second and one but last rows of B_2 are -11 instead of -13.

2.3.9 A SEVENTEEN-POINT FORMULA.

Following the analysis given in the preceding sections a seventeen-point stencil involving central differencing can be formulated by taking the first four terms from equation (2.2.19). Adding to its

y-analogue and expanding by Taylor's series produces:

$$\frac{1}{h^2}\left(\delta_x^2 - \frac{1}{12}\delta_x^4 + \frac{1}{90}\delta_x^6 - \frac{1}{560}\delta_x^8\right)u_{i,j} + \frac{1}{k^2}\left(\delta_y^2 - \frac{1}{12}\delta_y^4 + \frac{1}{90}\delta_y^6 - \frac{1}{560}\delta_y^8\right)u_{i,j}$$

$$= u_{i,j}^{(2,)} + u_{i,j}^{(,2)} - \frac{h^8}{3150}u_{i,j}^{(10,)} - \frac{k^8}{3150}u_{i,j}^{(,10)} + \ldots$$

This indicates a local truncation error of order $O(h^8) + O(k^8)$.

A similar analysis as in the previous cases can be carried out for adjustment at the boundaries of this stencil. Adjustments are required at points along the mesh-lines: $x_0 + h$, $x_0 + 2h$, $x_0 + 3h$, $x_{n+1} - 3h$, $x_{n+1} - 2h$, $x_{n+1} - h$, and the corresponding mesh-lines parallel to the x-axis. The approximations for the second derivative at points along $x_0 + h$, $x_0 + 2h$ and $x_0 + 3h$ obtained from equation (2.2.16), (2.2.17) and (2.2.18) are

$$u_{xx} \simeq \frac{1}{h^2}\left(\Delta_x^2 - \frac{1}{12}\Delta_x^4 + \frac{1}{12}\Delta_x^5 - \frac{13}{180}\Delta_x^6 + \frac{11}{180}\Delta_x^7 - \frac{29}{560}\Delta_x^8\right)u_{i,j-1} ,$$

$$u_{xx} \simeq \frac{1}{h^2}\left(\Delta_x^2 + \Delta_x^3 - \frac{1}{12}\Delta_x^4 + \frac{1}{90}\Delta_x^6 - \frac{1}{90}\Delta_x^7 + \frac{47}{5040}\Delta_x^8\right)u_{i,j-2} ,$$

and

$$u_{xx} \simeq \frac{1}{h^2}\left(\Delta_x^2 + 2\Delta_x^3 + \frac{11}{12}\Delta_x^4 - \frac{1}{12}\Delta_x^5 + \frac{1}{90}\Delta_x^6 - \frac{1}{560}\Delta_x^8\right)u_{i,j-3} .$$

Adding to these the respective y-approximations and subsequent Taylor's series expansions, indicates, as in section (2.3.4), that the error will vary from order $O(h^7) + O(k^7)$ to $O(h^8) + O(k^8)$. For points near the other boundaries, a truncation error of this order also occurs.

For equation (1.2.2), this formula gives rise to the matrix equation (1.3.3) where V, W are given in pages 42 and 43. The matrices A and F are as before. The matrices B_1 and B_2 are given in pages 44 and 45.

$$V = \frac{1}{5040h^2}\begin{pmatrix}
-128 & -5616 & 684 & -128 & 9 & & & & & & & \\
20916 & 9268 & -7308 & 1008 & -128 & . & & & & & & \\
-38556 & -1008 & 13216 & -8064 & 1008 & . & . & & & & & \\
37030 & -5670 & -6930 & 14350 & -8064 & . & . & . & & & & \\
-23688 & 4144 & 252 & -8064 & 14350 & . & . & . & & & & \\
9828 & -1764 & 196 & 1008 & -8064 & . & . & . & & & & \\
-2396 & 432 & -72 & -128 & 1008 & . & . & . & & & & \\
261 & -47 & 9 & 9 & -128 & . & . & . & 9 & -47 & 261 \\
 & & & & 9 & . & . & . & -72 & 432 & -2396 \\
 & & & & & . & . & . & 196 & -1764 & 9828 \\
 & & & & & . & . & . & 252 & 4144 & -23688 \\
 & & & & & & . & . & -6930 & -5670 & 37030 \\
 & & & & & & & & 13216 & -1008 & -38556 \\
 & & & & & & & . & -7308 & 9268 & 20916 \\
 & & & & & & & & 684 & -5616 & -128
\end{pmatrix}_{n\times n} .$$

$$W = \frac{1}{5040k^2}\begin{pmatrix}
-128 & 20916 & -38556 & 37030 & -23688 & 9828 & -2396 & 261 & & & & \\
-5616 & 9268 & -1008 & -5670 & 4144 & -1764 & 432 & -47 & & & & \\
684 & -7308 & 13216 & -6930 & 252 & 196 & -72 & 9 & & & & \\
-128 & 1008 & -8064 & 14350 & -8064 & 1008 & -128 & 9 & & & & \\
9 & -128 & 1008 & -8064 & 14350 & -8064 & 1008 & -128 & 9 & & & \\
 & . & . & . & . & . & . & . & . & . & & \\
 & & . & . & . & . & . & . & . & . & . & \\
 & & . & . & . & . & . & . & . & . & . & \\
 & & & & 9 & -72 & 196 & 252 & -6930 & 13216 & -7308 & 684 \\
 & & & & -47 & 432 & -1764 & 4144 & -5670 & -1008 & 9268 & -5616 \\
 & & & & 261 & -2396 & 9828 & -23688 & 37030 & -38556 & 20916 & -128
\end{pmatrix}_{m\times m},$$

$$B_1 = \frac{1}{5040h^2}\begin{pmatrix} 3267u_{1,0} & -261u_{1,0} & 47u_{1,0} & -9u_{1,0} & & -9u_{1,m+1} & 47u_{1,m+1} & -261u_{1,m+1} & 3267u_{1,m+1} \\ 3267u_{2,0} & -261u_{2,0} & 47u_{2,0} & -9u_{2,0} & & -9u_{2,m+1} & 47u_{2,m+1} & -261u_{2,m+1} & 3267u_{2,m+1} \\ & & & & & \cdot & & & \cdot \\ & & & & 0 & & & & \cdot \\ & & & & & & & & \cdot \\ 3267u_{n,0} & -261u_{n,0} & 47u_{n,0} & -9u_{n,0} & & -9u_{n,m+1} & 47u_{n,m+1} & -261u_{n,m+1} & 3267u_{n,m+1} \end{pmatrix}$$

$$B_2 = \frac{1}{5040k^2}\begin{pmatrix}
3267u_{0,1} & 3267u_{0,2} & 3267u_{0,3} & \cdots & 3267u_{0,m} \\
-261u_{0,1} & -261u_{0,2} & -261u_{0,3} & \cdots & -261u_{0,m} \\
47u_{0,1} & 47u_{0,2} & 47u_{0,3} & \cdots & 47u_{0,m} \\
-9u_{0,1} & -9u_{0,2} & -9u_{0,3} & \cdots & -9u_{0,m} \\
 & & 0 & & \\
-9u_{n+1,1} & -9u_{n+1,2} & -9u_{n+1,3} & \cdots & -9u_{n+1,m} \\
47u_{n+1,1} & 47u_{n+1,2} & 47u_{n+1,3} & \cdots & 47u_{n+1,m} \\
-261u_{n+1,1} & -261u_{n+1,2} & -261u_{n+1,3} & \cdots & -261u_{n+1,m} \\
3267u_{n+1,1} & 3267u_{n+1,2} & 3267u_{n+1,3} & \cdots & 3267u_{n+1,m}
\end{pmatrix}$$

2.3.10 A TWENTY ONE-POINT FORMULA.

Using the same approach as previously, the matrices on pages 47 through 50 are obtained in discretizing equation (1.2.2) by a twenty one-point formula. The local truncation error due to the use of central differencing formula is $O(h^{10}) + O(k^{10})$ and that due to the use of one-sided differencing for the same number of points as in central differencing is $O(h^9) + O(k^9)$. However, in both the seventeen and twenty one-point formula the order of accuracy at all points over the solution region could be made identical by improving the accuracy near the boundaries as in the procedure outlined in section (2.3.5).

It can be noted that the number of non-zero rows and columns in B_1 and B_2 increases with the increase in number of points used in the formula.

$$V = \frac{1}{25200h^2}\begin{pmatrix}
-20295 & -24840 & 2645 & -560 & 125 & -8 & & & & & & \\
188010 & 32615 & -33255 & 4680 & -1000 & 125 & . & & & & & \\
-401880 & 29280 & 57860 & -39360 & 6000 & -1000 & . & . & & & & \\
527660 & -84420 & -21210 & 70070 & -42000 & 6000 & . & . & . & & & \\
-501354 & 83216 & -13734 & -38304 & 73766 & -42000 & . & . & . & -8 & 37 & -153 & 962 \\
344820 & -56910 & 12530 & 3360 & -42000 & 73766 & . & . & . & 80 & -415 & 1720 & -10735 \\
-167560 & 27360 & -6420 & 320 & 6000 & -42000 & . & . & . & -315 & 2115 & -8830 & 54630 \\
54630 & -8830 & 2115 & -315 & -1000 & 6000 & . & . & . & 320 & -6420 & 27360 & -167560 \\
-10735 & 1720 & -415 & 80 & 125 & -1000 & . & . & . & 3360 & 12530 & -56910 & 344820 \\
962 & -153 & 37 & -8 & -8 & 125 & . & . & . & -38304 & -13734 & 83216 & -501354 \\
 & & & & & -8 & . & . & . & 70070 & -21210 & -84420 & 527660 \\
 & & & & & & . & . & . & -39360 & 57860 & 29280 & -401880 \\
 & & & & & & & . & . & 4680 & -33255 & 32615 & 188010 \\
 & & & & & & & & . & -560 & 2645 & -24840 & -20295
\end{pmatrix}_{n\times n}$$

$25200k^2W = Z$, where

$$Z = \begin{pmatrix}
-20295 & 188010 & -401880 & 527660 & -501354 & 344820 & -167560 & 54630 & -10735 & 962 & & \\
-24840 & 32615 & 29280 & -84420 & 83216 & -56910 & 27360 & -8830 & 1720 & -153 & & \\
2645 & -33255 & 57860 & -21210 & -13734 & 12530 & -6420 & 2115 & -415 & 37 & & \\
-560 & 4680 & -39360 & 70070 & -38304 & 3360 & 320 & -315 & 80 & -8 & & \\
125 & -1000 & 6000 & -42000 & 73766 & -42000 & 6000 & -1000 & 125 & -8 & & \\
-8 & 125 & -1000 & 6000 & -42000 & 73766 & -42000 & 6000 & -1000 & 125 & -8 & \\
 & . & . & . & . & . & . & . & . & . & . & \\
 & & . & . & . & . & . & . & . & . & & \\
 & & . & . & . & . & . & . & . & . & & \\
 & & -8 & 80 & -315 & 320 & 3360 & -38304 & 70070 & -39360 & 4680 & -560 \\
 & & 37 & -415 & 2115 & -6420 & 12530 & -13734 & -21210 & 57860 & -33255 & 2645 \\
 & & -153 & 1720 & -8830 & 27360 & -56910 & 83216 & -84420 & 29280 & 32615 & -24840 \\
 & & 962 & -10735 & 54630 & -167560 & 344820 & -501354 & 527660 & -401880 & 188010 & -20295
\end{pmatrix}_{m \times m}$$

$$25200h^2 B_1 = Z \text{ , where}$$

$$Z = \begin{pmatrix} 14258u_{1,0} & -962u_{1,0} & 153u_{1,0} & -37u_{1,0} & 8u_{1,0} & \cdots & 8u_{1,m+1} & -37u_{1,m+1} & 153u_{1,m+1} & -962u_{1,m+1} & 14258u_{1,m+1} \\ 14258u_{2,0} & -962u_{2,0} & 153u_{2,0} & -37u_{2,0} & 8u_{2,0} & \cdots & 8u_{2,m+1} & -37u_{2,m+1} & 153u_{2,m+1} & -962u_{2,m+1} & 14258u_{2,m+1} \\ \vdots & & & & & & & & & & \\ 14258u_{n,0} & -962u_{n,0} & 153u_{n,0} & -37u_{n,0} & 8u_{n,0} & \cdots & 8u_{n,m+1} & -37u_{n,m+1} & 153u_{n,m+1} & -962u_{n,m+1} & 14258u_{n,m+1} \end{pmatrix}$$

$$B_2 = \frac{1}{25200k^2}$$

$$\begin{pmatrix}
14258u_{0,1} & 14258u_{0,2} & 14258u_{0,3} & \cdots & 14258u_{0,m} \\
-962u_{0,1} & -962u_{0,2} & -962u_{0,3} & \cdots & -962u_{0,m} \\
153u_{0,1} & 153u_{0,2} & 153u_{0,3} & \cdots & 153u_{0,m} \\
-37u_{0,1} & -37u_{0,2} & -37u_{0,3} & \cdots & -37u_{0,m} \\
8u_{0,1} & 8u_{0,2} & 8u_{0,3} & \cdots & 8u_{0,m} \\
 & & 0 & & \\
8u_{n+1,1} & 8u_{n+1,2} & 8u_{n+1,3} & \cdots & 8u_{n+1,m} \\
-37u_{n+1,1} & -37u_{n+1,2} & -37u_{n+1,3} & \cdots & -37u_{n+1,m} \\
153u_{n+1,1} & 153u_{n+1,2} & 153u_{n+1,3} & \cdots & 153u_{n+1,m} \\
-962u_{n+1,1} & -962u_{n+1,2} & -962u_{n+1,3} & \cdots & -962u_{n+1,m} \\
14258u_{n+1,1} & 14258u_{n+1,2} & 14258u_{n+1,3} & \cdots & 14258u_{n+1,m}
\end{pmatrix}$$

CHAPTER 3

SOME RECENT DEVELOPMENTS FOR THE DISCRETE SOLUTION OF ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS

3.1 INTRODUCTION.

Many physical problems require the solution of elliptic partial differential equations of the form (1.2.1). In solving such an equation by finite difference methods, one usually encounters a large system of linear algebraic equations, which in composite matrix formulation can be represented by

$$MX = Y \tag{3.1.1}$$

where M is an $n \times n$ matrix of block tridiagonal form, viz.

$$M = \begin{pmatrix} A_1 & C_1 & & & \\ B_2 & A_2 & C_2 & & \\ & \ddots & \ddots & \ddots & \\ & & & C_{n-1} & \\ & & B_n & A_n & \end{pmatrix} \tag{3.1.2}$$

The matrices A_i, B_i and C_i are of order p.

Define $\underline{x}_i$ to be the vector whose components comprise the i-th vertical line of the array X,

$$\underline{x}_i = \begin{pmatrix} X_{i1} \\ X_{i2} \\ \vdots \\ X_{ip} \end{pmatrix}, \quad 1 \leq i \leq n .$$

The block vector X can be written with x_i as components,

$$X = \begin{pmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \underline{x}_3 \\ \vdots \\ \underline{x}_n \end{pmatrix} .$$

The vector Y has the same block form as X.

The usual Gaussian elimination method is not always satisfactory for such a system (Forsythe and Wasow [10], § 21.2-3). In the sequel some recent fast direct methods for the solution of system of the form (3.1.1) are reviewed.

3.2 THE CYCLIC ODD-EVEN REDUCTION AND FACTORIZATION ALGORITHM.

This method is taken from Buzbee et al [6].

Consider the system of equation (3.1.1) where

$$M = \begin{bmatrix} A & -T & & & \\ -T & A & -T & & \\ & \ddots & \ddots & \ddots & \\ & & & -T & \\ & & -T & A & \end{bmatrix} \qquad (3.2.1)$$

with the assumptions,

(i) $TA = AT$, A and T are of order p ,

(ii) $n = 2^{k+1} - 1$ where k is any positive integer

Then the system (3.1.1) with (3.2.1) may be written as

$$\begin{aligned} A\underline{x}_1 - T\underline{x}_2 &= \underline{y}_1 \ , \\ -T\underline{x}_{j-1} + A\underline{x}_j - T\underline{x}_{j+1} &= \underline{y}_j \ , \ j = 2, 3, \ldots, n-1 \ , \\ -T\underline{x}_{n-1} + A\underline{x}_n &= \underline{y}_n \ . \end{aligned} \qquad (3.2.2)$$

Multiplying the 1st and 3rd equations by T and adding them to A times the 2nd, multiplying the 3rd and 5th equations by T and adding them to A times the 4th equation, and continuing in this fashion, the system in (3.2.2) can be reduced to two lower order systems

$$\begin{bmatrix} A^{(1)} & -T^{(1)} & & & \\ -T^{(1)} & A^{(1)} & -T^{(1)} & & \\ & \ddots & \ddots & \ddots & \\ & & -T^{(1)} & A^{(1)} & -T^{(1)} \\ & & & -T^{(1)} & A^{(1)} \end{bmatrix} \begin{bmatrix} \underline{x}_2 \\ \underline{x}_4 \\ \cdot \\ \cdot \\ \cdot \\ \underline{x}_{n-1} \end{bmatrix} = \begin{bmatrix} T\underline{y}_1 + A\underline{y}_2 + T\underline{y}_3 \\ T\underline{y}_3 + A\underline{y}_4 + T\underline{y}_5 \\ \cdot \\ \cdot \\ \cdot \\ T\underline{y}_{n-2} + A\underline{y}_{n-1} + T\underline{y}_n \end{bmatrix} \tag{3.2.3}$$

where

$$A^{(1)} = A^2 - 2T^2 \tag{3.2.4}$$

$$T^{(1)} = T^2 ,$$

and

$$\begin{bmatrix} A & 0 & & & \\ 0 & A & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & 0 & A & 0 \\ & & & 0 & A \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_3 \\ \cdot \\ \cdot \\ \cdot \\ \underline{x}_n \end{bmatrix} = \begin{bmatrix} \underline{y} + T\underline{x}_2 \\ \underline{y}_3 + T\underline{x}_2 + T\underline{x}_4 \\ \cdot \\ \cdot \\ \cdot \\ \underline{y}_n + T\underline{x}_{n-1} \end{bmatrix} \tag{3.2.5}$$

Since $n = 2^{k+1} - 1$, the new system in (3.2.3) is of block dimension $(2^k - 1)$ involving $\underline{x}$ with even indices, and the system in (3.2.5) is of block dimension 2^k. The system (3.2.3) is also block tridiagonal and of the form (3.2.1). So the same reduction can be repeated until only one block remains. The process of reducing the system of equations in the above fashion is called cyclic reduction.

Define sequences,

$$A^{(0)} = A$$

$$T^{(0)} = T$$

$$\underline{y}_j^{(0)} = \underline{y}_j \ , \ j = 1, 2, \ldots, n \ ,$$

$$A^{(r+1)} = -2(T^{(r)})^2 + (A^{(r)})^2 \ .$$

for (3.2.6)

$$r = 0, 1, 2, \ldots k \ ,$$

$$T^{(r+1)} = (T^{(r)})$$

$$\underline{y}_j^{(r+1)} = T^{(r)}\underline{y}_{j-2h}^{(r)} + A^{(r)}\underline{y}_j^{(r)} + T^{(r)}\underline{y}_{j+2h}^{(r)} \ ,$$

$$j = i \cdot 2h \ , \ i = 1, 2, 3, \ldots, 2^{k+1-r} - 1 \ ,$$

where

$$h = 2^{r-1} \ .$$

After r reduction the new system of equations is

$$\begin{pmatrix} A^{(r)} & -T^{(r)} & & & & \\ -T^{(r)} & A^{(r)} & -T^{(r)} & & & \\ & \ddots & \ddots & \ddots & & \\ & & & & -T^{(r)} & \\ & & & -T^{(r)} & A^{(r)} & \end{pmatrix} \begin{pmatrix} \underline{x}_{2h} \\ \underline{x}_{2\cdot 2h} \\ \vdots \\ \underline{x}_{j\cdot 2h} \\ \vdots \end{pmatrix} = \begin{pmatrix} \underline{y}_{2h}^{(r)} \\ \underline{y}_{2\cdot 2h}^{(r)} \\ \vdots \\ \underline{y}_{j\cdot 2h}^{(r)} \\ \vdots \end{pmatrix}$$

and

$$\begin{bmatrix} A^{(r-1)} & 0 & & & \\ 0 & A^{(r-1)} & 0 & & \\ & \ddots & \ddots & \ddots & 0 \\ & & 0 & & A^{(r-1)} \end{bmatrix} \begin{bmatrix} \underline{x}_h \\ \underline{x}_{3h} \\ \vdots \\ \underline{x}_{(2j-1)h} \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} \underline{y}_h^{(r-1)} + T\underline{x}_{2h}^{(r-1)} \\ \underline{y}_{3h}^{(r-1)} + T\left(\underline{x}_{2\cdot 2h}^{(r-1)} - \underline{x}_{2h}^{(r-1)}\right) \\ \underline{y}_{(2j-1)}^{(r-1)} + T\left(\underline{x}_{j\cdot 2h}^{(r-1)} - \underline{x}_{(j-1),2h}^{(r-1)}\right) \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \end{bmatrix}$$

which are of block dimensions $2^{k+1-r} - 1$ and 2^{k+1-r} respectively. After k steps (3.2.2) reduces to the single $p \times p$ matrix equation

$$A^{(k)}\underline{x}_{2^k} = \underline{y}_{2^k}^{(k)} \tag{3.2.7}$$

From (3.2.4), it follows that $A^{(1)}$ is a polynomial of degree 2 in A and T. By induction it can be shown that $A^{(r)}$ in (3.2.6) is a polynomial of degree $\leq r$ in A and T.

A linear factorization of $A^{(r)}$ produces (Buzbee et al [6])

$$A^{(r)} = \prod_{j=1}^{2^r} (A - \underline{x}_j(r)T) , \tag{3.2.8}$$

where

$$\underline{x}_j(r) = 2 \cos \frac{2j - 1}{2^{r+1}} \pi , \quad j = 1, 2, \ldots, 2^r .$$

Set

$$G_j^{(k)} = A - \underline{x}_j(k)T$$

Then, to solve (3.2.7) put

$$\underline{z}_1 = - \underline{y}_{2^k}^{(k)}$$

and respectively solve

$$G_j^{(k)} \underline{z}_{j+1} = \underline{z}_j \quad \text{for} \quad j = 1, 2, \ldots, 2^k \ .$$

Thus,

$$\underline{x}_{2^k} = \underline{z}_{2^k+1} \ .$$

The numerical calculation of $\underline{y}_j^{(r)}$ in (3.2.6) is subject to considerable round off errors in many cases of interest (Buzbee et al [6]). Buzbee et al give Buneman variants of cyclic odd-even reduction and factorization (CORF).

3.3 BUNEMAN VARIANT TWO OF CORF.

The difference between the Buneman algorithm and CORF algorithm lies in the way that the right hand side is calculated at each stage of the reduction.

Assume,

$$T = I_p \ , \tag{3.3.1}$$

the identity matrix of order p in the system (3.2.2).

The Buneman variant two of CORF consists of three phases: preprocessing, reduction and backsubstitution.

The matrices $A^{(r)}$ are computed from $A^{(0)} = A$ using the recurrence

$$A^{(r)} = (A^{(r-1)})^2 - 2I$$

and then using identities in (3.2.8).

The vectors $\underline{q}_j^{(r)}$ are computed starting with

$$\underline{q}_j^{(0)} = \underline{y}_j , \qquad j = 1, 2, \ldots, n$$

and

$$\underline{q}_j^{(1)} = \underline{q}_{j-1}^{(0)} + \underline{q}_{j+1}^{(0)} + 2A^{-1}\underline{q}_j^{(0)} , \qquad j = 1, 2, \ldots, n-1 .$$

The remaining $\underline{q}_j^{(r)}$ are determined for $r = 2, \ldots, k$ and

$$j = 2^r , 2\cdot 2^r , \ldots, 2^{k+1} - 2^r$$

using

$$\underline{q}_j^{(r)} = \underline{q}_{j-2h}^{(r-1)} - \underline{q}_{j-h}^{(r-2)} + \underline{q}_j^{(r-1)} - \underline{q}_{j+h}^{(r-2)} + \underline{q}_{j+2h}^{(r-1)}$$

$$+ (A^{(r-1)})^{-1}\left[-\underline{q}_{j-3h}^{(r-2)} + \underline{q}_{j-2h}^{(r-1)} - \underline{q}_{j-h}^{(r-2)} + 2\underline{q}_j^{(r-1)}\right]$$

$$- \underline{q}_{j+h}^{(r-2)} + \underline{q}_{j+2h}^{(r-1)} - \underline{q}_{j+3h}^{(r-2)}$$

where,

$$h = 2^{r-2} .$$

Define

$$\underline{q}_{j-\frac{1}{2}}^{(-1)} + \underline{q}_{j+\frac{1}{2}}^{(-1)} + \underline{q}_j^{(0)} = \underline{x}_0 = \underline{x}_n = 0 \ ,$$

then the solution vectors $\underline{x}_j$ are given for $r = k, \ldots, 0$ and

$$j = 2^r \ , \ 3 \cdot 2^r \ , \ \ldots, \ 2^{k+1} - 2^r$$

by

$$\underline{x}_j = \frac{1}{2}\left(\underline{q}_j^{(r)} - \underline{q}_{j-2h}^{(r-1)} - \underline{q}_{j+2h}^{(r-1)}\right)$$

$$+ (A^{(r)})^{-1}\left(\underline{q}_j^{(r)} + \underline{x}_{j-4h} + \underline{x}_{j+4h}\right)$$

Here $\underline{x}_{j+4h}$ and $\underline{x}_{j-4h}$ are computed at a previous step in the back-substitution.

All matrix computation can be performed using the factored form of $A^{(r)}$.

Define an operation as consisting of a multiplication or division plus an addition or subtraction and considering only those computations which contribute to the asymptotic count, then the operation count for CORF for an $n \times n$ mesh in (Dorr [9])

$$\frac{9}{2} n^2 \log_2 n \ ,$$

whereas for the Buneman variant two of CORF it is (Swarztrauber [22])

$$3n^2 \log_2 n \ .$$

3.4 THE MARCHING ALGORITHM.

This is taken from Bank and Rose [4].

Consider, for simplicity, $p = n$ in (3.1.1) with (3.2.1) and (3.3.1). Premultiplication of this linear system by an $n^2 \times n^2$ permutation matrix P yields the partitioned system,

$$\left[\begin{array}{ccccccc|c} -I & A & -I & & & & & \\ & -I & A & -I & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & -I & A & -I & & \\ & & & & -I & A & & \\ \hline A & -I & & & & & 0 & \end{array}\right] \left[\begin{array}{c} \underline{x}_1 \\ \underline{x}_2 \\ \vdots \\ \underline{x}_{n-1} \\ \hline \underline{x}_n \end{array}\right] = \left[\begin{array}{c} \underline{y}_2 \\ \underline{y}_3 \\ \vdots \\ \underline{y}_n \\ \hline \underline{y}_1 \end{array}\right] \qquad (3.4.1)$$

For convenience, this may be written as

$$PM = \left[\begin{array}{c|c} B & C \\ \hline R & 0 \end{array}\right] \left[\begin{array}{c} \tilde{\underline{x}} \\ \hline \underline{x}_n \end{array}\right] = \left[\begin{array}{c} \tilde{\underline{y}} \\ \hline \underline{y}_1 \end{array}\right]$$

where the symbols B, C, R, $\tilde{\underline{x}}$, $\tilde{\underline{y}}$ are used to denote their corresponding submatrices in (3.4.1). Using the modified Chebyshev polynomials,

$$S_0(\alpha) = 1,\ S_1(\alpha) = \alpha,\ S_\ell(\alpha) = \alpha S_{\ell-1}(\alpha) - S_{\ell-2}(\alpha),\ \ell \geq 2$$

the factorization of PM is

$$\left[\begin{array}{c|c} B & C \\ \hline R & 0 \end{array}\right] = \left[\begin{array}{c|c} I & 0 \\ \hline RB^{-1} & I \end{array}\right]\left[\begin{array}{c|c} B & C \\ \hline 0 & S_n(A) \end{array}\right] \tag{3.4.3}$$

Here

$$RB^{-1} = [-S_1(A), -S_2(A), \ldots, -S_{n-1}(A)]^T$$

The block solution of (3.4.3) is carried out as follows:

$$\left[\begin{array}{c|c} I & 0 \\ \hline RB^{-1} & I \end{array}\right]\left[\begin{array}{c} \tilde{v} \\ \hline v_n \end{array}\right] = \left[\begin{array}{c} \tilde{y} \\ \hline y_1 \end{array}\right] \tag{3.4.4}$$

and

$$\left[\begin{array}{c|c} B & C \\ \hline 0 & S_n(A) \end{array}\right]\left[\begin{array}{c} \tilde{x} \\ \hline x_n \end{array}\right] = \left[\begin{array}{c} \tilde{v} \\ \hline v_n \end{array}\right] \tag{3.4.5}$$

From (3.4.4),

$$\tilde{v} = \tilde{y} \tag{3.4.6}$$

and

$$v_n = y_1 - RB^{-1}\tilde{v} = y_1 - RB^{-1}\tilde{y} \tag{3.4.7}$$

Note that R and B are sparse whereas RB^{-1} is not, hence it is advantageous to solve first

$$B\tilde{\underline{w}} = \tilde{\underline{y}} \ , \ \tilde{w} = [\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_{n-1}] \tag{3.4.8}$$

and then

$$\underline{v}_n = \underline{y}_1 - R\tilde{\underline{w}} = -(A\underline{w}_1 - \underline{w}_2 - \underline{y}_1)$$

$$\equiv - \underline{w}_0 \tag{3.4.9}$$

Equation (3.4.5) yields,

$$B\tilde{\underline{x}} = \tilde{\underline{v}} - C\underline{x}_n \tag{3.4.10}$$

and

$$S_n(A)\underline{x}_n = \underline{v}_n$$

$$= - \underline{w}_0 \tag{3.4.11}$$

Computation of (3.4.11) can be simplified with the use of the identity (Bank and Rose [4]),

$$S_n(A) = \prod_{j=1}^{n} (A - r_n(j)I) \ ,$$

$$r_n(j) = 2 \cos \frac{\pi j}{n + 1} \ .$$

The algorithm (3.4.6) to (3.4.11) may be summarized as follows:

$$\underline{w}_{n-1} = -\underline{y}_n \qquad \text{(using (3.4.8))}$$

$$\underline{w}_{n-2} = A\underline{w}_{n-1} - \underline{y}_{n-1}$$

$$\underline{w}_{n-j} = A\underline{w}_{n-j-1} - \underline{w}_{n-j+2} - \underline{y}_{n-j+1} \ , \ 3 \le j \le n$$

$$\underline{z}_0 = \underline{w}_0$$

$$(A - r_n(j)I)\underline{z}_j = \underline{z}_{j-1} \ , \ 1 \le j \le n$$

$$\underline{x}_n = \underline{z}_n$$

$$\underline{x}_{n-1} = A\underline{x}_n - \underline{y}_n \qquad \text{(using (3.4.10))}$$

$$\underline{x}_{n-j} = A\underline{x}_{n-j+1} - \underline{x}_{n-j+2} - \underline{y}_{n-j+1} \ , \ 2 \le j \le n - 1$$

The asymptotic operation count for this algorithm is [4]

$$O(n^2 \log_2(n/k))$$

for an $n \times n$ mesh where

$$n = k2^{\ell} - 1 \ , \ \ell \ge 1 \ .$$

3.5 A DIRECT METHOD FOR THE DISCRETE SOLUTION OF SEPARABLE ELLIPTIC EQUATIONS.

The following is taken from Swarztrauber [21].

If equation (1.2.1) with Dirichlet or Neuman boundary conditions is discretized using the five-point formula a linear system as in equation (3.1.1) arise where

$$M = \begin{bmatrix} A_1 & C_1 & & & & \\ B_2 & A_2 & C_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & & B_{n-1} & A_{n-1} & C_{n-1} \\ & & & & B_n & A_n \end{bmatrix} \qquad (3.5.1)$$

and vectors X and Y are as in section (3.1.1). The block size n is assumed to be of the form $2^k - 1$. Each of the blocks of M in (3.5.1) is of order p and are of the following form,

$$B_i = b_i I \qquad (3.5.2)$$

$$A_i = A + a_i I \qquad (3.5.3)$$

$$C_i = c_i I \qquad (3.5.4)$$

where b_i, a_i, c_i are scalars and the matrix A is tridiagonal.

The reduction of the system is carried out as follows: eliminate the unknowns $\underline{x}_{i-1}$, $\underline{x}_{i+1}$ between the three block equations corresponding to block rows $i - 1$, i, $i + 1$. Multiplying these rows by matrices $\overline{\Theta}_i$, $\overline{\phi}_i$, $\overline{\psi}_i$ (yet to be determined) and add, then

$$\begin{aligned} &\overline{\Theta}_i B_{i-1}\underline{x}_{i-2} + (\overline{\Theta}_i A_{i-1} + \overline{\phi}_i B_i)\underline{x}_{i-1} + (\overline{\Theta}_i C_{i-1} + \overline{\phi}_i A_i + \overline{\psi}_i B_{i+1})\underline{x}_i \\ &\quad + (\overline{\phi}_i C_i + \overline{\psi}_i A_{i+1})\underline{x}_{i+1} + \overline{\psi}_i C_{i+1}\underline{x}_{i+2} \\ &\quad = \overline{\Theta}_i \underline{y}_{i-1} + \overline{\phi}_i \underline{y}_i + \overline{\psi}_i \underline{y}_{i+1} \end{aligned} \qquad (3.5.5)$$

In order to eliminate $\underline{x}_{i-1}$, $\underline{x}_{i+1}$, choose $\bar{\theta}_i$, $\bar{\phi}_i$, $\bar{\psi}_i$ such that

$$\bar{\theta}_i A_{i-1} + \bar{\phi}_i B_i = 0 \tag{3.5.6}$$

$$\bar{\phi}_i C_i + \bar{\psi}_i A_{i+1} = 0 \tag{3.5.7}$$

Since the matrices A_i, B_i, C_i commute, this system has an infinite number of solutions. For simplicity select,

$$\bar{\theta}_i = A_{i+1} B_i \tag{3.5.8}$$

$$\bar{\phi}_i = -A_{i-1} A_{i+1} \tag{3.5.9}$$

$$\bar{\psi}_i = C_i A_{i-1} \tag{3.5.10}$$

Substitution of these equations in (3.5.5) yields

$$B_i^{(1)} \underline{x}_{i-2} + A_i^{(1)} \underline{x}_i + C_i^{(1)} \underline{x}_{i+2} = y_i^{(1)} \tag{3.5.11}$$

where

$$B_i^{(1)} = B_i A_{i+1} B_{i-1} \tag{3.5.12}$$

$$A_i^{(1)} = B_i A_{i+1} C_{i-1} - A_{i-1} A_{i+1} A_i + C_i A_{i-1} B_{i+1} \tag{3.5.13}$$

$$C_i^{(1)} = C_i A_{i-1} C_{i+1} \tag{3.5.14}$$

and

$$\underline{y}_i^{(1)} = B_i A_{i+1} \underline{y}_{i-1} - A_{i-1} A_{i+1} \underline{y}_i + C_i A_{i-1} \underline{y}_{i+1} \tag{3.5.15}$$

The system in (3.5.11) is block tridiagonal and has about half $(2^{k-1} - 1)$ of the unknown vectors $\underline{x}_i$ for $i = 2, 4, \ldots, 2^k - 2$. The general algorithm is as follows:

Define $b_1 = c_n = 0$ and for $i = 1, 2, \ldots, n$,

$$B_i^{(0)} = b_i I \tag{3.5.16}$$

$$A_i^{(0)} = A + a_i I \tag{3.5.17}$$

$$C_i^{(0)} = c_i I \tag{3.5.18}$$

and

$$\underline{y}_i^{(0)} = \underline{y}_i \tag{3.5.19}$$

From (3.5.17) and (3.5.13) it can be observed that $A_i^{(0)}$ is linear in A and $A_i^{(1)}$ is a cubic polynomial in A. The degree of $A_i^{(r)}$ would triple at each step of reduction. Therefore to reduce the degree of the polynomial and consequently the amount of computation, define for $r = 0, 1, \ldots, k - 2$ and

$$i = 4h, 2 \cdot 4h, \ldots, (2^{k-r-1} - 1) \cdot 4h,$$

where

$$h = 2^{r-1},$$

$$B_i^{(r+1)} = (G_i^{(r+1)})^{-1} B_i^{(r)} A_{i+2h}^{(r)} B_{i-2h}^{(r)} \tag{3.5.20}$$

$$A_i^{(r+1)} = (G_i^{(r+1)})^{-1}(B_i^{(r)}A_{i+2h}^{(r)}C_{i-2h}^{(r)} - A_{i-2h}^{(r)}A_{i+2h}^{(r)}A_i^{(r)} + C_i^{(r)}A_{i-2h}^{(r)}B_{i+2h}^{(r)}) \tag{3.5.21}$$

$$C_i^{(r+1)} = (G_i^{(r+1)})^{-1}C_i^{(r)}A_{i-2h}^{(r)}C_{i+2h}^{(r)} \tag{3.5.22}$$

and

$$\underline{y}_i^{(r+1)} = (G_i^{(r+1)})^{-1}(B_i^{(r)}A_{i+2h}^{(r)}B_{i-2h}^{(r)} - A_{i-2h}^{(r)}A_{i+2h}^{(r)}\underline{y}_i^{(r)} + A_{i-2h}^{(r)}C_i^{(r)}\underline{y}_{i+2h}^{(r)}) \tag{3.5.23}$$

where

$$G_i^{(r+1)} = A_{i-h}^{(r-1)}A_{i+h}^{(r-1)} \tag{3.5.24}$$

is a common divisor of the right-hand sides of (3.5.20),(3.5.21) and (3.5.22).

Also define $\underline{x}_0 = \underline{x}_{4h} = 0$. Then for each r and

$$i = 2h,\ 2\cdot 2h,\ \ldots,\ (2^{k-r} - 1)\cdot 2h,$$

the block tridiagonal system,

$$B_i^{(r)}\underline{x}_{i-2h} + A_i^{(r)}\underline{x}_i + C_i^{(r)}\underline{x}_{i+2h} = y_i^{(r)} \tag{3.5.25}$$

takes the form

$$A_{2^{k-1}}^{(k-1)}\underline{x}_{2^{k-1}} = \underline{y}_{2^{k-1}}^{(k-1)} \tag{3.5.26}$$

when $r = k - 1$.

Now solve for $\underline{x}_{2^{k-1}}$ from (3.5.26) and for $r = k - 2, k - 3, \ldots, 0$ and

$$i = 2h,\ 3\cdot 2h,\ 5\cdot 2h,\ \ldots\ (2^{k-r} - 1)\cdot 2h \quad .$$

The remaining unknowns are evaluated using (3.5.25):

$$\underline{x}_i = (A_i^{(r)})^{-1}(\underline{y}_i^{(r)} - B_i^{(r)}\underline{x}_{i-2h} - C_i^{(r)}\underline{x}_{i+2h}) \qquad (3.5.27)$$

The vectors $\underline{x}_{i-2h}$, $\underline{x}_{i+2h}$ on the right hand side are known from a previous step in the back-substitution process.

As r increases the matrices $A_i^{(r)}$, $B_i^{(r)}$, $C_i^{(r)}$ fill rapidly which can be expensive. These matrices can be expressed as polynomials in the single matrix A and instead of storing the matrices, compute and store the zeros of the polynomial that represent them.

Define

$$G_i^{(1)} = I\ ,\ A_i^{(-1)} = I\ ,$$

then, in the preprocessing phase, zeros are computed from the polynomial,

$$A_i^{(r+1)} = (A_{i-h}^{(r-1)}A_{i+h}^{(r-1)})^{-1}(\bar{\alpha}_i^{(r)}\bar{\gamma}_{i-2h}A_{i+h}^{(r-1)}A_{i+2h}^{(r)}A_{i-3h}^{(r-1)}$$

$$- A_{i-2h}^{(r)}A_i^{(r)}A_{i+2h}^{(r)} + \bar{\alpha}_{i+2h}^{(r)}\bar{\gamma}_i^{(r)}A_{i-h}^{(r-1)}A_{i-2h}^{(r)}A_{i+3h}^{(r-1)}) \ , \qquad (3.5.28)$$

$$r = 0, 1, 2, \ldots, k - 2,\ i = 4h, 2\cdot 4h, \ldots, (2^{k-r-1} - 1)\cdot 4h$$

where

$$\bar{\alpha}_i^{(r)} = \prod_{j=1-2h+1}^{i} a_i \tag{3.5.29}$$

$$\bar{\gamma}_i^{(r)} = \prod_{j=1}^{i+2h-1} c_j \tag{3.5.30}$$

The reduction phase is:

$$\underline{y}_i^{(r+1)} = (A_{i-h}^{(r-1)} A_{i+h}^{(r-1)})^{-1} (\bar{\alpha}_i^{(r)} A_{i+h}^{(r-1)} A_{i+2h}^{(r)} \underline{y}_{i-2h}^{(r)}$$

$$- A_{i-2h}^{(r)} A_{i+2h}^{(r)} \underline{y}_i^{(r)} + \bar{\gamma}_i^{(r)} A_{i-h}^{(r-1)} A_{i-2h}^{(r-1)} \underline{y}_{i+2h}^{(r)}) \tag{3.5.31}$$

and the back substitution phase is

$$\underline{x}_i = (A_i^{(r)})^{-1} (\underline{y}_i^{(r)} - \bar{\alpha}_i^{(r)} A_{i+h}^{(r-1)} \underline{x}_{i-2h} - \bar{\gamma}_i^{(r)} A_{i-h}^{(r)} \underline{x}_{i+2h}) \tag{3.5.32}$$

The algorithm so far is unstable. It may be stabilized by writing the reduction phase as

$$\underline{p}_i^{(r+1)} = \bar{\alpha}_i^{(r)} (A_{i-h}^{(r-1)})^{-1} \underline{q}_{i-2h}^{(r)} + \bar{\gamma}_i^{(r)} (A_{i+h}^{(r-1)})^{-1} \underline{q}_{i+2h}^{(r)} - \underline{p}_i^{(r)} \tag{3.5.33}$$

where

$$\underline{p}_i^{(r)} = (A_{i-h}^{(r-1)} A_{i+h}^{(r-1)})^{-1} \underline{y}_i^{(r)}$$

$$\underline{q}_i^{(r)} = (A_i^{(r)})^{-1} A_{i-h}^{(r-1)} A_{i+h}^{(r-1)} \underline{p}_i^{(r)}$$

for $r = 0, 1, \ldots, k-2$, and $i = 4h, 2 \cdot 4h, \ldots, (2^{k-r-1} - 1) \cdot 4h$

and the back substitution phase as:

$$\underline{x}_i = (A_i^{(r)})^{-1} A_{i-h}^{(r-1)} A_{i+h}^{(r-1)} \{ p_i^{(r)} - \bar{\alpha}_i^{(r)} (A_{i-h}^{(r-1)})^{-1} \underline{x}_{i-2h} - \bar{\gamma}_i^{(r)} (A_{i+h}^{(r-1)})^{-1} \underline{x}_{i+2h} \} \qquad (3.5.34)$$

For an $n \times n$ mesh, Swarztrauber [21] finds the asymptotic operation count to be

$$O(n^2 \log_2 n) \ .$$

3.6 A CYCLIC REDUCTION ALGORITHM FOR SOLVING TRIDIAGONAL SYSTEM OF ARBITRARY DIMENSIONS.

The following is taken from a paper by R. A. Sweet [20].

Consider the following system

$$\begin{pmatrix} A & -I & & & & \\ -I & A & -I & & & \\ & -I & A & -I & & \\ & & -I & A & -I & \\ & & & -I & A & -I \\ & & & & -I & A \end{pmatrix} \begin{pmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \underline{x}_3 \\ \underline{x}_4 \\ \underline{x}_5 \\ \underline{x}_6 \end{pmatrix} = \begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \underline{y}_3 \\ \underline{y}_4 \\ \underline{y}_5 \\ \underline{y}_6 \end{pmatrix} \qquad (3.6.1)$$

Similar linear combinations as in section (3.2) except with

$$T = I$$

yield,

$$\begin{bmatrix} A^{(1)} & -I & \\ -I & A^{(1)} & -I \\ & -I & B^{(1)} \end{bmatrix} \begin{bmatrix} \underline{x}_2 \\ \underline{x}_4 \\ \underline{x}_6 \end{bmatrix} = \begin{bmatrix} A\underline{y}_2 + \underline{y}_1 + \underline{y}_3 \\ A\underline{y}_4 + \underline{y}_3 + \underline{y}_5 \\ A\underline{y}_6 + \underline{y}_5 \end{bmatrix} = \begin{bmatrix} \underline{y}_2^{(1)} \\ \underline{y}_4^{(1)} \\ \underline{y}_6^{(1)} \end{bmatrix} \qquad (3.6.2)$$

where

$$A^{(1)} = A^2 - 2I \ , \ B^{(1)} = A^2 - I \ .$$

From this example, the following two distinct cases become evident.

Define

$$h = 2^r \ , \ J_r = n_r h$$

where n_r is the block size at r reduction and

$$A^{(0)} = B^{(0)} = A$$

$$C^{(0)} = I$$

$$J_0 = n_0 = n \ .$$

Case I: n_r is an even number. In this case the unknowns $\underline{x}_{J_r}$ will not be eliminated. The new equation for $\underline{x}_{J_r}$ is obtained by multiplying the last equation by $A^{(r)}$ and addition of the last but one equation to it.

Case II: n_r is an odd number. Here $\underline{x}_{J_r}$ is to be eliminated at this step while $\underline{x}_{J_r-h}$ will not be eliminated. The new system for $\underline{x}_{J_r-h}$ is obtained by multiplying the last but one equation by $A^{(r)}$ and adding to it the preceding equation plus $(B^{(r)})^{-1}C^{(r)}A^{(r)}$ times the last equation.

Consider the system (3.1.1) with (3.2.1) and (3.3.1).

Above considerations along with

$$\underline{p}_j^{(0)} = 0$$
$$\underline{q}_j^{(0)} = \underline{y}_j \ , \ j = 1, 2, \ldots, n \ ,$$

yield at the (r+1) lst step the reduced system of equations:

$$\begin{bmatrix} A^{(r+1)} & -I & & & & \\ -I & A^{(r+1)} & -I & & & \\ & \ddots & \ddots & \ddots & & \\ & & & -I & A^{(r+1)} & -I \\ & & & & -I & B^{(r+1)}(C^{(r+1)})^{-1} \end{bmatrix} \begin{bmatrix} \underline{x}_{2h} \\ \underline{x}_{4h} \\ \\ \\ \underline{x}_{J_{r+1}-2h} \\ \underline{x}_{J_{r+1}} \end{bmatrix}$$

$$= \begin{bmatrix} A^{(r+1)}\underline{p}_{2h}^{(r+1)} + \underline{q}_{2h}^{(r+1)} \\ A^{(r+1)}\underline{p}_{4h}^{(r+1)} + \underline{q}_{4h}^{(r+1)} \\ \vdots \\ A^{(r+1)}\underline{p}_{J_{r+1}-2h}^{(r+1)} + \underline{q}_{J_{r+1}-2h}^{(r+1)} \\ B^{(r+1)}(C^{(r+1)})^{-1}\underline{p}_{J_{r+1}}^{(r+1)} + \underline{q}_{J_{r+1}}^{(r+1)} \end{bmatrix} \qquad (3.6.3)$$

where

$$A^{(r+1)} = (A^{(r)})^2 - 2I \tag{3.6.4}$$

$$\underline{p}_j^{(r+1)} = \underline{p}_j^{(r)} + (A^{(r)})^{-1}(\underline{q}_j^{(r)} + \underline{p}_{j-h}^{(r)} + \underline{p}_{j+h}^{(r)}) \tag{3.6.5}$$

$$\underline{q}_j^{(r+1)} = \underline{q}_{j-h}^{(r)} + \underline{q}_{j+h}^{(r)} + 2\underline{p}_j^{(r+1)} \quad , \; j = 2h, 4h, \ldots, J_{r+1} - 2h \tag{3.6.6}$$

and, in case I,

$$B^{(r+1)} = A^{(r)}B^{(r)} - C^{(r)} \quad , \; C^{(r+1)} = C^{(r)} \tag{3.6.7}$$

$$\underline{p}_{J_{r+1}}^{(r+1)} = \underline{p}_{J_r}^{(r)} + (B^{(r)})^{-1}C^{(r)}(\underline{q}_{J_r}^{(r)} + \underline{p}_{J_r-h}^{(r)}) \tag{3.6.8}$$

$$\underline{q}_{J_{r+1}}^{(r+1)} = \underline{q}_{J_r-h}^{(r)} + \underline{p}_{J_{r+1}}^{(r+1)} \quad , \; J_{r+1} = J_r \; , \tag{3.6.9}$$

while, in case II,

$$B^{(r+1)} = A^{(r)}(A^{(r)}B^{(r)} - C^{(r)}) - B^{(r)} \quad , \; C^{(r+1)} = B^{(r)} \tag{3.6.10}$$

$$\underline{p}_{J_{r+1}}^{(r+1)} = \underline{p}_{J_r-h}^{(r)} + (A^{(r)})^{-1}(\underline{q}_{J_r-h}^{(r)} + \underline{p}_{J_r}^{(r)} + \underline{p}_{J_r-2h}^{(r)}) \tag{3.6.11}$$

$$\underline{q}_{J_{r+1}}^{(r+1)} = \underline{q}_{J_r-2h}^{(r)} + \underline{p}_{J_{r+1}}^{(r+1)} + (B^{(r)})^{-1}C^{(r)}A^{(r)}(\underline{q}_{J_r}^{(r)} + \underline{p}_{J_{r+1}}^{(r+1)}) \tag{3.6.12}$$

$$J_{r+1} = J_r - h$$

Using this general reduction scheme the original system under consideration may be reduced at step $r = s$ to the single equation.

$$B^{(s)}(C^{(s)})^{-1}\underline{x}_{J_s} = B^{(s)}(C^{(s)})^{-1}\underline{p}_{J_s}^{(s)} + \underline{q}_{J_s}^{(s)}$$

or,
$$B^{(s)}(\underline{x}_{J_s} - \underline{p}_{J_s}^{(s)}) = C^{(s)}\underline{q}_{J_s}^{(s)} \qquad (3.6.13)$$

It appears from (3.6.4), (3.6.7) and (3.6.10) that $A^{(r)}$, $B^{(r)}$, $C^{(r)}$ are polynomials in the original matrix, A. It has been shown in section (3.2) that $A^{(r)}$ has degree 2^r. Suppose $B^{(r)}$ has degree K_r and $C^{(r)}$ has degree ℓ_r. Then from (3.6.7) and (3.6.10),

$$K_r = \begin{cases} K_r + 2^r, & \text{case I} \\ K_r + 2^{r+1}, & \text{case II} \end{cases} \qquad (3.6.14)$$

and

$$\ell_r = \begin{cases} \ell_r, & \text{case I} \\ K_r, & \text{case II} \end{cases} \qquad (3.6.15)$$

Now substituting $A = 2\cos\theta$, it can be shown that

$$A^{(r)} = 2T_{2^r}(\tfrac{1}{2}A),$$

$$B^{(r)} = U_{K_r}(\tfrac{1}{2}A),$$

$$C^{(r)} = U_{\ell_r}(\tfrac{1}{2}A),$$

where,

$$K_r = 2^r + \ell_r \quad \text{and} \quad 0 \le \ell_r \le 2^r - 1, \quad \text{and}$$

$$T_m(\alpha) = \prod_{i=1}^{m}(2\alpha - 2\cos\frac{(2i-1)\pi}{2m}),$$

$$u_m(\alpha) = \prod_{i=1}^{m}(2\alpha - 2\cos\frac{i\pi}{m+1})$$

denote respectively the Chebyshev polynomial of the first and second kinds.

Equation (3.6.13) can be written as

$$\prod_{i=1}^{k_s} (A - \lambda_i^{(s)} I)(\underline{x}_{J_s} - \underline{p}_{J_s}^{(s)}) = \prod_{i=1}^{\ell_s} (A - \mu_i^{(s)} I)\underline{q}_{J_s}^{(s)}$$

where

$$\lambda_i^{(s)} = 2 \cos \frac{i}{k_s + 1} \pi \quad ,$$

$$\mu_i^{(s)} = 2 \cos \frac{i}{\ell_s + 1} \pi \quad .$$

To avoid the matrix multiplication of the form

$$(A - \mu I)\underline{q}$$

a technique suggested by Swarztrauber [21, page 1143] can be used. The following algorithm follows from the analysis.

1. Set $\underline{z}_0 = \underline{q}_{J_s}^{(s)}$

2. Solve the linear system

$$(A - \lambda_i^{(s)} I)\tilde{\underline{z}}_i = (\lambda_i^{(s)} - \mu_i^{(s)})\underline{z}_{i-1} \quad ,$$

$$\underline{z}_i = \underline{z}_{i-1} + \tilde{\underline{z}}_i \quad , \quad \text{for} \quad i = 1, 2, \ldots, \ell_s \quad ,$$

3. Solve the linear system

$$(A - \lambda_i^{(s)})\underline{z}_i = \underline{z}_{i-1} \quad , \text{ for} \quad i = \ell_{s+1}, \ell_{s+2}, \ldots, k_s \quad ,$$

4. $\underline{x}_{J_s} = \underline{p}_{J_s} + \underline{z}_{K_s}$.

Observe that λ_i and μ_i should be selected so that $\mu_i^{(s)}$ is as close as possible to one of $\lambda_i^{(s)}$ which reduces considerably the accumulation of round off errors.

The computation of the last part of (3.6.12) is done by a similar algorithm. The remaining unknowns are then computed by the usual back substitution process.

The asymptotic operation count for this algorithm is (Sweet [20])

$$O(n^2 \log_2 n)$$

for an $n \times n$ mesh.

3.7 THE NUMERICAL SOLUTION OF THE MATRIX EQUATION $XA + AY = G$.

Hoskins et al [14] presented an iterative method for solving the matrix equation

$$XA + AY = G \tag{3.7.1}$$

where X, Y, G are known matrices of orders $m \times m$, $n \times n$, $m \times n$ respectively. The algorithm is:

Step 1. While $X \neq I$ and $Y \neq I$ execute steps 2 to 4.

Step 2. Set $G = \frac{1}{2}(G + X^{-1}GY^{-1})$

Step 3. Set $X = \frac{1}{2}(X + X^{-1})$

Step 4. Set $Y = \frac{1}{2}(Y + Y^{-1})$

Step 5. $A = \frac{1}{2} G$.

After s application of steps 2 to 4 of the above algorithm the following equation is obtained:

$$X_s A + A Y_s = G_s \tag{3.7.2}$$

where

$$X_s = \frac{1}{2}(X_{s-1} + X_{s-1}^{-1}) \ , \ X_0 = X \ ,$$

$$Y_s = \frac{1}{2}(Y_s + Y_{s-1}^{-1}) \ , \ Y_0 = Y \ ,$$

$$s = 1, 2, \ldots \ .$$

Multiplication of equation (3.7.2) on the left by X_s^{-1} and on the right by Y_s^{-1} produces

$$X_s^{-1} A + A Y_s^{-1} = X_s^{-1} G_s Y_s^{-1} \tag{3.7.3}$$

Addition of equation (3.7.2) and (3.7.3) and division by 2 gives,

$$\frac{1}{2}(X_s + X_s^{-1})A + \frac{1}{2} A(Y_s + Y_s^{-1}) = \frac{1}{2}(G_s + X_s^{-1} G_s Y_s^{-1})$$

from which it is clear that

$$G_{s+1} = \frac{1}{2}(G_s + X_s^{-1} G_s Y_s^{-1})$$

converges to $2A$ whenever

$$X_{s+1} = \frac{1}{2}(X_s + X_s^{-1})$$

$$Y_{s+1} = \frac{1}{2}(Y_s + Y_s^{-1}) \tag{3.7.4}$$

converges to the identity matrix I. It can be shown (Hoskins, et al [14]) that convergence occurs when the eigenvalues of either X and Y or $-X$ and $-Y$ have positive real parts.

Let

$$\tilde{B}_s = \frac{1}{||X_s^{-1}||}$$

and

$$\tilde{C}_s = ||X_s|| \quad .$$

The iteration (3.7.4) can be generalized to

$$X_{s+1} = \tilde{\alpha}_s X_s + \tilde{\beta}_s X_s^{-1} \quad , \; s = 0, 1, 2, \ldots$$

where

$$\tilde{\alpha}_s = \frac{2\tilde{B}_s}{(\tilde{B}_s + \sqrt{(\tilde{B}_s \tilde{C}_s)})^2} \quad ,$$

$$\tilde{\beta}_s = \tilde{B}_s \tilde{C}_s \tilde{\alpha}_s$$

In the case that X and Y have real spectra, $\tilde{B}_{s+1}$, $\tilde{C}_{s+1}$ can be found from

$$\tilde{B}_{s+1} = 1 - \tilde{E}_s, \; \tilde{C}_{s+1} = 1 + \tilde{E}_s \quad ,$$

where

$$\tilde{E}_s = \left(\frac{\tilde{B}_s - \sqrt{\tilde{B}_s \tilde{C}_s}}{\tilde{B}_s + \sqrt{\tilde{B}_s \tilde{C}_s}} \right)^2 \quad .$$

The operation count for this algorithm is [14]

$$O(n^3)$$

for an $n \times n$ system.

CHAPTER 4

SOLUTION OF MATRIX EQUATIONS ARISING FROM HIGHER ORDER DISCRETIZATIONS

4.1 INTRODUCTION.

It has been mentioned earlier that discretization of equation (1.2.2) on the uniform rectangular mesh (1.3.2) leads to a matrix equation of the form (1.3.3) which can also be written in composite or block form as (Bickley and McNamee [5], Mitchell [16], page 102, Varga [24], page 196-197):

$$M\underline{x} = \underline{y} \tag{4.1.1}$$

where

$$M = \begin{bmatrix} W+v_{11}I & v_{21}I & v_{31}I & \cdots & v_{n1}I \\ v_{12}I & w+v_{22}I & v_{32}I & \cdots & v_{n2}I \\ v_{13}I & v_{23} & w+v_{33}I & \cdots & v_{n3}I \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ v_{1n}I & v_{2n}I & v_{3n}I & \cdots & w+v_{nn}I \end{bmatrix}, \tag{4.1.2}$$

$$\underline{x} = [a_{11}, a_{12}, \ldots, a_{1n}; a_{21}, a_{22} \cdots a_{2n}; a_{31}, \cdots ; a_{m1}, a_{m2} \ldots, a_{mn}]$$

$$\underline{y} = [f_{11}, f_{12}, \ldots, f_{1n}; f_{21}, f_{22}, \ldots, f_{2n}; f_{31}, \ldots; f_{m1}, f_{m2}, \ldots, f_{mn}]$$

The matrix A is a numerical approximation of the discretized solution of equation (1.2.2) at the internal points of (1.3.2). The matrix V is of dimension $n \times n$, W is $m \times m$ and I is an identity matrix of order $m \times m$.

In this chapter attempts are made to generalize the methods [4], [6], [20] and [21] for the solution of the matrix equation (4.1.1) where (4.1.2) arise from a higher order finite difference approximation to the equation (1.2.2) with Dirichlet boundary condition.

Finite difference approximations to equation (1.2.2) with Dirichlet boundary conditions using the standard five-point formula on uniform rectangular meshes produce matrix equations of the form (4.1.1) where

$$M = \begin{bmatrix} A & -I & & & \\ -I & A & -I & & \\ & \ddots & \ddots & \ddots & \\ & & & & -I \\ & & & -I & A \end{bmatrix} \qquad (4.1.3)$$

The square diagonal submatrices

$$A = \begin{bmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & & & -1 \\ & & & -1 & 4 \end{bmatrix}$$

are of order n and the I's are $n \times n$ identity matrices.

In the special case mentioned above the matrices V and W are

$$V = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & & -1 & \\ & & & -1 & 2 \end{bmatrix}_{n\times n},$$

$$W = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & & -1 & \\ & & & -1 & 2 \end{bmatrix}_{m\times m}.$$

However, it may be observed that the fast direct methods of Chapter 3 are designed for the solution of the matrix equation (4.1.1) where M is tridiagonal and usually of the form (4.1.3).

4.2 HIGHER ORDER DISCRETIZATION AND FAST DIRECT METHODS.

Consider the following system

$$\begin{bmatrix} A & -D & I & & & & & & & \\ -I & A & -D & I & & & & & & \\ I & -I & A & -D & I & & & & & \\ & I & -I & A & -D & I & & & & \\ & & I & -I & A & -D & I & & & \\ & & & I & -I & A & -D & I & & \\ & & & & I & -I & A & -D & I & \\ & & & & & I & -I & A & -D & \\ & & & & & & I & -I & A & \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \underline{x}_3 \\ \underline{x}_4 \\ \underline{x}_5 \\ \underline{x}_6 \\ \underline{x}_7 \\ \underline{x}_8 \\ \underline{x}_9 \\ \underline{x}_{10} \end{bmatrix} = \begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \underline{y}_3 \\ \underline{y}_4 \\ \underline{y}_5 \\ \underline{y}_6 \\ \underline{y}_7 \\ \underline{y}_8 \\ \underline{y}_9 \\ \underline{y}_{10} \end{bmatrix} \qquad (4.2.1)$$

where matrix

$D = dI$, d is a scalar,

A is any quin-diagonal matrix,

I is an $n \times n$ identity matrix.

An attempt to generalize the reduction process of section 3 is as follows: Multiply the third and fourth equations respectively by A and D and add the first, second and fifth equations to them, multiply the fifth and sixth equations respectively by A and D

and add third, fourth and seventh equations to them and continue the process, then the system of equation (4.2.1) may be written as

$$\begin{bmatrix} 2A-I & A^2-2D+2I & 2A-D^2 & I & \\ I & 2A-I & A^2-2D+2I & 2A-D^2 & I \\ & I & 2A-I & A^2-2D+2I & 2A-D^2 \\ & & I & 2A-I & A^2-2D+2I \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_3 \\ \underline{x}_5 \\ \underline{x}_7 \\ \underline{x}_9 \end{bmatrix} = \begin{bmatrix} \tilde{\underline{y}}_1 \\ \tilde{\underline{y}}_3 \\ \tilde{\underline{y}}_5 \\ \tilde{\underline{y}}_7 \end{bmatrix} \qquad (4.2.2)$$

where

$$\tilde{\underline{y}}_1 = A\underline{y}_3 + D\underline{y}_4 + \underline{y}_1 + \underline{y}_2 + \underline{y}_5 ,$$

$$\tilde{\underline{y}}_3 = A\underline{y}_5 + D\underline{y}_6 + \underline{y}_3 + \underline{y}_4 + \underline{y}_7 ,$$

$$\tilde{\underline{y}}_5 = A\underline{y}_7 + D\underline{y}_8 + \underline{y}_5 + \underline{y}_6 + \underline{y}_9 ,$$

$$\tilde{\underline{y}}_7 = A\underline{y}_9 + D\underline{y}_{10} + \underline{y}_7 + \underline{y}_8 .$$

The reduced system of equation in (4.2.2) is no longer of the form (4.2.1) and therefore, cannot be reduced further in the same way.

Higher order finite difference approximations, for example, a nine-point approximation to equation (1.2.2) in a special case as above gives rise to a matrix M of the following form:

$$
M = \begin{pmatrix}
A_1 & B_1 & C_1 & D & & & & & & \\
B & A & B & C & & & & & & \\
C & B & A & B & C & & & & & \\
 & C & B & A & B & C & & & & \\
 & & \ddots & \ddots & \ddots & \ddots & \ddots & & & \\
 & & & C & B & A & B & C & & \\
 & & & & C & B & A & B & & \\
 & & & & D & C_1 & B_1 & A_1 &
\end{pmatrix}
$$

where

$$B_1 = -\frac{1}{2} I \ ,$$

$$C_1 = -\frac{1}{3} I \ ,$$

$$D = \frac{1}{12} I = C \ ,$$

$$B = -\frac{16}{12} I \ ,$$

$$A_1 = W_1 + \frac{20}{12} I \ ,$$

$$A = W_1 + \frac{30}{12} I \ ,$$

$$W_1 = k^2 W$$

and the matrix W is as in section (2.3.4) with the order of the identity matrices the same as the order of W. It can be observed that the matrix M and its diagonal submatrices have bandwidth greater than 5. The use of the nine-point formula of section (2.3.5) will add one more element in each of the first and last rows of the matrix M and its diagonal submatrices. It appears, therefore, that higher order formulae and their corresponding modification will increase the bandwidth of the matrix M and adversely affect the usefulness of the cyclic reduction algorithm for a system as in (4.2.1). The system of equation (4.1.1) where matrix M, as above, obtained using a higher order formula cannot be solved using the marching algorithm since it also takes advantage of the special block structure of the matrix M.

However, an interesting result can be obtained with the use of an alternate nine-point approximation of section (2.3.6).

If $m = n$, the equation (2.3.16) can also be rearranged as (4.1.1) where M is of the form (3.2.1) with

$$A = \begin{bmatrix} 20 & -4 & & & \\ -4 & 20 & -4 & & \\ & \ddots & \ddots & \ddots & \\ & & & -4 & \\ & & & -4 & 20 \end{bmatrix}, \qquad (4.2.3)$$

and

$$T = \begin{bmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & & & 1 \\ & & & 1 & 4 \end{bmatrix} \qquad (4.2.4)$$

The linear system (4.1.1) may be written, using the notation of chapter 3, as follows:

$$\begin{bmatrix} A & -T & & & \\ -T & A & -T & & \\ & \ddots & \ddots & \ddots & \\ & & & & -T \\ & & & -T & A \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \vdots \\ \underline{x}_{n-1} \\ \underline{x}_n \end{bmatrix} = \begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_{n-1} \\ \underline{y}_n \end{bmatrix} \qquad (4.2.5)$$

Multiplication of (4.2.5) by block matrix $\mathrm{Diag}[T^{-1}]$ (Bank [3], page 4-16) yields,

$$\begin{bmatrix} T^{-1}A & -I & & & \\ -I & T^{-1}A & -I & & \\ & \ddots & \ddots & \ddots & \\ & & & & -I \\ & & & -I & T^{-1}A \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \vdots \\ \underline{x}_{n-1} \\ \underline{x}_n \end{bmatrix} = \begin{bmatrix} \underline{z}_1 \\ \underline{z}_2 \\ \vdots \\ \underline{z}_{n-1} \\ \underline{z}_n \end{bmatrix} \qquad (4.2.6)$$

where

$$T\underline{z}_i = \underline{y}_i \ , \quad 1 \le i \le n \ .$$

The matrix M of (4.2.6) has the special form as in section (3.4) except when $T^{-1}A$ is likely to be full. The solution of (4.2.6) may be carried out (Bank [3], page 4-17) using the generalized marching algorithm. Here PMP^T is dealt with instead of PM ; P is the permutation matrix and P^T its transpose.

Suppose n is of the form $2^{k+1} - 1$, $k \ge 0$. Since the matrices (4.2.3) and (4.2.4) are symmetric tridiagonal and

$$AT = TA$$

the odd-even cyclic reduction and its Buneman variant can be applied for the solution of the linear system (4.2.5).

The algorithm of section (3.7), in general, can be applied for the solution of linear systems which arise from any finite difference approximation of elliptic partial differential equations with Dirichlet boundary conditions on a rectangular region. The algorithm works for any pair of matrices V and W which may have complex spectra provided the real parts of their eigenvalues are positive. It appears to be numerically stable even when V and W are not too well conditioned (Walton [25], page 90), and there is no significant change in either complexity of implementation or number of operations when used for matrix equations which arise from higher order discretizations of elliptic partial differential equations.

CHAPTER 5

NUMERICAL ILLUSTRATIONS

5.1 INTRODUCTION.

In this chapter, some model problems are considered for numerical illustration. Throughout this chapter the region of solution is taken to be a unit square. A uniform mesh is used for convenience. It has $n = 15$ internal mesh-lines parallel to each axis. The spacing, h, between mesh lines, is given by

$$h = \frac{1}{n + 1} = \frac{1}{16} .$$

The maximum absolute actual error, that is, maximum deviation of the numerical solution of the problem from the analytic solution is determined in absolute value. The relative errors are also tabulated. Let

$$U = [u_{ij}] \quad \begin{matrix} i = 1(1)m , \\ j = 1(1)n , \end{matrix}$$

where

$$u_{i,j} = u(x_j, y_i) ,$$

be the analytic and

$$A = [a_{ij}] \quad \begin{matrix} i = 1(1)m , \\ j = 1(1)n \end{matrix}$$

be the numerical solution of a problem at the internal points of (1.3.2) respectively, then the maximum absolute error, e_a is given by

$$e_a = \max_i \max_j |u_{i,j} - a_{i,j}| \quad \begin{matrix} i = 1(1)m \ , \\ j = 1(1)n \ , \end{matrix} \qquad (5.1.1)$$

and the maximum relative error, e_r is given by

$$e_r = \max_i \max_j \left| \frac{u_{i,j} - a_{i,j}}{u_{i,j}} \right| , \quad \begin{matrix} i = 1(1)m \ , \\ j = 1(1)n \ . \end{matrix} \qquad (5.1.2)$$

It has been indicated that a solution correct to seven decimal places can be achieved for matrices up to 63×63 in fewer than five iterations when using the algorithm in section (3.7) (Hoskins et al [14]). The following examples illustrate that higher order discretization formulae yield a higher order of accuracy as was anticipated by using a more accurate Taylor's series expansion. The accuracy, indicated in chapter 2, due to the use of formulae from section (2.3.1) to section (2.3.8) can be achieved in five or less iterations whereas 6 and 7 iterations are required respectively for the 17-point and 21-point formulae. For large n, the operation count remains $O(n^3)$.

The condition numbers of the matrices of chapter 2 are tabulated since the behaviour of the matrices with respect to the inverse is correlated with their condition number (Todd [23], page 45). If M is the matrix, the condition number is (Todd [23], page 44)

$$\chi(M) = ||M||\ ||M^{-1}|| \qquad (5.1.3)$$

where the norm used is the maximum absolute row sum. In the following

tables the nine-point formula of section (2.3.4) is referred to as 9-point (a), the alternate 9-point formula as 9-point (b), and the 13-point formulae of sections (2.3.7) and (2.3.8) as 13-point (a) and 13-point (b) respectively.

The algorithm SOLVEXAAY for 9-point (b) was implemented using both equations (2.3.16) and (2.3.17).

The calculations summarized in the following sections were performed using double precision arithmetic in APL on an IBM/360 model 50 computer.

5.2 EXAMPLE 1.

Consider the Dirichlet problem

$$\frac{\partial^2}{\partial x^2} u(x, y) + \frac{\partial^2}{\partial y^2} u(x, y) = 2e^{x+y} \quad \text{in} \quad R : 0 \leq x \leq 1, \; 0 \leq y \leq 1 ,$$

$$u(x, y) = \begin{cases} e^{x} , & y = 0 , \\ e^{y} , & x = 0 , \\ e^{1+y} , & x = 1 , \\ e^{x+1} , & y = 1 ; \end{cases}$$

which has the analytic solution

$$u(x, y) = e^{x+y} .$$

The maximum absolute actual and relative errors for this problem are summarized in tables (5.1) and (5.2).

Table 5.1

DISCRETIZATION	NUMBER OF ITERATIONS			
	4	5	6	7
5-point	1.7715×10^{-4}	1.3995×10^{-4}		
9-point (a)	5.3253×10^{-5}	5.2339×10^{-7}	5.2347×10^{-7}	
9-point (b) Equation (2.3.16)	5.2689×10^{-5}	1.0934×10^{-7}	1.0970×10^{-7}	
9-point (b) Equation (2.3.17)	5.2592×10^{-5}	9.7817×10^{-10}	2.8016×10^{-11}	2.8015×10^{-11}
13-point (a)	1.8379×10^{-4}	7.9606×10^{-9}	1.6559×10^{-9}	
13-point (b)	1.3303×10^{-3}	4.1547×10^{-7}	9.8259×10^{-11}	9.8211×10^{-11}
17-point	6.0683×10^{-3}	4.6013×10^{-6}	8.7987×10^{-12}	5.5768×10^{-12}
21-point	3.5367×10^{-2}	2.4464×10^{-4}	4.6630×10^{-8}	1.4677×10^{-13}

GREATEST ABSOLUTE ERROR FOR EXAMPLE 1.

Table 5.2

DISCRETIZATION	NUMBER OF ITERATIONS			
	4	5	6	7
5-point	5.8603×10^{-5}	5.1454×10^{-5}		
9-point (a)	2.8215×10^{-5}	1.0084×10^{-7}	1.0076×10^{-7}	
9-point (b) Equation (2.3.16)	2.2689×10^{-5}	4.0265×10^{-8}	4.0324×10^{-8}	
9-point (b) Equation (2.3.17)	2.2428×10^{-5}	2.4732×10^{-10}	1.0291×10^{-11}	1.0291×10^{-11}
13-point (a)	5.0024×10^{-5}	1.2208×10^{-9}	3.3613×10^{-10}	3.3613×10^{-11}
13-point (b)	0.2040×10^{-3}	6.3715×10^{-8}	3.3019×10^{-11}	3.3019×10^{-11}
17-point	9.4786×10^{-4}	7.0563×10^{-7}	3.4950×10^{-12}	1.3197×10^{-12}
21-point	5.4237×10^{-3}	3.7517×10^{-5}	7.1510×10^{-9}	5.9274×10^{-14}

GREATEST RELATIVE ERROR FOR EXAMPLE 1.

5.3 EXAMPLE 2.

Consider the problem

$$\frac{\partial^2}{\partial x^2} u(x, y) + \frac{\partial^2}{\partial y^2} u(x, y) = 2\{x(x - 1) + y(y - 1)\} \quad \text{in}$$

$$R : 0 \leq x \leq 1 \ , \ 0 \leq y \leq 1 \ ,$$

$$u(x, y) = 0 \quad \text{on} \quad \partial R$$

which has the solution

$$u(x, y) = x(x - 1)y(y - 1) \ .$$

The problem was discretized using the different schemes in Chapter 2 with $h = 1/16$. Note that this model problem has no truncation error. The errors in the numerical solution, A , are due to round off and are summarized in tables (5.3) and (5.4).

Table 5.3

DISCRETIZATION	NUMBER OF ITERATIONS			
	4	5	6	7
5-point	4.6736×10^{-7}	7.9283×10^{-13}	5.6205×10^{-16}	5.8286×10^{-16}
9-point (a)	6.2303×10^{-7}	4.5848×10^{-12}	5.5511×10^{-16}	5.8286×10^{-16}
9-point (b) Equation (2.3.16)	5.9591×10^{-7}	9.0132×10^{-12}	1.3877×10^{-15}	1.4155×10^{-15}
9-point (b) Equation (2.3.17)	5.9591×10^{-7}	9.0132×10^{-12}	1.3877×10^{-15}	1.4155×10^{-15}
13-point (a)	6.5003×10^{-7}	1.8103×10^{-11}	2.4286×10^{-16}	2.4980×10^{-16}
13-point (b)	1.3691×10^{-6}	1.5610×10^{-10}	3.5388×10^{-16}	4.0939×10^{-16}
17-point	8.6702×10^{-6}	4.8188×10^{-10}	1.2975×10^{-15}	1.2836×10^{-15}
21-point	2.5968×10^{-8}	2.1316×10^{-12}	4.8511×10^{-15}	4.8494×10^{-15}

GREATEST ABSOLUTE ERROR FOR EXAMPLE 2.

Table 5.4

DISCRETIZATION	NUMBER OF ITERATIONS			
	4	5	6	7
5-point	8.0218×10^{-6}	1.2685×10^{-11}	1.4416×10^{-14}	1.4628×10^{-14}
9-point (a)	9.9685×10^{-6}	7.8042×10^{-11}	1.3356×10^{-14}	1.3780×10^{-14}
9-point (b) Equation (2.3.16)	9.5845×10^{-6}	1.5170×10^{-10}	3.2132×10^{-14}	3.2369×10^{-14}
9-point (b) Equation (2.3.17)	9.5845×10^{-6}	1.5170×10^{-10}	3.2132×10^{-14}	3.2369×10^{-14}
13-point (a)	1.0400×10^{-5}	2.9917×10^{-10}	9.979×10^{-15}	1.0231×10^{-14}
13-point (b)	3.4367×10^{-5}	3.1190×10^{-9}	9.4739×10^{-15}	1.088×10^{-14}
17-point	1.3872×10^{-4}	1.3423×10^{-8}	8.8423×10^{-14}	4.1179×10^{-14}
21-point	4.2440×10^{-4}	5.2905×10^{-7}	2.9036×10^{-10}	1.4130×10^{-12}

GREATEST RELATIVE ERROR FOR EXAMPLE 2.

5.4 EXAMPLE 3.

Consider the model example

$$\frac{\partial^2}{\partial x^2} u(x, y) + \frac{\partial^2}{\partial y^2} u(x, y) = -2\pi^2 \sin \pi x \sin \pi y \quad \text{in}$$

$$R : 0 \le x \le 1 , \; 0 \le y \le 1 ,$$

$$u(x, y) = 1 \qquad , y = 0 ,$$

$$= 1 + \sin \pi \sin \pi y \; , x = 1 ,$$

$$= 1 + \sin \pi \sin \pi x \; , y = 1 ,$$

$$= 1 \qquad , x = 0$$

which has the solution

$$u(x, y) = 1 + \sin \pi x \sin \pi y .$$

The results of the experiment with this example are summarized in tables (5.5) and (5.6).

Table 5.5

DISCRETIZATION	NUMBER OF ITERATIONS			
	4	5	6	7
5-point	3.2281×10^{-3}	3.2189×10^{-3}		
9-point (a)	1.5414×10^{-5}	2.1245×10^{-5}		
9-point (b) Equation (2.3.16)	1.0430×10^{-5}	2.4869×10^{-5}		
9-point (b) Equation (2.3.17)	2.6412×10^{-5}	6.2637×10^{-8}	6.2830×10^{-8}	
13-point (a)	5.5076×10^{-5}	7.0755×10^{-7}	7.0721×10^{-7}	
13-point (b)	5.5076×10^{-5}	7.0755×10^{-7}	7.0721×10^{-7}	
17-point	9.0444×10^{-4}	6.4840×10^{-7}	2.1832×10^{-8}	2.1831×10^{-8}
21-point	5.1726×10^{-3}	3.5413×10^{-5}	6.2987×10^{-9}	6.3961×10^{-10}

GREATEST ABSOLUTE ERROR FOR EXAMPLE 3.

Table 5.6

DISCRETIZATION	NUMBER OF ITERATIONS			
	4	5	6	7
5-point	1.6140×10^{-3}	1.6094×10^{-3}		
9-point (a)	8.0230×10^{-6}	1.7777×10^{-5}	1.7777×10^{-5}	
9-point (b) Equation (2.3.16)	7.4630×10^{-6}	1.2434×10^{-5}	1.2434×10^{-5}	
9-point (b) Equation (2.3.17)	1.3462×10^{-5}	3.1318×10^{-8}	3.1415×10^{-8}	
13-point (a)	3.2859×10^{-5}	5.9205×10^{-7}	5.9177×10^{-7}	
13-point (b)	3.2859×10^{-5}	5.9205×10^{-7}	5.9177×10^{-7}	
17-point	8.7128×10^{-4}	6.2462×10^{-7}	1.8268×10^{-8}	1.2867×10^{-8}
21-point	4.9830×10^{-3}	3.4115×10^{-5}	6.0677×10^{-9}	5.3520×10^{-10}

GREATEST RELATIVE ERROR FOR EXAMPLE 3.

5.5 EXAMPLE 4.

Finally consider the problem

$$\frac{\partial^2}{\partial x^2} u(x, y) + \frac{\partial^2}{\partial y^2} u(x, y) = 0 \quad \text{in} \quad R : 0 \le x \le 1, \; 0 \le y \le 1 ,$$

$$u(x, 0) = u(x, 1) = \sin \pi x ,$$

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

which has the analytic solution

$$u(x, y) = \operatorname{sech} \frac{\pi}{2} \cosh \pi(y - \frac{1}{2}) \sin \pi x .$$

Since the differential equation is Laplacian, both of equation (2.3.16) and (2.3.17) are essentially the same. The experimental results are tabulated in tables (5.7) and (5.8).

Table 5.7

DISCRETIZATION	NUMBER OF ITERATIONS			
	4	5	6	7
5-point	1.8393×10^{-3}	1.8392×10^{-3}		
9-point (a)	2.2169×10^{-5}	1.3958×10^{-5}		
9-point (b)	7.7471×10^{-6}	5.4303×10^{-9}	5.4400×10^{-9}	
13-point (a)	2.1397×10^{-5}	4.2493×10^{-7}	4.2476×10^{-7}	
13-point (b)	1.0434×10^{-4}	1.3041×10^{-7}	1.0998×10^{-7}	
17-point	4.0068×10^{-4}	3.3425×10^{-7}	9.9760×10^{-9}	9.9761×10^{-9}
21-point	2.4226×10^{-3}	1.7674×10^{-5}	2.9069×10^{-9}	3.7630×10^{-10}

GREATEST ABSOLUTE ERROR FOR EXAMPLE 4.

Table 5.8

DISCRETIZATION	NUMBER OF ITERATIONS			
	4	5	6	7
5-point	4.6152×10^{-3}	4.6149×10^{-3}		
9-point (a)	1.0192×10^{-4}	1.1136×10^{-4}		
9-point (b)	1.9070×10^{-5}	1.3625×10^{-8}	1.3649×10^{-8}	
13-point (a)	5.6988×10^{-5}	4.2488×10^{-6}	4.2484×10^{-6}	
13-point (b)	6.9285×10^{-7}	6.7804×10^{-7}		
17-point	4.7808×10^{-4}	3.9882×10^{-7}	1.0613×10^{-7}	
21-point	2.8907×10^{-3}	2.1090×10^{-5}	4.3754×10^{-9}	3.7743×10^{-9}

GREATEST RELATIVE ERROR FOR EXAMPLE 4.

5.6 CONDITION NUMBERS OF THE MATRICES WHICH ARISE IN DISCRETIZATION.

The condition numbers of matrices encountered during the use of different discretization formulae are calculated using formula (5.1.3). Those for matrices V of section (2.3.1) through section (2.3.10) are shown in table (5.9) and those for matrices W of above sections in table (5.10) for different orders. It can be observed that for $m = n$,

$$W = V^T .$$

The results of tables (5.9) and (5.10) are shown graphically in figures (5.1) through (5.4).

Table 5.9

DISCRETIZATION	ORDER OF MATRIX			
	10×10	15×15	20×20	25×25
5-point	60	128	220	338
9-point (a)	80	170.66	293.33	450.66
13-point (a)	105.25	224.53	385.91	592.90
13-point (b)	208.83	445.51	765.72	1176.42
17-point	395.24	843.19	1449.23	2226.55
21-point	2280.41	3454.32	5478.57	8074.93

Condition numbers of the matrices V of Chapter 2.

Table 5.10

DISCRETIZATION	ORDER OF MATRIX			
	10×10	15×15	20×20	25×25
5-point	60	128	220	338
9-point	83.29	178.2	306.61	471.32
13-point (a)	123.35	258.69	445.13	684.24
13-point (b)	185.78	321.6	553.43	850.73
17-point	350.04	526.09	797.46	1225.86
21-point	1698.55	2374.95	3265.38	4155.93

Condition numbers of the matrices W of Chapter 2.

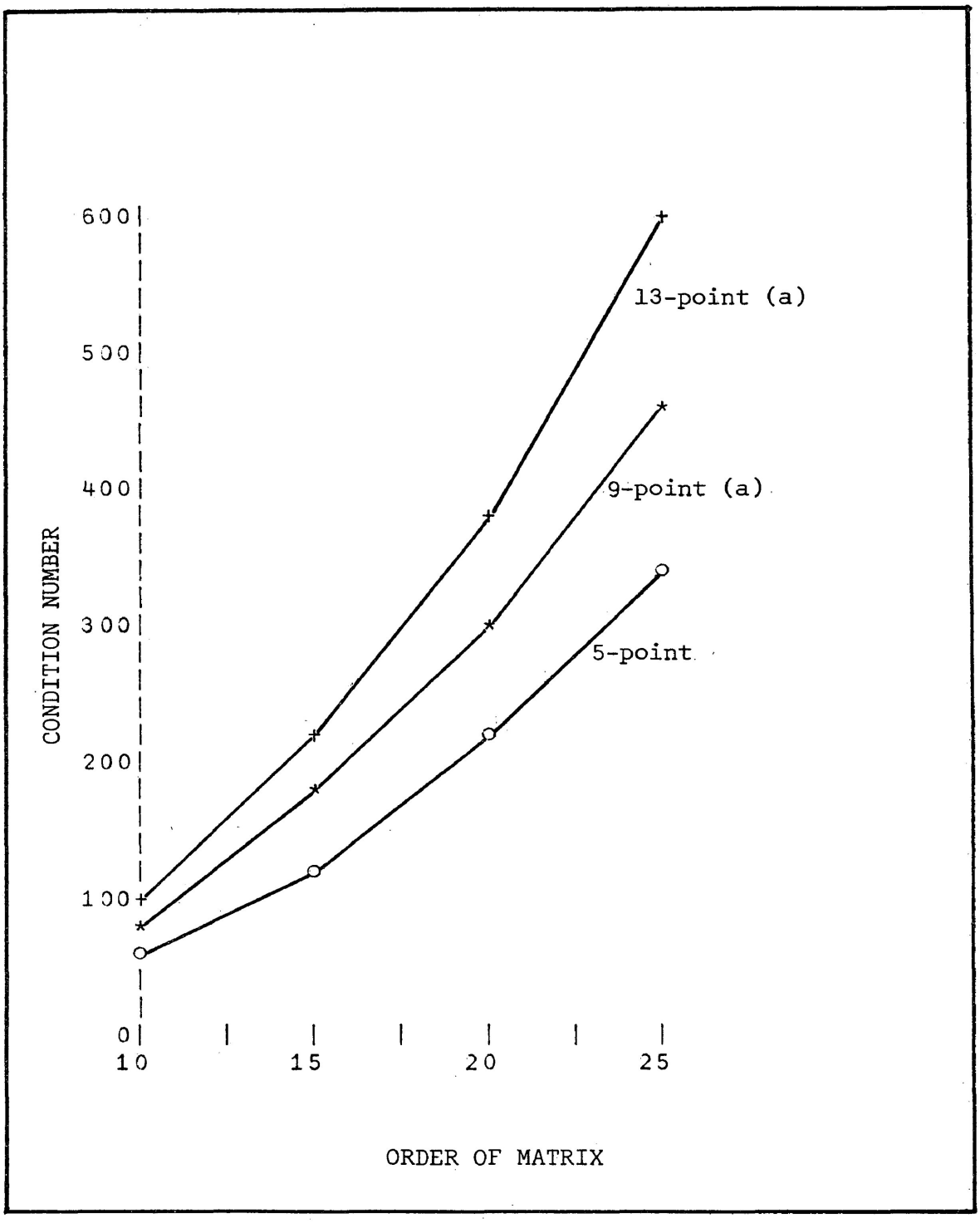


Figure 5.1

Condition number of matrices V of Chapter 2.

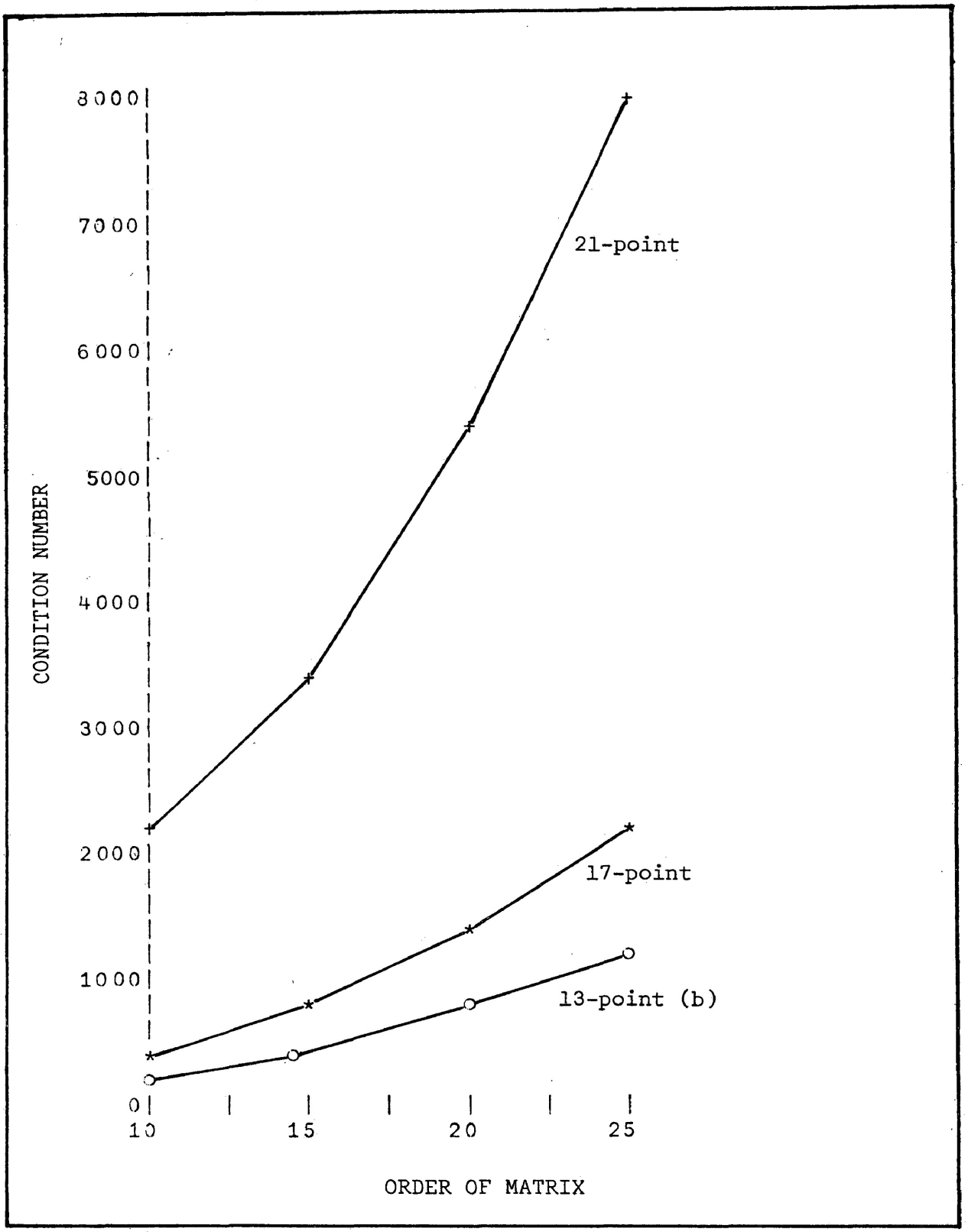


Figure 5.2

Condition number of matrices V of Chapter 2.

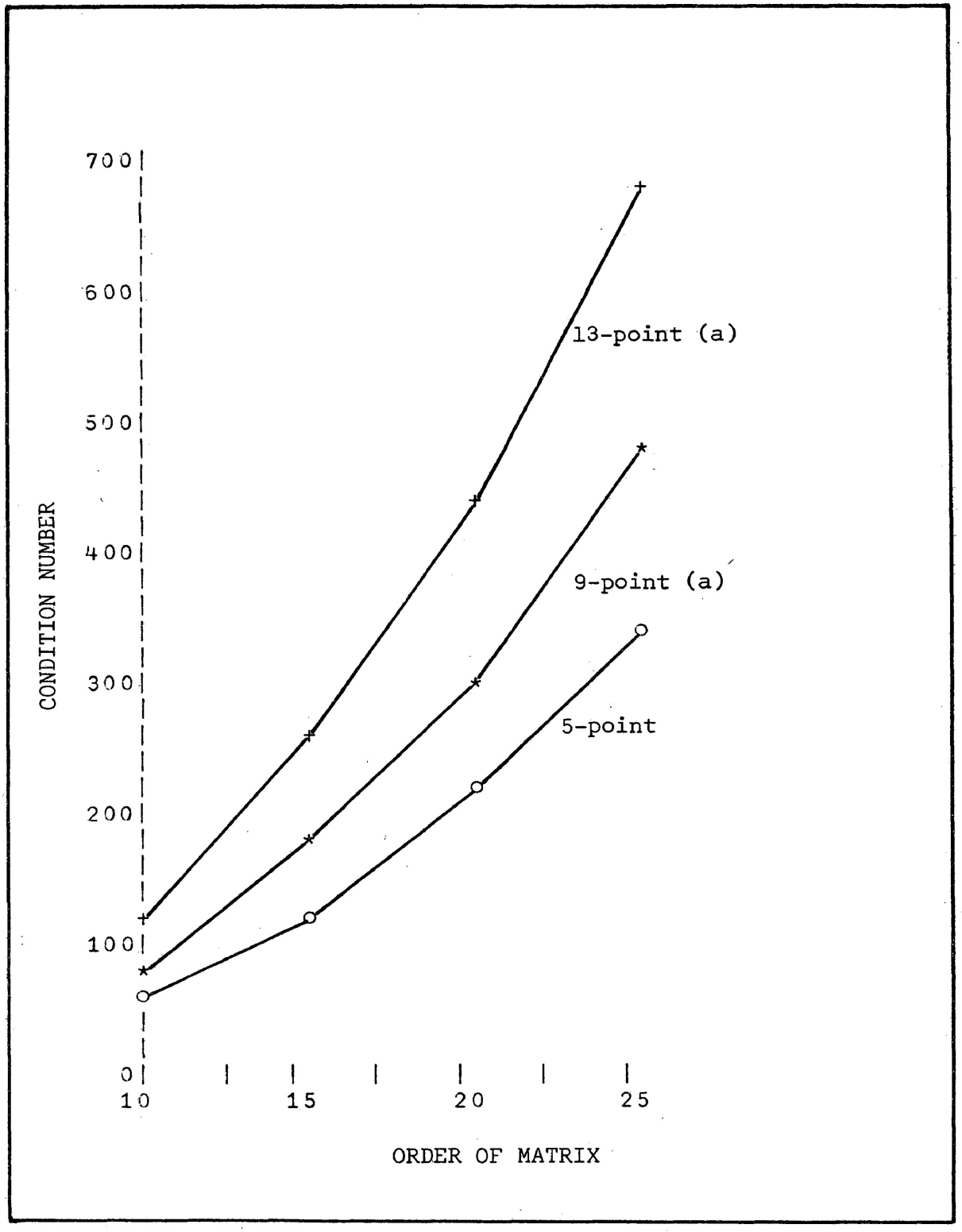


Figure 5.3

Condition number of matrices W of Chapter 2.

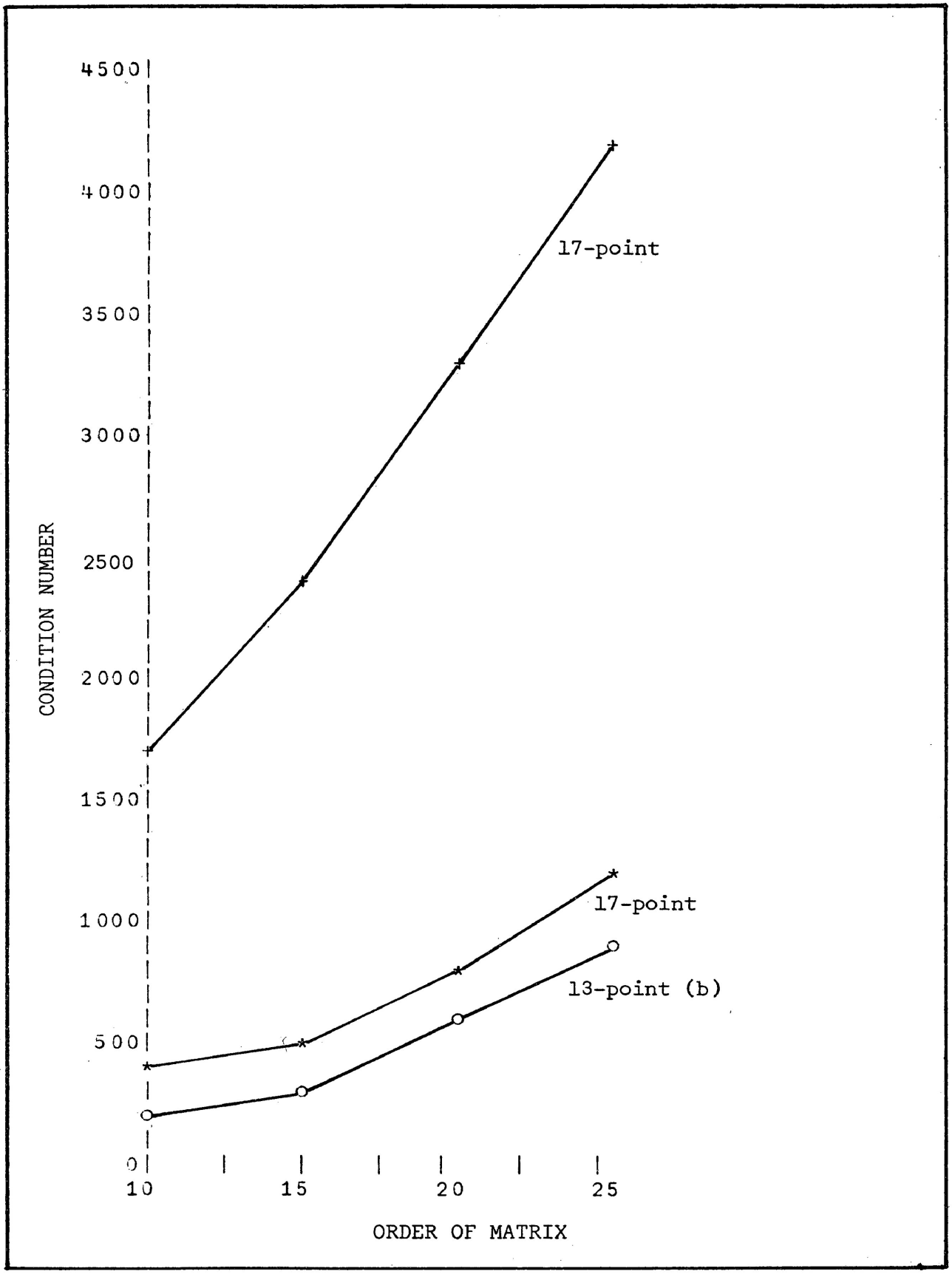


Figure 5.4

Condition number of matrices W of Chapter 2.

CHAPTER 6

SUMMARY AND CONCLUSIONS

The results of Chapter 5 indicate that very accurate numerical approximations to elliptic partial differential equations can be obtained when using higher order discretization formulae provided the analytic solutions are sufficiently smooth. A particular desired accuracy can also be achieved using substantially fewer internal mesh-points when applying a higher order discretization. A discretization using the five-point formula on a uniform mesh with a spacing of

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

gives an accuracy of

$$O(h^2) = O\left(\frac{1}{n^2}\right) .$$

In general, the accuracy obtained using a (4p+1)-point formula with a uniform mesh-spacing

$$k = \frac{1}{m}$$

is

$$O(k^{2p}) = O\left(\frac{1}{m^{2p}}\right) .$$

The truncation errors involved are of the same order provided

$$m = \sqrt[p]{n} \quad . \tag{6.1}$$

The relationship is illustrated in Table (6.1) assuming that the desired accuracy for a given problem can be achieved theoretically by using the five-point formula on a mesh of 161051 × 161051 internal points.

Table 6.1

DISCRETIZATION	NUMBER OF INTERNAL POINTS	CONDITION NUMBER OF MATRIX	
		V	W
5-point	161051×161051	1.2968×10^{10}	
9-point (a)	401×401		
13-point (a)	54×54		
17-point	20×20	1449.23	797.46
21-point	11×11	2597.62	2497.06

Tabulation of number of internal points and condition numbers of matrices for the same accuracy using various formulae.

The condition number of the matrices V and W, of chapter 2, for different discretization formulae are given in the last column of table (6.1). The condition numbers for V and W for the five-point formula are calculated using formula (5.1.3) and

$$||M^{-1}|| \leq \frac{(n+1)^2}{8},$$

with equality when n is odd. The norm used is the maximum absolute row sum (Rutherford [18]). It appears that condition numbers of matrices V and W for the 17-point formula are different since the matrices are different. A similar result appears for the 21-point formula. To obtain the condition numbers for the 20×20 and 11×11 matrices involved in the 17- and 21-point formulae respectively, the machine inverses of the matrices were used. Due to machine storage limitations, similar results are not available for the 401×401 and 54×54 matrices corresponding to the 9-point (a) and 13-point (a) entries of Table (6.1).

Since the mesh-spacing affects the discretization and round-off errors in the opposite sense (Ames [2], page 24, Ralston [17], page 80), the results for the five-point formula due to the use of a mesh-size as indicated above will be subject to severe round-off errors. The theoretical implication of relation (6.1) are, therefore, not desirable for large n when using a five-point formula. However, the results illustrate that a desired accuracy can be achieved with higher order formulae using substantially fewer points. The maximum relative error in case of example 1 of Chapter 5 are compared pairwise for different

discretization formulae for a practical illustration of relation (6.1) and presented in tables (6.2) to (6.5). Due to machine limitations, corresponding results using a five-point formula on a mesh with a sufficient number of internal points for a comparable accuracy are not available in all the cases. However, an approximate size of such a mesh as given by relation (6.1) is indicated for each of the tables (6.3) to (6.5).

Table 6.2

DISCRETIZATION	ORDER OF MATRIX	NUMBER OF ITERATIONS		
		4	5	6
5-point	25×25	6.8403×10^{-5}	1.9537×10^{-5}	1.9539×10^{-5}
9-point	5×5	1.2389×10^{-5}	1.2404×10^{-5}	1.2404×10^{-5}

Maximum relative error.

Table 6.3

DISCRETIZATION	ORDER OF MATRIX	NUMBER OF ITERATIONS		
		4	5	6
9-point (a)	18×18	4.4636×10^{-5}	4.3960×10^{-8}	4.3779×10^{-8}
13-point (a)	7×7	2.5550×10^{-5}	4.3862×10^{-8}	4.6898×10^{-8}

Maximum relative errors for accuracy comparable to solution when using a five-point formula on an $n \times n$ mesh where $n \approx 340$.

Table 6.4

DISCRETIZATION	ORDER OF MATRIX	NUMBER OF ITERATIONS		
		4	5	6
13-point (a)	21×21	1.1896×10^{-4}	6.3173×10^{-9}	3.5143×10^{-11}
17-point	10×10	3.4992×10^{-4}	2.4575×10^{-7}	4.1145×10^{-11}

Maximum relative errors for accuracy comparable to solution when using a five-point formula on an $n \times n$ mesh where $n \simeq 10^4$.

Table 6.5

DISCRETI-ZATION	ORDER OF MATRIX	NUMBER OF ITERATIONS			
		4	5	6	7
17-point	20×20	8.2773×10^{-4}	2.7019×10^{-6}	6.2131×10^{-12}	1.1742×10^{-13}
21-point	11×11	3.7227×10^{-3}	4.1666×10^{-5}	3.8520×10^{-9}	1.4754×10^{-13}

Maximum relative error for accuracy comparable to that of five-point formula on an $n \times n$ mesh where $n \simeq 16 \times 10^4$.

It appears from table (6.1) that matrices corresponding to higher order discretization formulae are better conditioned than matrices corresponding to the five-point formula when used to obtain the same order of accuracy.

The operation count for the methods cited in Chapter 3 for the solution of elliptic partial differential equations with Dirichlet boundary conditions using the five-point formula is

$$O(n^2 \log_2 n) ,$$

for an $n \times n$ mesh, whereas that for the method of Hoskins et al [14] for any order of discretization is

$$O(m^3)$$

for an $m \times m$ mesh.

For $p \geq 2$ and $m > 1$,

$$m^{2p} > m^3$$

$$\text{i.e. } pm^{2p} \log_2 m > m^3$$

$$\text{i.e. } m^{2p} \log_2 m^p > m^3$$

hence $n^2 \log_2 n > m^3$, using relation (6.1).

Therefore, it appears that the speed of $O(m^3)$ Poisson solvers based on higher order discretizations compares favourably with the speed of fast, i.e. $O(n^2 \log_2 n)$, Poisson solvers based on a five point discretization formula.

Although more work is required initially to set up the discretization matrices for higher order formulae, they need only be set up once and may be used for any problem when the same mesh is used. This is a small price to pay for the increase in accuracy and speed obtained when using higher order discretization formulae with the algorithm SOLVEXAAY.

BIBLIOGRAPHY

[1] M. Abramowitz and I. Stegun, Handbook of Mathematical Functions, Dover Publications, New York, 1972.

[2] W. F. Ames, Numerical Methods for Partial Differential Equations, Academic Press, New York, 1977.

[3] R. E. Bank, Ph.D. Thesis, pp. 4-16-18, Harvard 1975.

[4] R. E. Bank and D. J. Rose, "Marching algorithm for elliptic boundary value problems, I: The constant coefficient case", SIAM J. Numer. Anal., Vol. 14 No. 5, pp. 792-829, Oct. 1977.

[5] W. G. Bickley and J. McNamee, "Matrix and other direct methods for the solution of system of linear difference equations", Philos, Trans. Roy. Soc. London, Ser A, 252 (1960), pp. 69-131.

[6] B. L. Buzbee, G. H. Golub and C. W. Nielson, "On direct methods for solving Poisson's equations", SIAM J. Numer. Anal., Vol. 7, No. 4, pp. 627-656, Dec. 1970.

[7] L. Collatz, The Numerical Treatment of Differential Equations, Springer-Verlag, New York, 1966.

[8] R. Courant and D. Hilbert, Methodsof Mathematical Physics Vol. 2 Interscience, New York, 1966.

[9] F. W. Dorr, "The discrete solution of the direct Poisson's equation on a rectangle", SIAM Review, Vol. 12, No. 2, pp. 248-263, April 1970.

[10] G. E. Forsythe and W. R. Wasow, Finite Difference Methods for Partial Differential Equations, John Wiley & Sons, New York, 1965.

[11] L. Fox, Numerical Solution of Ordinary and Partial Differential Equations, Pergamon Press, Oxford, 1962.

[12] S. Goldberg, Introduction to Difference Equations, New York, Wiley, 1958.

[13] R. W. Hockney, "A fast direct solution of Poisson's equation using fourier analysis", J. Assoc. Comput. Mach., 12 (1965), pp. 95-113.

[14] W. D. Hoskins, D. S. Meek and D. J. Walton, "The numerical solution of the matrix equation XA + AY = F", BIT 17 (1977), pp. 184-190.

[15] L. Kantorovich and V. Krylov, Approximate Methods in Higher Analysis, Noordhoff (Interscience) 1958.

[16] A. R. Mitchell, Computational Methods in Partial Differential Equations, John Wiley & Sons, Toronto, 1969.

[17] A. Ralston, A First Course in Numerical Analysis, McGraw-Hill Book Company, New York, 1965.

[18] D. E. Rutherford, "Some continuant determinants arising in Physics and Chemistry, II", Proc. Roy. Soc. Edinburgh Sect. A, V. 63, 1952, pp. 232-241, MR 15, 495.

[19] G. D. Smith, Numerical Solution of Partial Differential Equations, Oxford University Press, 1969.

[20] R. A. Sweet, "A cyclic reduction algorithm for solving block tridiagonal system of arbitrary dimension", SIAM J. Numer. Anal. Vol. 14, No. 4, pp. 706-720, Sept. 1977.

[21] P. N. Swarztrauber, "A direct method for the discrete solution of separable elliptic equations", SIAM J. Numer. Anal. Vol. 11, No. 6, pp. 1136-1150, Dec. 1974.

[22] _____, "The methods of cyclic reduction, Fourier analysis and the FACR algorithm for the discrete solution of Poisson's equation on a rectangle", SIAM Review, Vol. 19, No. 3, pp. 490-501, July 1977.

[23] J. Todd, Basic Numerical Mathematics, Vol. 2, Numerical Algebra, Birkhäuser Verlag Basel Und Stuttgart, 1977.

[24] R. S. Varga, Matrix Interative Analysis, Prentice Hall, Englewood Cliffs, New Jersey, 1962.

[25] D. J. Walton, "Some new methods for the solution of matrix equations arising from discretized partial differential equations", Ph.D. dissertation, University of Manitoba, Winnipeg, Canada, 1978.