PRECONDITIONED EXPLICIT DECOUPLED GROUP METHODS FOR SOLVING ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS

by

ABDULKAFI MOHAMMED SAEED AHMED

Thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

April 2011

ACKNOWLEDGEMENTS

In the name of Allah, Most Gracious, Most Merciful

First and foremost, all thank and praise to Allah for blessing me well, the health and the resources to complete this thesis. I believe such success granted by Allah primarily refers to the supplications of my parents for me.

My appreciation and sincere gratitude go to my supervisor, Associate Professor Norhashidah Hj Mohd Ali for her constant guidance, continual support, patience, comments and suggestions throughout the completion of this thesis. I have benefited enormously from her moral support, insightful suggestions and deepest experience in the field of partial differential equations. My debt to her is enormous.

I would also like to take this opportunity to thank the Dean of the School of Mathematical Sciences, Universiti Sains Malaysia, Professor Ahmad Izani Md. Ismail, lecturers and staff of the department for their kind advice and support which have helped me to complete my Ph.D. thesis.

I also wish to extend my gratitude to Professor Jun Zhang for his useful discussions and valuable comments during his visit to the School of Mathematical Sciences, Universiti Sains Malaysia.

Special thanks to all my friends, especially those who have provided helpful suggestions and encouragements.

Last but not least, my heartfelt thanks to my mother, my wife, my son, my daughters, my brother and my sisters, for their unconditional love and support. Their encouragement was undoubtedly priceless.

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

$\|A\|$	Norm of matrix A
$\det(A)$	Determinant of A
$\|\lambda\|$	The modulus of λ
$\rho(A)$	Spectral radius of matrix A
$R_k(G)$	Average rate of convergence after k iterations
$\underset{\sim}{0}$	(2×2) Null matrix
I	Unit matrix
a_{ij}	Entry of matrix A located on the i^{th} row and j^{th} column
$[a,b]$	Closed interval $a \le x \le b$
(a,b)	Open interval $a < x < b$
$0(h)$	Order of truncation error
$\sum$	Summation
Re	Reynolds number

LIST OF ABBREVIATIONS

PDEs	:	Partial Differential Equations
i=1(1) n	:	i varies from 1 to n by intervals of 1, i.e. i=1,2,3,...,n-1,n
SOR	:	Successive Over Relaxation
BSOR	:	Block Successive Over Relaxation
EG	:	Explicit Group
EDG	:	Explicit Decoupled Group
MEG	:	Modified Explicit Group
MEDG	:	Modified Explicit Decoupled Group
EG SOR	:	Explicit Group Successive Over Relaxation
EDG SOR	:	Explicit Decoupled Group Successive Over Relaxation
MEG SOR	:	Modified Explicit Group Successive Over Relaxation
MEDG SOR	:	Modified Explicit Decoupled Group Successive Over Relaxation
CFD	:	Computational Fluid Dynamics
CO	:	Consistently Ordered
CO *(q, r)*	:	*(q, r)*-Consistently Ordered matrix
GCO (q,r)	:	Generalized *(q,r)*-Consistently Ordered
IDDM	:	Irreducibly Diagonally Dominant Matrix
SDD	:	Strictly Diagonally Dominant
GDD	:	Generalized Diagonally Dominant
DDD	:	Doubly Diagonally Dominant

KAEDAH KUMPULAN NYAHPASANGAN TAK TERSIRAT BERPRASYARAT UNTUK MENYELESAIKAN PERSAMAAN PEMBEZAAN SEPARA ELIPTIK

ABSTRAK

Perkembangan yang pesat bagi kaedah beza hingga adalah didorong oleh keperluan untuk mengatasi masalah yang kompleks hari ini dalam sains dan teknologi. Keperluan terkini bagi penyelesaian lebih cepat dan untuk menyelesaikan masalah saiz besar yang muncul dalam pelbagai aplikasi dalam bidang sains, seperti pemodelan, simulasi sistem yang besar dan dinamik bendalir. Oleh kerana itu, kajian yang berkaitan dengan teknik pemecutan telah dilakukan untuk mencapai keperluan tersebut. Terdapat beberapa teknik pendiskretan yang boleh digunakan untuk membina persamaan anggaran bagi menganggarkan persamaan pembezaan separa (PPS) seperti beza terhingga, elemen terhingga dan isipadu terhingga. Pendekatan persamaan ini akan digunakan untuk menghasilkan sistem persamaan linear yang bersepadan yang biasanya besar dan jarang. Kaedah lelaran menjadi lebih cekap berbanding dengan kaedah yang lain kerana ruangan simpanan yang diperlukan untuk penyelesaian lelaran pada komputer kurang ketika matriks pekali dari sistem ini adalah jarang. Kaedah Kumpulan lelaran tak tersirat berdasarkan anggaran beza sehingga putaran telah ditunjukkan jauh lebih cepat daripada kaedah yang berdasarkan pada rumus lima titik piawai dalam menyelesaikan PPS yang disebabkan oleh kompleksiti pengiraan keseluruhan yang lebih rendah kaedah tersebut. Terdapat beberapa pendekatan alternatif baru terhadap tujuan meningkatkan kadar penumpuan dalam menyelesaikan sistem linear besar akibat pendiskretan kaedah ini.

Teknik Berprasyarat menyediakan pendekatan alternatif baru layak dalam mencapai tujuan ini. Motivasi utama dari penyelidikan ini adalah untuk

membangunkan prasyarat terhadap kaedah lelaran berkumpulan tak tersirat dalam menyelesaikan beberapa jenis PPS umum yang eliptik dan PPS Navier-Stokes keadaan mantap. Kerja ini berkaitan dengan pelaksanaan prasyarat jenis pemisahan tertentu dalam perumusan blok yang diterapkan pada sistem yang asli yang diperolehi dari kaedah Kumpulan Nyahpasangan Tak Tersirat (KNTT) empat titik dan kaedah Kumpulan Nyahpasangan Tak Tersirat Terubahsuai empat titik (KNTTT) bagi menyelesaikan PPS eliptik dan persamaan Navier-Stokes keadaan mantap. Ujikaji berangha dijalankan ke atas setiap skema berprasyarat dan tidak berprasyarat yang dibangunkan bagi tujuan perbandingan. Keputusan menunjukkan bahawa terdapat pembaikan pada kadar penumpuan dan kecekapan skema lelaran berprasyarat yang baru diformulasi. Selanjutnya, analisis teoritis kaedah berprasyarat ini dilakukan untuk membuktikan bahawa prasyarat-prasyarat yang dicadangkan memenuhi beberapa sifat penumpuan teoritis yang meningkatkan kadar penumpuan skema lelaran kumpulan tak tersirat yang asal.

PRECONDITIONED EXPLICIT DECOUPLED GROUP METHODS FOR SOLVING ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS

ABSTRACT

The highly concern development of finite difference methods was stimulated by the need to cope with today's complex problems in science and technology. The current requirement for faster solutions and for solving large size problems arises in a variety of applications in science, such as modeling, simulation of large systems and fluid dynamics. Therefore, studies regarding several accelerated techniques have been carried out to achieve these requirements. There are several discretisation techniques that can be used to construct approximation equations for approximating partial differential equations (PDEs) such as finite difference, finite element and finite volume. These approximation equations will be used to generate the corresponding systems of linear equations which are normally large and sparse. The iterative methods are more efficient compared to the other methods since the storage space required for iterative solutions on a computer is less when the coefficient matrix of the system is sparse. Group explicit iterative methods based on the rotated finite difference approximations have been shown to be much faster than the methods based on the standard five-point formula in solving PDEs which are due to the formers' overall lower computational complexities. There are some new alternative approaches towards increasing the rate of convergence in solving large linear system resulting from the discretisation of these methods.

Preconditioning techniques provide a new feasible alternate approach in achieving this aim. The primary motivation of this research is to develop

preconditioners to the group explicit iterative methods in solving several common types of PDEs which are elliptic PDEs and steady state Navier-Stokes equations. This work is concerned with the application of a specific splitting-type preconditioner in block formulation applied to the original system obtained from the four point Explicit Decoupled Group (EDG) method and four point Modified Explicit Decoupled Group (MEDG) method for solving the elliptic PDEs and steady state Navier-Stokes equations. Numerical experiments are conducted on each developed non-preconditioned and preconditioned schemes for comparison purposes. The results reveal that there are improvements on the convergence rate and the efficiency of the newly formulated preconditioned iterative schemes. Furthermore, a theoretical analysis of these preconditioned methods is performed to prove that the proposed preconditioners satisfy some theoretical convergence properties which increase the convergence rate of the original group explicit iterative schemes.

CHAPTER 1

PRELIMINARIES

1.1 Introduction

Many physical phenomena in engineering, fluid dynamics and static field problems particularly in the electromagnetic field and the incompressible potential flow field are described by partial differential equations (PDEs) such as elliptic PDEs. These PDEs however, are usually difficult to solve analytically so that approximation methods become the alternate means of solutions. These approximation methods did not become a useful and popular proposition in its early days of introduction. The appearance of high speed computers was the impulse to the change in sentiment.

There are various numerical methods which can be used to solve PDEs. The methods include finite difference method, finite element method and finite volume method. Among these approximation methods, finite difference method is one of the more frequently used method due to their simplicity and universal applicability, plus, being one of the oldest method available (Ibrahim, 1993; Ali, 1998). When solved by the finite difference methods, the PDEs lead to a large and sparse system of linear equations which may be solved either by direct or indirect methods. Direct methods, however, usually involve rather complicated algorithms which yield the exact solutions in a finite number of steps. On the other hand, iterative method is one type of indirect methods which involve repetition of simple algorithms which lead to better approximation successively so that the exact answer is obtained as a limit of a sequence (Smith, 1985). In the following section we will display the motivation of this work. Moreover, the research problems, objectives and scope will also be presented.

1.2 The Motivation of This Research

It is known that iterative methods require less amount of storage space when the sparse matrix (many of its element are zeroes) is involved. Therefore, iterative method is more suitable in solving a large and sparse linear system.

When a linear system involved is getting larger, it will require more time to get a precise solution since the iterations are increasing too. Hence, preconditioned methods are introduced to increase the rate of convergence for the iterative methods. Roughly speaking, preconditioned methods are any form of modifying the original linear system so that it decreases the number of iterations needed to converge without changing its exact solution. Therefore we can define a preconditioner as a matrix that transforms the linear system into one that is equivalent in the sense that it has the same solution, but that has more favorable spectral properties.

For this thesis, new preconditioned iterative methods in solving several types of PDEs are formulated to accelerate their rates of convergence.

1.3 Research Problems

Group iterative methods based on the finite difference approximations have been shown to be much faster than the point iterative methods based in solving the PDEs which is due to the formers' overall lower computational complexities. Improved techniques using explicit group methods derived from the standard and skewed (rotated) finite difference operators have been developed over the last few years in solving the linear system that arise from the discretization of these PDEs (Yousif and

Evans, 1986; Evans and Yousif, 1990; Abdullah, 1991; Yousif and Evans, 1995; Othman and Abdullah, 2000; Ali et al., 2004; Ali and Ng, 2007). The rate of convergence of these group explicit iterative methods can be improved by using preconditioning techniques.

In this thesis, a second-order finite difference scheme derived from rotated discretisation formula is employed in conjunction with a preconditioner to obtain highly accurate and fast numerical solution of the two-dimensional elliptic partial differential equation and steady-state Navier-Stokes equation. We consider a more general form of the two dimensional steady-state Navier-Stokes equations which consisting of a coupled system of elliptic PDEs. The construction of a specific splitting-type preconditioner in block formulation applied to a class of group relaxation iterative methods derived from these rotated (skewed) finite difference approximations will be investigated to improve the convergence rates of these methods for solving the above types of equation. This preconditioned version of these iterative methods will be shown to have much better convergence rates than the regular version. In addition, the convergence properties of the proposed preconditioners which applied to the linear systems resulted from the explicit decoupled group iterative schemes in solving elliptic PDE and steady-state Navier-Stokes equation will be given in this research.

1.4 Research Objectives

The objectives of this thesis are as follows:

i) To derive a suitable preconditioner for the Explicit Decoupled Group (EDG) iterative method due to Abdullah (1991) which is able to accelerate the rate of convergence of this method for solving the elliptic PDEs.

ii) To formulate a suitable preconditioner for the Modified Explicit Decoupled Group (MEDG) iterative method due to Ali and Ng (2007) which is able to improve the rate of convergence of this method for solving the elliptic PDEs.

iii) To improve the acceleration of the Explicit Decoupled Group (EDG) iterative method due to Ali and Abdullah (1999) for solving a two-dimensional steady-state Navier-Stokes equation by using a suitable preconditioning technique.

iv) To enhance the convergence rate of the Modified Explicit Decoupled Group (MEDG) for the solution of the steady-state Navier-Stokes equation.

v) To compare the performance of these preconditioned methods with their unpreconditioned counterparts through numerical experiments.

vi) To establish the theoretical convergence properties of the proposed preconditioned methods.

The main goal of this work is to formulate new suitable preconditioners and apply them to EDG and MEDG iterative methods to accelerate the convergence rate of these methods. The details of these formulations will be given in Chapters 4 and 5. In addition to these formulations of the new preconditioned methods, the convergence analysis of these proposed methods will be introduced in this work and new convergence theorems will be established to verify the results in chapter 6.

1.5 Research Scope

From the discretisation of the elliptic PDEs, large sparse linear system of the following form will be resulted:

$$A\tilde{u} = \tilde{b} \tag{1.1}$$

where A is an unstructured large sparse matrix of order n. It is common belief that, for solving very large sparse linear systems, iterative methods are becoming the method of choice, due to their more favorable memory and computational costs, compared to the direct solution methods. A common strategy to enhance the convergence rates of iterative methods is to exploit preconditioning techniques by transforming Equation (1.1) into:

$$M^{-1}A\tilde{u} = M^{-1}\tilde{b}, \qquad (1.2)$$

in which M is a nonsingular matrix of the same order of the matrix A. It is obvious that the Equations (1.1) and (1.2) are equivalent and have the same solution. The matrix M is called the preconditioning matrix or preconditioner. The usefulness of a preconditioner depends very much on how much it can reduce the spectral value of the coefficient matrix and decrease the time needed to solve the linear system with an iterative method. If the choice of the preconditioning matrix is near to A, then the matrix $M^{-1}A$ will be near to identity matrix. This guarantees that the eigenvalue of matrix $M^{-1}A$ is near to 1. Therefore, it will converge faster with any iterative method.

Gunawardena, *et.al* (1991) was one of the early researchers of the preconditioned method. Their research applied the preconditioner P which eliminates the elements of the first upper codiagonal of A in Equation (1.1), where $P = I + \bar{S}$, I is the identity matrix which have the same dimension with A while $\bar{S}$ is the elements of the first upper diagonal of A,

$$\bar{S} = \begin{bmatrix} 0 & -a_{12} & 0 & \cdots & 0 \\ 0 & 0 & -a_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & -a_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

and the system become

$$(I+\bar{S})A\tilde{u}=(I+\bar{S})\tilde{b} \ . \tag{1.3}$$

This preconditioner improved the convergence rate of Gauss-Seidel iterative method. Such work had been further enhanced by Usui et al. (1995). Martins et al. (2001) analyzed and verified the superiority of the preconditioner proposed by Usui et al. (1995) theoretically. In Lee (2006), preconditioners have been successfully applied on the standard five point formula in solving the Poisson problem with the Dirichlet boundary conditions and the numerical experiments yield very encouraging results. As an extension of the preconditioner P in (1.3), we can modify and formulate new preconditioners which will be suitable to be applied to the coefficient matrices resulted from the class of explicit decoupled group methods. Due to the specific structures of the coefficient matrices resulted from the EG and MEG methods, the formulated preconditioners are found to be unsuitable for these iterative schemes and therefore will not be discussed in this thesis. The preconditioned system will be further discussed in Chapters 4, 5 and 6.

1.6 Organization of the Thesis

The thesis commences with the general foundation and the fundamentals of numerical solutions of PDEs by the finite difference methods. A general introduction to preconditioned method is also presented. The second chapter includes a review on basic concept PDEs and methods for solving systems of equations which are direct methods and indirect methods included point and group iterative methods. This chapter ends with a brief survey of the preconditioning methods currently available for the solution of linear system arising from the discretisation of the PDEs.

The development of formulas for group iterative methods such as Explicit Group (EG), Explicit Decoupled Group (EDG), Modified Explicit Group (MEG) and Modified Explicit Decoupled Group (MEDG) will be discussed in Chapter 3.

The formulation of new preconditioned for EDG SOR and MEDG SOR iterative methods in the solution of both elliptic PDEs and Navier-Stokes equation, is the main concern of Chapters 4 and 5. In the next chapter, the comparison theorems on the proposed preconditioned iterative methods are made to confirm the superiority of these new methods and to evaluate the efficiency of these proposed methods. Chapter 7 concludes the thesis and presents thesis summary, limitations and future work.

CHAPTER 2

BASIC CONCEPT OF PDEs AND METHODS FOR SOLVING SYSTEMS OF EQUATIONS

2.1 Introduction

It has been affirmed that the discretisation of PDEs using finite difference schemes normally yield a system of linear equations, which are large and sparse in nature. Iterative methods are usually used to solve these types of systems since these methods need less storage and are capable of preserving the sparsity property of the large system. The advantages of iterative methods are the simplicity and uniformity of the operations to be performed, which make them well suited for use on computers. Direct method is preferable if the coefficient matrix is dense. If the matrix is sparse, the use of direct methods requires a lot of storage space due to the problem of fill-in of the coefficient matrix. That is during the elimination process, entries in the coefficient matrix that were previously zero become nonzero. In this chapter, we will overview on basic concept of PDEs. In addition to that we will discuss some of the well known direct and iterative methods to solve a linear system of equations of the form:

$$A\tilde{u} = \tilde{b}, \tag{2.1}$$

where $A = (a_{ij}) \in R^{n \times n}$ is an $n \times n$ non-singular sparse matrix.

2.2 Classifications of Partial Differential Equations and Types of Boundary Conditions

A PDE can be defined as an equation that consists of one or more partial derivatives of an unknown function with respect to two or more independent variables. In general, a PDE for the dependent variable u and independent variables x and y can be written in the form as below:

$$F(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial x \partial y}, \frac{\partial^2 u}{\partial y^2}, \frac{\partial^3 u}{\partial x^3}, \frac{\partial^3 u}{\partial x^2 \partial y}, ...) = 0. \quad (2.2)$$

The order of a PDE is determined by the order of the highest partial derivative that occurs in the equation. The general form of second-order PDE in two independent variables can be expressed as

$$a\frac{\partial^2 u}{\partial x^2} + b\frac{\partial^2 u}{\partial x \partial y} + c\frac{\partial^2 u}{\partial y^2} + d\frac{\partial u}{\partial x} + e\frac{\partial u}{\partial y} + fu + g = 0, \quad (2.3)$$

where the equation is said to be linear if a, b, c, d, e, f and g are independent of u or its derivatives. The linear second-order PDE can be further distinguished according to their mathematical forms which are elliptic, parabolic and hyperbolic. Depending on the coefficients of the second derivative in (2.3) the equation is elliptic if $b^2 - 4ac < 0$, parabolic if $b^2 - 4ac = 0$ and hyperbolic if $b^2 - 4ac > 0$.

In general, elliptic PDEs govern steady-state or equilibrium problems and this thesis mainly deals with this elliptic problems. Examples of the known elliptic equations are

Poisson equation: $$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = g \quad (2.4a)$$

and Laplace equation: $$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \quad (2.4b)$$

Examples of parabolic PDEs are

Heat equation:

$$\frac{\partial u}{\partial t} = \alpha^2 \frac{\partial^2 u}{\partial x^2}, \quad \alpha^2 \text{ is a physical constant,} \quad (2.5)$$

and two dimensional diffusion equation

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}. \tag{2.6}$$

The simplest example of a hyperbolic PDE is the wave equation which may be written as

$$\frac{\partial^2 u}{\partial t^2}-\frac{\partial^2 u}{\partial x^2}=0. \tag{2.7}$$

Elliptic PDEs are usually classified as boundary value problems since boundary conditions are specified around region as shown in Figure 2.1.

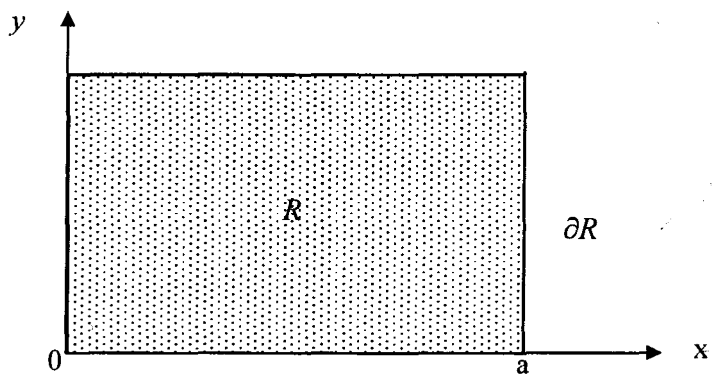

Figure 2.1 Computational domain for an elliptic PDE

For parabolic equation, initial boundary values are supplied on the sides of the open region, and the solutions march forwards the open side as shown in Figure 2.2.

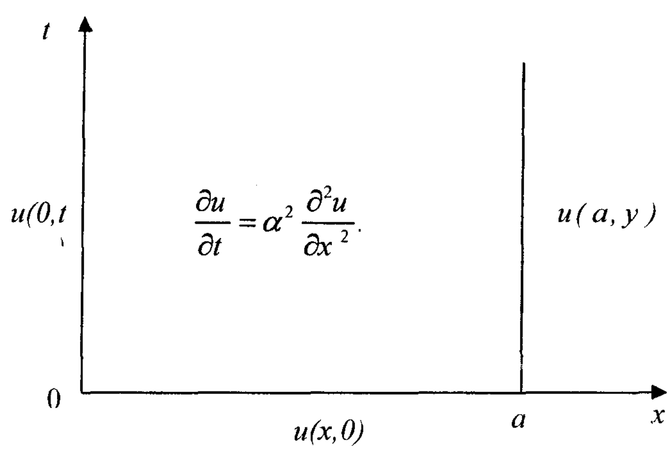

Figure 2.2 Domain of parabolic PDE.

To completely solving a PDE which describes a physical problem, the conditions required to determine the unique solution of a PDE are usually the boundary conditions and/or the initial conditions. These initial and boundary conditions can be classified into three different classes as below:

i) Dirichlet condition, where the condition $u=f$ is specified at each point of the boundary ∂R.

ii) Neumann condition, where the values of the normal derivatives, $\frac{\partial u}{\partial n}$, are given on ∂R, such that: $\frac{\partial u}{\partial n}$, denotes the directional derivative of u along the outward normal to ∂R.

iii) Robin's condition, where a linear combination of function u and its derivatives are given along the boundary ∂R. i.e. $\frac{\partial u}{\partial n} + ku = f$ on ∂R ($k>0$).

The physical meaning of the above three boundary value problems can be illustrated by the problem of steady-state temperature distribution.

The general approach to nonlinear equations is still the "linearize and iterate" approach. In this case consider that some initial approximation is known to the solution and an improved approximation to the solution is desired.

2.3 Basic Mathematical Concepts

Normally when finite difference methods are applied to the numerical solution of PDEs, a system of m simultaneous equations with n unknowns are usually involved in its solution process. In this section, basic mathematical concepts of matrix algebra definitions and theories relevant to the study of numerical methods are outlined and presented.

2.3.1 Matrix Algebra

An arbitrary system of n linear equations in n unknowns can be written as:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\quad \vdots \quad &\quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \tag{2.8}$$

where $x_1, x_2, x_3, \cdots, x_n$ are the unknowns and the subscripted a's and b's denote constants.

This system can be rewritten in the matrix form as:

$$A\tilde{x} = \tilde{b} \tag{2.9}$$

where A is the matrix of order $n \times n$ while $\tilde{x}$ and $\tilde{b}$ are row vectors of n order such that:

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{11} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \tilde{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \tilde{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}. \tag{2.10}$$

The entries of A are represented by a_{ij}, with i and j representing row and column respectively.

When A and $\tilde{b}$ are known, the solution of system (2.9) is the vector $\tilde{x}$. This system has a unique solution $\tilde{x} = A^{-1}\tilde{b}$ provided A is non-singular ($det\ A \neq 0$). However, if the size of the matrix is large, it would be very difficult to use this definition in finding the solution. In these cases, properties of the coefficient matrix A, such as diagonal dominance, positive definiteness and consistently ordered, can help decide the solvability of the system.

In this thesis, all matrices are assumed to be square matrices with order n unless stated otherwise. All matrices will be represented by capital letters and all small letters denotes the entries of the matrices.

Two matrices A and B are defined to be equal if they have the same size and their corresponding entries are equal. Mathematically, it means $a_{ij} = b_{ij}$ for $1 \leq i,j \leq n$.

Definition 2.3.1

A matrix $A = [a_{ij}]$ is said to be positive ($A > 0$) if $a_{ij} > 0$ for $1 \leq i,j \leq n$. However, the matrix

A is non negative ($A \geq 0$) if $a_{ij} \geq 0$ for $1 \leq i,j \leq n$ (Berman and Plemmons, 1994).

Definition 2.3.2

i) A matrix A is called a zero (null) matrix if all the entries are zero.

ii) A matrix $A = [a_{ij}]$ is called an identity matrix if

$$\{a_{ii} = 1 \text{ for all } 1 \leq i \leq n \text{ and } a_{ij} = 0 \text{ for all } 1 \leq i, j \leq n \text{ where } i \neq j\}$$

The following discusses several useful properties of a matrix due to Golub and Van Loan (1983) and Mitchell (1969).

The matrix $A = [a_{ij}]$ of order n is:

i) Symmetric, if $A = A^T$.

ii) Skew-symmetric matrix, if $A = -A^T$.

iii) Positive definite matrix, if $\tilde{x}^T A \tilde{x} > 0$ for $\tilde{x} \neq 0$, $\tilde{x} \in \mathbb{R}^n$.

iv) Diagonal, if $a_{ij} = 0$ for all $1 \le i, j \le n$ where $i \neq j$.

v) Diagonally dominant, if $|a_{ii}| \geq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$ for all $1 \le i \le n$.

vi) Band matrix, if $a_{ij} = 0$ for $|i-j| > q$, where $2q+1$ is the bandwidth of A.

vii) Tridiagonal matrix, if $q=1$ and it has the form as in Figure 2.3

$$A = \begin{bmatrix} a & b & 0 & \cdots & 0 \\ c & a & b & \ddots & \vdots \\ 0 & c & a & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & b \\ 0 & \cdots & 0 & c & a \end{bmatrix}$$

Figure 2.3 Tridiagonal Matrix

viii) Lower triangular, if $a_{ij} = 0$ for $i \le j$ and strictly lower triangular if $a_{ij} = 0$ for $i < j$.

ix) Upper triangular, if $a_{ij} = 0$ for $i \geq j$ and strictly upper triangular if $a_{ij} = 0$ for $i > j$.

x) Sparse matrix, if most of the entries elements are zeroes.

xi) Dense matrix, if most of the entries elements are nonzeroes.

The determinant of a matrix A is denoted as *det* (A) or $|A|$. For a matrix A with only a single entry, the determinant of A is the value of the single entry itself. If matrix A is of order 2, for example $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then $|A| = ad - bc$. Minor of an element a_{ik}

is the determinant of the sub matrix in matrix A. It is denoted as M_{ik}. The cofactor of the element a_{ik} can be obtained from $C_{ik} = (-1)^{i+k} M_{ik}$. Therefore the determinant of A is given by

$$|A| = \sum_{k=1}^{n} M_{ik}, \qquad 1 \leq i \leq n. \tag{2.11}$$

Definition 2.3.3

A matrix A is said to be

1) Block Diagonal, if

$$A = \begin{bmatrix} D_1 & & & & \\ & D_1 & & & \\ & & D_1 & & \\ & & & \ddots & \\ & & & & D_1 \end{bmatrix}$$

Figure 2.4 Block Diagonal Matrix

2) Block Tridiagonal, if

$$A = \begin{bmatrix} D_1 & U_1 & & & & \\ L_2 & D_2 & U_2 & & & \\ & L_3 & D_3 & U_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & L_{n-1} & D_{n-1} & U_{n-1} \\ & & & & L_n & D_n \end{bmatrix}$$

Figure 2.5 Block Tridiagonal Matrix

where D_i ,$1 \leq i \leq n$ are square matrices, whereas U_i*'s* and L_i*'s* are rectangular matrices (Evans, 1997).

If the D_i*'s* are square diagonal matrices, Young (1971) referred to this type of matrix as T-matrix.

Definition 2.3.4

A matrix $A = [a_{ij}]$ of order $n > 1$ is said to be irreducible if for any two non-empty disjoint subsets S and T of $W = \{1,2,\ldots,n\}$ where $S + T = W$, there exists $i \in S$ and $j \in T$ such that $a_{ij} \neq 0$.

The following definition of irreducibly diagonally dominant matrix is due to Berman and Plemmons (1994).

Definition 2.3.5

A matrix $A= [a_{ij}]$ is an irreducibly diagonally dominant matrix (IDDM) iff A is irreducible, $|a_{ii}| \geq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$ for all $1 \leq i \leq n$ and there is at least a strict inequality holds in this inequalities.

Theorem 2.3.1

If A is an irreducible diagonally dominant matrix, then $\det(A) \neq 0$ with nonvanishing diagonal elements.

Since the topic of eigenvalues and eigenvectors play an important role in the convergence theorems of iterative methods, which will be widely discussed later, the following presents a brief discussion on these themes.

2.3.2 Eigenvalues and eigenvectors

The eigenvalue of the matrix A of order n is a real (or complex) number, λ which satisfy the equation

$$A\tilde{x} = \lambda \tilde{x} \tag{2.12}$$

where $\tilde{x}$ is a non zero vector. The vector $\tilde{x}$ is called the eigenvector of the corresponding eigenvalue λ . Equation (2.12) can be rewritten as

$$(A - \lambda I)\tilde{x} = \tilde{0}. \tag{2.13}$$

A nontrivial solution to Equation (2.13) exists if and only if the matrix $(A - \lambda I)$ is singular, which means

$$\det(A - \lambda I) = 0. \tag{2.14}$$

Equation (2.14) is called the characteristic equation and its roots λ_i constitute the eigenvalues of the matrix A.

The characteristic equation of the degree n will give n numbers of eigenvalues for A. However, not all eigenvalues are needed. Usually, only the largest of the moduli of the eigenvalues known as spectral radius will be considered.

Definition 2.3.6

Given a matrix A of order n with eigenvalues λ_i, $1 < i < n$, then the spectral radius $\rho(A)$ is given by

$$\rho(A) = \max|\lambda_i|. \tag{2.15}$$

In Smith (1985), the eigenvalues of some common matrices are formulated as the following:

The eigenvalues of the $(n \times n)$ matrix

$$A = \begin{bmatrix} a & b & 0 & \cdots & 0 \\ c & a & b & \ddots & \vdots \\ 0 & c & a & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & b \\ 0 & \cdots & 0 & c & a \end{bmatrix},$$

are given by

$$\lambda_i = a + 2\sqrt{bc}\,cos(\frac{k\pi}{n+1}), \qquad k=1, 2, \ldots, n. \tag{2.16}$$

where *a, b* and c may be real or complex.

If *A* is an $(n \times n)$ cyclic tridiagonal matrix, i.e.,

$$A = \begin{bmatrix} a & b & & & & c \\ c & a & b & & & \\ & \ddots & \ddots & \ddots & 0 & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & c & a & b \\ b & & & & c & a \end{bmatrix}$$

then the eigenvalues are given by

$$\lambda_i = a + 2\sqrt{bc}\,cos(\frac{2k\pi}{n}), \qquad k=1, 2, \ldots, n. \tag{2.17}$$

Theorem 2.3.2

A real matrix is positive (non negative) definite if and only if it is symmetric and all its eigenvalues are positive (non negative, with at least one eigenvalue equal to zero) (Evans, 1997).

Two matrices are called commutative if $AB = BA$. They then possess the same set of eigenvectors.

For the purpose of analyzing the errors, the approximate methods are often associated with some vectors and matrices of which their magnitudes are measurable as non negative scalars. Such a measuring concept is called a norm.

Definition 2.3.7

Let the vector x be given by $x^T = [x_1, x_2, ..., x_n]$, the following scalars are defined as the 1, 2, and ∞ norm of a vector x:

$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n| \tag{2.18}$$

$$\|x\|_2 = (\sum_{i=1}^{n} |x_i|^2)^{\frac{1}{2}}, \quad \|x\|_\infty = \sup_{1 \le i \le n} |x_i|. \tag{2.19}$$

In general L_k-norms are given by

$$\|x\|_k = (\sum_{i=1}^{n} |x_i|^k)^{\frac{1}{k}}, \quad 1 \le k \le \infty . \tag{2.20}$$

A matrix norm $\|A\|$ is said to be compatible with a vector norm $\|x\|$ if:

$$\|Ax\| \le \|A\| \|x\|, \quad \text{for all non zero } x. \tag{2.21}$$

Theorem 2.3.3

If A is a matrix of order n, then

$$\|A\| \ge \rho(A). \tag{2.22}$$

Proof. See Evans (1997).

2.3.3 Property A and Consistently Ordered Matrices

In this section we will discuss three important properties which play important roles in the theoretical analysis of successive-over-relaxation (SOR) iterative methods which are, property A, consistently ordered (CO) and generalized consistently ordered (GCO) properties.

Definition 2.3.8

A matrix A of order n has property A if there exists two disjoint subsets S and T of

$W = \{1, 2, ..., n\}$ such that if $i \neq j$ and if either $a_{ij} \neq 0$ and $a_{ji} \neq 0$, then $i \in S$ and $j \in T$ else $i \in T$ and $j \in S$ (Martins et al, 2002).

Definition 2.3.9

A matrix A of order n is consistently ordered if for some t there exist disjoint subsets $S_1, S_2, ..., S_t$ of $W = \{1,2,...,n\}$ such that $\sum_{k=1}^{t} S_k = W$ and such that if i and j are associated, then $j \in S_{k+1}$ if $j > i$ and $j \in S_{k-1}$ if $j < i$, where S_k is the subset containing i (Martins et al, 2002).

Moreover, if A is consistently ordered, then the matrix A has property A.

An accurate analysis of convergence properties of the block SOR method is possible if the matrix A is consistently ordered in the following sense (Saridakis, 1986).

Definition 2.3.10

For given positive integers q and r, the matrix A of ordered n is a (q, r)-consistently ordered matrix (a CO (q, r) - matrix) if for some t, there exist disjoint subsets $S_1, S_2, ..., S_t$ of $W = \{1,2,...,N\}$ such that $\sum_{k=1}^{t} S_k = W$ and such that: if $a_{i,j} \neq 0$ and $i < j$, then $i \in S_1 + S_2 + ... + S_{t-r}$ and $j \in S_{k+r}$, where S_k is the subset containing i; if $a_{i,j} \neq 0$ and $i > j$, then $i \in S_{q+1} + S_{q+2} + ... + S_t$ and $j \in S_{k-q}$ where S_k is the subset containing i.

Definition 2.3.11

A matrix A is a generalized (q,r)-consistently ordered matrix (a $GCO(q,r)$-matrix) if: $\Delta = det(\alpha^q E + \alpha^{-r} F - kD)$ is independent of α for all $\alpha \neq 0$ and for all k. Here

$D=diag\ A$ and E and F are strictly lower and strictly upper triangular matrices, respectively, such that: $A=D-E-F$.

The following properties usually used in the theoretical analysis of group iterative methods and referred to Martins et al. (2002).

An ordered grouping π of $W=\{1,2,\ldots,n\}$ is a subdivision of W into disjoint subsets $R_1, R_2, \ldots, R_q$ such that $R_1+R_2+\ldots+R_q = W$.

Given a matrix A and an ordered grouping π we define the submatrices $A_{m,n}$ for $m,n = 1,2,\ldots q$ as follows: $A_{m,n}$ is formed from A deleting all rows except those corresponding to R_m and all columns except those corresponding to R_n .

Definition 2.3.12

Let π be an ordered grouping with q groups. A matrix A has property $A^{(\pi)}$ if the $q\times q$ matrix $Z=(z_{r,s})$ defined by $z_{r,s}=\{\ 0\ if\ A_{r,s}=0\ \ or\ \ 1\ if\ A_{r,s}\neq 0\ \}$ has property A.

Note that a matrix A is a π-consistently ordered matrix if the matrix Z is consistently ordered.

In iterative methods, the order of the internal points in which the $(k+1)^{th}$ iterates are evaluated can be referred to an ordering. There are several common ways mesh points can be ordered, row-wise, column wise (both are often referred to as natural ordering), diagonal ordering and red-black ordering. For a problem arising from the solution of the five-point difference equation on a square mesh, all the orderings just described are consistent orderings (Young, 1971). The concept of consistent ordering is central to the theory of the SOR iterative method and Group iterative methods for solving the system (2.1) because at present the calculation of the optimum acceleration parameter is possible only for consistently ordered matrices (Smith, 1985).

2.3.4 L, M Matrices and Some Subclasses of H-matrices

The following Matrices definitions are due to Martins (1982).

Definition 2.3.13

a): An $n\times n$ matrix A is an L-matrix if $a_{ii}>0, 1\leq i \leq n$ and $a_{ij}\leq 0,\ 1\leq i \leq n, 1\leq j \leq n, i\neq j$.

b): An $n\times n$ matrix A is an M-matrix, if $a_{ij} \leq 0,\ i\neq j,\ 1\leq i,j \leq n$ and $A^{-1}\geq 0$.

c): A matrix $A=[a_{ij}]\in\mathbb{C}^{n,n}$ is called an H-matrix if its comparison matrix $M(A)=[m_{ij}]$ defined by $m_{ii}=|a_{ii}|$, $m_{ij}=-|a_{ij}|$, $i, j=1, 2,\ldots, n, i\neq j$ is an M-matrix, i.e. $M(A)^{-1}\geq 0$.

Definition 2.3.14

i): A matrix $A=[a_{ij}]\in\mathbb{C}^{n,n}$ is called an SDD (Strictly Diagonally Dominant) matrix

if $|a_{ii}| > t_i(A),\ i = 1, 2, \ldots, n$. where for each nonempty subset S of indices $N := \{1,2,\ldots,n\}$, we denote $t_i(A) := \sum_{k\in S\setminus\{i\}} |a_{ik}|$.

ii): A matrix A is called a generalized diagonally dominant (GDD) by rows (columns), if there is a scaling on columns (rows) of A by multipliers non-nulls, such that the obtained matrix A is strictly diagonally dominant by rows(columns).

Definition 2.3.15

A matrix $A=[a_{ij}]\in\mathbb{C}^{n,n}$ is called a DDD (Doubly Diagonally Dominant) matrix if $|a_{ii}||a_{jj}| > t_i(A)t_j(A)$ with $i, j=1,2,\ldots,n,\ i\neq j$ (see Gao and Huang (2006)).

It is known that the two classes SDD, DDD are subclasses of H-matrices. Moreover, as it has been seen in Xiang and Zhang (2006), the above subclasses of H-matrix have specific relation, i.e. every SDD matrix is a DDD matrix too. Finally, we recall the most useful property of the class of H-matrix for considerations that follow: a given matrix A is an H-matrix if and only if there exists a positive diagonally matrix G such that AG is an SDD matrix.

2.4 Solution of PDEs by Finite Difference Methods

As mentioned earlier, a number of approaches have been developed for the treatment of PDEs. The most widely used of these approaches is the method of finite difference. In the following discussion a basic approach will be taken to introduce the finite difference method.

2.4.1 Finite Difference Approximations of Derivatives

In the finite difference method, finite difference approximations are used to replace the derivatives in the PDEs. In the first step of that, the solution domain must be divided into discrete points before applying any numerical methods. This strategy is called discretisation of the solution domain. Divide the solution domain into squares by grid lines parallel to the x-axis (uniform length Δx) and grid lines parallel to the y-axis (uniform length Δy) such that: $\Delta x = \Delta y = h$ as shown in Figure 2.6. The finite difference approximations are basically formed by Taylor series expansion. Taylor series for a function $u(x, y)$ expanded about (x, y) at $(x_i + h)$ and $(x_i - h)$ are respectively,

$$u(x+h,y) = u(x,y) + \frac{h}{1!}u_x(x,y) + \frac{h^2}{2!}u_{xx}(x,y) + \frac{h^3}{3!}u_{xxx}(x,y) + \dots \quad (2.23a)$$

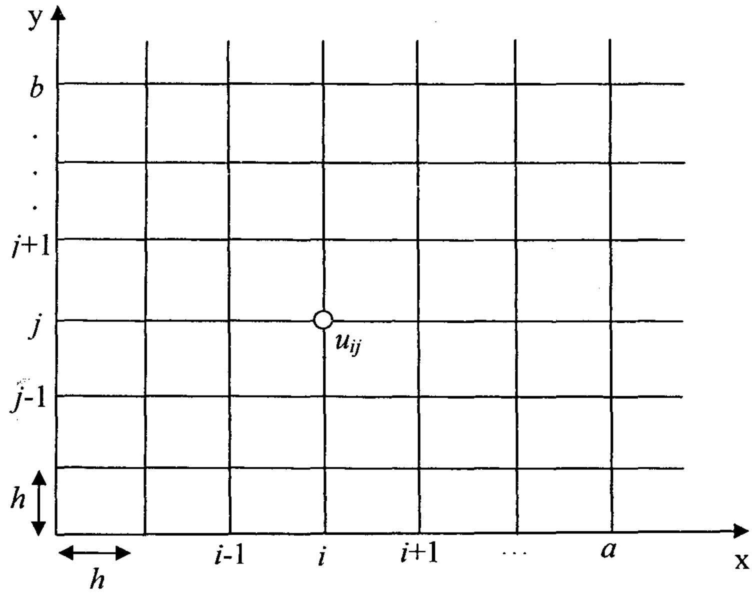

Figure 2.6 Discretisation of the solution domain.

$$u(x-h,y)=u(x,y)-\frac{h}{1!}u_x(x,y)+\frac{h^2}{2!}u_{xx}(x,y)-\frac{h^3}{3!}u_{xxx}(x,y)+\ldots \qquad (2.23b)$$

where h is the grid size, which is sufficiently small for the series to be convergent.

We can rewrite Equations (2.23a) and (2.23b) by using the double subscript notation in which the first subscript denotes the x-position and the second subscript denotes the y-position as the following

$$u_{i+1,j}=u_{i,j}+h(u_x)_{i,j}+\frac{h^2}{2!}(u_{xx})_{i,j}+\frac{h^3}{3!}(u_{xxx})_{i,j}+\ldots \qquad (2.24a)$$

$$u_{i-1,j}=u_{i,j}-h(u_x)_{i,j}+\frac{h^2}{2!}(u_{xx})_{i,j}-\frac{h^3}{3!}(u_{xxx})_{i,j}+\ldots \qquad (2.24b)$$

Equation (2.24a) can be written as,

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

$$=\frac{u_{i+1,j}-u_{i,j}}{h}+O(h). \qquad (2.25a)$$