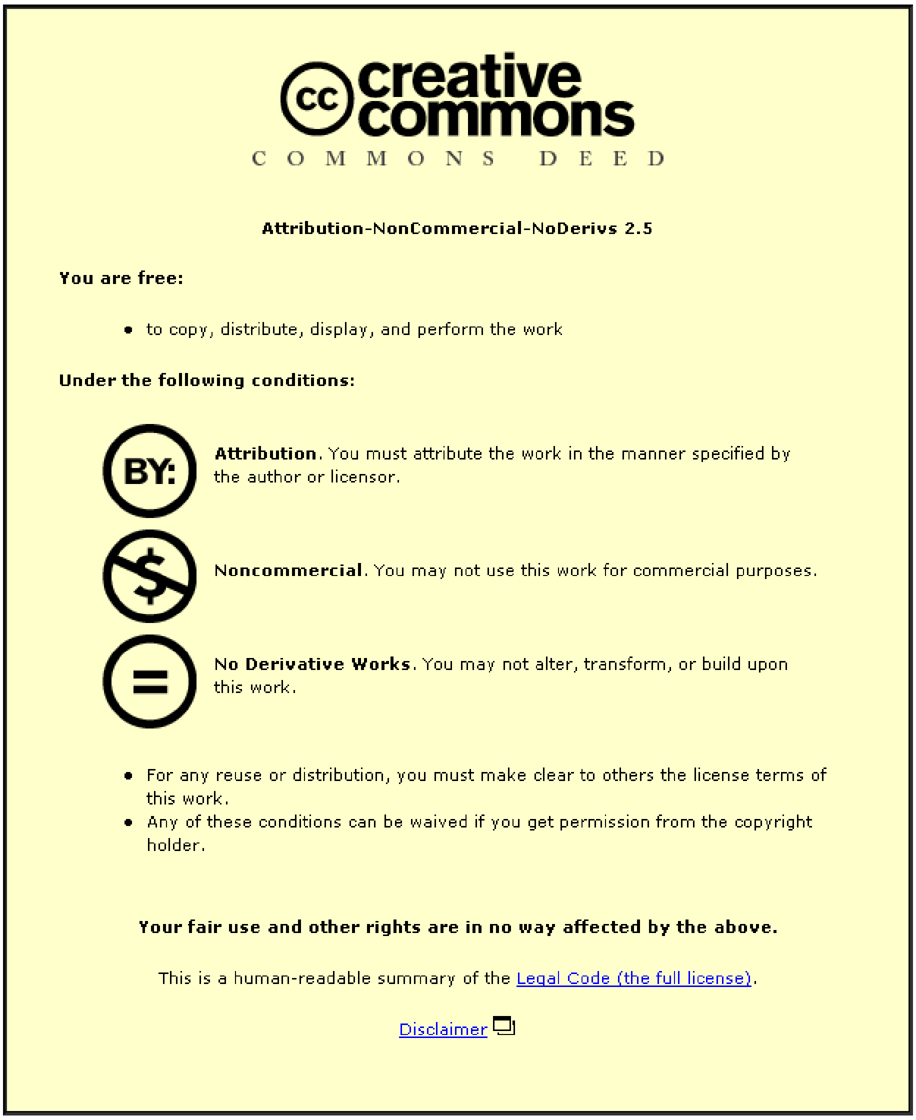


BLLID No: D 68018/86

NUMERICAL METHODS FOR SOLVING HYPERBOLIC AND PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

BY

MOHD SALLEH SAHIMI, B.SC., M.A.

A Doctoral Thesis
Submitted in partial fulfilment of the requirements
for the award of Doctor of Philosophy
of the Loughborough University of Technology
March, 1986

Supervisor: PROFESSOR D.J. EVANS, Ph.D., D.Sc.,
Department of Computer Studies

بسم الله الرحمن الرحيم

In the name of God, Most Gracious, Most Merciful.

"Not a Sign comes to them from among
the Signs of their Lord, but they turn away therefrom".

(Yaa-Sin, verse 46).

ACKNOWLEDGEMENTS

All praise be to God the Almighty with whose consent has bestowed me the desire, will and perseverance to complete the work of this thesis. I am greatly indebted to my supervisor, Professor D.J. Evans, who has rendered me valuable guidance and incessant encouragement and concern during this period of endurance.

I would also like to extend my thanks to the Universiti Kebangsaan Malaysia and the Jabatan Perkhidmatan Awam for giving me the opportunity to continue my studies and for their financial support.

ABSTRACT

The main object of this thesis is a study of the numerical solution of hyperbolic and parabolic partial differential equations. The introductory chapter deals with a general description and classification of partial differential equations. Some useful mathematical preliminaries and properties of matrices are outlined.

Chapters Two and Three are concerned with a general survey of current numerical methods to solve these equations. By employing finite differences, the differential system is replaced by a large matrix system. Important concepts such as convergence, consistency, stability and accuracy are discussed with some detail.

The group explicit (GE) methods as developed by Evans and Abdullah on parabolic equations are now applied to first and second order (wave equation) hyperbolic equations in Chapter 4. By coupling existing difference equations to approximate the given hyperbolic equations, new GE schemes are introduced. Their accuracies and truncation errors are studied and their stabilities established.

Chapter 5 deals with the application of the GE techniques on some commonly occurring examples possessing variable coefficients such as the parabolic diffusion equations with cylindrical and spherical symmetry. A complicated stability analysis is also carried out to verify the stability, consistency and convergence of the proposed scheme.

In Chapter 6 a new iterative alternating group explicit (AGE) method with the fractional splitting strategy is proposed to solve various linear and non-linear hyperbolic and parabolic problems in one dimension. The AGE algorithm with its PR (Peaceman Rachford) and DR

(Douglas Rachford) variants is implemented on tridiagonal systems of difference schemes and proved to be stable. Its rate of convergence is governed by the acceleration parameter and with an optimum choice of this parameter, it is found that the accuracy of this method, in general, is better if not comparable to that of the GE class of problems as well as other existing schemes.

The work on the AGE algorithm is extended to parabolic problems of two and three space dimensions in Chapter 7. A number of examples are treated and the DR variant is used because of consideration of stability requirement. The thesis ends with a summary and recommendations for future work.

CONTENTS

CHAPTER ONE

INTRODUCTION AND MATHEMATICAL PRELIMINARIES

1.1 INTRODUCTION

Most of the problems dealt with in physics and engineering fall into one of three physical categories: *equilibrium problems*, *eigenvalue problems* and *propagation problems*.

Equilibrium problems are problems of steady state in which we determine the equilibrium configuration ϕ in a domain D by solving the differential equation

$$L[\phi] = f \,, \tag{1.1.1}$$

within D, subject to certain boundary conditions,

$$B_i[\phi] = g_i \,, \tag{1.1.2}$$

on the boundary of D. This is illustrated in Fig. 1.1.1 below in which the integration domain D is generally closed and bounded.

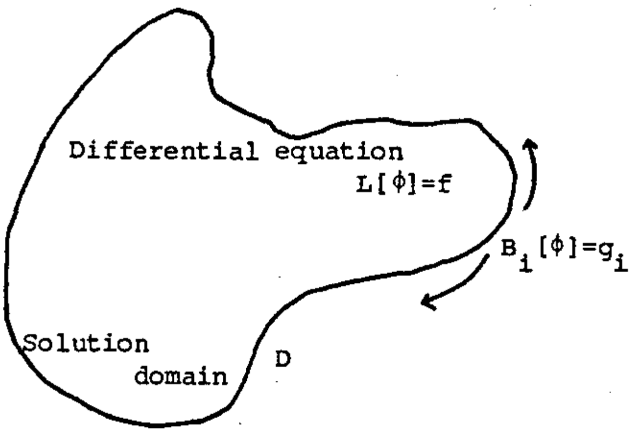


FIGURE 1.1.1: Representation of the general equilibrium problem

Such problems are known as *boundary value problems*. Typical physical examples include steady viscous flow, steady temperature distributions, equilibrium stresses in elastic structures and steady voltage distributions. The governing equations for this category of problems are known to be *elliptic*.

As extensions to equilibrium problems we have another category known as eigenvalue problems wherein critical values of certain parameters are to be determined in addition to the corresponding steady-state configurations. Mathematically the problem is to find one or more constants (μ) and the corresponding functions (ϕ) such that the differential equation

$$L[\phi] = \mu M[\phi] \tag{1.1.3}$$

is satisfied within D and the boundary conditions

$$B_i[\phi] = \mu E_i[\phi] \tag{1.1.4}$$

hold on the boundary of D. Typical physical examples include buckling and stability of structures, resonance in electric circuits and acoustics, natural frequency problems in vibrations and others. The operators L and M are again of elliptic type.

Propagation problems are initial value problems that have an unsteady state or transient nature. One wishes to predict the subsequent behaviour of a system given the initial state. This is to be done by solving the differential equation

$$L[\phi] = f \tag{1.1.5}$$

within the domain D when the initial state is prescribed as

$$I_i[\phi] = h_i \tag{1.1.6}$$

and subject to prescribed conditions

$$B_i[\phi] = g_i \tag{1.1.7}$$

on the boundaries. The integration domain D is open. The general propagation problem is illustrated in Fig. 1.1.2 below. Such problems are called *initial boundary value problems*. Typical physical examples include the propagation of pressure waves in a fluid, propagation of stresses and displacements in elastic systems, propagation of heat and the development of self-excited vibrations. The governing equations for propagation problems are *parabolic* or *hyperbolic*.

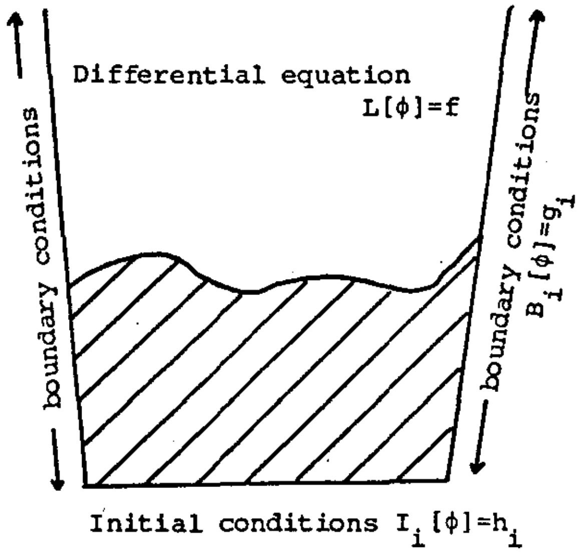


FIGURE 1.1.2: Representation of the general propagation problem

The differential equations mentioned above generally involve partial derivatives of the dependent variable with respect to the independent variables and they are therefore referred to as *partial differential equations*. As examples, we list below a few fundamentally important physical situations together with their corresponding equations:

(1) The small transverse vibrations y of a tightly stretched string (Fig. 1.1.3).

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{a^2}\frac{\partial^2 y}{\partial t^2} \quad . \tag{1.1.8}$$

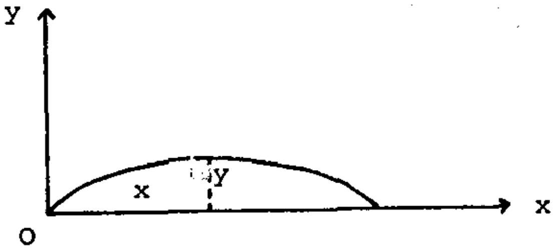


FIGURE 1.1.3

(2) The small vibrations z of a circular membrane with no twisting motion (Fig. 1.1.4).

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial z}{\partial r}\right) = \frac{1}{a^2}\frac{\partial^2 z}{\partial t^2} \quad . \qquad (1.1.9)$$

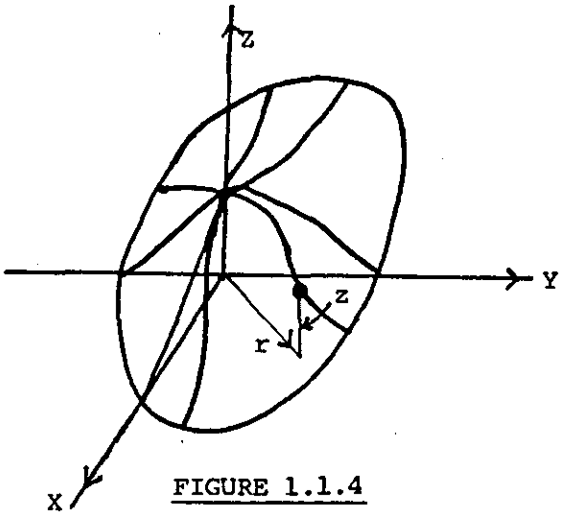


FIGURE 1.1.4

(3) The motion in a vertical section of a wave surface in a bay open to the sea in the X direction (Fig. 1.1.5).

$$\frac{\partial^2 y}{\partial t^2} = \frac{g}{b}\frac{\partial}{\partial x}\left(hb\frac{\partial y}{\partial x}\right) \quad , \qquad (1.1.10)$$

where b is breadth and h the depth of the bay.

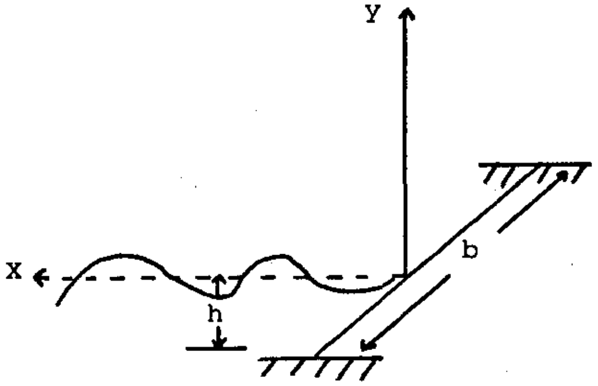


FIGURE 1.1.5

(4) The distribution of temperature V in a long narrow rectangular plate (one dimensional flow) (Fig. 1.1.6).

$$\frac{\partial^2 V}{\partial x^2} = \frac{1}{a^2}\frac{\partial V}{\partial t} \qquad (1.1.11)$$

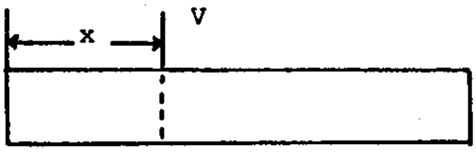


FIGURE 1.1.6

(5) The distribution of temperature V within a sphere (Fig. 1.1.7).

$$\frac{\partial^2(rV)}{\partial r^2} = \frac{1}{a^2}\frac{\partial(rV)}{\partial t} . \qquad (1.1.12)$$

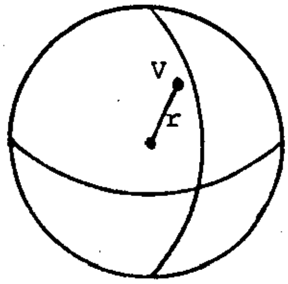


FIGURE 1.1.7

(6) The distribution of temperature V in a cylinder (Fig. 1.1.8).

$$\frac{\partial^2 V}{\partial r^2} + \frac{1}{r}\frac{\partial V}{\partial r} = \frac{1}{a^2}\frac{\partial V}{\partial t} . \qquad (1.1.13)$$

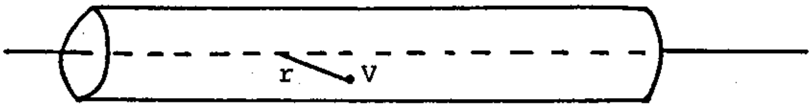


FIGURE 1.1.8

All of the above differential equations may be derived mathematically. For the purpose of demonstration, however, we shall only attempt to derive the heat equation (1.1.11) and the vibration equation (1.1.8).

(a) Derivation of eqn.(1.1.11). We shall assume:

(i) that heat flows from high to low temperatures,

(ii) that Q, the quantity of heat, is jointly proportional to the body mass and to its temperature V,

(iii) that the rate at which heat flows across an area is jointly proportional to the area and to the rate $\partial V/\partial x$ of change of temperature measured perpendicular to the area.

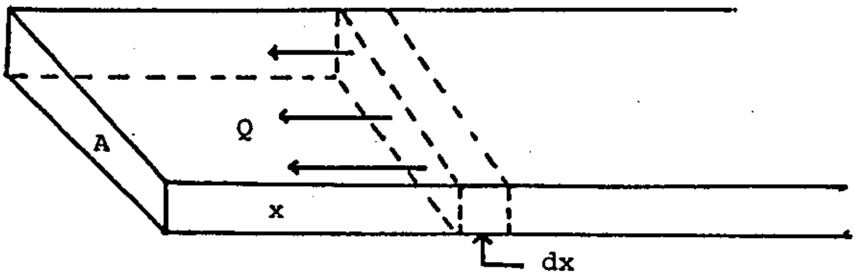


FIGURE 1.1.9

Consider the homogeneous bar (Fig. 1.1.9) with sectional area A whose sides are well insulated and whose uniform density is ρ. The mass of the differential section is

$$\rho A dx.$$

Thus, by assumption (ii),

$$Q = k(\rho A dx) V \text{ cal.}$$

and (taking derivative with respect to t)

$$\frac{\partial Q}{\partial t} = k(\rho A dx)\frac{\partial V}{\partial t} \text{ cal./sec.} \qquad (1.1.14)$$

This last expression is proportional to the rate at which heat flows into the differential volume minus the rate at which it flows out; that is, it is proportional to

$$-A[\frac{\partial V}{\partial x}]_x + A[\frac{\partial V}{\partial x}]_{x+dx}$$

with the signs satisfying assumption (i). That is, with proportionality factor k_1,

$$(k\rho Adx)\frac{\partial V}{\partial t} = k_1 A[(\frac{\partial V}{\partial x})_{x+dx} - (\frac{\partial V}{\partial x})_x] \ . \qquad (1.1.15)$$

Now, by Taylor's series, $(\frac{\partial V}{\partial x})_{x+dx}$ is expanded in powers of dx and terms involving higher powers of dx are discarded. Accordingly,

$$k\rho Adx \ \frac{\partial V}{\partial t} = k_1 A(\frac{\partial V}{\partial x} + \frac{\partial^2 V}{\partial x^2} dx - \frac{\partial V}{\partial x})$$

or

$$\frac{\partial V}{\partial t} = a^2 \ \frac{\partial^2 V}{\partial x^2}$$

where a^2 represents the quantity k_1/k_ρ.

(b) Derivation of eqn.(1.1.8). We consider a free body section of the string as shown in Fig. 1.1.10 below,

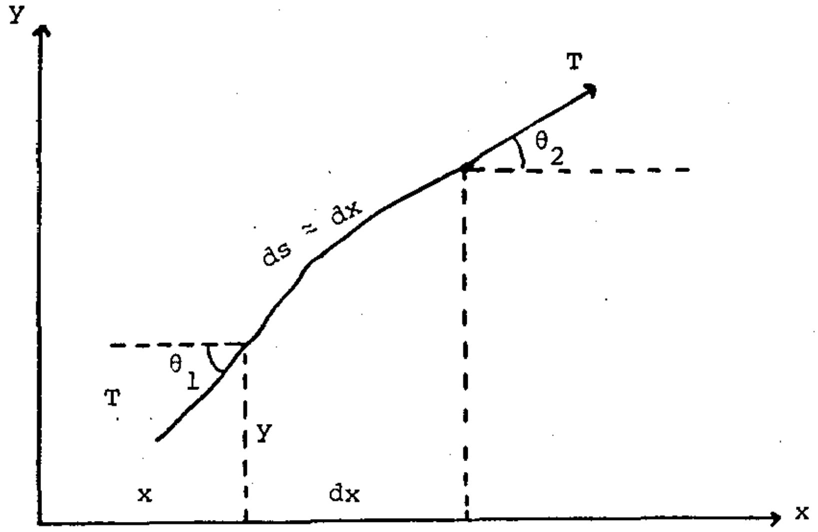


FIGURE 1.1.10

Let F_y represent the sum of the forces acting in the y direction. Then by Newton's law,

$$F_y = (mdx)\,\frac{\partial^2 y}{\partial t^2} \tag{1.1.16}$$

where m is the mass per unit length of the string. In this it should be noted that since $y'^2 << 1$ then,

$$ds = \sqrt{1+y'^2}\,dx \approx dx$$

But for any section as shown, since assumptions provide that θ_1, θ_2 are small enough to replace $\sin\theta$ by $\tan\theta$, we have

$$\begin{aligned} F_y &= T\sin\theta_2 - T\sin\theta_1 \approx T(\tan\theta_2 - \tan\theta_1) \\ &= T[(\frac{\partial y}{\partial x})_{x+dx} - (\frac{\partial y}{\partial x})_x] \ . \end{aligned}$$

The notation $(\frac{\partial y}{\partial x})_{x+dx}$ stands for the slope of the curve at the point whose abscissa is x+dx. Let $g(x,t) = \frac{\partial y}{\partial x}$. Then,

$$(\frac{\partial y}{\partial x})_{x+dx} = g(x+dx,t) \ .$$

Now by Taylor's series, a function $g(x+dx)$ may be expanded in powers of dx; that is,

$$g(x+dx) = g(x) + g'(x)dx + \frac{1}{2!}\,g''(x)dx^2 + \ldots$$

(t being considered constant, we omit writing it explicitly here).

So here,

$$(\frac{\partial y}{\partial x})_{x+dx} = (\frac{\partial y}{\partial x}) + (\frac{\partial^2 y}{\partial x^2})\,dx + \ldots$$

With this, we may rewrite the last expression for F_y as

$$F_y = T[(\frac{\partial y}{\partial x}) + (\frac{\partial^2 y}{\partial x^2})dx + \ldots - (\frac{\partial y}{\partial x})]$$

or

$$F_y \approx T\,\frac{\partial^2 y}{\partial x^2}\,dx \tag{1.1.17}$$

wherein terms involving $dx^2, dx^3, \ldots$ are discarded. Equating the two expressions (1.1.16) and (1.1.17) for F_y,

$$m\frac{\partial^2 y}{\partial t^2} = T\frac{\partial^2 y}{\partial x^2}$$

which is of the form (1.1.8).

We shall now formally define a partial differential equation for a dependent variable u(x,y,...) as a relation of the form

$$F(x,y,\ldots,u,\frac{\partial u}{\partial x},\frac{\partial u}{\partial y},\ldots,\frac{\partial^2 u}{\partial x^2},\frac{\partial^2 u}{\partial x\partial y},\ldots) = 0 \qquad (1.1.18)$$

where F is a given function of the independent variables x,y,... and the "unknown" function u and of a finite number of its partial derivatives. The order of the equation is equal to the order of the highest partial differential coefficient occurring in it. For example, the equations,

$$3y^2\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 2u \qquad (1.1.19)$$

$$\frac{\partial^2 u}{\partial x^2} + f(x,y)\frac{\partial^2 u}{\partial y^2} = 0 \qquad (1.1.20)$$

(where f(x,y) is any given function) are typical partial differential equations of the first and second orders respectively, x and y being the independent variables, and u=u(x,y) being the dependent variable whose form is to be found by solving the appropriate equation. Likewise the equation,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0 \qquad (1.1.21)$$

(known as Laplace's equation in three variables) is an equation of second order for u(x,y,z) where x,y and z are independent variables. We note that equations (1.1.19), (1.1.20) and (1.1.21) are all *linear*. In the language of linear operators, we say that the operator L is linear when it satisfies the following relations:

$$\cancel{L[av] = aL[v]} \qquad (1.1.22)$$

$$\cancel{L[v_1+v_2] = L[v_1] + L[v_2]} \qquad (1.1.23)$$

$$L[av_1+bv_2] = aL[v_1] + bL[v_2] \ , \qquad (1.1.24)$$

where a and b are constants. Differential operators such as $\frac{\partial}{\partial t}$, $\frac{\partial^2}{\partial x^2}$ and so on can be readily verified to be linear. To demonstrate the importance of this linearity assumption, let us consider the heat conduction equation,

$$\frac{\partial^2 u}{\partial x^2} = a \frac{\partial u}{\partial t} \ . \qquad (1.1.25)$$

Let u_1 and u_2 be solutions of equation (1.1.25), that is

$$\frac{\partial^2 u_1}{\partial x^2} = a \frac{\partial u_1}{\partial t} \text{ and } \frac{\partial^2 u_2}{\partial x^2} = a \frac{\partial u_2}{\partial t} \ . \qquad (1.1.26)$$

By virtue of the linearity of the operators $\frac{\partial^2}{\partial x^2}$ and $\frac{\partial}{\partial t}$, it follows that u_1+u_2 (in fact au_1+bu_2, where a and b are arbitrary constants) is also a solution.

Let us, on the other hand, consider the more general heat conduction equation,

$$\frac{\partial}{\partial x} [k(u) \frac{\partial u}{\partial x}] = c \frac{\partial u}{\partial t} \qquad (1.1.27)$$

where c is a constant and k depends upon the temperature u, $k(u)=k_o u$ say, where k_o is a constant. Now equation (1.1.27) becomes

$$u \frac{\partial^2 u}{\partial x^2} + (\frac{\partial u}{\partial x})^2 = \frac{c}{k_o} \frac{\partial u}{\partial t} \ . \qquad (1.1.28)$$

Suppose u_1 and u_2 are solutions of equation (1.1.28). We wish to know whether u_1+u_2 is also a solution of the same equation. By substituting u_1+u_2 into equation (1.1.28) we find after some algebra that,

$$[u_1 \frac{\partial^2 u_1}{\partial x^2} + (\frac{\partial u_1}{\partial x})^2 - \frac{c}{k_o} \frac{\partial u_1}{\partial t}] + [u_2 \frac{\partial^2 u_2}{\partial x^2} + (\frac{\partial u_2}{\partial x})^2 - \frac{c}{k_o} \frac{\partial u_2}{\partial t}]$$

$$+ u_1 \frac{\partial^2 u_2}{\partial x^2} + u_2 \frac{\partial^2 u_1}{\partial x^2} + 2 \frac{\partial u_1}{\partial x} \frac{\partial u_2}{\partial x} = 0 . \tag{1.1.29}$$

The first two bracketed terms vanish by virtue of equation (1.1.28). The last three terms are not zero. Therefore u_1+u_2 is not a solution implying that equation (1.1.27) is not linear.

When a partial differential equation is not linear, it can either be *quasilinear* or *nonlinear*. If the coefficients, such as k of equation (1.1.27) are functions only of the dependent variable and not of its derivatives, then the equation is quasilinear. Otherwise, the equations are nonlinear.

A linear equation is said to be *homogeneous* if each term contains either the dependent variable or one of its derivatives. For example, Laplace's equation in two-dimensions (that is, two independent variables),

$$\nabla^2 u = 0 , \tag{1.1.30}$$

where ∇^2 is the two-dimensional Laplace operator (defined in rectangular Cartesian coordinates (x,y) by $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$) is homogeneous. However, the two-dimensional Poisson equation,

$$\nabla^2 u = f(x,y) , \tag{1.1.31}$$

where f(x,y) is any given (non-zero) function, is termed *inhomogeneous* (or non-homogeneous). If $u_1,u_2,\ldots,u_n$ where n may be finite or non-finite, are n different solutions of a linear homogeneous partial differential equation in some given domain then,

$$u = c_1u_1 + c_2u_2 + \ldots + c_nu_n \tag{1.1.32}$$

is also a solution in the same domain, where the coefficients $c_1,c_2,\ldots,c_n$ are arbitrary constants.

Whereas the general solution of a linear ordinary differential equation contains arbitrary constants of integration, the general solution of a linear partial differential equation contains arbitrary functions. To illustrate this point we consider the problem of the formation of partial differential equations from given functions. For example, if,

$$u = yf(x) \tag{1.1.33}$$

where f(x) is an arbitrary function of x, then differentiating with respect to y we have,

$$\frac{\partial u}{\partial y} = f(x) \ . \tag{1.1.34}$$

On eliminating f(x) between (1.1.33) and (1.1.34) we find that,

$$y \frac{\partial u}{\partial y} = u \tag{1.1.35}$$

which is a first-order linear partial differential equation whose general solution is given by (1.1.33). The significant point here is that the solution of (1.1.35) as given by (1.1.33) contains an arbitrary function.

1.2 CLASSIFICATION OF EQUATIONS AND BOUNDARY CONDITIONS

The general two-dimensional second-order equation is given by

$$A \frac{\partial^2 u}{\partial x^2} + B \frac{\partial^2 u}{\partial x \partial y} + C \frac{\partial^2 u}{\partial y^2} + D \frac{\partial u}{\partial x} + E \frac{\partial u}{\partial y} + Fu + G = 0 \qquad (1.2.1)$$

where A,B,C,D,E,F and G may be functions of the independent variables x and y and of the dependent variable u. Special cases of this equation occur more frequently than any other in *field problems* (that is, problems in science and engineering involving partial differential equations). One such form is the quasilinear equation

$$A \frac{\partial^2 u}{\partial x^2} + B \frac{\partial^2 u}{\partial x \partial y} + C \frac{\partial^2 u}{\partial y^2} + G = 0 \qquad (1.2.2)$$

where the coefficients may be functions of x,y,u, $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial y}$ but not of $\frac{\partial^2 u}{\partial x^2}$, $\frac{\partial^2 u}{\partial x \partial y}$ and $\frac{\partial^2 u}{\partial y^2}$. Some examples serve to illustrate this:

(1) The wave equation,

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} \qquad (1.2.3)$$

is obtained by taking A=1, $C=-\frac{1}{c^2}$, B=G=0 and associating the variable y with the time variable t.

(2) The Laplace equation in two variables,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \qquad (1.2.4)$$

is obtained from (1.2.2) by choosing A=1, C=1, B=G=0.

(3) Putting A=1, B=C=0 and $G= - k \frac{\partial u}{\partial t}$(where we associate y with t) we get the familiar heat equation,

$$\frac{\partial^2 u}{\partial x^2} = k \frac{\partial u}{\partial t} . \qquad (1.2.5)$$

It will be shown that at every point in the x-y plane there are two directions in which the integration of the partial differential equation reduces to the integration of an equation involving total differentials only. Furthermore it will be seen that this leads to a natural classification of partial differential equations.

Let us denote the first and second derivatives by

$$\frac{\partial u}{\partial x} = p, \quad \frac{\partial u}{\partial y} = q, \quad \frac{\partial^2 u}{\partial x^2} = r, \quad \frac{\partial^2 u}{\partial x \partial y} = s \text{ and } \frac{\partial^2 u}{\partial y^2} = t \tag{1.2.6}$$

Let C be a curve in the x-y plane on which the values of u,p and q are such that they and the second-order derivatives r,s and t derivable from them, satisfy equation (1.2.2). Therefore, the differentials of p and q in directions tangential to C satisfy the equations,

$$dp = \frac{\partial p}{\partial x} dx + \frac{\partial p}{\partial y} dy = rdx + sdy \tag{1.2.7}$$

and

$$dq = \frac{\partial q}{\partial x} dx + \frac{\partial q}{\partial y} dy = sdx + tdy \tag{1.2.8}$$

where

$$Ar + Bs + Ct + G = 0 \tag{1.2.9}$$

and $\frac{dy}{dx}$ is the slope of the tangent to C. Eliminating r and t from equation (1.2.9) by means of equations (1.2.7) and (1.2.8) gives us,

$$\frac{A}{dx}\,(dp-sdy) + Bs + \frac{C}{dy}(dq-sdx) + G = 0,$$

i.e.,

$$s\left\{A\left(\frac{dy}{dx}\right)^2 - B\left(\frac{dy}{dx}\right) + C\right\} - \left\{A\,\frac{dp}{dx}\,\frac{dy}{dx} + C\,\frac{dq}{dx} + G\,\frac{dy}{dx}\right\} = 0. \tag{1.2.10}$$

Now choose the curve C so that the slope of the tangent at every point on it is a root of the equation,

$$A\left(\frac{dy}{dx}\right)^2 - B\left(\frac{dy}{dx}\right) + C = 0 , \tag{1.2.11}$$

so that s is also eliminated. By (1.2.10) it follows that in this direction, we have,

$$A \frac{dp}{dx} \frac{dy}{dx} + C \frac{dq}{dx} + G \frac{dy}{dx} = 0 \; . \qquad (1.2.12)$$

This shows that at every point P(x,y) of the solution domain there are two directions, given by the roots of (1.2.11), along which there is a relationship, given by (1.2.12), between the total differentials dp and dq.

The directions given by the roots of equation (1.2.11) are called the *characteristic directions* and the partial differential equation is said to be hyperbolic, parabolic or elliptic according to whether these roots are real and distinct, equal or complex respectively, that is, according to whether $D=B^2-4AC$ is positive, zero or negative.

Returning to our examples above it is clear that the wave equation (1.2.3) is of hyperbolic type, the Laplace equation is of elliptic type and the heat equation is parabolic. It is also important to note that equations with variable coefficients may change their type on passing from one region of the xy-plane to another. For example, as shown in Fig. 1.2.1, the equation,

$$y \frac{\partial^2 u}{\partial x^2} + 2x \frac{\partial^2 u}{\partial x \partial y} + y \frac{\partial^2 u}{\partial y^2} = 0 \; , \qquad (1.2.13)$$

is elliptic in the region where $y^2-x^2>0$, parabolic along the lines $y^2-x^2=0$, and hyperbolic in the region where $y^2-x^2<0$.

A similar but more complicated classification can be carried out for equations in three or more independent variables. The classification of partial differential equations into these three categories is valuable because the basic analytical and numerical methods for treating field problems are inherently different for the three types of equations.

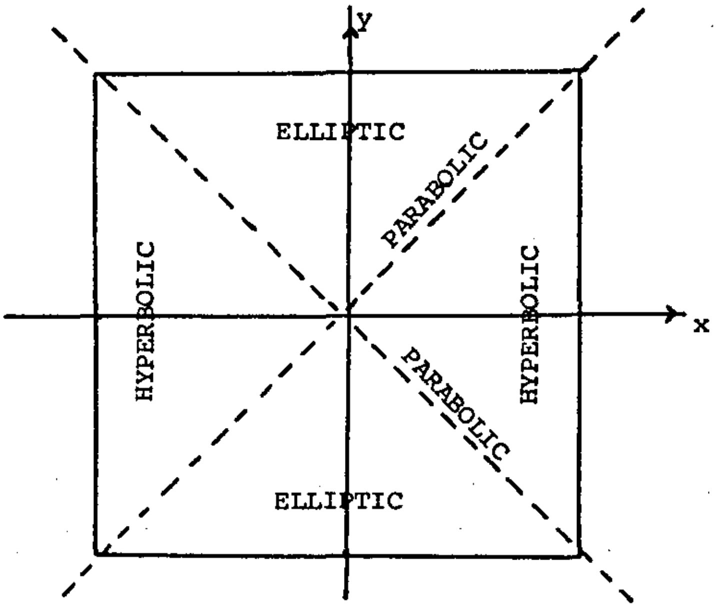


FIGURE 1.2.1

The partial differential equations encountered in practical problems are often very complicated. There is a need, therefore, for the problems to be correctly formulated and properly posed. In the sense of Hadamard, a problem is said to be *well posed* or *properly posed* provided two criteria are satisfied. Firstly, the solution should be unique. Secondly the solution obtained should be *stable* and depends continuously on the boundary data. In other words, a small change in the given boundary conditions should produce only a correspondingly small change in the solution. This is vital since, when the boundary conditions are arrived at by experiment, certain small observational errors in their values will always exist and these errors should not lead to large changes in the solution (a similar situation arises with sets of linear algebraic equations, where under certain circumstances the equations may be ill-conditioned - that is, small changes in the coefficients may produce large changes in the solutions). As a formal example, we consider Laplace's equation in two dimensions,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \ . \tag{1.2.14}$$

We now search for a solution which satisfies the boundary conditions

$$u(x,0) = \frac{\sin nx}{n} \ , \ \left(\frac{\partial u}{\partial y}\right)_{y=0} = 0 \ , \tag{1.2.15}$$

where n is some parameter. Such a solution is easily verified to be

$$u(x,y) = \frac{1}{n} \sin nx \cosh ny \ . \tag{1.2.16}$$

However, as $n \to \infty$ the boundary conditions converge to

$$u(x,0) = 0 \ , \ \left(\frac{\partial u}{\partial y}\right)_{y=0} = 0 \ , \tag{1.2.17}$$

and these, together with (1.2.14) imply, by Taylor's series, $u(x,y)=0$. However, as $n \to \infty$, $u(x,y)$ as given by (1.2.16) becomes infinitely large. The problem defined by (1.2.14) and (1.2.15) is not, therefore, well posed and could not be associated with a physical phenomenon. Much work has been carried out to determine the types of boundary conditions, which, when imposed on linear partial differential equations, lead to unique and stable solutions. There are several types of boundary conditions which arise frequently in the description of physical phenomena. If we wish to solve the equation $L[u]=f$ over a region D whose boundary is denoted by B then the specifications along B are of the following types:

(1) *Dirichlet condition or the first boundary condition.*

Here u satisfies the condition,

$$u = g \text{ on } B \ , \tag{1.2.18}$$

with g being a prescribed continuous function on B.

An example of a problem using Dirichlet conditions for its solution is the task of finding the temperature distribution in the interior of a plate when the temperature is prescribed at all points on the boundary of the plate.

(2) *Neumann condition or the second boundary condition.*

The normal derivative $\frac{\partial u}{\partial n}$ satisfies the condition

$$\frac{\partial u}{\partial n} = g \text{ on } B , \tag{1.2.19}$$

where g is a prescribed function continuous on B. The symbol $\frac{\partial u}{\partial n}$ denotes the directional derivative of u along the outward normal to B.

We note that the boundary conditions of (1.2.18) and (1.2.19) are inhomogeneous. An inhomogeneous boundary condition can be further classified as *time varying* and *time invariant*. For example, boundary conditions such as,

$$u(0,t) = 0, \ \alpha \frac{\partial u}{\partial x} (1,t) = 0$$

are homogeneous, while,

$$u(0,t) = 3, \ \alpha \frac{\partial u}{\partial x} (1,t) = e^{-t}$$

are examples of inhomogeneous boundary conditions, the first being time invariant and the second time varying.

(3) *Combined or third boundary conditions.*

The Dirichlet and Neumann types of boundary conditions are combined and a typical example is

$$u(x_0,t) + \alpha \frac{\partial u}{\partial x} (x_0,t) = \gamma(t) \ , \ \alpha \neq 0 \tag{1.2.20}$$

(4) *Mixed or fourth boundary condition.*

Here, a Dirichlet type of boundary condition is required to be satisfied on a portion B_1 of the boundary B, and a Neumann type of boundary condition on the rest B_2 of B. That is, the boundary conditions take the form,

$$u = g_1 \text{ on } B_1 \ ,$$

$$\frac{\partial u}{\partial n} = g_2 \text{ on } B_2$$

and $B = B_1 \cup B_2$.

(5) *Cauchy condition.*

If one of the independent variables is t (time, say), and the values of both u and $\frac{\partial u}{\partial t}$ on a boundary t=0 (that is, the critical values of u and $\frac{\partial u}{\partial t}$) are given, then the boundary conditions are of Cauchy type with respect to the variable t.

As an example, we consider the case of the one-dimensional wave equation,

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \ , \qquad (1.2.21)$$

representing, say, the transverse oscillations of a stretched string. The Cauchy conditions correspond to giving the initial values of both the transverse displacement u and the transverse velocity $\frac{\partial u}{\partial t}$ of the string, that is,

$$u(x,t_0) = f(x)$$

and

$$\frac{\partial u}{\partial t}(x,t_0) = g(x) \ . \qquad (1.2.22)$$

These conditions can be shown to be necessary and sufficient for the existence of a unique solution.

(6) *Periodic boundary condition.*

In this case a solution is sought such that it satisfies the periodicity conditions like,

$$u\big|_x = u\big|_{x+\ell} \ , \ \frac{\partial u}{\partial n}\bigg|_x = \frac{\partial u}{\partial n}\bigg|_{x+\ell} \qquad (1.2.23)$$

where ℓ is the period.

1.3 NUMERICAL APPROXIMATION METHODS OF SOLUTIONS

As has been stated in the previous section, the partial differential equations encountered in practical problems are often very complicated. Usually, they have variable coefficients, non-linearities, irregular boundaries and occur in coupled systems of differing types (say, parabolic and hyperbolic). It is generally impossible to obtain analytical solutions to the characterising equations. Even if these solutions exist, their evaluations are often a laborious task, as can be seen by inspecting the solution by Fourier series of the torsion problem for a rectangular cross-section defined by x=±a, y=±b, namely,

$$\phi(x,y) = b^2-y^2-32b^2\pi^{-3} \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^3} \operatorname{sech} \frac{(2n+1)\pi a}{2b} \cosh \frac{(2n+1)\pi x}{2b} \cos \frac{(2n+1)\pi y}{2b} .$$

Numerical methods are therefore used almost exclusively to treat such problems. Prior to the 1960's, analogue simulation methods were widely used. In recent years, however, these have been almost completely replaced by the high-speed, large-memory digital computer. A number of approaches to derive suitable digital computer algorithms have been developed. The most widely used of these are the *finite difference method* and the *finite element method*.

The aim of both these methods is to reduce a given continuum problem into a discrete mathematical model suitable for computation. Toward this end various discretization schemes have been proposed both by engineers and mathematicians. Among the various discretization strategies, the finite difference and the finite element methods are perhaps the most general and therefore the most popular. The finite difference method, developed primarily by mathematicians, is a very general procedure and can be used directly if the governing differential equations of a physical system are available.

Early development of the finite element method, however, was to a large extent contributed by engineers rather than mathematicians in the mid 1950s. As many physical systems of engineering interest can be viewed as interconnected components, the engineers attempted to create anologies between finite portions of a continuum with discrete components or "finite elements" of the physical system. Thus, in contrast with finite difference methods, the starting point of the traditional finite element method was the physical system itself, and the discretization process was intuitive and based on physical arguments. Nevertheless, the finite element method almost always led to acceptable numerical solutions.

Both the finite difference and the finite element methods are considered to be *total discretization schemes* because derivatives no longer appear in the discretized model. In addition to these total discretization schemes, some semi-discretization methods also play an important role particularly in the development of general-purpose software packages for solving partial differential equations. These schemes, called the method of lines, discretize all but one independent variable converting a given partial differential equation into a system of ordinary differential equations. These equations take the form of initial or boundary value problems depending upon the type of the original partial differential equation and the manner in which time and space variables are treated. These ordinary differential equations are then solved using one of the standard methods of solution.

Although finite element methods have been briefly introduced above the purpose of this thesis is to focus solely on finite difference methods and develop new strategies to improve them.

1.4 REVIEW OF FUNDAMENTAL MATHEMATICAL RESULTS

In this section we shall briefly include some basic mathematical results, mainly from the calculus and linear algebra, that are useful in the development and investigation of numerical procedures in subsequent chapters.

(A) CALCULUS

Theorem 1.1 - Taylor's Theorem

If f(x) is a continuous, single-valued function of x with continuous derivatives $f'(x), f''(x), \ldots$ up to and including $f^{(n)}(x)$ in a given interval $a \leq x \leq b$, and if $f^{(n+1)}(x)$ exists in $0<x<b$, then,

$$f(x) = f(a) + \frac{(x-a)}{1!} f'(a) + \frac{(x-a)^2}{2!} f''(a) + \ldots + \frac{(x-a)^n}{n!} f^{(n)}(a) + E_n(x) \tag{1.4.1}$$

where,

$$E_n(x) = \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(\xi) \tag{1.4.2}$$

and $a<\xi<x$.

An alternative form of (1.4.1) may be obtained by changing x to a+x. Then,

$$f(a+x) = f(a) + \frac{x}{1!} f'(a) + \frac{x^2}{2!} f''(a) + \ldots + \frac{x^n}{n!} f^{(n)}(a) + E_n(x), \tag{1.4.3}$$

where now from (1.4.2),

$$E_n(x) = \frac{x^{n+1}}{(n+1)!} f^{(n+1)}(a+\theta x) \tag{1.4.4}$$

and $0<\theta<1$.

Theorem 1.2 - Taylor's Theorem for Functions of Two Independent Variables

If f(x,y) is defined on a region R of the xy-plane and all its partial derivatives of orders up to and including the (n+1)th are continuous in R, then for any point (a,b) in this region

$$f(a+h,b+k) = f(a,b) + {*}Df(a,b) + \frac{1}{2!} {*}D^2 f(a,b) + \ldots + \frac{1}{n!} {*}D^n f(a,b) + E_n, \tag{1.4.5}$$

where *D is the differential operator

$$*D = h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \tag{1.4.6}$$

and $*D^r f(a,b)$ means $(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y})^r f(x,y)$ (1.4.7)

evaluated at the point (a,b). The Lagrange error term E_n is given by,

$$E_n = \frac{1}{(n+1)!} {*}D^{n+1} f(a+\theta h, b+\theta k) \tag{1.4.8}$$

where $0<\theta<1$.

Theorem 1.3 - Fundamental Theorem of Algebra

Let $p(x) \equiv a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$ be an nth degree polynomial with $n \geq 1$, $a_n \neq 0$. Then there exist uniquely n constants, $\{r_j\}_{j=1}^n$, ("zeros of p(x)") such that

$$p(x) = a_n (x-r_1)(x-r_2)\ldots(x-r_n) \tag{1.4.9}$$

where the r_j's need not all be distinct or all real.

A quadratic equation is of the form

(a) $ax^2+bx+c = 0$ $(a \neq 0)$ or

(b) $x^2+px+q = 0$ (reduced form)

Definition 1.4.1

The discriminant of the equation (a) is the number $D=b^2-4ac$ and that of the equation (b) is the number $D=p^2-4q$.

Theorem 1.4 - Roots of a Quadratic Equation

For $D \neq 0$, the equation has two distinct roots;

$\lambda^2 +$

for D=0, the equation has one double root.

If the coefficients of the equation are real, then:

for D>0, it has two distinct real roots;

for D<0, it has two complex conjugate roots;

for D=0, it has only one real (double) root.

The solution can be found by:

1. Factorization into linear factors:

$$ax^2+bx+c = a(x-x_1)(x-x_2)$$

$$a(x_1+x_2) = -b$$

$$a(x_1x_2) = c$$

or

$$x^2+px+q = (x-x_1)(x-x_2)$$

$$x_1+x_2 = -p$$

$$x_1x_2 = q$$

2. The formula,

$$x_{1,2} = \frac{-b\pm\sqrt{(b^2-4ac)}}{2a}$$

for equation $ax^2+bx+c = 0$.

3. The formula,

$$x_{1,2} = -\frac{p}{2} \pm\sqrt{(\frac{p^2}{4} - q)}$$

for equation $x^2+px+q=0$.

A cubic equation is of the form,

$$ax^3+bx^2+cx+d = 0 \ , \quad a\neq 0. \qquad (1.4.10)$$

By the substitution $x=y-b/3a$ and dividing by a, the equation (1.4.10) becomes,

$$y^3+3py+2q = 0 \ , \qquad (1.4.11)$$

where,

$$3p = \frac{3ac-b^2}{3a^2}\ ,\ 2q = \frac{2b^{\overset{\overset{4}{\uparrow}}{3}}}{27a^3} - \frac{bc}{3a^2} + \frac{d}{a} \qquad (1.4.12)$$

Definition 1.4.2

The discriminant of the equation (1.4.11) is the number $D=-p^3-q^2$.

Theorem 1.5 - Roots of a Cubic Equation

For $D\neq 0$, equation (1.4.11) has three distinct roots; for $D=0$, equation (1.4.11) has either a double root (if $p^3=-q^2\neq 0$) or a triple zero root (if $p=q=0$).

If the coefficients of equation (1.4.11) are real, then: for $D\geq 0$, it has three real roots which are distinct if $D>0$; for $D<0$, it has one real and two complex conjugate roots.

Methods of solution are:

1. By factorization into linear factors:

$$ax^3+bx^2+cx+d = 0; \quad a(x-x_1)(x-x_2)(x-x_3) = 0$$

(x_1, x_2, x_3 are the roots);

$$x_1+x_2+x_3 = -\frac{b}{a}; \quad x_1x_2+x_1x_3+x_2x_3 = \frac{c}{a}; \quad x_1x_2x_3 = -\frac{d}{a}$$

2. By solving algebraically.

The roots y_1, y_2, y_3 of equation (1.4.11) are,

$$y_1 = u+v, \quad y_2 = \varepsilon_1 u+\varepsilon_2 v, \quad y_3 = \varepsilon_2 u+\varepsilon_1 v$$

where,

$$\varepsilon_{1,2} = -\tfrac{1}{2} \pm i\,\frac{\sqrt{3}}{2}\ ,\ u = \sqrt[3]{[-q+ (q^2+p^3)]}\ ,\ v = \sqrt[3]{[-q-\sqrt{(q^2+p^3)}]};$$

here we choose the cube roots so that $uv=-p$. (The cube roots are obtained by the method of solving the binomial equation (1.4.16) below). This method is not suitable if (1.4.11) has real coefficients and $D>0$, since the real roots y_1, y_2, y_3 are expressed in terms of roots of complex numbers (the irreducible case).

3. By solving trigonometrically.

Let the coefficients p,q of equation (1.4.11) be real and different from zero. Denote the roots by y_1, y_2, y_3. Put $r=\varepsilon\sqrt{|p|}$ where $\varepsilon=1$ if $q>0$ and $\varepsilon=-1$ if $q<0$. Then the roots can be determined by means of the trigonometric or hyperbolic functions according to Table 1.4.1.

If q=0 in equation (1.4.11), then the equation has the common factor y and can be solved easily.

If p=0 in equation (1.4.11), then (1.4.11) is a binomial equation of the form (1.4.16),

$p<0$		$p>0$	Check
$p^3+q^2\leq 0$	$p^3+q^2>0$		
$\cos\phi = \frac{q}{r^3}$	$\cosh\phi = \frac{q}{r^3}$	$\sinh\phi = \frac{b}{r^3}$	
$y_1=-2r\cos\phi/3$ $y_2=2r\cos(60°-\frac{\phi}{3})$ $y_3=2r\cos(60°+\frac{\phi}{3})$	$y_1=-2r\cosh\phi/3$ $y_2=r\cosh\frac{\phi}{3}+ i\sqrt{3}\, r\sinh\frac{\phi}{3}$ $y_3=r\cosh\frac{\phi}{3} - i\sqrt{3}\, r\sinh\frac{\phi}{3}$	$y_1=-2r\sinh\phi/3$ $y_2=r\sinh\frac{\phi}{3} + i\sqrt{3}\, r\cosh\frac{\phi}{3}$ $y_3=r\sinh\frac{\phi}{3} - i\sqrt{3}\, r\cosh\frac{\phi}{3}$	$y_1+y_2+y_3=0$

TABLE 1.4.1

Note that ϕ lies in the interval (0,90°).

A quartic (or biquadratic) equation has the form

$$ax^4+bx^3+cx^2+dx+e = 0 \ , \quad a\neq 0 \ . \qquad (1.4.13)$$

By the substitution $x=y-b/4a$ and dividing by a, equation (1.4.13) becomes,

$$y^4+py^2+qy+r = 0 \ , \qquad (1.4.14)$$

where,

$$p = -\frac{3b^2}{8a^2} + \frac{c}{a},\ q = \frac{b^3}{8a^3} - \frac{bc}{2a^2} + \frac{d}{a},\ r = -\frac{3b^4}{256a^4} + \frac{b^2c}{16a^3} - \frac{bd}{4a^2} + \frac{e}{a}. \tag{1.4.15}$$

Methods of solution are:

1. By factorization into linear factors,

$$ax^4+bx^3+cx^2+dx+e = 0,\ a(x-x_1)(x-x_2)(x-x_3)(x-x_4) = 0$$

$(x_1, x_2, x_3, x_4$ are the roots)

$$x_1+x_2+x_3+x_4 = -\frac{b}{a};\ x_1x_2+x_1x_3+x_1x_4+x_2x_3+x_2x_4+x_3x_4 = \frac{c}{a};$$

$$x_1x_2x_3+x_1x_2x_4+x_1x_3x_4+x_2x_3x_4 = -\frac{d}{a};\ x_1x_2x_3x_4 = \frac{e}{a}$$

2. By solving algebraically.

The roots y_1, y_2, y_3, y_4 of equation (1.4.14) are

$$y_1 = \sqrt{z}_1+\sqrt{z}_2+\sqrt{z}_3,\quad y_2 = \sqrt{z}_1-\sqrt{z}_2-\sqrt{z}_3$$

$$y_3 = -\sqrt{z}_1+\sqrt{z}_2-\sqrt{z}_3,\quad y_4 = -\sqrt{z}_1-\sqrt{z}_2+\sqrt{z}_3$$

where z_1, z_2, z_3 are the roots of the equation (the reducing cubic)

$$z^3 + \frac{p}{2}z^2 + \left(\frac{p^2}{16} - \frac{r}{4}\right)z - \frac{q^2}{64} = 0 .$$

Here the roots $\sqrt{z}_1$, $\sqrt{z}_2$, $\sqrt{z}_3$ should be chosen (the roots are obtained by the method of solving the binomial equation (1.4.16) below) such that,

$$\sqrt{z}_1\ \sqrt{z}_2\ \sqrt{z}_3 = -\frac{q}{8}.$$

Definition 1.4.3

An equation of the form,

$$x^n-\alpha = 0, \tag{1.4.16}$$

where α is a non-zero complex number is called a binomial equation.

Definition 1.4.4

The roots of equation (1.4.16) are said to be the nth roots of the number α and are denoted by $\sqrt[n]{\alpha}$. Thus $\sqrt[n]{\alpha}$ stands for any of the n roots of equation (1.4.16).

Theorem 1.6 - Roots of a Binomial Equation

Equation (1.4.16) has n simple roots $x_1,\ldots,x_n$ given by $x_{k+1} = \sqrt[n]{(r)}\ (\cos\frac{\phi+2k\pi}{n} + i\sin\frac{\phi+2k\pi}{n})$ $(k=0,1,\ldots,n-1)$ where $\alpha=r(\cos\phi+i\sin\phi)$ is the trigonometric form of the number α, $\sqrt[n]{r}>0$.

(B) LINEAR ALGEBRA

The numerical solution of partial differential equations generally requires us:

(i) to solve the matrix system,

$$A\underline{u} = \underline{f} \ , \qquad (1.4.17)$$

where A has m rows and columns with the elements a_{ij} $(i,j=1,2,\ldots,m)$. The vectors $\underline{u}$ and $\underline{f}$ have m components.

or

(ii) to find all possible *eigenvalue-eigenvector* $\mu,\underline{v}$ pairs of matrix A which satisfy the relation,

$$A\underline{v} = \mu\underline{v} \ , \qquad (1.4.18)$$

for some scalar μ and a non-zero column vector $\underline{v}$.

A review of notation and properties for a square matrix A of order m which is relevant to the solution of equation (1.4.17) or (1.4.18) is now given. The following notations are used:

$A=[a_{ij}]$	square matrix of order m.
a_{ij}	the element in the ith row and jth column of the matrix A.
$A^{-1}=[\alpha_{ij}]$	inverse of A.
$A^T=[a_{ji}]$	transpose of A.
$A^H=[\bar{a}_{ji}]$	conjugate hermitian transpose of A.
det(A)	determinant of A.
$\rho(A)$	spectral radius of A.
I	identity matrix of order m.
O	null matrix.
$\underline{u}$	column vector with elements u_i $(i=1,2,\ldots,m)$.
$\underline{u}^T$	row vector with elements u_j $(j=1,2,\ldots,m)$.
$\bar{\underline{u}}$	complex conjugate of $\underline{u}$.

$\|\|A\|\|$	norm of A.
$\|\|\underline{u}\|\|$	norm of $\underline{u}$.
π	permutation matrix which has entries of zeros and ones only, with one non-zero entry in each row and column.

Definitions 1.4.5

The matrix A is:

non-singular if $\det(A)\neq 0$,

symmetric if $A=A^T$,

orthogonal if $A^{-1}=A^T$,

null if $a_{ij}=0$ $(i,j=1,2,\ldots,m)$,

diagonal if $a_{ij}=0$ for $i\neq j$ and $a_{ii}=\alpha_i$ (we usually write $A=\text{diag}\{\alpha_1,\alpha_2,\ldots,\alpha_m\}=\text{diag}\{\alpha_i\}$).

diagonally dominant if $|a_{ii}|\geq\sum_{j\neq i}|a_{ij}|$ for all i.

banded with band width $w=2p+1$ if $a_{ij}=0$ for $|i-j|>p$ (for $p=1$, A becomes tridiagonal and for $p=2$, A is ~~pen~~tadiagonal).

quin

block diagonal if

$$A=\begin{bmatrix} B_1 & & & \\ & B_2 & & O \\ & & \ddots & \\ & O & & B_s \end{bmatrix}$$

where each B_k $(k=1,2,\ldots,s)$ is a square matrix, not necessarily of the same order.

upper triangular if $a_{ij}=0$, $i>j$.

lower triangular if $a_{ij}=0$, $j>i$.

irreducible if there exists no permutation transformation $\pi A\pi^{-1}$ which reduces A to the form

$$\begin{bmatrix} P & O \\ R & Q \end{bmatrix}$$

where P and Q are square submatrices of order p and q respectively $(p+q=m)$ and O is a $p\times q$ null matrix.

1.5 EIGENVALUES AND EIGENVECTORS

Definition 1.5.1

The characteristic equation of A is $\det(A-\mu I)=0$. The roots μ_i $(i=1,2,\ldots,m)$ of this equation are the eigenvalues of A and they are related to the coefficients a_{ij} of A by

$$\text{(i)} \quad \mu_1+\mu_2+\ldots+\mu_m = a_{11}+a_{22}+\ldots+a_{mm} = \text{trace}(A); \tag{1.5.1}$$

$$\text{(ii)} \quad \mu_1\mu_2\ldots\mu_m = \det(A). \tag{1.5.2}$$

Definition 1.5.2

Two matrices A and B are similar if $B=H^{-1}AH$ for some non-singular matrix H. $H^{-1}AH$ is a similarity transformation of A. A and B commute if $AB=BA$.

Theorem 1.7

If A is an m×m matrix, H is a non-singular matrix of the same order and $B=H^{-1}AH$, then A and B have the same eigenvalues. If $\underline{v}$ is any eigenvector of A then B has a corresponding eigenvector $\underline{w}=H^{-1}\underline{v}$.

Proof:

Let μ be any eigenvalue of A and $\underline{v}$ be the corresponding eigenvector. Then,

$$A\underline{v} = \mu\underline{v}\ .$$

But

$$B = H^{-1}AH.$$

Thus if $\underline{w}=H^{-1}\underline{v}$, we obtain

$$B\underline{w} = H^{-1}AH(H^{-1}\underline{v}) = H^{-1}A\underline{v} = H^{-1}(\mu\underline{v})$$

$$= \mu H^{-1}\underline{v} = \mu\underline{w}\ .$$

Hence μ is also an eigenvalue of B and $\underline{w}$ is the corresponding eigenvector. Conversely, we can easily prove that if γ is an eigenvalue of B, with corresponding eigenvector $\underline{w}$, then γ is an eigenvalue of A corresponding to $\underline{v}=H\underline{w}$.

Definition 1.5.3

Vectors $\underline{u}_1,\underline{u}_2,\ldots,\underline{u}_m$ are said to be linearly dependent if non-zero constants $c_1,c_2,\ldots,c_m$ exist such that

$$c_1\underline{u}_1 + c_2\underline{u}_2 +\ldots+c_m\underline{u}_m = \underline{O} . \qquad (1.5.3)$$

If this equation holds only for $c_1=c_2=\ldots=c_m=0$, however, the vectors $\underline{u}_1,\underline{u}_2,\ldots,\underline{u}_m$ are said to be linearly independent.

Theorem 1.8

If A has distinct eigenvalues $\mu_1,\mu_2,\ldots,\mu_m$ then the corresponding eigenvectors $\underline{v}_1,\underline{v}_2,\ldots,\underline{v}_m$ are linearly independent.

Proof:

Suppose, if possible, that the eigenvectors $\underline{v}_1,\underline{v}_2,\ldots,\underline{v}_m$ are linearly dependent. Without loss of generality we may assume that $\underline{v}_1$, $\underline{v}_2,\ldots,\underline{v}_k$ are linearly independent while $\underline{v}_1,\underline{v}_2,\ldots,\underline{v}_k,\underline{v}_{k+1}$ form a linearly dependent set. Then for some constants $\alpha_1,\alpha_2,\ldots,\alpha_k$ we must have,

$$\underline{v}_{k+1} = \sum_{i=1}^{k} \alpha_i\underline{v}_i . \qquad (1.5.4)$$

Multiplying eqn.(1.5.4) by matrix A gives,

$$\mu_{k+1}\underline{v}_{k+1} = A\underline{v}_{k+1} = \sum_{i=1}^{k} \alpha_i A\underline{v}_i = \sum_{i=1}^{k} \alpha_i\mu_i\underline{v}_i . \qquad (1.5.5)$$

Multiplying eqn.(1.5.4) by μ_{k+1} gives,

$$\mu_{k+1}\underline{v}_{k+1} = \sum_{i=1}^{k} \alpha_i\mu_{k+1}\underline{v}_i . \qquad (1.5.6)$$

Subtracting (1.5.5) from (1.5.6) shows that

$$\sum_{i=1}^{k} \alpha_i(\mu_{k+1}-\mu_i)\underline{v}_i = \underline{0} .$$

Since $\mu_{k+1}-\mu_i \neq 0$, for any $i=1,\ldots,k$ and $\underline{v}_1,\underline{v}_2,\ldots,\underline{v}_k$ are independent, we must have $\alpha_i=0$ for all i. Hence $\underline{v}_{k+1}=0$ which is impossible.

Corollary 1.8

If an m×m matrix A has distinct eigenvalues $\mu_1,\mu_2,\ldots,\mu_m$ then A is similar to the diagonal matrix $\mathrm{diag}\{\mu_1,\mu_2,\ldots,\mu_m\}$.

Proof:

Since A has distinct eigenvalues, it has a set $\underline{v}_1,\underline{v}_2,\ldots,\underline{v}_m$ of linearly independent eigenvectors. Let H be the matrix whose column vectors are $\underline{v}_1,\underline{v}_2,\ldots,\underline{v}_m$, i.e. $H=(\underline{v}_1,\underline{v}_2,\ldots,\underline{v}_m)$. Since the columns of H are linearly independent, H is a non-singular matrix and has an inverse H^{-1}.

Now $H^{-1}AH=H^{-1}A(\underline{v}_1,\underline{v}_2,\ldots,\underline{v}_m)=H^{-1}(A\underline{v}_1,A\underline{v}_2,\ldots,A\underline{v}_m)$. But for each i, $A\underline{v}_i=\mu_i\,\underline{v}_i$.

Hence,

$$\begin{aligned} H^{-1}AH &= H^{-1}(\mu_1\underline{v}_1,\mu_2\underline{v}_2,\ldots,\mu_m\underline{v}_m) \\ &= H^{-1}(\underline{v}_1,\underline{v}_2,\ldots,\underline{v}_m)\ \mathrm{diag}(\mu_1,\mu_2,\ldots,\mu_m) \\ &= I\ \mathrm{diag}(\mu_1,\mu_2,\ldots,\mu_m) \\ &= \mathrm{diag}(\mu_1,\mu_2,\ldots,\mu_m) \end{aligned}$$

where I is the identity matrix of order m.

Theorem 1.9

The matrices A and A^T always have the same eigenvalues.

Proof:

For any square matrix M, we have $\det(M)=\det(M^T)$. Hence,

$$\det(A-\mu I) = \det((A-\mu I)^T) = \det(A^T-\mu I)$$

so that A and A^T have the same characteristic equation and therefore the same eigenvalues.

We note that although A and A^T have the same eigenvalues they will not generally have the same eigenvectors. For any eigenvalue μ, the corresponding eigenvector of A is a solution of

$$A\underline{v} = \mu\underline{v} \ , \qquad (1.5.7)$$

whereas the eigenvector of A^T is a solution of

$$A^T\underline{w} = \mu\underline{w} \ . \qquad (1.5.8)$$

The eigenvectors of A^T are sometimes referred to as the left-hand eigenvectors of A, since transposing equation (1.5.8) gives $\underline{w}^TA=\mu\underline{w}^T$.

Theorem 1.10

The eigenvalues of a real symmetric matrix are all real.

Proof:

Let μ be an eigenvalue and $\underline{v}$ an associated eigenvector of the symmetric matrix A. Since μ might be complex, we should assume that, in general, $\underline{v}=\underline{x}+i\underline{y}$ where $\underline{x}$ and $\underline{y}$ are real vectors and $i=\sqrt{-1}$, the imaginary number.

Now,
$$A\underline{v} = \mu\underline{v} \ , \qquad (1.5.9)$$

and so
$$\underline{v}^HA\underline{v} = \mu\underline{v}^H\underline{v} \ , \qquad (1.5.10)$$

with
$$\underline{v}^H\underline{v} = (\bar{v}_1,\bar{v}_2,\ldots,\bar{v}_m)(v_1,v_2,\ldots,v_m)^T$$
$$= \sum_{i=1}^{m} |v_i|^2 \ , \text{ which is real and positive.}$$

Taking the complex conjugate of eqn.(1.5.9), transposing, and post-multiplying by $\underline{v}$ gives,

$$(\underline{v}^HA^H)\underline{v}=\bar{\mu}\underline{v}^H\underline{v} \ .$$

But, since A is real and symmetric, $A^H = A$ and so

$$\underline{v}^H A\underline{v} = \bar{\mu}\underline{v}^H\underline{v} \ . \qquad (1.5.11)$$

Comparing eqns. (1.5.10) and (1.5.11) shows that

$$\mu\underline{v}^H\underline{v} = \bar{\mu}\underline{v}^H\underline{v} \ .$$

Since $\underline{v}^H\underline{v} \neq 0$, we obtain immediately $\mu = \bar{\mu}$. We note here that since all the eigenvalues of a real symmetric matrix are real it follows that the eigenvectors are also real.

Theorem 1.11 - Orthogonal Property

If $\underline{v}_1$ and $\underline{v}_2$ are eigenvectors of a real symmetric matrix A, corresponding to distinct eigenvalues μ_1 and μ_2, then $(\underline{v}_1, \underline{v}_2) = \underline{v}_1^T\underline{v}_2 = 0$.

Theorem 1.12 - Diagonalisation Theorem

If A is a real symmetric matrix, then there exists a real orthogonal matrix P such that $P^T AP = D$ is a diagonal matrix.

Theorem 1.13 - Gerschgorin's First Theorem

The largest of the moduli of the eigenvalues of the square matrix A cannot exceed the largest sum of the moduli of the elements along any row or any column.

Proof:

Let μ_i be an eigenvalue of the (m×m) matrix A and $\underline{v}_i$ the corresponding eigenvector with components $v_1, v_2, \ldots, v_m$. Then, from the equation $A\underline{v}_i = \mu_i\underline{v}_i$ we have,

$$\begin{aligned} a_{11}v_1 + a_{12}v_2 + \ldots + a_{1m}v_m &= \mu_i v_1 \\ a_{21}v_1 + a_{22}v_2 + \ldots + a_{2m}v_m &= \mu_i v_2 \\ &\vdots \end{aligned}$$

$$a_{s1}v_1 + a_{s2}v_2 + \ldots + a_{sm}v_m = \mu_i v_s$$

$$\vdots$$

Let v_s be the largest in modulus of $v_1, v_2, \ldots, v_m$. Select the s^{th} equation and divide by v_s, giving,

$$\mu_i = a_{s1}\left(\frac{v_1}{v_s}\right) + a_{s2}\left(\frac{v_2}{v_s}\right) + \ldots + a_{sm}\left(\frac{v_m}{v_s}\right)$$

Therefore,

$$|\mu_i| \leqslant |a_{s1}| + |a_{s2}| + \ldots + |a_{sm}|$$

since $\left|\frac{v_i}{v_s}\right| \leqslant 1$ for $i=1,2,\ldots,m$.

In particular this holds for $|\mu_i| = \max|\mu_s|$, $s=1,2,\ldots,m$. Since the eigenvalues of the transpose of A are the same as those of A the theorem is also true for columns.

Theorem 1.14 - Gerschgorin's Circle Theorem or Brauer's Theorem

Each eigenvalue of a matrix A satisfies at least one of the inequalities,

$$|\mu - a_{ii}| \leqslant r_i \quad , r_i = \sum_{j=1}^{m}{}' |a_{ij}| , \ i=1,2,\ldots,m \qquad (1.5.12)$$

where the prime indicates that the term i=j in the sum is omitted. In words, every eigenvalue of A lies in at least one of the circles with centre Q_{ii} and radius r_i in the complex μ-plane.

Proof:

If μ is an eigenvalue of A, and $\underline{v}$ the corresponding eigenvector, then $A\underline{v} = \mu\underline{v}$, which implies,

$$(\mu - a_{ii})v_i = \sum_{j=1}^{m}{}' a_{ij}v_j \ , \ i=1,\ldots,m \qquad (1.5.13)$$

where the prime indicates that the term i=j in the sum has been omitted.

Suppose that v_k is the largest element of $\underline{v}$. Then $\left|\frac{v_j}{v_k}\right| \leq 1$ for all j, and,

$$|\mu - a_{kk}| \leq \sum_{j=1}^{m}{}' |a_{kj}| \left|\frac{v_j}{v_k}\right| \leq \sum_{j=1}^{m}{}' |a_{kj}| .$$

Since this is true for any eigenvalue, this proves the theorem.

In practice, we often wish to use Gerschgorin's Circle Theorem to estimate the eigenvalues of a matrix C where the off-diagonal elements of C are much smaller than the diagonal elements. Instead of applying Gerschgorin's theorem directly to C, much more accurate bounds can often be found by first applying a simple similarity transformation $Q^{-1}CQ$, where Q is diagonal. We illustrate the procedure in the 3×3 case. If Q is the matrix obtained by multiplying the first row of the unit matrix by k, we have,

$$Q^{-1}CQ = \begin{bmatrix} c_{11} & \frac{c_{12}}{k} & \frac{c_{13}}{k} \\ kc_{21} & c_{22} & c_{23} \\ kc_{31} & c_{32} & c_{33} \end{bmatrix} \qquad (1.5.14)$$

The Gerschgorin circles are given by,

Centres	c_{11}	c_{22}	c_{33}
Radii	$\frac{\lvert c_{12}\rvert + \lvert c_{13}\rvert}{k}$	$k\lvert c_{21}\rvert + \lvert c_{23}\rvert$	$k\lvert c_{31}\rvert + \lvert c_{32}\rvert$

Suppose that for k=1 the three circles are disjoint. As k increases, the radius of the first circle will decrease, whereas the radii of the other two circles will increase. Clearly there will be an optimum value of k for which the radius of the first circle will be as small as possible, while still being disjoint from the others.

As an example, we estimate the eigenvalues of,

$$C = \begin{pmatrix} 1 & -10^{-5} & 2\times10^{-5} \\ 4\times10^{-5} & 0.5 & -3\times10^{-5} \\ -10^{-5} & 3\times10^{-5} & 0.1 \end{pmatrix}$$

Direct application of the Gerschgorin's circle theorem shows that the eigenvalues μ_i of C satisfy

$$|1-\mu_1|\leqslant 3\times10^{-5}, \quad |0.5-\mu_2|\leqslant 7\times10^{-5}, \quad |0.1-\mu_2|\leqslant 4\times10^{-5}.$$

However, a much better bound on μ_1 can be obtained by taking $k=10^4$ in (1.5.14), which is the largest power of 10 such that the Gerschgorin circle for μ_1 is disjoint from the other two. (The nearest power of 10 is chosen for convenience). This gives,

$$Q^{-1}CQ = \begin{pmatrix} 1 & -10^{-9} & 2\times10^{-9} \\ 4\times10^{-1} & 0.5 & -3\times10^{-5} \\ -10^{-1} & 3\times10^{-5} & 0.1 \end{pmatrix}$$

so that,

$$|1-\mu_1|\leqslant 3\times10^{-9}, \quad |0.5-\mu_2|\leqslant 7\times10^{-9}, \quad |0.1-\mu_3|\leqslant 4\times10^{-9}.$$

A Note on Eigenvalues and Eigenvectors

Let $\underline{v}$ be an eigenvector of the matrix A corresponding to the eigenvalue μ. Then $A\underline{v}=\mu\underline{v}$. Hence $A(A\underline{v})=A^2\underline{v}=\mu A\underline{v}=\mu^2 v$, showing that the matrix A^2 has an eigenvalue μ^2 corresponding to the eigenvector $\underline{v}$. Similar results hold for $A^p\underline{v}=\mu^p\underline{v}$, p=3,4,...

(i) If $f(A) = a_pA^p+a_{p-1}A^{p-1}+\ldots+a_0I$ is a polynomial in A with scalar coefficients $a_p,\ldots,a_0$, then $f(A)\underline{v}=(a_p\mu^p+\ldots+a_0)\underline{v}=f(\mu)\underline{v}$ showing that f(A) has an eigenvalue f(μ) corresponding to the eigenvector $\underline{v}$.

(ii) The eigenvalue of $[f_1(A)]^{-1}f_2(A)$ corresponding to the eigenvector $\underline{v}$ is $f_2(\mu)/f_1(\mu)$, where $f_1(A)$ and $f_2(A)$ are polynomials in A. The proof is as follows:

By (i) we have,

$$f_1(A)\underline{v} = f_1(\mu)\underline{v} \quad \text{and} \quad f_2(A)\underline{v} = f_2(\mu)\underline{v}.$$

Premultiply both equations by $[f_1(A)]^{-1}$ and write as

$$[f_1(A)]^{-1}\underline{v} = \underline{v}/f_1(\mu) \quad \text{and} \quad [f_1(A)]^{-1}f_2(A)\underline{v} = f_2(\mu)[f_1(A)]^{-1}\underline{v}.$$

Then the elimination of $[f_1(A)]^{-1}\underline{v}$ between these two equations shows that,

$$[f_1(A)]^{-1}f_2(A)\underline{v} = \{f_2(\mu)/f_1(\mu)\}\underline{v}$$

which states, by the definition of an eigenvalue, that $f_2(\mu)/f_1(\mu)$ is an eigenvalue of $[f_1(A)]^{-1}f_2(A)$ corresponding to the eigenvector $\underline{v}$. In a similar manner the eigenvalue of $f_2(A)[f_1(A)]^{-1}$ corresponding to the eigenvector $\underline{v}$ is $f_2(\mu)/f_1(\mu)$.

(iii) All $m\times m$ Hermitian matrices, which includes real symmetric matrices have m linearly independent eigenvectors.

(iv) If the matrices A and B commute and have linear elementary divisors then they have a common system of eigenvectors. In particular, all matrices with distinct eigenvalues, all Hermitian and therefore all real symmetric matrices, have linear elementary divisors.

(v) Let A and B be matrices with a common system of eigenvectors. Let η and μ be the eigenvalues of A and B respectively corresponding to the common eigenvector $\underline{v}$. Then $\underline{v}$ is an eigenvector of AB and $A^{-1}B$ and the corresponding eigenvalues are $\eta\mu$ and $\eta^{-1}\mu$ respectively. These results are easily proved. By hypothesis $B\underline{v}=\mu\underline{v}$ and $A\underline{v}=\eta\underline{v}$. Therefore $AB\underline{v}=\mu A\underline{v}=\mu\eta\underline{v}$. Also $A^{-1}B\underline{v}=\mu A^{-1}\underline{v}=(\mu/\eta)\underline{v}$.

1.6 EIGENVALUES OF SOME COMMONLY ARISING MATRICES

The eigenvalues of the (m×m) tridiagonal matrix A,

$$A = \begin{bmatrix} a & b & & & & \\ c & a & b & & \bigcirc & \\ & c & a & b & & \\ & & \ddots & \ddots & \ddots & \\ & \bigcirc & & c & a & b \\ & & & & c & a \end{bmatrix} \qquad (1.6.1)$$

where a,b and c may be real or complex numbers are given by,

$$\mu_i = a+2b\sqrt{c/b}\ \cos\left(\frac{i\pi}{m+1}\right), \quad i=1,2,\ldots,m. \qquad (1.6.2)$$

If A takes the cyclic tridiagonal form,

$$A = \begin{bmatrix} a & b & & & & c \\ c & a & b & & \bigcirc & \\ & c & a & b & & \\ & & \ddots & \ddots & \ddots & \\ & \bigcirc & & c & a & b \\ b & & & & c & a \end{bmatrix} \qquad (1.6.3)$$

then its eigenvalues are given by,

$$\mu_i = a+2\sqrt{cb}\ \cos\left(\frac{2i\pi}{m}\right), \quad i=0,1,\ldots,m-1. \qquad (1.6.4)$$

1.7 POSITIVE DEFINITE MATRICES

Definition 1.7.1

The matrix A in this definition is always hermitian. (This means symmetric in the real case). The quadratic form $(\underline{u},A\underline{u})$ is said to be positive definite if $(\underline{u},A\underline{u})>0$ for all $\underline{u}\neq 0$, and non-negative if $(\underline{u},A\underline{u})\geq 0$ for all $\underline{u}$. The matrix A is said to be positive or non-negative definite if the corresponding quadratic form is positive or non-negative definite. A quadratic form is said to be indefinite if it is positive for some $\underline{u}$, negative for others.

Definition 1.7.2

A Stieltjes matrix is a real positive definite matrix with all its off-diagonal elements non-positive. If the properties of irreducibility and diagonal dominance are added, the matrix is often referred to as an S-matrix. An S-matrix has the following properties.

(i) $a_{ij} = a_{ji}$

(ii) $a_{ij}\leq 0$ for $i\neq j$

(iii) $|a_{ii}|\geq \sum_{j\neq i} |a_{ij}|$ with strict inequality for at least one i,

(iv) $S\equiv[a_{ij}]$ is irreducible

(v) S is positive definite; and

(vi) the elements of S^{-1} are positive.

Such matrices occur repeatedly in the finite difference solution of partial differential equations.

Theorem 1.15

If $(\underline{u},A\underline{u})$ is a positive definite quadratic form and $B=S^{H}AS$, where S is a non-singular square matrix, then $(\underline{u},B\underline{u})$ is also a positive definite quadratic form. Similarly for non-negative definite form, etc.

Theorem 1.16

(i) $(\underline{u},A\underline{u})$ is positive definite if and only if every eigenvalue of A is positive.

(ii) $(\underline{u},A\underline{u})$ is non-negative definite if and only if all the eigenvalues of A are non-negative.

(iii) $(\underline{u},A\underline{u})$ is indefinite if A has both positive and negative eigenvalues and conversely.

Theorem 1.17

A real matrix is positive (non-negative) definite if and only if it is symmetric and all its eigenvalues are positive (non-negative).

Theorem 1.18

Necessary conditions for a hermitian matrix A to be positive definite are:

(i) The diagonal elements of A must be positive.

(ii) $a_{ii}a_{jj}>|a_{ij}|^2$, $i\neq j$.

(iii) The element of A of largest absolute value must lie on the diagonal.

(iv) $\det(A)>0$ (A is non-singular).

Proof:

In the quadratic form $(\underline{u},A\underline{u})$ choose all the u_j to be zero except u_i. Then $(\underline{u},A\underline{u})=a_{ii}|u_i|^2$, and since $u_i\neq 0$, we must have $a_{ii}>0$. To prove (ii), choose all the u_j to be zero except u_i and u_k. Then,

$$(\underline{u},A\underline{u}) = a_{ii}|u_i|^2+a_{ij}\bar{u}_iu_j+\bar{a}_{ij}u_i\bar{u}_j+a_{jj}|u_j|^2$$

$$= a_{ii}|u_i + \frac{a_{ij}u_j}{a_{ii}}|^2 + \frac{\{a_{ii}a_{jj}-|a_{ij}|^2\}|u_j|^2}{a_{ii}} \qquad (1.7.1)$$

By choosing $u_i=-a_{ij}u_j/a_{ii}$ in this expression, we see that, since $a_{ii}>0$, a necessary condition for $(\underline{u},A\underline{u})$ to be positive is that $a_{ii}a_{jj}-|a_{ij}|^2>0$. To prove (iii), suppose that for some i,j we have $|a_{ij}|\geq a_{ii}$, $|a_{ij}|\geq a_{jj}$ where from (i), $a_{ii}>0$, $a_{jj}>0$. In this case $|a_{ij}|^2\geq a_{ii}a_{jj}$ which contradicts (ii). To prove (iv), we know from eqn.(1.5.2) of Definition 1.5.1 that det(A) is equal to the product of the eigenvalues. From Theorem 1.16, the eigenvalues of a positive definite matrix are all positive so that det(A)>0 which proves (iv).

Theorem 1.19

Either of the following sets of conditions is necessary and sufficient for $(\underline{u},A\underline{u})$ to be positive definite (A is hermitian):

(i) Reduce A to row-echelon form working systematically along the main diagonal. Then all the pivots are positive.

(ii) The principal minors consisting of the determinants of the k×k matrices in the top left-hand corner of A (k=1 to m) are all positive:

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \ldots \qquad (1.7.2)$$

As an example to demonstrate the application of Theorem 1.19, we consider the following quadratic form,

$$F = 2u_1^2+u_2^2+6u_3^2+2u_1u_2+u_1u_3+4u_2u_3 \ .$$

We write down the matrix of the quadratic form, and reduce it to an upper triangular form but not dividing the resulting rows by the pivots, so that the pivots appear on the diagonal of the final matrix,

$$A = \begin{bmatrix} 2 & 1 & \frac{1}{2} \\ 1 & 1 & 2 \\ \frac{1}{2} & 2 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1\frac{3}{4} \\ 0 & 0 & -\frac{1}{4} \end{bmatrix} .$$

The pivots are 2,½ and -1/4 and since one of them is negative, the quadratic form is not positive definite.

1.8 VECTOR AND MATRIX NORMS

Definition 1.8.1

The vector norm of $\underline{u}$ is a non-negative number denoted by $||\underline{u}||$, associated with $\underline{u}$ satisfying:

(a) $||\underline{u}||>0$ for $\underline{u}\neq\underline{0}$, $||\underline{u}||=\underline{0}$, implies $\underline{u}=\underline{0}$.

(b) $||k\underline{u}||=|k|\ ||\underline{u}||$ for any scalar k.

(c) $||\underline{u}+\underline{v}||\leq||\underline{u}||+||\underline{v}||$ (the triangle inequality).

Definition 1.8.2

Each of the following quantities defines a vector norm:

$$||\underline{u}||_1 = |u_1| + |u_2|+ \dots +|u_m|, \tag{1.8.1}$$

$$||\underline{u}||_2 = (|u_1|^2 + |u_2|^2 + \dots+|u_m|^2)^{\frac{1}{2}} . \tag{1.8.2}$$

$$||\underline{u}||_\infty = \max_i|u_i| . \tag{1.8.3}$$

We call these the 1,2 and ∞ norms and they are special cases of the p-norm,

$$||\underline{u}||_p = (|u_1|^p+|u_2|^p + \dots+|u_m|^p)^{1/p} .$$

Definition 1.8.3

The matrix norm of a square matrix A is a non-negative number denoted by $||A||$, associated with A such that:

(a) $||A||>0$ for $A\neq0$, $||A||=0$ implies $A=0$.

(b) $||kA|| = |k|\ ||A||$ for a scalar k .

(c) $||A+B||\leq||A|| + ||B||$.

(d) $||AB||\leq||A||\ ||B||$.

A matrix norm is said to be compatible with a vector norm $||\underline{u}||$ if

(e) $||A\underline{u}||\leq||A||\ ||\underline{u}||$.

Definition 1.8.4

A matrix norm that is compatible with the vector norm $||\underline{u}||$ is defined as,

$$||A|| = \sup_{\underline{u}\neq 0} \frac{||A\underline{u}||}{||\underline{u}||} \qquad (1.8.4)$$

where sup denotes the least upper bound for all $\underline{u}\neq 0$. If we introduce $\underline{z} = \frac{\underline{u}}{||\underline{u}||}$, then $||\underline{z}||=1$ and $\frac{||A\underline{u}||}{||\underline{u}||} = ||A\underline{z}||$ so that (1.8.4) is equivalent to,

$$||A|| = \sup_{||\underline{z}||=1} ||A\underline{z}|| \ . \qquad (1.8.5)$$

If the maximum is attained, then,

$$||A|| = \max_{||\underline{z}||=1} ||A\underline{z}|| \ .$$

Theorem 1.20

The natural norms associated with the 1,2 and ∞ vector norms are:

$$||A||_1 = \max_j \sum_{i=1}^{m} |a_{ij}| \text{ (maximum absolute column sum).} \qquad (1.8.6)$$

$$||A||_2 = \{\text{maximum eigenvalue of } A^H A\}^{\frac{1}{2}} = [\rho(A^H A)]^{\frac{1}{2}} \ . \qquad (1.8.7)$$

$$||A||_\infty = \max_i \sum_{j=1}^{m} |a_{ij}| \text{ (maximum absolute row sum).} \qquad (1.8.8)$$

Theorem 1.21

If A is hermitian (symmetric), $||A||_2=\rho(A)$.

Proof:

Since A is hermitian (or symmetric in the real case) then $A^H=A$ and,

$$\begin{aligned} ||A||_2^2 &= \rho(A^H A) \\ &= \rho(A^2) \end{aligned}$$

$$= \rho^2(A)$$

and hence the result follows.

Theorem 1.22

For any natural norm $||A||$,

$$\rho(A) \leq ||A|| \; .$$

Proof:

Let $\underline{v}_i$ be any normalised eigenvector corresponding to an eigenvalue μ_i. Then,

$$\begin{aligned} ||A|| &= \max_{||\underline{z}||=1} ||A\underline{z}|| \\ &\geq ||A\underline{v}_i|| \\ &= ||\lambda_i \underline{v}_i|| \\ &= |\lambda_i| \; ||\underline{v}_i|| \\ &= |\lambda_i| \, . \end{aligned}$$

This holds for any eigenvalue, so the theorem follows.

Definition 1.8.5

If $||A||$ denotes any matrix norm then the convergence of a sequence of matrices in the sense of this norm is defined as follows:

$\lim\limits_{k\to\infty} A^{(k)}=A$ if and only if $\lim\limits_{k\to\infty}||A-A^{(k)}||=0$. Clearly, if $\lim\limits_{k\to\infty} A^{(k)}=A$ then $\lim\limits_{k\to\infty}||A^{(k)}||=||A||$.

Theorem 1.23

If $||A||<1$, then $\lim\limits_{k\to\infty} A^k=0$. (We say that A is convergent).

Proof:

$$||A^k|| = ||A \; A^{k-1}||$$

$$\leq ||A||\ ||A^{k-1}|| \text{ by Definition 1.8.3(d)}$$
$$\leq ||A||^2\ ||A^{k-2}||$$
$$\vdots$$
$$\leq ||A||^k \text{ and the result follows.}$$

Theorem 1.24

In order that $\lim_{k\to\infty} A^k=0$, it is necessary and sufficient that all eigenvalues μ_i of the matrix A should be less than 1 in modulus.

Proof:

Sufficiency follows at once from Theorem 1.23 and Theorem 1.22. The proof for the necessity can be constructed along the following lines:

When a matrix A has m distinct eigenvalues, then we know from Theorem 1.8 and its corollary that there always exists a similarity transformation which diagonalises A,

$$A \to H^{-1}AH = \operatorname{diag}(\mu_i).$$

If A has multiple eigenvalues then while it may not be possible to find a similarity transformation which reduces it to a diagonal form, it is always possible to reduce it to the *Jordan canonical form.* First we factorise the characteristic polynomial of A:

$$p(\mu) = \prod_{i=1}^{k} (\mu-\mu_i)^{m_i}$$

where m_i is the multiplicity of the eigenvalue of μ_i. A non-singular matrix M can then be found such that,

$$J \equiv M^{-1}AM = \left(\begin{array}{c:c:c:c} J_1 & & & \\ \hdashline & J_2 & & \\ \hdashline & & \ddots & \\ \hdashline & & & J_k \end{array}\right).$$

Here each of the blocks J_i is an $m_i \times m_i$ *Jordan submatrix* of the form,

$$J_i = \begin{pmatrix} \mu_i & e_1 & 0 \cdots\cdots & 0 & 0 \\ 0 & \mu_i & e_2 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \mu_i & e_{m_i-1} \\ 0 & 0 & 0 \cdots\cdots & 0 & \mu_i \end{pmatrix} .$$

Each of the quantities $e_1, e_2, \ldots, e_{m_i-1}$ appearing on the superdiagonal of J is either 0 or 1. Without loss of generality, we may assume in the argument which follows that all the quantities e_1, e_2, etc. take the value 1 in a typical submatrix J_i.

Now note that if

$$J = M^{-1}AM \quad \text{then}$$

$$A = MJM^{-1}$$

and

$$A^k = MJ^kM^{-1}.$$

To obtain the k^{th} power, J^k of the Jordan canonical form it is enough to examine the submatrices J_i. Typically we get:

$$J_i^2 = \begin{pmatrix} \mu_i^2 & 2\mu_i & 1 & 0 \cdots\cdots 0 \\ 0 & \mu_i^2 & 2\mu_i & 1 \cdots\cdots 0 \\ 0 & 0 & \mu_i^2 & 2\mu_i \cdots\cdots 0 \\ \cdots & \cdots & \cdots & \cdots\cdots\cdots \\ & & & \end{pmatrix} ,$$

$$J_i^3 = \begin{pmatrix} \mu_i^3 & 3\mu_i^2 & 3\mu_i & 1 & 0 \cdots\cdots 0 \\ 0 & \mu_i^3 & 3\mu_i^2 & 3\mu_i & 1 \cdots\cdots 0 \\ 0 & 0 & \mu_i^3 & 3\mu_i^2 & 3\mu_i \cdots\cdots 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots\cdots\cdots \\ & & & & \end{pmatrix}$$

and so on. It is then clear that $J^k \to 0$ (and so that $A^k \to 0$) as $k\to\infty$ if and only if $|\mu_i|<1$ for all i.

Theorem 1.25

(a) A necessary and sufficient condition for the series $I+A+A^2+A^3+\ldots$ to converge is that $\lim_{k\to\infty} A^k=0$.

(b) If A is convergent, then I-A is non-singular and $(I-A)^{-1}=I+A+A^2+A^3+\ldots$

Proof:

A necessary condition for the series in part (a) to converge is that $\lim_{k\to\infty} A^k=0$, i.e., that A be convergent. The sufficiency will follow from part (b).

Let A be convergent, where by Theorem 1.24 we know that $\rho(A)<1$. Since the eigenvalues of I-A are $1-\mu$, it follows that $\det(I-A)\neq 0$ and hence this matrix is non-singular. Now consider the identity,

$$(I-A)(I+A+A^2+\ldots+A^k) = I-A^{k+1}$$

which is valid for all integers k. Since A is convergent, the limit as $k\to\infty$ of the right-hand side exists. The limit, after multiplying both sides on the left by $(I-A)^{-1}$, yields

$$(I+A+A^2+\ldots) = (I-A)^{-1}$$

and part (b) follows.

Corollary 1.25

If in some natural norm, $||A||<1$, then $I\pm A$ is non-singular and

$$\frac{1}{1+||A||} \leq ||(I\pm A)^{-1}|| \leq \frac{1}{1-||A||} \quad . \qquad (1.8.9)$$

Proof of Corollary 1.25

By Theorem 1.23 and Theorem 1.25, it follows that if $||A||<1$ then $(I-A)$ is non-singular. For a natural norm we note that $||I||=1$ and so taking the norm of the identity

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

yields,

$$\begin{aligned} 1 &\leq ||(I-A)||.||(I-A)^{-1}|| \\ &\leq (1+||A||)\ ||(I-A)^{-1}||. \end{aligned}$$

Thus the left-hand side inequality is established.

Now write the identity as

$$(I-A)^{-1} = I+A(I-A)^{-1}$$

and take the norm to get

$$||(I-A)^{-1}|| \leq 1+||A||.||(I-A)^{-1}||.$$

Since $||A||<1$ this yields,

$$||(I-A)^{-1}|| \leq \frac{1}{1-||A||} .$$

It should be observed that if A is convergent, so is (-A) and $||A||=||-A||$. Thus Theorem 1.25 and its corollary are immediately applicable to matrices of the form I+A.

Theorem 1.26

If $||A||<1$, then,

$$||(I-A)^{-1}-(I+A+A^2+\ldots+A^k)|| \leq \frac{||A||^{k+1}}{1-||A||} .$$

Proof:

$$(I-A)^{-1}-(I+A+\ldots+A^k) = A^{k+1} + A^{k+2}+ \ldots$$

which implies that,

$$\begin{aligned} ||(I-A)^{-1}-(I+A+\ldots+A^k)|| &\leq ||A||^{k+1} + ||A||^{k+2} + \ldots \\ &= \frac{||A||^{k+1}}{1-||A||} \quad \text{if } ||A||<1 \text{ and } ||I||=1. \end{aligned}$$

1.9 NUMERICAL ALGORITHMS TO SOLVE SOME COMMONLY OCCURRING SYSTEMS OF EQUATIONS

Let us consider the matrix system (1.4.17), $A\underline{u} = \underline{f}$.

(i) If A is tridiagonal then system (1.4.17) may be rewritten as:

$$c_i u_{i-1} + a_i u_i + b_i u_{i+1} = f_i \ , \qquad (1.9.1)$$

for $1 \leq i \leq m$ and $c_1 = b_m = 0$, with

$$A = \begin{pmatrix} a_1 & b_1 & & & \\ c_2 & a_2 & b_2 & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & c_{m-1} & a_{m-1} & b_{m-1} \\ & & & c_m & a_m \end{pmatrix} .$$

To ensure stability of the non-pivoting elimination process, it is assumed that A is diagonally dominant.

The algorithm to solve (1.9.1) is as follows:

First compute,
$$\beta_i = a_i - \frac{c_i b_{i-1}}{\beta_{i-1}} \quad \text{with } \beta_1 = a_1$$

and
$$\gamma_i = \frac{f_i - c_i \gamma_{i-1}}{\beta_i} \quad \text{with } \gamma_1 = \frac{f_1}{a_1} \ .$$

The values of the dependent variable are then computed from,

$$u_m = \gamma_m \quad \text{and } u_i = \gamma_i - \frac{b_i u_{i+1}}{\beta_i} \ .$$

This method to solve tridiagonal systems is called *the Thomas algorithm.*

(ii) If A is cyclic tridiagonal then system (1.4.17) may be rewritten as,

$$\begin{aligned} &a_1 u_1 + b_1 u_2 + c_1 u_m = f_1 \\ &c_i u_{i-1} + a_i u_i + b_i u_{i+1} = f_i \ , \ i=2,3,\ldots,m-1 \qquad (1.9.2) \\ &b_m u_1 + c_m u_{m-1} + a_m u_m = f_m \end{aligned}$$

with

$$A = \begin{pmatrix} a_1 & b_1 & & & & c_1 \\ c_2 & a_2 & b_2 & & O & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & O & & c_{m-1} & a_{m-1} & b_{m-1} \\ b_m & & & & c_m & a_m \end{pmatrix} .$$

To ensure stability of the non-pivoting elimination process, it is assumed that A is diagonally dominant.

The algorithm to solve (1.9.2) is as follows:

Let,

$$g_1 = \frac{b_1}{a_1} \quad ; \quad g_i = \frac{b_i}{a_i - c_i g_{i-1}}$$

$$h_1 = \frac{c_1}{a_1} \quad ; \quad h_i = \frac{c_i h_{i-1}}{a_i - c_i g_{i-1}}$$

$$k_1 = \frac{f_1}{a_1} \quad ; \quad k_i = \frac{f_i + c_i h_{i-1}}{a_i - c_i g_{i-1}} \qquad i=2,3,\ldots,m-1$$

$$G_1 = b_m \quad ; \quad G_i = g_{i-1} G_{i-1}$$

$$H_1 = c_m \quad ; \quad H_i = H_{i-1} - G_{i-1} h_{i-1}$$

$$F_1 = f_m \quad ; \quad F_i = F_{i-1} + G_{i-1} k_{i-1} \quad , \ i=2,3,\ldots,m-1$$

$$g_m = h_m = F_m = G_m = 0 \ ,$$

$$H_m = H_{m-1} - (G_{m-1} + c_m)(g_{m-1} + h_{m-1})$$

and

$$k_m = F_{m-1} + (G_{m-1} + a_m) k_{m-1}$$

The components u_i of the solution vector $\underline{u}$ are then given recursively by,

$$u_m = \frac{k_m}{H_m} \quad ; \qquad u_i = k_i + g_i u_{i+1} + h_i u_m , \quad i=m-1,m-2,\ldots,1.$$

CHAPTER TWO

SURVEY OF CURRENT METHODS TO SOLVE HYPERBOLIC PARTIAL DIFFERENTIAL EQUATIONS

2.1 FINITE DIFFERENCES AND COMPUTATIONAL MOLECULES

The numerical solution of partial differential equations by finite differences refers to the process of replacing the partial derivatives by finite difference quotients and then obtaining solutions of the resulting system of algebraic equations. In this section we shall examine some first and second order difference quotients and determine the *orders of accuracy* for them. We shall do this for functions of two variables x and y. An extension of the discussion to three-dimensional systems (and reduction to one-dimensional system) should be obvious.

Let the problem under consideration be the boundary value problem,

$$L[U] = f, \quad U=U(x,y) \tag{2.1.1}$$

in a domain D subject to certain boundary conditions on the boundary of D. Let the points P_{ij} form a discrete approximation for D with uniform spacing $\Delta x, \Delta y$. A first derivative can be defined as a limit in several ways, as follows,

$$\frac{\partial U}{\partial x}(x,y) = \lim_{\Delta x \to 0} \frac{U(x+\Delta x,y)-U(x,y)}{\Delta x}, \tag{2.1.2}$$

$$\frac{\partial U}{\partial x}(x,y) = \lim_{\Delta x \to 0} \frac{U(x,y)-U(x-\Delta x,y)}{\Delta x}, \tag{2.1.3}$$

and
$$\frac{\partial U}{\partial x}(x,y) = \lim_{\Delta x \to 0} \frac{U(x+\Delta x,y)-U(x-\Delta x,y)}{2\Delta x}, \tag{2.1.4}$$

Now if we replace a derivative by a difference quotient, we are interested to know how good an approximation it is. For this, we use Taylor's Theorem 1.2. For example,

$$U(x+\Delta x,y) = U(x,y) + \Delta x \frac{\partial U}{\partial x}(x,y) + \frac{(\Delta x)^2}{2} \frac{\partial^2 U}{\partial x^2}(x^*,y)$$

where $x \leq x^* \leq x+\Delta x$. Upon solving for $\frac{\partial U}{\partial x}$ leads to,

$$\frac{\partial U}{\partial x}(x,y) = \frac{U(x+\Delta x,y)-U(x,y)}{\Delta x} - \frac{\Delta x}{2}\frac{\partial^2 U}{\partial x^2}(x^*,y) \qquad (2.1.5)$$

The first term on the right-hand side is called a *forward-difference quotient* and it is a first-order replacement for the derivative $\frac{\partial U}{\partial x}$ because the error ε is of the order of Δx, i.e. $\varepsilon = O(\Delta x)$. In the index notation, eqn.(2.1.5) would be written as,

$$\left.\frac{\partial U}{\partial x}\right|_{i,j} = \frac{1}{\Delta x}[U_{i+1,j}-U_{ij}] + O(\Delta x) \ . \qquad (2.1.6)$$

As an alternative to the forward difference approximation of (2.1.6) a backward difference is obtained in a similar manner. The Taylor series expansion for $U(x-\Delta x,y)$ about (x,y) is,

$$U(x-\Delta x,y) = U(x,y) - \Delta x\frac{\partial U}{\partial x}(x,y) + \frac{(\Delta x)^2}{2}\frac{\partial^2 U}{\partial x^2}(x^{**},y) ; \ x-\Delta x \leq x^{**} \leq x.$$

Therefore,

$$\left.\frac{\partial U}{\partial x}\right|_{i,j} = \frac{(U_{ij}-U_{i-1,j})}{\Delta x} + O(\Delta x) \ , \qquad (2.1.7)$$

so that the *backward difference quotient* of (2.1.3) is also first order.

On the other hand, eqn.(2.1.4) involves a *central difference quotient*. To find its error, we need to use Taylor series expansions that are carried one term further as follows,

$$U(x+\Delta x,y) = U(x,y) + \Delta x\frac{\partial U}{\partial x}(x,y) + \frac{(\Delta x)^2}{2!}\frac{\partial^2 U}{\partial x^2}(x,y) + \frac{(\Delta x)^3}{3!}\frac{\partial^3 U}{\partial x^3}(x',y)$$

and

$$U(x-\Delta x,y) = U(x,y) - \Delta x\frac{\partial U}{\partial x}(x,y) + \frac{(\Delta x)^2}{2!}\frac{\partial^2 U}{\partial x^2}(x,y) - \frac{(\Delta x)^3}{3!}\frac{\partial^3 U}{\partial x^3}(x'',y) \ ,$$

where $x \leq x' \leq x+\Delta x$ and $x-\Delta x \leq x'' \leq x$.

If we now subtract the two equations, we find that,

$$U(x+\Delta x,y)-U(x-\Delta x,y) = 2\Delta x\frac{\partial U}{\partial x}(x,y) + \frac{(\Delta x)^3}{6}\frac{\partial^3 U}{\partial x^3}(x',y) + \frac{(\Delta x)^3}{6}\frac{\partial^3 U}{\partial x^3}(x'',y) \ ,$$

That is,

$$U(x+\Delta x,y)-U(x-\Delta x,y) = 2\Delta x \frac{\partial U}{\partial x}(x,y) + \frac{(\Delta x)^3}{3} \frac{\partial^3 U}{\partial x^3}(x''',y)$$

where $x-\Delta x \leq x''' \leq x+\Delta x$. On solving for $\frac{\partial U}{\partial x}$ we get,

$$\frac{\partial U}{\partial x} = \frac{U(x+\Delta x,y)-U(x-\Delta x,y)}{2\Delta x} - \frac{(\Delta x)^2}{6} \frac{\partial^3 U}{\partial x^3}(x''',y)$$

Hence at the points P_{ij}, we have

$$\left.\frac{\partial U}{\partial x}\right|_{i,j} = \frac{U_{i+1,j}-U_{i-1,j}}{2\Delta x} + O[(\Delta x)^2] \ . \tag{2.1.8}$$

Thus we see that the central difference is of higher order since the *leading error term* is $O[(\Delta x)^2]$ instead of $O(\Delta x)$. However, as we shall see later, this does not mean that its application will always give rise to a more useful numerical technique than equation (2.1.6). Which form is preferable frequently depends on the particular problem.

Second difference quotients can be obtained by expanding $U(x+\Delta x,y)$ and $U(x-\Delta x,y)$ in Taylor series to remainders involving $\frac{\partial^4 U}{\partial x^4}$ and then adding the two expansions to obtain

$$\frac{\partial^2 U}{\partial x^2} = \frac{U(x+\Delta x,y)-2U(x,y)+U(x-\Delta x,y)}{(\Delta x)^2} - \frac{(\Delta x)^2}{12} \frac{\partial^4 U}{\partial x^4}(x''',y) \ ,$$

where $x-\Delta x \leq x''' \leq x+\Delta x$. At the points P_{ij}, this central difference quotient becomes,

$$\left.\frac{\partial^2 U}{\partial x^2}\right|_{ij} = \frac{1}{(\Delta x)^2}(U_{i+1,j}-2U_{ij}+U_{i-1,j}) + O[(\Delta x)^2] \ . \tag{2.1.9}$$

There are other difference expressions for $\frac{\partial^2 U}{\partial x^2}$ but equation (2.1.9) is consistently used.

All of the above difference quotients at the points P_{ij} may be pictorially represented by their *stencils* or *computational molecules*.

Some examples are illustrated in Fig. 2.1.1. The numbers, in the various positions, represent the multipliers that are to be applied to the values of u at these stations.

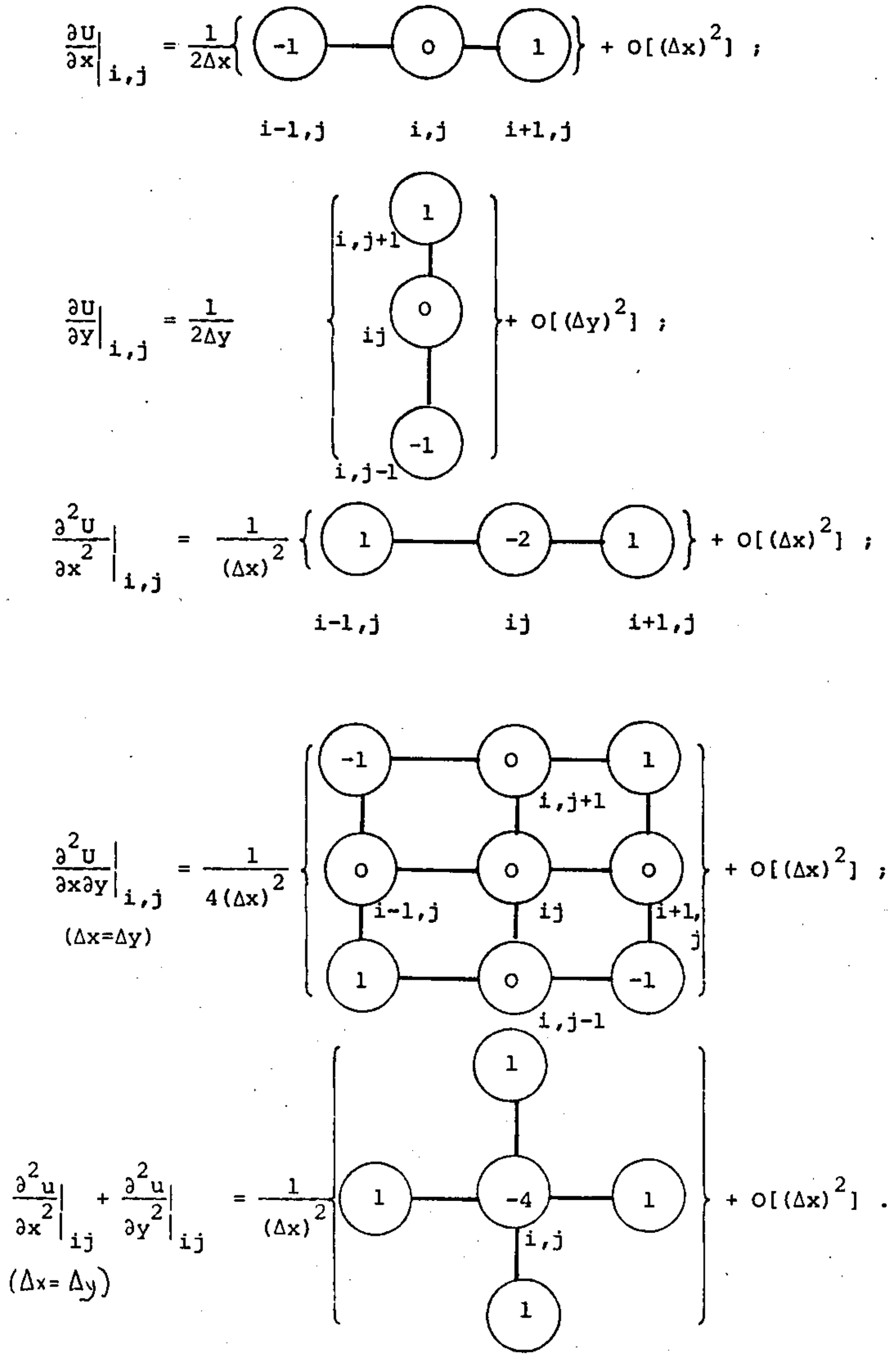
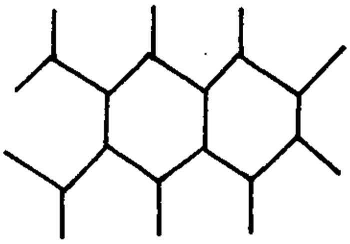


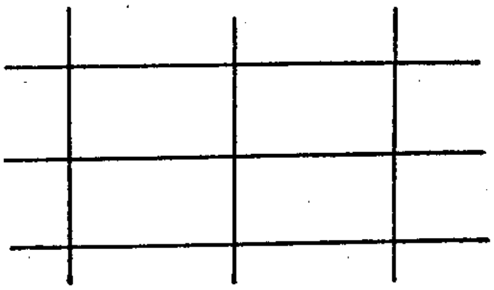
FIGURE 2.1.1

2.2 CHOICE OF DISCRETE NETWORK

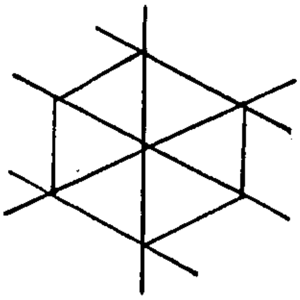
A regular spacing of the mesh points $P_{ij}=P(x_i,y_j)=P(i\Delta x,j\Delta y)$ is generally favoured. This is because if irregular nets are used, the form of the approximating difference equations changes from zone to zone creating tedious programming details. Secondly, computers execute with great speed if simple net structures are used. The only regular networks which can completely fill the x-y plane are rectangles, triangles and hexagons as shown in Fig. 2.2.1.



(a) Hexagons



(b) Rectangles



(c) Triangles

FIGURE 2.2.1

Throughout the course of this thesis, however, the rectangular network will be used. The solution region shall consist of a rectangle which is divided up by a grid system as shown in Fig. 2.2.2. The mesh point P_{ij} lies at the intersection of each grid line. There are m grid lines in the x-direction and n grid lines in the y-direction. The mesh or grid is therefore specified by the sequences $x_0, x_1, \ldots, x_{m-1}$ and $y_0, y_1, \ldots, y_{n-1}$, and

$$x_{i+\frac{1}{2}} = \tfrac{1}{2}(x_i + x_{i+1}) \text{ and } y_{j+\frac{1}{2}} = \tfrac{1}{2}(y_j + y_{j+1}) \quad . \qquad (2.2.1)$$

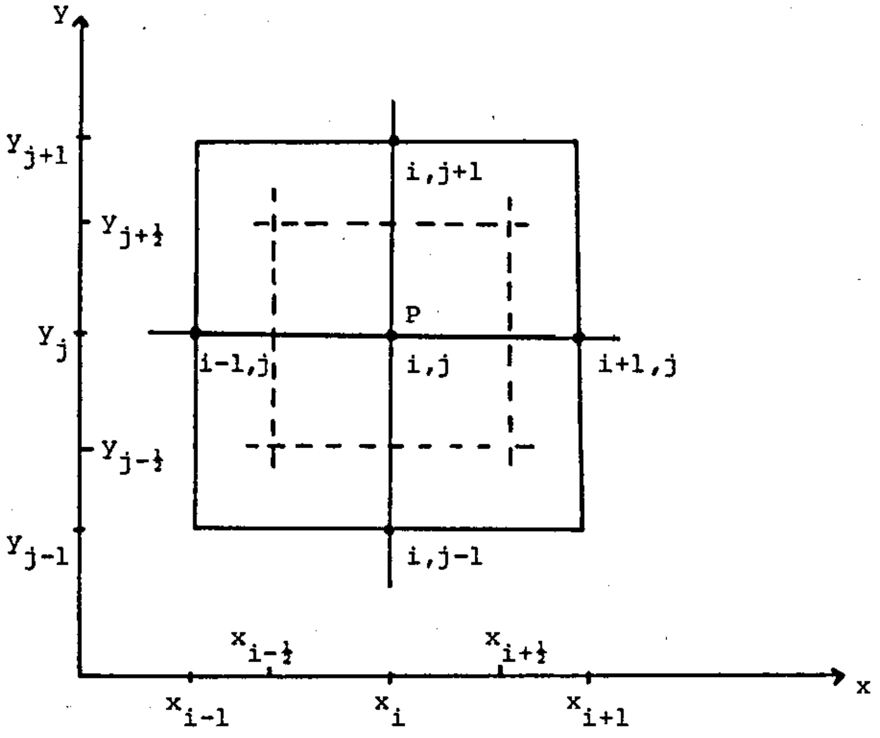


FIGURE 2.2.2

2.3 CONSISTENCY, ACCURACY, EFFICIENCY AND STABILITY

When a partial differential equation (say in one space dimension x and time t) is approximated by a finite difference analogue, one naturally expects that the difference scheme indeed represents the differential system in some sense. By this, we mean that a difference system is consistent with a differential system when the former becomes identical with the latter in the limit as $\Delta x, \Delta t \to 0$. Clearly, *consistency* is a fundamental requirement.

Accuracy of a numerical solution depends on two major classes of errors, i.e. round-off and truncation errors. *Round-off errors* characterise the differences between the solution furnished by the computer and the exact solution of the difference equations. We would obtain the exact solution if it were possible to carry out all calculations to an infinite number of decimal places. In practice, however, each calculation is carried out to a finite number of decimal places or significant figures, a procedure that introduces a rounding error every time it is used. The real danger results from the fact that solutions obtained in one cycle of calculations are used as "initial conditions" for obtaining the values of u for the subsequent time increments. There therefore exists the danger of *error propagation* and *error growth* as the solution proceeds over a large number of steps. This phenomenon is closely related to computational *stability*.

Truncation errors are caused by the approximations involved in representing differential equations. Truncation errors depend upon the spatial grid size (Δx) and step size (Δt). Intuitively, one would assume that the accuracy of any finite difference solution could be decreased by increasing the grid sizes, since, evidently, if the grid spacing is reduced to zero, the discretized equivalent becomes identical to the continuous field. Magnitude of truncation errors can be estimated using Taylor series expansion, as was seen in Section 2.1.

Efficiency refers to the amount of computational work done by the computer in solving a problem over unit time-length. In recommending a numerical method, we need to strike a balance between efficiency and accuracy. This is because a method might have incurred more work (particularly so in an iterative process) to attain accuracy. On the other hand, one may settle for a less accurate method in favour of its simplicity and computing cost effectiveness.

The final concept to be studied is stability. If $u(x,t)$ is the exact solution and u_{ij} is the solution of the finite difference equations, the error of the approximation at the point (i,j) is $(u_{ij}-u(i\Delta x,j\Delta t))$. One is interested to know the behaviour of $|u_{ij}-u(i\Delta x,j\Delta t)|$ as $j\to\infty$ for fixed $\Delta x,\Delta t$. That is, whether the solution is *bounded (stable)* as the time index $j\to\infty$. Also of interest is the behaviour of $|u_{ij}-u(i\Delta x,j\Delta t)|$ as $\Delta t,\Delta x\to 0$ for a fixed value of $j\Delta t$. That is, whether the difference scheme is *convergent*.

It is clear in both cases that as the number of cycles of calculations become large there is a possibility for unlimited amplification of errors. If errors are amplified from time step to time step, the total accumulated error will quickly swamp the solution rendering it worthless. It can therefore be said that a numerical method is stable if a small error at any stage produces a smaller cumulative error.

In the important *Lax equivalence theorem* below, is given a relation between consistency, stability and convergence of the approximations of linear initial value problems by finite difference equations.

Theorem 2.1: Lax Equivalence Theorem

Given a properly posed initial boundary value problem and a finite difference approximation to it that satisfies the consistency condition, then stability is the necessary and sufficient condition for convergence.

The concept of stability may be better understood by first considering a single ordinary differential equation. We now define an error ε_j which occurs at time level j. In stability analysis, one is interested in the amplification of the error at time level j+1. This can be done by writing

$$\varepsilon_{j+1} = \gamma\varepsilon_j \tag{2.3.1}$$

where γ is an *amplification factor* that is related to the integration scheme employed and therefore to the truncation errors of the scheme. To ensure there is no build up of errors, we must have for stability,

$$|\varepsilon_{j+1}| \leqslant |\varepsilon_j| \tag{2.3.2}$$

and by virtue of (2.3.1), this implies

$$|\gamma\varepsilon_j| \leqslant |\varepsilon_j| \tag{2.3.3}$$

or

$$|\gamma| \leqslant 1 \quad . \tag{2.3.4}$$

This idea can obviously be generalised to a system of m first-order ordinary differential equations. The corresponding equation is

$$\underline{\varepsilon}_{j+1} = \Gamma\underline{\varepsilon}_j \quad . \tag{2.3.5}$$

where $\underline{\varepsilon}_j$ is an m-vector and Γ an *amplification matrix* of size (m×m). If equation (2.3.5) is diagonalised, we get

$$\varepsilon_{j+1,\mu} = \gamma_\mu \varepsilon_{j,\mu} \tag{2.3.6}$$

where γ_μ is an eigenvalue of Γ, and ε_μ is the associated eigenvector. The condition for stability must now be applied separately to the amplitude of each error eigenvector which leads to

$$|\varepsilon_{j+1,\mu}| \leqslant |\varepsilon_{j,\mu}| \quad \text{for all } \mu ,$$

or

$$|\gamma_\mu| \leqslant 1 \text{ for all } \mu . \tag{2.3.7}$$

The condition for stability, therefore, reduces to the requirement that the *spectral radius of the amplification matrix* Γ must be less than or equal to one. In general, the eigenvalues of the amplification matrix may be complex, in which case

$$|\gamma_\mu| = (\gamma_\mu \overline{\gamma}_\mu)^{\frac{1}{2}} \tag{2.3.8}$$

where $\overline{\gamma}_\mu$ is the complex conjugate of γ_μ.

Various techniques are available for a quantitative treatment of stability of finite difference schemes. Among those commonly used are the linearised Fourier analysis method, the matrix method, the maximum principle and the energy method. In the *Fourier analysis method*, we examine the propagating effect of a single row of errors, say along the line t=0. These are represented by a finite Fourier series in which the number of terms is equal to the number of mesh points on the line. Usually the effect of a single term is analysed and the complete effect is then obtained by linear superposition. This method is also known as the *von Neumann criterion* for stability and it enables us to derive an expression for the amplification factor. The criterion ignores the boundary conditions, and hence would be truly valid only for pure initial value problems. However, properly posed boundary conditions have little effect on stability. On the other hand, the second technique called the *matrix method* for analysing stability will automatically include the effects of the boundaries. By incorporating the boundary values, the finite difference schemes generate a matrix form of equations. The eigenvalues of the amplification matrix are examined for stability in the same manner as was demonstrated in the case of a system of ordinary differential equations above. In the *maximum principle* we define the difference between the exact solution and the discrete approximation at each of the mesh points. The boundedness of the solution can be established by using the maximum operation on both sides of this equation and assuming the positivity of the coefficients as well as the attainment of a maximum value of their sum. It is worth noting that the above procedure, in a slightly modified form, can be applied even if the coefficients of the differential equation are variable. The final method known as the *energy method* is a name given to a group of techniques based on the use of certain energy-like quantities. In some cases, the energy-like quantities, do, in fact correspond to the physical energy in the system. In mathematical terminology, these quantities are called norms. The key property that norms have in common with energy is that both are positive definite. The use of the energy method clearly depends mainly on one's ability to find a norm for a given difference scheme and then obtain a boundedness condition of the solution u_j as $j\to\infty$ which in turn is the requirement of stability. The above concepts of consistency, stability, convergence and accuracy will be dealt with in greater mathematical detail wherever appropriate throughout this thesis.

2.4 FIRST ORDER HYPERBOLIC DIFFERENTIAL EQUATIONS AND CHARACTERISTICS

In this section, we examine, the solution of the first-order quasi-linear differential equation of the form,

$$a \frac{\partial U}{\partial x} + b \frac{\partial U}{\partial y} = c \qquad (2.4.1)$$

where a,b and c are functions of U,x and y but do not involve the partial derivatives of U. It is customary to use the notation $\frac{\partial U}{\partial x} = p$ and $\frac{\partial U}{\partial y} = q$ and to write the above equation as

$$ap + bq = c \ . \qquad (2.4.2)$$

Let us assume that values of U are prescribed on a certain curve C in the (x,y) plane. We would then be interested to find uniquely the solution values U elsewhere. In other words, we would expect to determine a surface U=U(x,y) on which the differential equation (2.4.1) or (2.4.2) is satisfied. This is analogous to finding the derivatives p and q on the curve C on which U is given in such a way that the differential equation is satisfied. For this purpose, we consider a parameter s which is the arc length on the curve C. The arc rate of change of U along C is known and is related to the values of p and q satisfying the differential equation by

$$\frac{dU}{ds} = \frac{\partial U}{\partial x}\frac{dx}{ds} + \frac{\partial U}{\partial y}\frac{dy}{ds} \ ,$$

i.e.,

$$\frac{dx}{ds}p + \frac{dy}{ds}q = \frac{dU}{ds} \ . \qquad (2.4.3)$$

It is seen that (2.4.2) and (2.4.3) are two simultaneous equations in the two unknowns p and q given by,

$$\frac{dx}{ds}p + \frac{dy}{ds}q = \frac{dU}{ds}$$

and ap+bq = c,

which will in general bear the unique solution given by,

$$\frac{p}{\Delta_1} = \frac{q}{\Delta_2} = \frac{1}{\Delta} \quad , \tag{2.4.4}$$

where

$$\Delta = \begin{vmatrix} \frac{dx}{ds} & \frac{dy}{ds} \\ a & b \end{vmatrix} = (b\frac{dx}{ds} - a\frac{dy}{ds}) \;; \tag{2.4.5}$$

$$\Delta_1 = \begin{vmatrix} \frac{dU}{ds} & \frac{dy}{ds} \\ c & b \end{vmatrix} = (b\frac{dU}{ds} - c\frac{dy}{ds}) \tag{2.4.6}$$

and

$$\Delta_2 = \begin{vmatrix} \frac{dx}{ds} & \frac{dU}{ds} \\ a & c \end{vmatrix} = (c\frac{dx}{ds} - a\frac{dU}{ds}) \;, \tag{2.4.7}$$

provided $\Delta \neq 0$. However, when $\Delta = 0$, the values of p and q will usually be infinite in which case the known values of U on C will not satisfy the differential equation. If Δ_1 and Δ_2 also likewise vanish then p and q can be finite and satisfy equation (2.4.2). In such a case, we have from (2.4.5)

$$ady - bdx = 0 \;, \tag{2.4.8}$$

from (2.4.6)

$$bdU - cdy = 0 \;, \tag{2.4.9}$$

and from (2.4.7),

$$cdx - adU = 0 \;. \tag{2.4.10}$$

Equation (2.4.8) is a differential equation for the curve C and equations (2.4.9) and (2.4.10) are differential equations for the solution values of U on C. The curve C relating x and y is called a *characteristic curve* or *characteristic*. The three equations above may be written as,

$$\frac{dx}{a} = \frac{dy}{b} = \frac{dU}{c} \quad , \tag{2.4.11}$$

which are known as the *subsidiary equations*.

2.5 ROLE OF THE CHARACTERISTICS

The characteristics play a very important role in developing solutions for hyperbolic equations. This is related to the extent to which initial and boundary conditions determine unique solutions in a certain region or regions. Although our prime concern is on first order equations, the above idea is perhaps best illustrated by first examining the solution of a Cauchy problem involving differential equations of second order. We consider the well-known simple wave equation,

$$\frac{\partial^2 U}{\partial x^2} - \frac{\partial^2 U}{\partial t^2} = 0 . \tag{2.5.1}$$

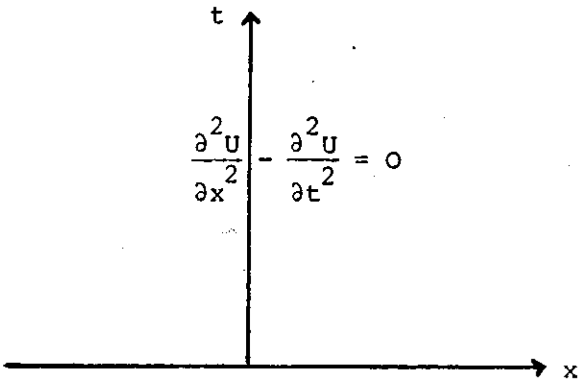
Our Cauchy problem for (2.5.1) will be in the form of an initial value problem in which we seek a function $U(x,t)$ which is defined and continuous for $-\infty<x<\infty$, $t\geq 0$; which satisfies equation (2.5.1) for $-\infty<x<\infty$, $t>0$ and which satisfies the initial conditions,

$$U(x,0) = f_1(x) \ , \quad -\infty<x<\infty \ , \tag{2.5.2a}$$

and

$$\frac{\partial U}{\partial t}(x,0) = f_2(x) \ , \quad -\infty<x<\infty, \tag{2.5.2b}$$

where $f_1(x)$ and $f_2(x)$ are given functions of x. Obviously, this Cauchy problem is defined on a half-plane as is shown in the following figure,



$U(x,0)=f_1(x)$; $\frac{\partial U}{\partial t}(x,0) = f_2(x)$

FIGURE 2.5.1

The solution to the Cauchy problem is given by the D'Alembert formula,

$$U(x,t) = \tfrac{1}{2}\{f_1(x+t)+f_1(x-t) + \int_{x-t}^{x+t} f_2(r)dr\} \qquad (2.5.3)$$

which may be derived in the following manner:

The change to variables,

$$\xi = x+t \ , \eta = x-t \ , \qquad (2.5.4)$$

changes equation (2.5.1) to

$$\frac{\partial^2 U}{\partial\xi\partial\eta} = 0 \ . \qquad (2.5.5)$$

On integrating this equation, gives us

$$\frac{\partial U}{\partial\xi} = F_1(\xi) \ ,$$

and
$$U = \int_0^\xi F_1(r)dr + G_2(\eta)$$

where F_1 and G_2 are arbitrary differentiable functions. Setting

$$G_1(\xi) = \int_0^\xi F_1(r)dr$$

yields,

$$U = G_1(\xi) + G_2(\eta) \ , \qquad (2.5.6)$$

from which we find,

$$\frac{\partial U}{\partial t} = \frac{\partial G_1}{\partial\xi}\frac{\partial\xi}{\partial t} + \frac{\partial G_2}{\partial\eta}\frac{\partial\eta}{\partial t} = \frac{\partial G_1(\xi)}{\partial\xi} - \frac{\partial G_2(\xi)}{\partial\eta} \ . \qquad (2.5.7)$$

By using (2.5.6) and (2.5.7) together with equations (2.5.2a),(2.5.2b) and (2.5.4), it follows that,

$$U(x,0) = G_1(x) + G_2(x) = f_1(x) \ , \qquad (2.5.8)$$

and
$$\frac{\partial U}{\partial t}(x,0) = G_1'(x)-G_2'(x) = f_2(x) \ . \qquad (2.5.9)$$

On differentiating equation (2.5.8) gives us

$$G_1'(x) + G_2'(x) = f_1'(x) \ . \qquad (2.5.10)$$

Hence, on solving equations (2.5.9) and (2.5.10) for $G_1'(x)$ and $G_2'(x)$, leads us to,

$$G_1'(x) = \tfrac{1}{2}\{f_1'(x)+f_2(x)\};\; G_2'(x) = \tfrac{1}{2}\{f_1'(x)-f_2(x)\}. \quad (2.5.11)$$

This implies from integration that

$$G_1(x) = \tfrac{1}{2}\{f_1(x)+\int_0^x f_2(r)dr\};\; G_2(x) = \tfrac{1}{2}\{f_1(x)-\int_0^x f_2(r)dr\}. \quad (2.5.12)$$

From equations (2.5.4), (2.5.6) and (2.5.12) we deduce that

$$U(x,t) = \tfrac{1}{2}\{f_1(x+t)+\int_0^{x+t} f_2(r)dr\} + \tfrac{1}{2}\{f_1(x-t)-\int_0^{x-t} f_2(r)dr\}$$

i.e. $$U(x,t) = \tfrac{1}{2}\{f_1(x+t)+f_1(x-t)+\int_0^{x+t} f_2(r)dr + \int_{x-t}^{0} f_2(r)dr\}$$

or $$U(x,t) = \tfrac{1}{2}\{ f_1(x+t)+f_1(x-t) + \int_{x-t}^{x+t} f_2(r)dr\}$$

which is known as the *D'Alembert formula.*

A number of observations can be made from this formula. Let us suppose that we wish to solve the Cauchy problem at a point (x_0,t_0) as shown in Figure 2.5.2 below.

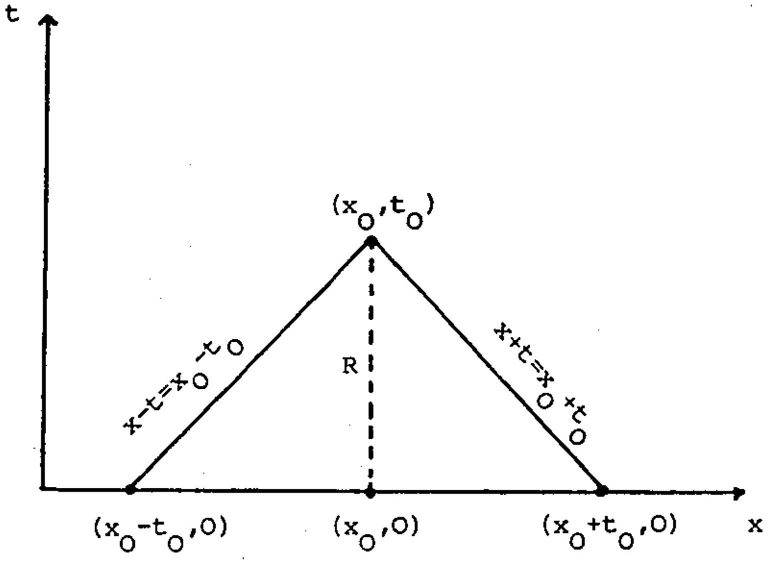


FIGURE 2.5.2

It follows immediately from equation (2.5.3) that the solution at the particular point (x_0,t_0), i.e. $U(x_0,t_0)$ is completely determined by the initial data on that segment of the x-axis cut out by the lines $x-t=x_0-t_0$ and $x+t=x_0+t_0$ (i.e. determined by a knowledge of f_1 and f_2 only between the points $(x_0-t_0,0)$ and $(x_0+t_0,0)$ on the x-axis) as is shown in Fig. 2.5.2. This segment $x_0-t_0 \leq x \leq x_0+t_0$ is therefore called the *interval of dependence* of the point (x_0,t_0). The region R interior to the triangle with vertices (x_0,t_0), $(x_0+t_0,0)$ and $(x_0-t_0,0)$ is called the *region of dependence*. The lines $x-t=x_0-t_0$ and $x+t=x_0+t_0$ are the characteristics of the wave equation through (x_0,t_0).

If, for example, the initial conditions are given only on $0 \leq x \leq a$, then the solution is found in the region of dependence determined by the point $(x_0,t_0) = (\frac{a}{2},\frac{a}{2})$, that is the triangle bounded by the initial curve Γ (interval of dependence) and the characteristics $t=x$ and $t=-x+a$ inclined at an angle of $\theta_1=\tan^{-1}1$ and $\theta_2=\tan^{-1}-1$ respectively with the x-axis. It is clear that the solution U cannot be determined at points lying outside this region of dependence. Thus we see the importance of the characteristics in determining the solutions of hyperbolic equations.

2.6 DETERMINATION OF THE UNIQUE SOLUTION

We demonstrate the importance of the characteristics in determining unique solutions of hyperbolic initial and boundary value problems by considering a system of two simultaneous first-order quasi-linear equations,

$$a_1 \frac{\partial U}{\partial x} + b_1 \frac{\partial U}{\partial y} + c_1 \frac{\partial V}{\partial x} + d_1 \frac{\partial V}{\partial y} = f_1 \ ,$$

$$a_2 \frac{\partial U}{\partial x} + b_2 \frac{\partial U}{\partial y} + c_2 \frac{\partial V}{\partial x} + d_2 \frac{\partial V}{\partial y} = f_2 \ . \tag{2.6.1}$$

There are three typical situations in which a *unique solution* for the system may be found. It is assumed that in all cases, the hyperbolic system (2.6.1) gives rise to two families of characteristic curves which we identify as the α- *and* β-*characteristics*. The continuously differentiable values of U or V may be specified on one of the characteristics or on both of them or they may not be prescribed on either characteristics.

(a) In the first case (Fig. 2.6.1), the initial curve is given by the continuously differentiable non-characteristic curve CD on which values of U and V are prescribed. A unique solution would then lie in a region CDE bounded by CD, the α-characteristic (DE) and the β-characteristic (CE).

(b) In the second situation (Fig. 2.6.2) we have CD as the non-characteristic curve which is continuously differentiable and along which values of U or V are given. The values of U or V are also given along the α-characteristic (CF). To ensure continuity at C, U and V are known at this point and must be compatible with the α-, β-characteristics. The uniqueness region would then be CDEF.

(c) In the last case (Fig. 2.6.3), a unique solution is determined in the region CDFE where EF and FD are characteristics (α- and β-characteristics). The values of U and V are known at C and continuously differentiable values of U or V are given along the characteristics CE and CD. The values at C must be compatible with the characteristics for continuity.

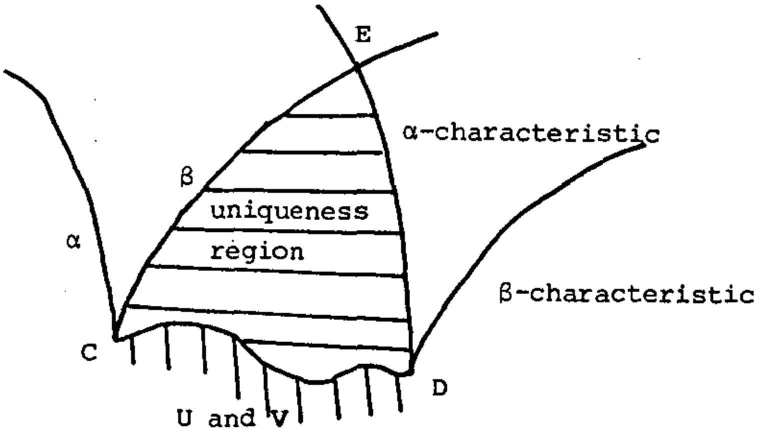


FIGURE 2.6.1

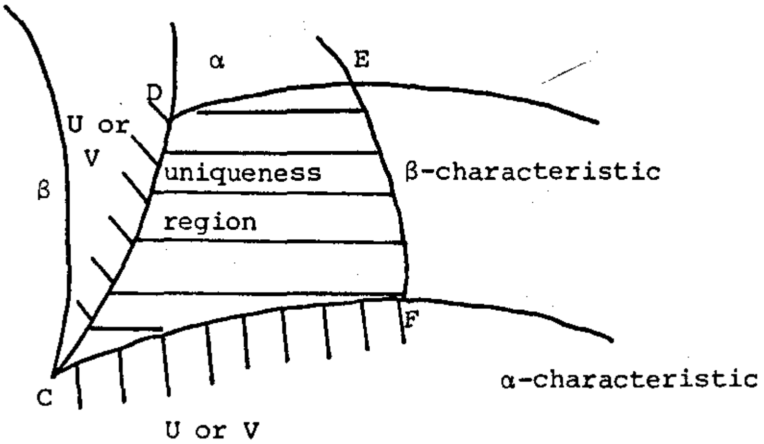


FIGURE 2.6.2

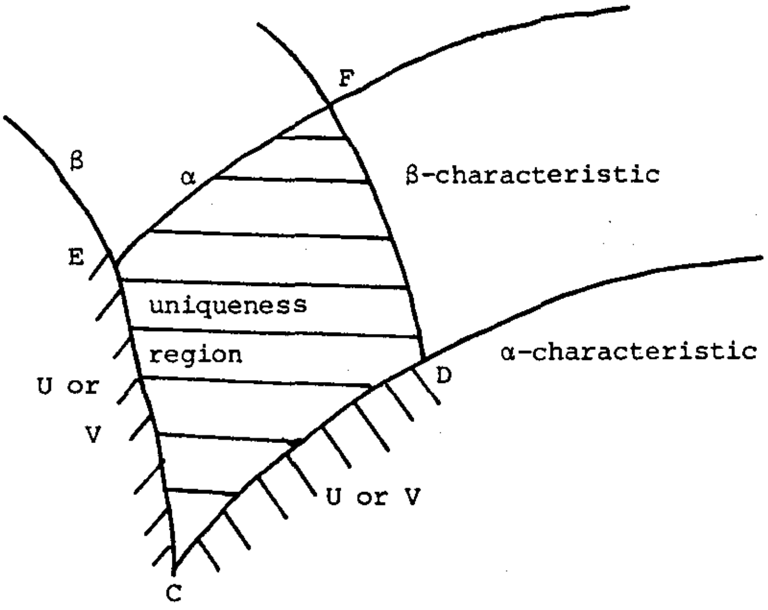


FIGURE 2.6.3

Some simple examples may serve to demonstrate the underlying principle of uniqueness and the influence of characteristics to the solution of first-order hyperbolic equations.

Example 2.1

We consider the equation,

$$y \frac{\partial U}{\partial x} + \frac{\partial U}{\partial y} = 2 \ , \qquad (2.6.2)$$

where the initial data $U(x_i,0)=U_i$ for U are specified on the non-characteristic line segment Γ defined by $y=0$, $0\leq x\leq 1$.

The subsidiary equations are given by,

$$\frac{dx}{y} = \frac{dy}{1} = \frac{dU}{2} \ . \qquad (2.6.3)$$

From the first of equation (2.6.3), we find that the characteristics are the parabolas $x=y^2/2+\alpha$, α being a characteristic constant. Along this family of characteristics, we have from the last equation of (2.6.3) $U=2y+\beta$. Since on the initial curve Γ, $U(x_i,0)=U_i$, $y=0$, $0\leq x\leq 1$, then the value of U for $y>0$ is obtained by integrating along the characteristics drawn from the points x_i on Γ. Thus

$$U(x,y) = 2y + U_i$$

on the parabolic curve $y^2=2(x-x_i)$ for each x_i, $0\leq x_i\leq 1$, $i=1,2,\ldots$ This solution is unique in the region bounded by, and including, the terminal characteristics $y^2=2x$ and $y^2=2(x-1)$ as is indicated by Fig. 2.6.4. $U(x,y)$ cannot be determined at points off this region of dependence.

Now suppose that the initial curve happens to be a characteristic, say the characteristic $y^2=2x$ which passes through the origin. Clearly, we would have only one solution $U=2y+U_0$ where U_0 is the specified initial value of $U(x,y)$ at the origin. This is in contrast with the original case when the initial curve Γ is a non-characteristic curve on which the initial values of U can be arbitrarily prescribed. Elsewhere the solution is not unique since we can take, for example,

$$U = 2y + U_0 + A(y^2-2x)$$

which obviously gives us back the solution $U=2y+U_0$ along $y^2=2x$ for any value of A. However, it is also a solution for points not on $y^2=2x$ (i.e. $y^2\neq 2x$) since by direct differentiation it satisfies the differential equation (2.6.2) which we want to solve. Consequently, since A is arbitrary we can have an infinite number of solutions. The solution is not defined uniquely at points off $y^2=2x$ essentially because the *terminal characteristics* are effectively coincidental.

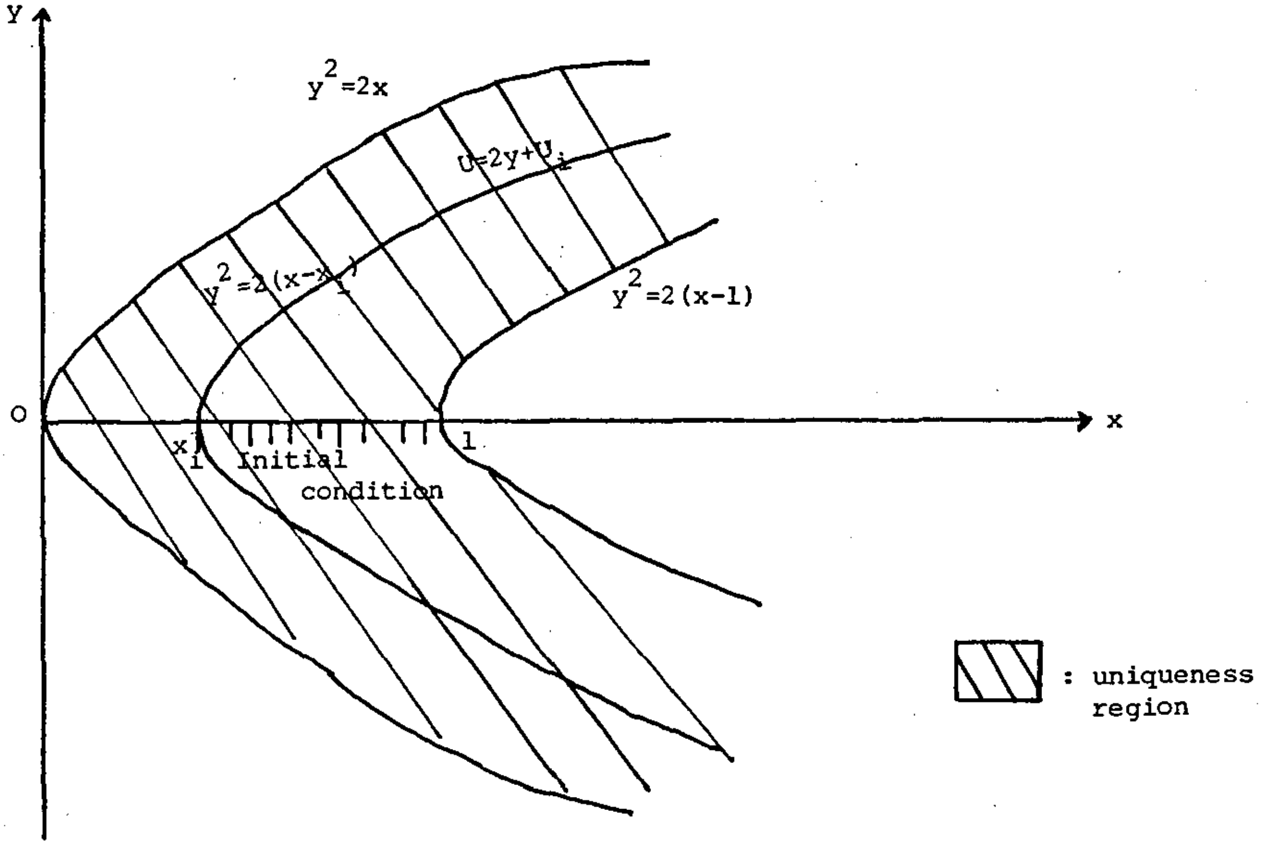


FIGURE 2.6.4

In general, when the curve on which initial values are given is itself a characteristic, the differential equation can have no solution unless the initial conditions fulfilled the necessary differential relationship for this characteristic. In the latter case, the solution will be unique along the initial curve and will not be so elsewhere.

Example 2.2

We now consider the initial-boundary value problem given by

$$\frac{\partial U}{\partial x} + \frac{\partial U}{\partial t} = 0, \ 0<x<\infty, t>0 \ , \qquad (2.6.4)$$

with the initial conditions,

$$\begin{aligned} U(x,0) &= x(x-2) \ , \ 0 \leqslant x \leqslant 2 \\ U(x,0) &= 2(x-2) \ , \ x \geqslant 2 \end{aligned} \tag{2.6.5a}$$

and the boundary condition,

$$U(0,t) = 2t, \ t>0 \ . \tag{2.6.5b}$$

From the given differential equation (2.6.4), its associated subsidiary equations are,

$$\frac{dt}{1} = \frac{dx}{1} = \frac{dU}{0} \ . \tag{2.6.6}$$

The first of these equations yields the characteristics,

$$t = x+\alpha \tag{2.6.7}$$

which are in fact straight lines with α being a parameter which is a constant for each characteristic. The curve Γ along which values of U are known is made up of the x- and t-axes. On the x-axis we have the given initial values (2.6.5a) whilst on the t-axis lies the known boundary values (2.6.5b). Hence we find that the first family of characteristics are drawn from points x_i on Γ. The second family of characteristics are in turn drawn from points t_j. In the first case, (2.6.7) takes the form $t=x-x_i$. Similarly, we obtain $t-t_j=x$ for the second case. We note from the subsidiary equations (2.6.6) that U is constant along each of these families of characteristic lines. Therefore if we denote $U(x,0)=\phi(x)$ and $U(0,t)=\psi(t)$ then the solutions $U(x,t)$ along the characteristics $t=x-x_i$ drawn from the points x_i are given by,

$$\begin{aligned} U(x,t) &= \phi(x_i) = \phi(x-t) \\ &= \begin{cases} (x-t)(x-t-2), & 0 \leqslant x_i \leqslant 2 \\ 2(x-t-2); & x_i \geqslant 2 \end{cases} \qquad i=1,2,\ldots \end{aligned}$$

In the same manner, we arrive at the solution

$$U(x,t) = \psi(t_j) = \psi(t-x)$$
$$= 2(t-x); \quad t_j>0, \ j=1,2,...$$

along the characteristics $t-t_j=x$ extended from the points t_j on Γ. The position of the characteristics and the solution domains of the problem are displayed in Fig. 2.6.5.

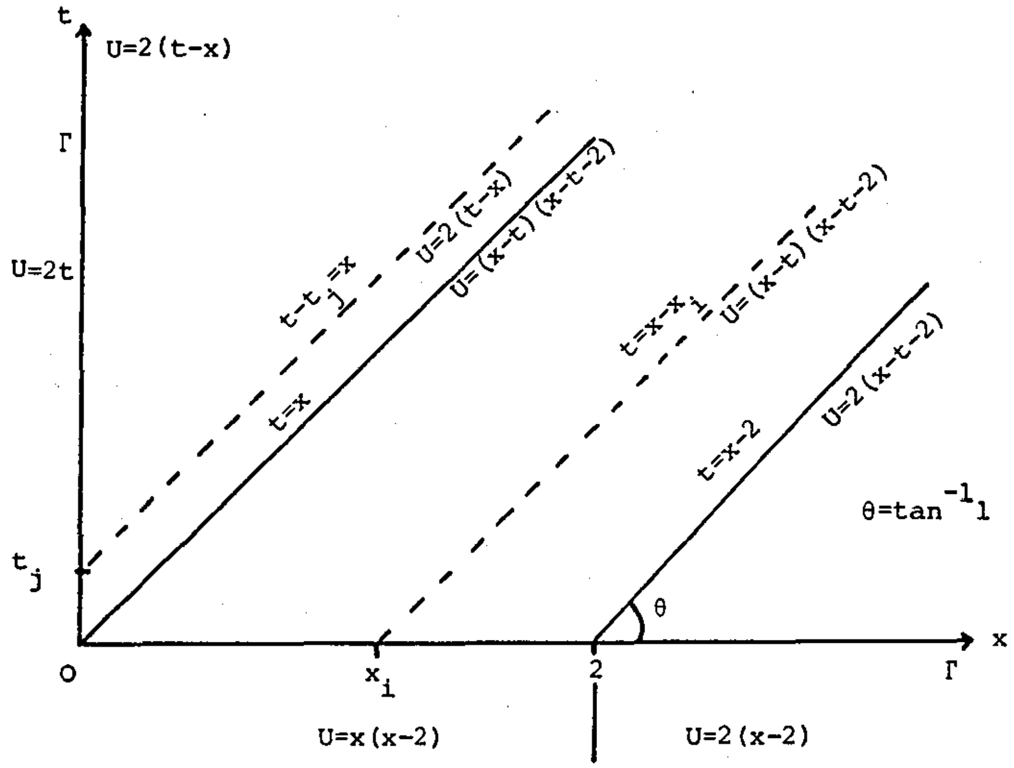


FIGURE 2.6.5

It deserves our attention here to observe that the known values of U on the initial line Γ are continuous at (0,0) and (2,0). Likewise, we also note that the derivatives $\frac{\partial U}{\partial x}$ and $\frac{\partial U}{\partial t}$ are continuous at the point (2,0). This question on the continuity of the initial values and derivatives has relevance to the discussion that is to follow in the subsequent sections of this chapter.

2.7 DISCONTINUITIES AND PROPAGATED DISCONTINUITIES

(a) *Discontinuous Initial Values*

An important point which we should take into account in dealing with hyperbolic problems is the possibility of *discontinuous initial values*. We will illustrate the significance of this and its effect through an example. We consider the problem of solving the equation,

$$2\frac{\partial U}{\partial x} + \frac{\partial U}{\partial y} = 1 \tag{2.7.1}$$

given that the data U are specified on the initial line Γ, y=0 as follows,

$$U(x,0) = \begin{cases} f(x); & 0<x<x_1, \\ \\ g(x); & x_1<x<1 \end{cases} \tag{2.7.2}$$

and it is assumed that U is double-valued at $x=x_1$ on Γ.

From the associated subsidiary equations

$$\frac{dx}{2} = \frac{dy}{1} = \frac{dU}{1} \tag{2.7.3}$$

the characteristic curve passing through any point x_i on Γ is,

$$y = \tfrac{1}{2}(x-x_i) \ , \ 0<x_i<1 \ .$$

By integrating along this characteristic using (2.7.3) again, we obtain the solution on this line as

$$U(x,y) = U_i + \tfrac{1}{2}(x-x_i) = U_i + y \ , \tag{2.7.4}$$

where $U_i=U(x_i,0)$, the initial data for U. Therefore, it is seen that if $0<x_j<x_1$ on Γ, then the equation of the characteristic extended from x_j is $y=\frac{1}{2}(x-x_j)$ and the corresponding solution along this characteristic is, from (2.7.4) and (2.7.2),

$$U(x,y) = U^- = f(x_j) + y \ . \tag{2.7.5}$$

By the same argument, if there is another point to the right of x_1, x_k

say then for $x_1<x_k<1$, the characteristic passing through it is $y=\frac{1}{2}(x-x_k)$ and the solution is

$$U(x,y) = U^+ = g(x_k) + y . \qquad (2.7.6)$$

The position of the characteristics and the solutions along them is shown in the following diagram.

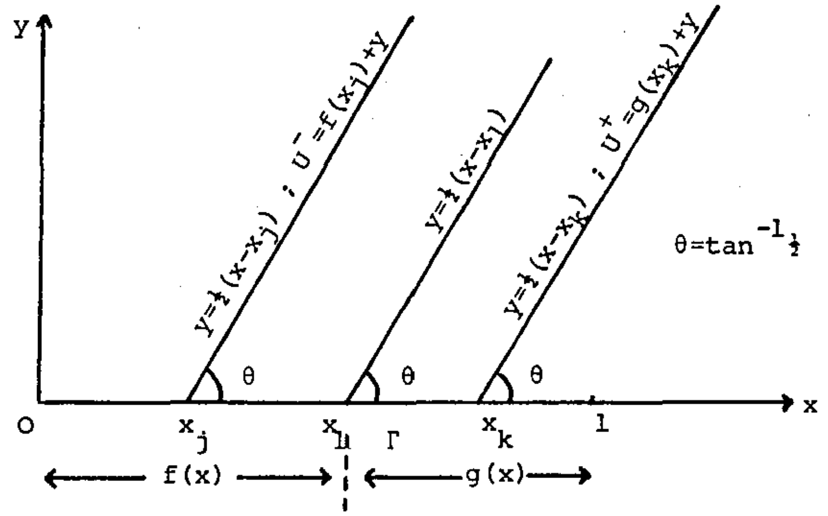


FIGURE 2.7.1

We deduce from (2.7.5) and (2.7.6) that for any particular value of y,

$$U^- - f(x_j) = U^+ - g(x_k)$$

or

$$U^- = U^+ + \{f(x_j)-g(x_k)\}.$$

Now as x_j tends to x_1 and x_k tends to x_1, $f(x_j)$ tends to $f(x_1)$ and $g(x_k)$ tends to $g(x_1)$. However U is double-valued at $x=x_1$ on Γ meaning that $U(x,0)$ is discontinuous at $x=x_1$ and $f(x_1)\neq g(x_1)$. This implies that U^- does not tend to U^+ on the characteristic $y=\frac{1}{2}(x-x_1)$. The double valued nature of the initial values perpetuates all along the characteristic $y=\frac{1}{2}(x-x_1)$. The values of the solution to the left of this characteristic is determined by $U(x,0)=f(x)$ and to the right by $g(x)$. In short, a discontinuity of the initial values leads to a discontinuity of the solution as well as along the specific characteristic.

(b) *Discontinuous Initial Derivatives*

We shall now examine whether a similar phenomenon of *discontinuity of $U(x,y)$ and/or the derivatives* $p=\frac{\partial U}{\partial x}$ and $q=\frac{\partial U}{\partial y}$ also persist along the characteristic curves in the solution domain given that these derivatives are discontinuous at some point on the initial curve Γ. For this purpose, we consider again equation (2.7.1),

$$2\frac{\partial U}{\partial x}+\frac{\partial U}{\partial y}=1\ .$$

The initial conditions on Γ, $y=0$ are now, however, changed to

$$U(x,0)=\begin{cases}x^2 & \text{for } 0<x<\frac{1}{2}\\ -x+\frac{1}{2} & \text{for } \frac{1}{2}<x<1\ .\end{cases}\tag{2.7.7}$$

As before, the equation of the characteristic passing through the point $x_1=\frac{1}{2}$ on Γ is $y=\frac{1}{2}(x-\frac{1}{2})$. By taking the points x_j to the left of x_1 and the points x_k to the right we find that on Γ,

$$p=\left.\frac{\partial U}{\partial x}\right|_{x=x_j}=2x_j=p^-\ \text{say}$$

and

$$p=\left.\frac{\partial U}{\partial x}\right|_{x=x_k}=-1=p^+\ .$$

At these points, we are also able to evaluate $q=\frac{\partial U}{\partial y}$ since from equation (2.7.1), $q=1-2p$. Hence the corresponding values of q at these points are,

$$q^-=1-2p^-=1-4x_j$$

and

$$q^+=1-2p^+=3.$$

By allowing x_j,x_k to tend to $x_1=\frac{1}{2}$ results in p^- tending to 1, $p^+=-1$, q^- tending to -1 and $q^+=3$. Clearly, we see that both p and q are double valued at $x=x_1=\frac{1}{2}$ which give the implication that the derivatives

$p=\frac{\partial U}{\partial x}$ and $q=\frac{\partial U}{\partial y}$ are discontinuous at the point $(\frac{1}{2},0)$ on the initial curve Γ, $y=0$.

By the same reasoning, it is also found that U is discontinuous at the point $(\frac{1}{2},0)$ on Γ since it is double valued there.

We shall now investigate whether these discontinuities are propagated into the integration field. By integrating along the characteristics drawn from the points x_j and x_k on the initial line segment Γ, we obtain

$$U(x,y) = x_j^2 + y , \qquad (2.7.8)$$

as the solution of the differential equation on the line $y=\frac{1}{2}(x-x_j)$ to the left of the characteristic $y=\frac{1}{2}(x-\frac{1}{2})$ and

$$U(x,y) = \tfrac{1}{2}(x-3x_k+1) , \qquad (2.7.9)$$

as the solution on the line $y=\frac{1}{2}(x-x_k)$ to the right of $y=\frac{1}{2}(x-\frac{1}{2})$. Following the same line of argument as before, again we observe that there are discontinuities in the solution U(x,y) and the derivatives $\frac{\partial U}{\partial x}$, $\frac{\partial U}{\partial y}$ along the characteristic $y=\frac{1}{2}(x-\frac{1}{2})$ in the solution domain.

It is worth noting here that there are cases in which, given that the initial derivatives are discontinuous at some point on the initial curve Γ, the values of U (the solution) would still be continuous at the point of discontinuity and along the characteristic extended from that point into the solution domain. The initial discontinuities in the partial derivatives would, however, remain to be propagated undiminished along the characteristic across the solution domain. To illustrate this, let us take, for example, the differential equation

$$\frac{\partial U}{\partial x} + \frac{\partial U}{\partial y} = 1, \quad y\geq 0 , \ -\infty<x<\infty \qquad (2.7.10)$$

with the initial conditions,

$$\begin{aligned} U(x,0) &= 0, \quad -\infty<x\leq 0 \\ U(x,0) &= x, \quad 0<x<\infty \end{aligned} \qquad (2.7.11)$$

on Γ, $y=0$.

For this simple example, we find that from the initial conditions (2.7.11) and the relation p+q=1 (from 2.7.10), the derivatives $p=\frac{\partial U}{\partial x}$ and $q=\frac{\partial U}{\partial y}$ are discontinuous at the point (0,0) on Γ. It is also immediately evident that the initial data U are continuous at the same point.

The associated subsidiary equations of the given differential equation (2.7.10) are,

$$\frac{dx}{1} = \frac{dy}{1} = \frac{dU}{1} \quad .$$

Hence, the equation of the characteristic passing through $(x_i,0)$ on Γ is $y=x-x_i$ and the solution along it $U-U_i=y=x-x_i$. Clearly, the characteristic becomes y=x when $x_i=0$. To the left of this characteristic, the solution along $y=x-x_j$ drawn from the point $(x_j,0)$ where $x_j \leq 0$ is $U^-=y$. Similarly, to the right we have the solution along $y=x-x_k$ as $U^+=x$.

As x_j and x_k tend to 0, we see that the solution along y=x is continuous. But there are discontinuities of the derivatives p and q along the characteristic across the solution domain. Since the initial derivatives are themselves discontinuous at (0,0) we conclude that there is a persistent propagation of the discontinuities of the derivatives from the point of discontinuity along the characteristic which is drawn from it into the solution domain.

To exhibit numerically how the discontinuities in the derivatives propagate across the characteristics in the solution domain, we consider the following initial-boundary value problem,

$$\frac{\partial U}{\partial x} + \frac{\partial U}{\partial y} = 1, \ 0 \leq x < \infty, \ y > 0 \ , \qquad (2.7.12)$$

with the boundary condition $U(0,y) = 0, \ 0<y<\infty$,

and the initial conditions $U(x,0) = 0$ for $0 \leq x \leq 3$ (2.7.13)

and $U(x,0) = x-3$ for $3 \leq x < \infty$.

The subsidiary equations of the above problem are

$$\frac{dx}{1} = \frac{dy}{1} = \frac{dU}{1} . \tag{2.7.14}$$

Therefore, the equation of the characteristic from the point $(0,y_j)$, $y_j>0$ is

$$y = x+y_j \tag{2.7.15}$$

and the equation of the characteristic passing through $(x_i,0)$ with $x_i \geq 0$ is

$$y = x-x_i . \tag{2.7.16}$$

Hence we have, for example, at (0,0), the characteristic

$$y = x$$

and at (3,0), the characteristic

$$y = x-3.$$

By using (2.7.14) and (2.7.13), the solutions across these characteristics are worked out to be as follows:

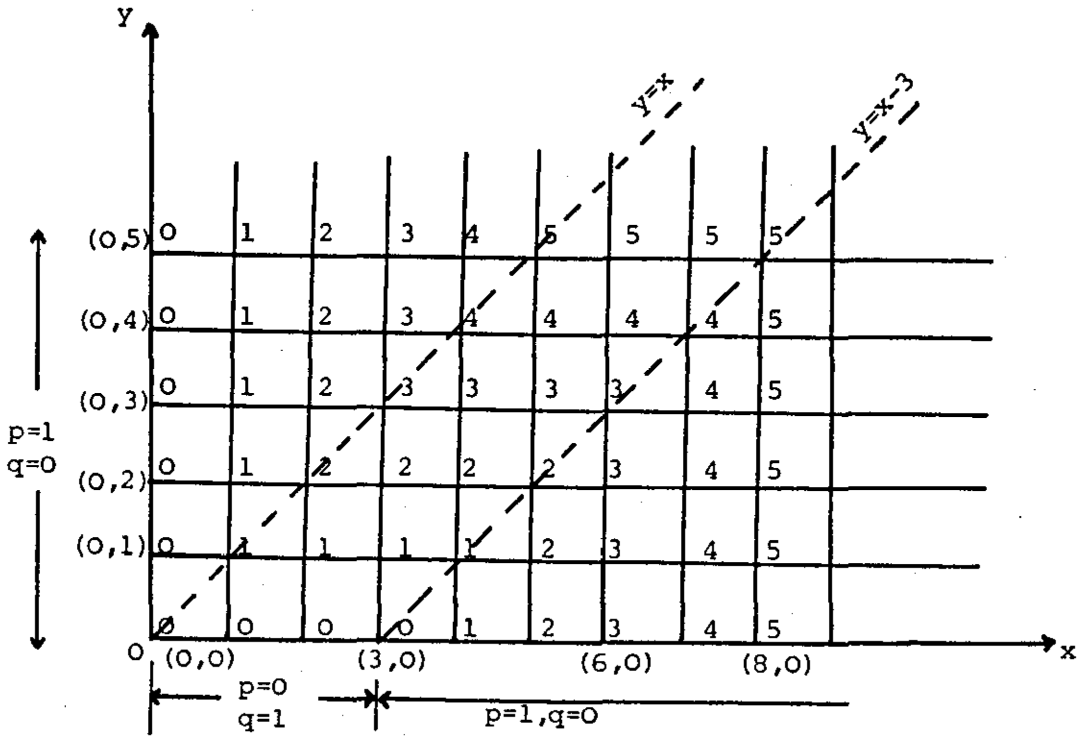
(i) $U=x$ on the characteristics $y=x+y_j$ passing through $(0,y_j)$, $y_j>0$;

(ii) $U=y$ on the characteristics $y=x-x_i$ from $(x_i,0)$ with $x_i \in [0,3]$ and

(iii) $U=x-3$ on the characteristics $y=x-x_i$ from $(x_i,0)$, $x_i \geq 3$.

We can easily infer from (2.7.13) that U is continuous at (0,0) and (3,0). However, at points $(0,y)$, $y>0$, $\frac{\partial U}{\partial y}=0$ and $\frac{\partial U}{\partial x}=1-\frac{\partial U}{\partial y}=1$ (from equation (2.7.12)) whilst at points $(x,0)$, $0 \leq x \leq 3$, $\frac{\partial U}{\partial x}=0$ and $\frac{\partial U}{\partial y}=1-\frac{\partial U}{\partial x}=1$. Also, at points $(x,0)$, $x \geq 3$ we find $\frac{\partial U}{\partial x}=1$ and $\frac{\partial U}{\partial y}=0$. Therefore, $\frac{\partial U}{\partial x}$ and $\frac{\partial U}{\partial y}$ are discontinuous at (0,0) and (3,0). Furthermore, we note that the solution U_1^- to the left of the characteristic $y=x$ is $U_1^-=x$ and the solution U_1^+ to the right is $U_1^+=y$. Thus $U_1^+=U_1^-$ along the characteristic $y=x$ implying that U is continuous along this characteristic. Similarly we obtain the solution to the left of $y=x-3$ as $U_2^-=y$ and $U_2^+=x-3$ to the right. Therefore $U_2^-=U_2^+$ across the characteristic y=x-3 and U is also continuous along this characteristic in the solution domain. On the other hand, it can be easily verified that the derivatives of these solutions are not continuous along each of these characteristics.

The manner in which this *propagation of discontinuity* in the derivatives is reflected by the analytical solutions for x=0(1)8 and y=0(1)5 in the integration field is shown in Table 2.7.1 below. It is seen that the characteristics are the *natural boundaries* between which lies the region where the discontinuities are propagated undiminished across it.



(0,5) 0	1	2	3	4	5	5	5	5
(0,4) 0	1	2	3	4	4	4	4	5
(0,3) 0	1	2	3	3	3	3	4	5
(0,2) 0	1	2	2	2	2	3	4	5
(0,1) 0	1	1	1	1	2	3	4	5
0	0	0	0	1	2	3	4	5

TABLE 2.7.1

2.8 THE METHOD OF CHARACTERISTICS FOR HYPERBOLIC EQUATIONS

In the preceding sections, we discuss some analytical methods to solve a given hyperbolic equation and arrive at exact solutions. In many cases, the analytical solutions cannot be found and even if they do exist their evaluation is often a laborious task. Thus, approximation methods which are normally numerical in character are the only means of solution. We would expect the solutions to be obtained, for example, by a *process of numerical integration*. For this purpose we consider again the quasi-linear first order equations (2.4.1) or (2.4.2)

$$a \frac{\partial U}{\partial x} + b \frac{\partial U}{\partial y} = c$$

or $$ap + bq = c.$$

Our aim is then to find the solutions of (2.4.2) by integrating in one direction only at each point of the x-y plane and that we are not concerned with derivatives in any other direction. Hence along this direction, the integration of (2.4.2) will transform to the integration of an ordinary differential equation. If P(x,y) is a point on the curve C in the x-y plane then along the direction of the tangent at P to the curve C we have,

$$dU = \frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy$$

i.e. $dU=pdx+qdy$ (2.8.1)

where $\frac{dy}{dx}$ is the slope of C at P.

From the differential equation (2.4.2) we can eliminate p to produce

$$adU - cdx + q(bdx-ady) = 0 . \qquad (2.8.2)$$

Obviously this equation is independent of p since a,b and c are functions of x,y and U only. We can also eliminate the effect of q by choosing the curve C so that its slope $\frac{dy}{dx}$ satisfies

$$bdx = ady . \qquad (2.8.3)$$

Therefore, equations (2.8.2) and (2.8.3) give us,

$$\frac{dx}{a} = \frac{dy}{b} = \frac{dU}{c} \quad ,$$

the original subsidiary equations. Our next step is to perform approximate numerical integration and at the appropriate points determine simultaneously the characteristics and the solution. For this purpose, we denote as usual Γ as the non-characteristic initial curve along which values of U are specified and through a known point $Q(x_0,y_0)$ on Γ passes the characteristic C. Let the adjacent point on C be P(x,y) such that $x-x_0$ is small as shown in Fig. 2.8.1 below.

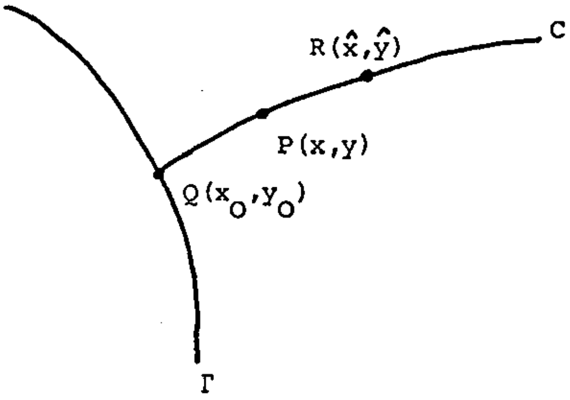


FIGURE 2.8.1

By assuming that x is given and starting from $Q(x_0,y_0)$, we attempt to find simultaneously improved values or approximations of y and U at the point P. This necessitates us to employ differences using (2.8.3) for our first approximation and our second and subsequent approximations are effected by some iterative means. Let $y^{(1)}, u^{(1)}$ be our first approximations and $y^{(2)}, u^{(2)}, \ldots$ for our second and subsequent approximations or iterates. Our process of approximation will now be described as follows:

First Approximation.

Using equations (2.8.2) and (2.8.3) we have,

$$a_0(y^{(1)}-y_0) = b_0(x-x_0) \text{ giving } y^{(1)} \qquad (2.8.4)$$

and

$$a_0(u^{(1)}-u_0) = c_0(x-x_0) \text{ giving } u^{(1)} \ . \qquad (2.8.5)$$

Second and Subsequent Approximations.

As the coefficients a,b and c are known, we now take their mean values over the arc PQ. Again using equations (2.8.2) and (2.8.3) we obtain,

$$\tfrac{1}{2}(a_0+a^{(1)})(y^{(2)}-y_0) = \tfrac{1}{2}(b_0+b^{(1)})(x-x_0) \text{ giving } y^{(2)} \qquad (2.8.6)$$

and

$$\tfrac{1}{2}(a_0+a^{(1)})(u^{(2)}-u_0) = \tfrac{1}{2}(c_0+c^{(1)})(x-x_0) \text{ giving } u^{(2)} \ . \qquad (2.8.7)$$

This procedure is repeated until successive iterates, say $u^{(n)}$ and $u^{(n-1)}$ fulfils a certain termination criterion, i.e.

$$|u^{(n)}-u^{(n-1)}| \leq \varepsilon$$

where ε is a preset tolerance.

The values of y and u at other grid points on the characteristic C can be derived in a similar way. Starting with the point P(x,y) the above approximation process is reinitiated and repeated to give us y and u at the next point $R(\hat{x},\hat{y})$ and so on. The approximation of course improves as the interval in x gets smaller which is desirable in any case for the minimization of the truncation error in our finite difference approximations.

The numerical application of the method of characteristics to the solution of a given first-order hyperbolic differential equation is illustrated in the following example.

Consider the problem,

$$x^2U \frac{\partial U}{\partial x} + e^{-y} \frac{\partial U}{\partial y} = - U^2 \qquad (2.8.8)$$

with the initial condition U=1 on $\Gamma = \{(x,y):0<x<\infty,\ y=0\}$.

If we compare with the equation (2.4.1) we obtain the coefficients,

$$a = x^2U,\ b = e^{-y} \text{ and } c = -U^2 . \qquad (2.8.9)$$

The subsidiary equations of (2.8.8) are

$$\frac{dx}{x^2U} = \frac{dy}{e^{-y}} = \frac{dU}{U^2} ,$$

from which we obtain the equation of the family of characteristics C passing through the point $(x_i,0)$ on Γ as

$$y = \frac{1}{x_i} - \frac{1}{x} \qquad (2.8.10)$$

and the solution along it

$$U = e^{-y} . \qquad (2.8.11)$$

We are interested to find approximations to the solution and to the value of y at the point $P(x,y)=P(1.1,y)$, $y>0$ on C which passes through $Q(x_0,y_0)=Q(1,0)$ on Γ.

First Approximation for $y^{(1)}$ and $u^{(1)}$.

From equation (2.8.9), we have at the point $Q(x_0,y_0)=Q(1,0)$

$$a_0 = 1,\ \ b_0 = 1 \text{ and } c_0 = -1.$$

Therefore with x=1.1 we obtain from equations (2.8.4) and (2.8.5),

$$y^{(1)} = 0.1$$

and $$u^{(1)} = 0.9$$

Second Approximation for $y^{(2)}$ and $u^{(2)}$.

By using the results in our first approximation, equations (2.8.9) yield the coefficients at the point $Q(x,y^{(1)})=Q(1.1,0.1)$ as

$$a^{(1)} = x^2u^{(1)} = 1.089,\ b^{(1)} = e^{-y^{(1)}} = 0.904837 \text{ and}$$
$$c^{(1)} = -(u^{(1)})^2 = -0.81.$$

Hence with x=1.1, equations (2.8.6) and (2.8.7) give us

$$y^{(2)} = 0.091184 \text{ and } u^{(2)} = 0.913356.$$

Of course, these approximations can be further improved by successive iterations. As a comparison, the analytical values of y and U from (2.8.10) and (2.8.11) worked out to be

$$y = 0.090909 \quad \text{and } U = 0.9131007.$$

The method of characteristics can be readily extended to second-order quasi-linear equations and systems of first-order equations. We recall from Section 1.2 of Chapter One that the slopes of the characteristic directions associated with the second-order equation,

$$a\frac{\partial^2 U}{\partial x^2} + b\frac{\partial^2 U}{\partial x \partial y} + c\frac{\partial^2 U}{\partial y^2} + e = 0 , \tag{2.8.12}$$

are given by the roots of the quadratic equation,

$$a\left(\frac{dy}{dx}\right)^2 - b\left(\frac{dy}{dx}\right) + c = 0 , \tag{2.8.13}$$

and along these characteristic directions the differentials dp and dq are related by the equation,

$$a\frac{dy}{dx}\frac{dp}{dx} + c\frac{dq}{dx} + e\frac{dy}{dx} = 0 . \tag{2.8.14}$$

By assuming that equation (2.8.12) is hyperbolic, the roots of equation (2.8.13) will be real and distinct. Let these be

$$\frac{dy}{dx} = f \quad \text{and} \quad \frac{dy}{dx} = g . \tag{2.8.15}$$

Let Γ be a non-characteristic curve along which initial values for U,p and q are known. Let P and Q be points on Γ that are close together and let the f-characteristic through P intersect the g-characteristic through Q at the point $R(x_R,y_R)$ as in Figure 2.8.2.

As a first approximation we may regard the arcs PR and QR as straight lines of slopes f_P and g_Q respectively. Then equation (2.8.15) can be approximated by

$$y_R - y_P = f_P(x_R - x_P) \tag{2.8.16}$$

and

$$y_R - y_Q = g_Q(x_R - x_Q) \tag{2.8.17}$$

giving two equations for the two unknowns x_R, y_R.

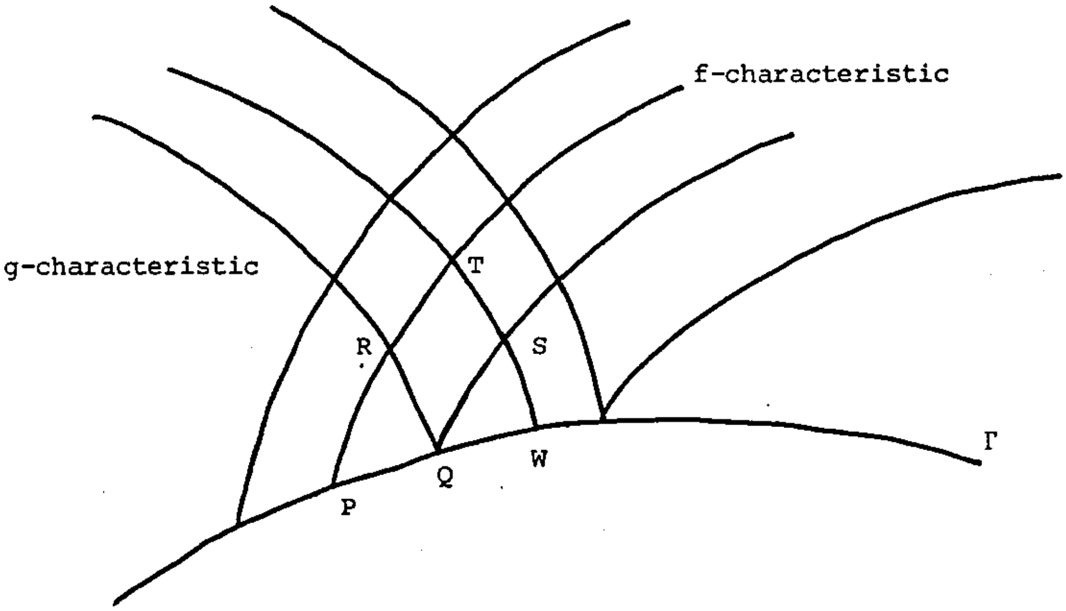


FIGURE 2.8.2

From equation (2.8.14) the differential relationships along the characteristics are

$$afdp + cdq + edy = 0 \tag{2.8.18}$$

and

$$agdp + cdq + edy = 0 \ . \tag{2.8.19}$$

The first one can be approximated along PR by the equation,

$$a_P f_P (p_R - p_P) + c_P(q_R - q_P) + e_P(y_R - y_P) = 0 \tag{2.8.20}$$

and the second along QR by the equation,

$$a_Q g_Q (p_R - p_Q) + c_Q(q_R - q_Q) + e_Q(y_R - y_Q) = 0 \ . \tag{2.8.21}$$

These are two equations for the two unknowns p_R, q_R as soon as x_R, y_R have been calculated from (2.8.16) and (2.8.17). The value of U at R can then be obtained from

$$dU = \frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy = pdx + qdy$$

by replacing the values of p and q along PR by their average values and approximating the last equation by

$$u_R - u_P = \tfrac{1}{2}(p_P+p_R)(x_R-x_P)+\tfrac{1}{2}(q_P+q_R)(y_R-y_P) \ . \tag{2.8.22}$$

This first approximation for u_R can now be improved by replacing the pivotal values of the various coefficients by their average values. Equations (2.8.16) and (2.8.17) for the improved values of x_R and y_R then become

$$y_R - y_P = \tfrac{1}{2}(f_P+f_R)(x_R-x_P) \tag{2.8.23}$$

and

$$y_R - y_Q = \tfrac{1}{2}(g_Q+g_R)(x_R-x_Q) \tag{2.8.24}$$

and equations (2.8.20), (2.8.21) for improved values of p_R, q_R become

$$\begin{aligned}&\tfrac{1}{2}(a_P+a_R)\tfrac{1}{2}(f_P+f_R)(p_R-p_P)+\tfrac{1}{2}(c_P+c_R)(q_R-q_P)\\&+\tfrac{1}{2}(e_P+e_R)(y_R-y_P) = 0\end{aligned} \tag{2.8.25}$$

and

$$\begin{aligned}&\tfrac{1}{2}(a_Q+a_R)\tfrac{1}{2}(g_Q+g_R)(p_R-p_Q)+\tfrac{1}{2}(c_Q+c_R)(q_R-q_Q)\\&+\tfrac{1}{2}(e_Q+e_R)(y_R-y_Q) = 0 \ .\end{aligned} \tag{2.8.26}$$

An improved value for u_R can then be found from equation (2.8.22). Repetition of this last cycle of operations will eventually yield u_R to the accuracy warranted by these difference approximations. Provided Q is close to P the number of iterations will usually be small. In this way, we can calculate solution values at the grid points R and S (see

Fig. 2.8.2), and then proceed to the grid point T and so on.

As an example, we consider the quasi-linear equation,

$$\frac{\partial^2 U}{\partial x^2} - U^2 \frac{\partial^2 U}{\partial y^2} = 0 .$$

We shall use the method of characteristics to derive a solution at the first characteristic grid point between x=0.2 and 0.3,y>0 with U satisfying the conditions,

$$U = 0.2 + 5x^2 \qquad \text{and} \qquad \frac{\partial U}{\partial y} = 3x,$$

along the initial line y=0, for $0 \le x \le 1$.

Since U is given as a continuous function of x along Ox (the x-axis) the initial value of $p = \frac{\partial U}{\partial x}$ is 10x. The slopes of the characteristics are the roots of the equation $m^2 - U^2 = 0$ where $m = \frac{dy}{dx}$. Hence,

$$f = U = -g .$$

In this example the characteristics depend on the solution and so the network of characteristics can be built up only as the solution has been worked out.

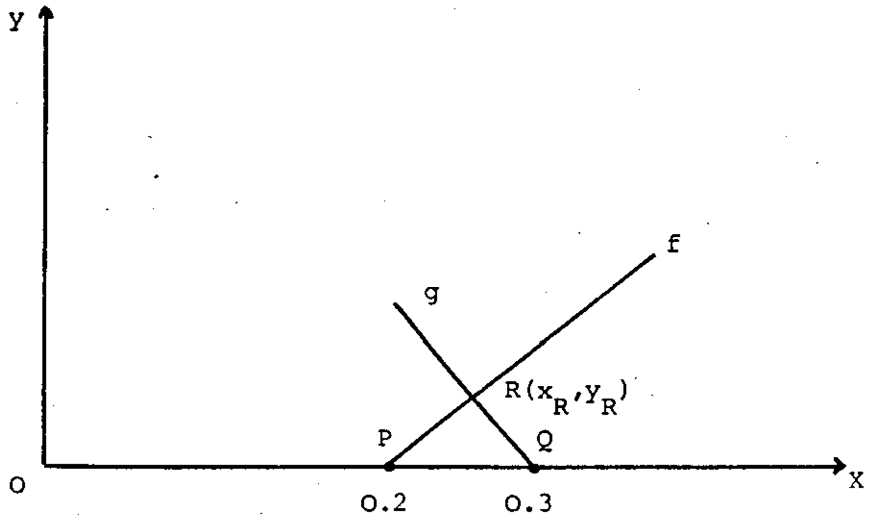


FIGURE 2.8.3

Initially,

$$u = 0.2 + 5x^2 = f$$

$$= -g$$

$$p = 10x \text{ and } q = 3x.$$

Also a=1, b=e=0, $c=-u^2$. Hence from Fig. 2.8.3, we have

$$f_P = 0.4,\ g_Q = -0.65,\ p_P = 2.0,\ p_Q = 3.0,\ u_P = 0.4,$$

$$u_Q = 0.65,\ g_P = 0.6,\ g_Q = 0.9,\ c_P = -0.16,\ c_Q = -0.4225.$$

From equations (2.8.16) and (2.8.17), we find that,

$$y_R = 0.4(x_R-0.2)$$

and

$$y_R = -0.65(x_R-0.3)$$

giving as a first approximation

$$x_R = 0.26190,\ y_R = 0.024762$$

to five significant figures.

The difference relationships along the characteristics are, from equations (2.8.20) and (2.8.21),

$$0.4(p_R-2.0)-0.16(q_R-0.6) = 0$$

and

$$-0.65(p_R-3.0)-0.4225(q_R-0.9) = 0 .$$

Their solution is,

$$p_R = 2.45524;\quad q_R = 1.73810.$$

By virtue of equation (2.8.22), we obtain,

$$u_R = 0.4+\tfrac{1}{2}(2.0+2.45524)(0.0619)+\tfrac{1}{2}(1.73810+0.6)(0.024762)$$

$$= 0.56684.$$

For the second approximation, we have

$$f_R = -g_R = u_R = 0.56684;\ c_R = -u_R^2 = -0.32131.$$

From equations (2.8.23) and (2.8.24), more accurate values for x_R, y_R are given by

$$y_R = \tfrac{1}{2}(0.4+0.56684)(x_R-0.2)$$

and $y_R = -\frac{1}{2}(0.65+0.56684)(x_R-0.3)$

from which

$$x_R = 0.25572 \text{ and } y_R = 0.026938.$$

By using equations (2.8.25) and (2.8.26) leads to

$$\tfrac{1}{2}(0.4+0.56684)(p_R-2.0)-\tfrac{1}{2}(0.16+0.32131)(q_R-0.6) = 0 ,$$

$$-\tfrac{1}{2}(0.65+0.56684)(p_R-3.0)-\tfrac{1}{2}(0.4225+0.32131)(q_R-0.9) = 0 .$$

These equations give the improved values,

$$p_R = 2.53117 \text{ and } q_R = 1.66700.$$

Hence, the second approximation to u_R, by means of equation (2.8.22) is $u_R = 0.4+\frac{1}{2}\{(2+2.53117)(0.05572)+(0.6+1.6670)(0.026938)\}$

$= 0.55677.$

We find that, for the next solution,

$$x_R = 0.25578, \quad y_R = 0.02668$$

$$p_R = 2.52876, \quad q_R = 1.67637$$

and $u_R = 0.55667$. Since to four decimal places,

$$u_R^{(1)} = 0.5668,\ u_R^{(2)} = 0.5568 \text{ and } u_R^{(3)} = 0.5567$$

it is obvious that the solution of the difference equations for u_R is 0.5567, to this degree of accuracy. A fourth iteration does, in fact, give u_R=0.55666 to five decimal places.

Finally we mention the case of simultaneous first-order equations of the form,

$$a\frac{\partial U}{\partial x} + b\frac{\partial U}{\partial y} + e\frac{\partial V}{\partial x} + f\frac{\partial V}{\partial y} = g \qquad (2.8.27a)$$

$$A\frac{\partial U}{\partial x} + B\frac{\partial U}{\partial y} + E\frac{\partial V}{\partial x} + F\frac{\partial V}{\partial y} = G \qquad (2.8.27b)$$

where the coefficients are functions of x,y and U but not of the derivatives, and for which U and V are given on the initial line. Following our standard procedure we look for the characteristics by examining the linear equations given by (2.8.27) and,

$$dU = \frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy \qquad (2.8.28a)$$

$$dV = \frac{\partial V}{\partial x} dx + \frac{\partial V}{\partial y} dy \qquad (2.8.28b)$$

for the determination of the four derivatives. The solutions are unique unless the determinant,

$$\begin{vmatrix} dx & dy & 0 & 0 \\ 0 & 0 & dx & dy \\ a & b & e & f \\ A & B & E & F \end{vmatrix} = 0 \, , \qquad (2.8.29)$$

the equation which defines the characteristics. On expanding this determinant, gives,

$$(dy)^2(eA-aE)+(dx)(dy)(bE-eB+aF-fA)+(dx)^2(fB-bF) = 0 \qquad (2.8.30)$$

and the equations form a hyperbolic system if the roots are real. We also have the relation,

$$\begin{vmatrix} dU & dx & dy & 0 \\ dV & 0 & 0 & dx \\ g & a & b & e \\ G & A & B & E \end{vmatrix} = 0 \, , \qquad (2.8.31)$$

between U,V,x and y along the characteristics, which we can use in the hyperbolic case to obtain a solution by integration along the characteristics. Two characteristics pass through every point and the computation is almost identical with that of the second-order case.

2.9 FINITE DIFFERENCE METHODS FOR HYPERBOLIC EQUATIONS

A question that one would naturally ask is whether we could develop discrete approximations to the solution of a given hyperbolic differential equation using finite differences which are suitable to the higher speed computer. It is known that finite difference methods have been widely used with great success in parabolic and elliptic problems. However, use of these methods in hyperbolic equations are in a sense quite restricted.

In previous sections of this chapter we have seen the segmenting of the solution domain by the characteristics. In other words, these characteristics play the role of natural boundaries in the solution domain. If, for example, there are more than one family of characteristics then the partition of the solution domain is influenced by the boundary conditions of these characteristics. We have also noted how the existence of discontinuities in the initial values and initial derivatives leads to discontinuities across the characteristic along its entire length from the point of discontinuity.

Finite difference methods, on the other hand, do not accommodate these possibilities. The central feature of these methods is the replacement or approximation of derivatives at a grid point of the finite difference network, say the rectangular (square) grid by some difference quotients over a small interval. This process of approximation is extended to all other grid points in the network over the area of integration of interest without taking into consideration such features as the role of the characteristics. Further difficulties arise if these characteristics are not straight lines or are dependent on the solution.

As an illustration we refer to the following problem,

$$U \frac{\partial U}{\partial x} + y \frac{\partial U}{\partial y} = 1$$

with $U=0$ on $\Gamma = \{(x,y): 0<x<a,\ y=1\}$.

The subsidiary equations are

$$\frac{dx}{U} = \frac{dy}{y} = \frac{dU}{1}$$

from which we obtain the solution,

$$U = \sqrt{2(x-x_i)}$$

along the charactersitic $y=e^U=e^{\sqrt{2(x-x_i)}}$ that passes through the point $(x_i,1)$ on Γ, $0<x_i<a$. The graph of these characteristics is shown in the following figure.

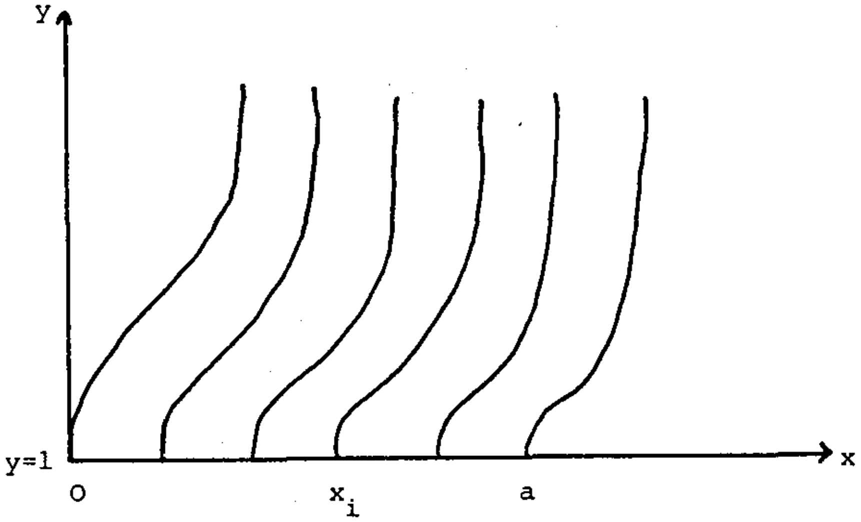


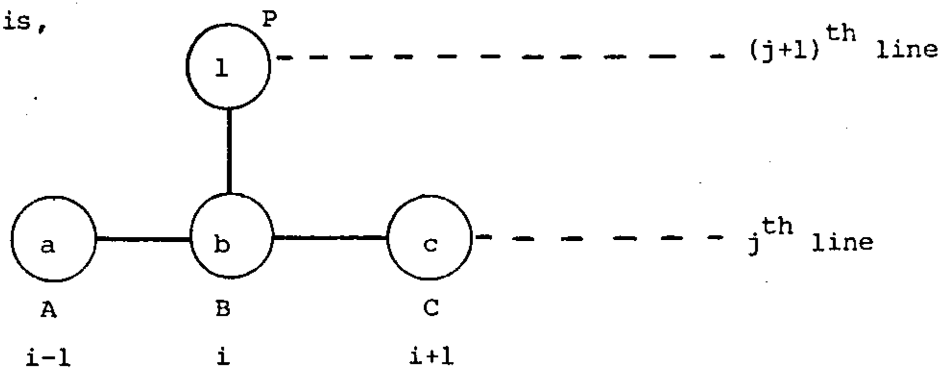
FIGURE 2.9.1

We note that the characteristics are dependent on the solution and they are not straight lines. Hence we see that if a rectangular grid is used in our finite difference method then the above characteristics do not pass through many of the grid points in which case the method may not be truly representative of the given problem. Should there be any propagated discontinuities along the characteristic then this would lead to further computational problems and inaccuracies.

With these inadequacies on finite difference approximations it is to be anticipated that the method of characteristics provide more accurate results. This is because solutions are progressed from the initial point along a characteristic that does not intersect the characteristic drawn from a point of discontinuity. Its programming for numerical evaluation on the digital computers, however, may prove to be difficult and bothersome, more so when it involves a set of simultaneous first-order equations. Therefore, if the differential equations that we are dealing with are of no great complexity and are known to possess well-behaved solutions, finite difference methods may still prove viable as a method of discrete approximation. We are then interested in the questions of accuracy, convergence and stability. These considerations may be directly linked with the limitations imposed by the characteristics themselves. A case in point is the *explicit* finite difference schemes which are of the form,

$$u_{i,j+1} = au_{i-1,j} + bu_{ij} + cu_{i+1,j} ,$$

The computational molecule of the above difference scheme in the rectangular grid is,



The spacing used by the grid points in the x and y direction is Δx and Δy and using known values of u at the preceding y level at the points A,B and C enables us to determine the solution at the present y level at the point P. Let Γ denote the initial line segment on which values of u are known. If the points A,B and C are on Γ then the solution u_P at P is determined by the values of u at these points. Hence u_P will change if changes are made on u_A, u_β and u_C. However the solution U_P of the differential equation at P is not affected by these changes since it lies on the characteristic drawn from the point D on Γ as depicted in the following figure:

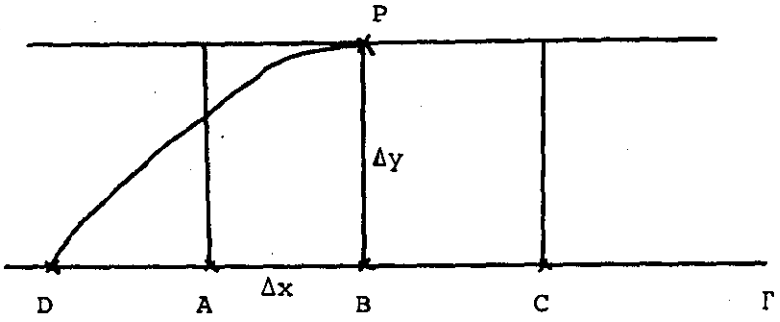


FIGURE 2.9.2

Hence *the finite difference solution* u_P will not converge to *the solution of the differential equation* U_P as $\Delta x, \Delta y$ tend to 0. Convergence is only assured if D lies between A and C. This is known as *the C.F.L. (Courant-Friedrichs-Lewy) condition* for convergence of the solution of difference equation to the solution of differential equation. Since we have a family of characteristics then this condition applies to all other points in the rectangular grid. In particular, if the characteristic PD is a straight line then for convergence, the slope of PD must be greater than or equal to the slope of PA. In some finite difference schemes, this also provides the stability condition as we march our solutions across its solution domain.

We shall now attempt to derive a *generalised two-time level finite difference approximation* to the *convection equation* in one space dimension of the form,

$$-a\frac{\partial U}{\partial x} = \frac{\partial U}{\partial t} \tag{2.9.1}$$

where a>0 and is constant. We shall also examine the stability and truncation errors of these approximations.

The space derivative $\frac{\partial U}{\partial x}$ may be approximated by

$$\frac{\partial U}{\partial x} \approx \frac{(u_{i+\frac{1}{2}}-u_{i-\frac{1}{2}})}{\Delta x} \quad . \tag{2.9.2}$$

$u_{i+\frac{1}{2}}$ may be chosen in a number of ways. The simplest option is the arithmetic averaging given by

$$u_{i+\frac{1}{2}} = (u_i+u_{i+1})/2 \quad . \tag{2.9.3}$$

Another option which is more commonly used is

$$u_{i+\frac{1}{2}} = u_i \quad . \tag{2.9.4}$$

A "distance-weighting" parameter w can therefore be used to characterise these options as follows:

backward-in-distance weighting: $w=1$,

centred-in-distance weighting: $w=\frac{1}{2}$,

forward-in-distance weighting: $w=0$,

and clearly equations (2.9.3) and (2.9.4) can be combined into the general equation,

$$u_{i+\frac{1}{2}} = wu_i + (1-w)u_{i+1} \quad . \tag{2.9.5}$$

Similarly, we have

$$u_{i-\frac{1}{2}} = wu_{i-1} + (1-w)u_i \ . \qquad (2.9.6)$$

We can use a similar parametric approach to the differentials, for example, the *explicit* , *implicit* and *centred-in-time* difference equations, wherein the spatial derivatives are evaluated, respectively, at t_j, t_{j+1} or $t_{j+\frac{1}{2}}$. If we use θ as the "time-weighting" parameter and prescribing the values,

implicit:
(backward-in-time) $\theta=1$,

centred-in-time: $\theta=\frac{1}{2}$,

explicit:
(forward-in-time) $\theta=0$,

then equation (2.9.1) can be replaced by the difference analogue,

$$-\frac{a}{\Delta x}[\theta(u_{i+\frac{1}{2},j+1}-u_{i-\frac{1}{2},j+1})+(1-\theta)(u_{i+\frac{1}{2},j}-u_{i-\frac{1}{2},j})] = \frac{u_{i,j+1}-u_{ij}}{\Delta t} \ , \qquad (2.9.7)$$

where we have employed the usual forward difference approximation for $\frac{\partial U}{\partial t}$.

By substituting equations (2.9.5) and (2.9.6) into equation (2.9.7), we arrive at the final, generalised difference analogue for the hyperbolic equation (2.9.1) as:

$$-\lambda[\theta\{(1-w)u_{i+1,j+1}+(2w-1)u_{i,j+1}-wu_{i-1,j+1}\}+(1-\theta)\{(1-w)u_{i+1,j}+ (2w-1)u_{i,j}-wu_{i-1,j}\}] = u_{i,j+1}-u_{ij} \ , \qquad (2.9.8)$$

where $\lambda = \frac{a\Delta t}{\Delta x}$ is the *mesh ratio*. Most of the well-known standard methods may be obtained from formula (2.9.8). For example, by putting $\theta=0$ and $w=(1+\lambda)/2\leq 1$ gives us,

$$u_{i,j+1} = \tfrac{1}{2}\lambda(1+\lambda)u_{i-1,j}+(1-\lambda^2)u_{ij} -\tfrac{1}{2}\lambda(1-\lambda)u_{i+1,j} \qquad (2.9.9)$$

which is called the *Lax-Wendroff explicit formula*. This is probably the most well known method for first-order hyperbolic equations and as we shall see later is second-order accurate. For $\theta=\frac{1}{2}$ and $w=\frac{1}{2}$, we get the implicit centred-in-distance and centred-in-time formula,

$$\frac{\lambda}{4}u_{i-1,j+1}-u_{i,j+1}-\frac{\lambda}{4}u_{i+1,j+1} = -\frac{\lambda}{4}u_{i-1,j}-u_{i,j}+\frac{\lambda}{4}u_{i+1,j} \qquad (2.9.10)$$

which is also second-order accurate in both space and time. Equation (2.9.10) is of the *Crank-Nicolson type* that is frequently encountered in *parabolic* problems and can be expressed as,

$$[1+\frac{1}{4}(\Delta_x+\nabla_x)]u_{i,j+1} = [1-\frac{1}{4}\lambda(\Delta_x+\nabla_x)]u_{ij} \qquad (2.9.11)$$

or
$$-\frac{a}{4\Delta x}(\Delta_x+\nabla_x)(u_{i,j+1}+u_{ij}) = \frac{u_{i,j+1}-u_{ij}}{\Delta t} \qquad (2.9.12)$$

where $\Delta_x u_{i,j} = u_{i+1,j}-u_{ij}$

and $\nabla_x u_{i,j} = u_{i,j}-u_{i-1,j}$.

Other schemes of interest are displayed in Table 2.9.1 below together with their computational molecules and stability conditions and truncation errors, both of which will be derived in the next two sections.

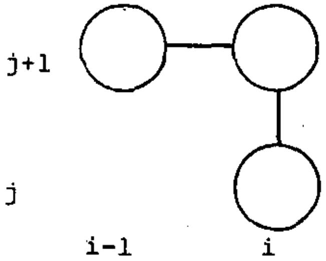
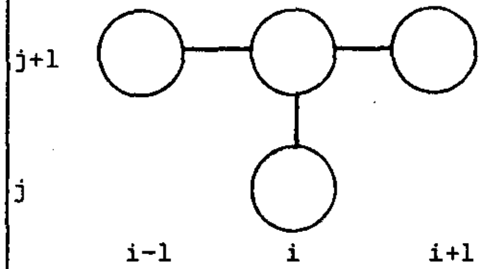
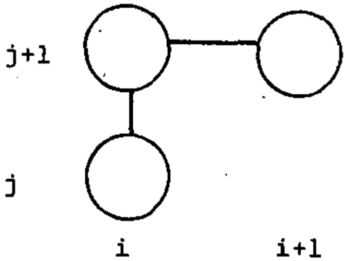
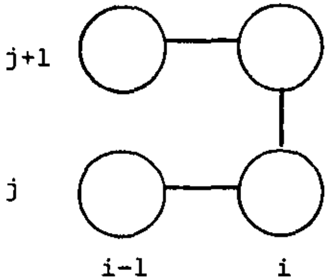
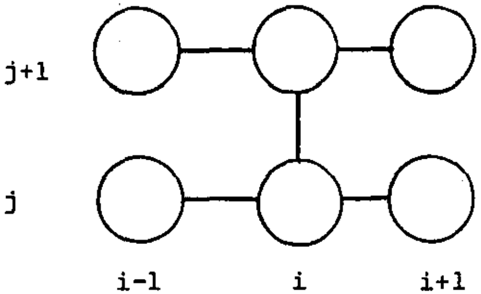
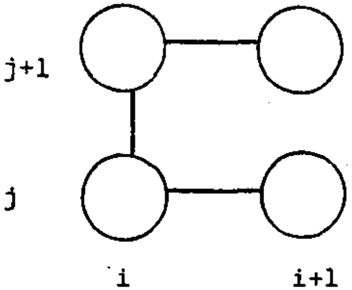
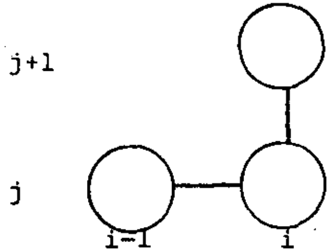
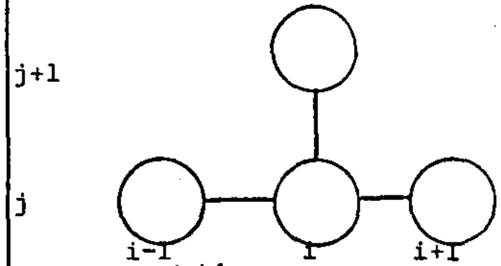
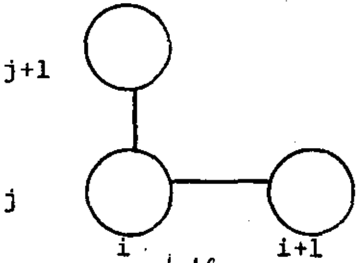
w θ	Backward-in-distance w=1	Centred-in-distance w=½	Forward-in-distance w=O
Backward-in-time θ=1 (implicit)	 Always stable. T=O(Δx)+O(Δt)	 Always stable. $T = O([\Delta x]^2)+O(\Delta t)$	 Stable if λ≥1. T = O(Δx)+O(Δt)
Centred-in-time θ=½	 Always stable $T = O(\Delta x)+O([\Delta t]^2)$	 Always (neutrally) stable. $T = O([\Delta x]^2+O([\Delta t]^2)$	 Always unstable. $T = O(\Delta x)+O([\Delta t]^2)$
Forward-in-time θ=O (explicit)	 Stable if λ≤1. T = O(Δx)+O(Δt)	 Always unstable. $T = O([\Delta x]^2)+O(\Delta t)$	 Always unstable. T = O(Δx)+O(Δt)

TABLE 2.9.1

2.10 STABILITY ANALYSIS OF THE GENERALISED FINITE DIFFERENCE METHODS

An analysis using *the von Neumann criterion* will be carried out to investigate the stability of (2.9.8). We shall assume that if u_{ij} is a solution to the difference equation at the point (x_i,t_j) then its perturbation $u_{ij}+\varepsilon_{ij}$ also satisfies the difference equation and we examine the possible growth of ε_{ij}. Specifically, if u_{ij} satisfies equation (2.9.8) then,

$$-\lambda[\theta\{(1-w)(u_{i+1,j+1}+\varepsilon_{i+1,j+1})+(2w-1)(u_{i,j+1}+\varepsilon_{i,j+1})-w(u_{i-1,j+1}+\varepsilon_{i-1,j+1})\}+(1-\theta)\{(1-w)(u_{i+1,j}+\varepsilon_{i+1,j})+(2w-1)(u_{ij}+\varepsilon_{ij})-w(u_{i-1,j}+\varepsilon_{i-1,j})\}]$$
$$= (u_{i,j+1}+\varepsilon_{i,j+1})-(u_{ij}+\varepsilon_{ij}) \quad . \qquad (2.10.1)$$

On subtracting equation (2.9.8) from (2.10.1) we get

$$-\lambda[\theta\{(1-w)\varepsilon_{i+1,j+1}+(2w-1)\varepsilon_{i,j+1}-w\varepsilon_{i-1,j+1}\}+(1-\theta)\{(1-w)\varepsilon_{i+1,j}+(2w-1)\varepsilon_{ij}-w\varepsilon_{i-1,j}\}] = \varepsilon_{i,j+1}-\varepsilon_{i,j} \quad . \qquad (2.10.2)$$

We see that the error equation (2.10.2) has exactly the same form as the original difference equation (2.9.8). It will generally be true that a difference equation and its error equation will be identical when the difference equation is linear and homogeneous.

The von Neumann stability analysis consists of expanding the error ε_{ij} in a Fourier series of the form

$$\varepsilon_{ij} = \sum_{\beta} \xi_{\beta}^{j} \exp(i_c \beta x_i) \qquad (2.10.3)$$

where $i_c=\sqrt{-1}$. This is followed by substituting the series into the error equation and solving for the amplification factor $\gamma_\beta=\xi_\beta^{j+1}/\xi_\beta^{j}$ for each component. For stability (as was seen in Section 2.3), the modulus of *the amplification factor* must be less than or equal to one for all the components. The analysis is somewhat simplified by omitting the subscript β and by taking $x_i = i\Delta x$. Equation (2.10.3) will then take the form,

$$\varepsilon_{ij} = \xi^j \exp(i_c \beta i \Delta x) \quad , \qquad (2.10.4)$$

and substituting this directly into the difference equation and cancelling the resulting common factor, $\xi^j \exp(i_c \beta i \Delta x)$, leads to an equation that must be satisfied by the parameters γ, β, Δx and Δt.

We now apply the criterion to equation (2.9.8). By substituting the error (2.10.4) into (2.9.8) and cancelling the common factor $\exp(i_c \beta i \Delta x)$, we obtain

$$-\lambda[\theta\xi^{j+1}\{(1-w)\exp(i_c\beta\Delta x)+(2w-1)-w\exp(-i_c\beta\Delta x)\}$$
$$+(1-\theta)\xi^j\{(1-w)\exp(i_c\beta\Delta x)+(2w-1)-w\exp(-i_c\beta\Delta x)\}] = \xi^{j+1}-\xi^j \; .$$

By using the identity,

$$\exp(\pm i_c\beta\Delta x) = \cos(\beta\Delta x) \pm i_c \sin(\beta\Delta x) \quad ,$$

we get,

$$-\lambda[\theta\xi^{j+1}+(1-\theta)\xi^j][(2w-1)\{1-\cos(\beta\Delta x)\} + i_c\sin(\beta\Delta x)] = \xi^{j+1}-\xi^j \; .$$

Hence,

$$\frac{\xi^{j+1}}{\xi^j} = \gamma$$
$$= \frac{1-\lambda(1-\theta)(2w-1)\{1-\cos(\beta\Delta x)\}-i_c\lambda(1-\theta)\sin(\beta\Delta x)}{1+\lambda\theta(2w-1)\{1-\cos(\beta\Delta x)\}+i_c\lambda\theta\sin(\beta\Delta x)} \qquad (2.10.5)$$

the amplification factor which is complex. Clearly, by taking the square of the modulus of λ we obtain,

$$|\gamma|^2 = \frac{[1-\lambda(1-\theta)(2w-1)\{1-\cos(\beta\Delta x)\}]^2+\lambda^2(1-\theta)^2\sin^2(\beta\Delta x)}{[1+\lambda\theta(2w-1)\{1-\cos(\beta\Delta x)\}]^2+\lambda^2\theta^2\sin^2(\beta\Delta x)} \; .$$

(2.10.6)

We now proceed to analyse, separately, the stability of

(a) *the Lax-Wendroff equation* (b) *centred-in-distance equations*

(c) *backward-in-distance equations* and (d) *forward-in-distance equations.*

(a) *Stability of the Lax-Wendroff Equation.*

With $w=(1+\lambda)/2$ and $\theta=0$, we have from (2.10.6)

$$|\gamma|^2 = [1-\lambda^2(1-\cos(\beta\Delta x))]^2 + \lambda^2\sin^2(\beta\Delta x)$$

$$= [1-2\lambda^2\sin^2(\frac{\beta\Delta x}{2})]^2 + \lambda^2\sin^2(\beta\Delta x)$$

$$= 1-4\lambda^2\sin^2(\frac{\beta\Delta x}{2})+4\lambda^4\sin^4(\frac{\beta\Delta x}{2}) + \lambda^2\sin^2(\beta\Delta x).$$

Since $\sin(\beta\Delta x) = 2\sin(\frac{\beta\Delta x}{2})\cos(\frac{\beta\Delta x}{2})$ we have

$$\lambda^2\sin^2(\beta\Delta x) = 4\lambda^2\sin^2(\frac{\beta\Delta x}{2})\cos^2(\frac{\beta\Delta x}{2})$$

$$= 4\lambda^2\sin^2(\frac{\beta\Delta x}{2})-4\lambda^2\sin^4(\frac{\beta\Delta x}{2}) .$$

Hence,

$$|\gamma| = [1-4\lambda^2(1-\lambda^2)\sin^4 \frac{\beta\Delta x}{2}]^{\frac{1}{2}} , \tag{2.10.7}$$

and for stability,

$$|\gamma|\leq 1 \text{ implies } 0\leq 4\lambda^2(1-\lambda^2) \text{ giving } 0<\lambda\leq 1.$$

Therefore the Lax-Wendroff method is conditionally stable for $0<\lambda\leq 1$.

(b) *Stability of the Centred-In-Distance Equations.*

For $w=\frac{1}{2}$, equation (2.10.6) reduces to

$$|\gamma|^2 = \frac{1+\lambda^2(1-\theta)^2\sin^2(\beta\Delta x)}{1+\lambda^2\theta^2\sin^2(\beta\Delta x)} . \tag{2.10.8}$$

The stability requirement $|\gamma|^2\leq 1$ leads to

$$1+\lambda^2(1-\theta)^2\sin^2(\beta\Delta x)\leq 1+\lambda^2\theta^2\sin^2(\beta\Delta x)$$

which simplifies to

$$(1-\theta)^2 \leq \theta^2 . \tag{2.10.9}$$

For $\theta=\frac{1}{2}$, $|\gamma|$ is exactly one.

For $\theta>\frac{1}{2}$, inequality (2.10.9) is always satisfied.

For $\theta<\frac{1}{2}$, inequality (2.10.9) is never satisfied.

We conclude that the centred-in-distance, centred-in-time difference equation is neutrally stable, the centred-in-distance, backward-in-time equation is unconditionally stable while the centred-in-distance forward-in-time equation is always unstable.

(c) *Stability of the Backward-in-Distance Equations.*

For $w=1$, equation (2.10.6) reduces to

$$|\gamma|^2 = \frac{[1-\lambda(1-\theta)\{1-\cos(\beta\Delta x)\}]^2+\lambda^2(1-\theta)^2\sin^2(\beta\Delta x)}{[1+\lambda\theta\{1-\cos(\beta\Delta x)\}]^2+\lambda^2\theta^2\sin^2(\beta\Delta x)} \; . \qquad (2.10.10)$$

It is immediately evident that for $\theta=1$, the numerator takes on exactly the value of one while the denominator is larger than one. Therefore, for this particular case, $|\gamma|<1$ and satisfies the stability requirement.

For the more general case, in order that $|\gamma|^2\leq 1$, the numerator of equation (2.10.10) must be less than or equal to the denominator. On expanding and after some manipulation, this leads to $-2\lambda\{1-\cos(\beta\Delta x)\}+\lambda^2(1-2\theta)\{1-\cos(\beta\Delta x)\}^2+\lambda^2(1-2\theta)\sin^2(\beta\Delta x)\leq 0$, which reduces to,

$$-2\lambda\{1-\cos(\beta\Delta x)\}[1-\lambda(1-2\theta)] \leq 0 .$$

Since $\lambda>0$ and $\cos(\beta\Delta x)\leq 1$, then we must have

$$1-\lambda(1-2\theta) \geq 0 \; , \qquad (2.10.11)$$

which is automatically satisfied for $\theta\geq\frac{1}{2}$ and for all $\lambda>0$. However, when $\theta<\frac{1}{2}$ we now have a restriction on λ for stability. That is, the equations are stable for $\lambda\leq\frac{1}{1-2\theta}$ and they are unstable otherwise $(\lambda>\frac{1}{1-2\theta})$. In particular, for $\theta=0$, the stability requirement is $\lambda\leq 1$. We conclude that both the backward-in-distance, backward-in-time and the backward-in-distance, centred-in-time formulae are always stable while the backward-in-distance, forward-in-time equation is conditionally stable for $\lambda\leq 1$.

(d) *Stability of the Forward-In-Distance Equations.*

For w=0, equation (2.10.6) becomes

$$|\gamma|^2 = \frac{[1+\lambda(1-\theta)\{1-\cos(\beta\Delta x)\}]^2+\lambda^2(1-\theta)^2\sin^2(\beta\Delta x)}{[1-\lambda\theta\{1-\cos(\beta\Delta x)\}]^2+\lambda^2\theta^2\sin^2(\beta\Delta x)} \qquad (2.10.12)$$

Again, the condition for stability is that the numerator be less than the denominator. This leads to

$$2\lambda\{1-\cos(\beta\Delta x)\}+\lambda^2(1-2\theta)\{1-\cos(\beta\Delta x)\}^2+\lambda^2(1-2\theta)\sin^2(\beta\Delta x)\leqslant 0$$

which reduces to

$$2\lambda\{1-\cos(\beta\Delta x)\}[1+\lambda(1-2\theta)] \leqslant 0.$$

Stability then requires that,

$$1+\lambda(1-2\theta) \leqslant 0 \ , \qquad (2.10.13)$$

which is never satisfied for any $\theta\leqslant\frac{1}{2}$.

When $\theta>\frac{1}{2}$, inequality (2.10.13) leads to the following restriction on λ,

$$\lambda \geqslant \frac{1}{2\theta-1} \ .$$

In particular, for $\theta=1$ (corresponding to the forward-in-distance, backward-in-time equation), the stability requirement is $\lambda\geqslant 1$ or

$$a\frac{\Delta t}{\Delta x} \geqslant 1 \ . \qquad (2.10.14)$$

In practice, it is difficult to satisfy the above stability condition at all mesh points as it entails excessively large time steps. We therefore conclude that none of the forward-in-distance equations is useful, either because of their unconditional instability (when $\theta=0$ or $\theta=\frac{1}{2}$), or because of the difficult stability restriction (when $\theta>\frac{1}{2}$).

2.11 TRUNCATION ERROR ANALYSIS OF THE GENERALISED FINITE DIFFERENCE METHODS

Let us consider the differential equation at the point $(i\Delta x, j\Delta t)$,

$$L(U_{ij}) = 0 \qquad (2.11.1)$$

where U is the *exact solution of the differential equation* at that point. For example, for equation (2.9.1) we have

$$L(U_{ij}) = (-a\frac{\partial U}{\partial x} - \frac{\partial U}{\partial t})_{ij} = 0 \ . \qquad (2.11.2)$$

The derivatives in this equation may be replaced exactly at the point $(i\Delta x, j\Delta t)$ by an appropriate infinite (difference) series. To derive a finite-difference equation approximating the differential equation, however, each series is truncated after a certain number of terms. Let us denote this approximating difference equation by

$$F(u_{ij}) = 0 \qquad (2.11.3)$$

where u is the *exact solution of the difference equation* at the point $(i\Delta x, j\Delta t)$.

The amount by which the exact solution U of the partial differential equation does not satisfy the differential equation at the point $(i\Delta x, j\Delta t)$ is called the *local truncation error* T_{ij}. Clearly,

$$T_{ij} = F(U_{ij}) \ . \qquad (2.11.4)$$

By means of a Taylor series expansion we will be able to find the principal part of the local truncation error and hence deduce the order of accuracy of the difference equation.

To study the accuracy of the generalised finite difference methods, we use equation (2.9.8),

$$-\frac{a}{\Delta x}[\theta\{(1-w)u_{i+1,j+1}+(2w-1)u_{i,j+1}-wu_{i-1,j+1}\}+ (1-\theta)\{(1-w)u_{i+1,j}+$$

$$(2w-1)u_{ij}-wu_{i-1,j}\}] - \frac{u_{i,j+1}-u_{ij}}{\Delta t} = F(u_{ij}) \ . \qquad (2.11.5)$$

We note that the expression,

$$(1-w)u_{i+1} + (2w-1)u_i - wu_{i-1}$$

can also be written as

$$\tfrac{1}{2}(u_{i+1}-u_{i-1})+(\tfrac{1}{2}-w)(u_{i+1}-2u_i+u_{i-1}) \ . \qquad (2.11.6)$$

Hence, using equations (2.11.4), (2.11.5) and (2.11.6) we can now write an expression for the local truncation error as,

$$T = -a[\theta\{\frac{U_{i+1,j+1}-U_{i-1,j+1}}{2\Delta x} +\Delta x(\tfrac{1}{2}-w)\frac{U_{i+1,j+1}-2U_{i,j+1}+U_{i-1,j+1}}{(\Delta x)^2}\}$$

$$+(1-\theta)\{\frac{U_{i+1,j}-U_{i-1,j}}{2\Delta x} + \Delta x(\tfrac{1}{2}-w)\ \frac{U_{i+1,j}-2U_{ij}+U_{i-1,j}}{(\Delta x)^2}\}]$$

$$-\frac{U_{i,j+1}-U_{ij}}{\Delta t} \ . \qquad (2.11.7)$$

From equations (2.1.8) and (2.1.9), we have

$$\frac{U_{i+1}-U_{i-1}}{2\Delta x} = (\frac{\partial U}{\partial x})_i + O([\Delta x]^2)$$

and

$$\frac{U_{i+1}-2U_i+U_{i-1}}{(\Delta x)^2} = (\frac{\partial^2 U}{\partial x^2})_i + O([\Delta x]^2) \ .$$

Substituting these into equation (2.11.7) gives us

$$T = -a[\theta(\frac{\partial U}{\partial x})_{i,j+1} + \theta\Delta x(\tfrac{1}{2}-w)(\frac{\partial^2 U}{\partial x^2})_{i,j+1} + (1-\theta)(\frac{\partial U}{\partial x})_{i,j}$$

$$+(1-\theta)\Delta x(\tfrac{1}{2}-w)(\frac{\partial^2 U}{\partial_x^2})_{i,j}] - \frac{U_{i,j+1}-U_{i,j}}{\Delta t} + O([\Delta x]^2) \ . \qquad (2.11.8)$$

We can now apply Taylor's series expansions on $U_{i,j+1}$ and U_{ij} about the point $(i\Delta x,(j+\theta)\Delta t)$ using the obvious relations

$$j+1 = (j+\theta)+(1-\theta) \qquad (2.11.9a)$$

and

$$j = (j+\theta)-\theta \qquad (2.11.9b)$$

to give us,

$$U_{i,j+1} = U_{i,j+\theta}+(1-\theta)\Delta t(\frac{\partial U}{\partial t})_{i,j+\theta}+ \frac{1}{2!}(1-\theta)^2[\Delta t]^2(\frac{\partial^2 U}{\partial t^2})_{i,j+\theta}+ O([\Delta t]^3)$$

and

$$U_{ij} = U_{i,j+\theta}-\theta\Delta t(\frac{\partial U}{\partial t})_{i,j+\theta}+ \frac{1}{2!}\theta^2[\Delta t]^2 (\frac{\partial^2 U}{\partial_t{}^2})_{i,j+\theta}+ O([\Delta t]^3).$$

Hence,

$$\frac{U_{i,j+1}-U_{ij}}{\Delta t} = (\frac{\partial U}{\partial t})_{i,j+\theta}+(\tfrac{1}{2}-\theta)\Delta t(\frac{\partial^2 U}{\partial t^2})_{i,j+\theta}+ O([\Delta t]^2) \ . \qquad (2.11.10)$$

In addition,

$$(\frac{\partial U}{\partial x})_{i,j+1} = (\frac{\partial U}{\partial x})_{i,j+\theta} +(1-\theta)\Delta t(\frac{\partial^2 U}{\partial t\partial x})_{i,j+\theta}+O([\Delta t]^2)$$

and

$$(\frac{\partial U}{\partial x})_{i,j} = (\frac{\partial U}{\partial x})_{i,j+\theta} -\theta\Delta t(\frac{\partial^2 U}{\partial t\partial x})_{i,j+\theta}+ O([\Delta t]^2) \ .$$

Using these two equations, we obtain

$$\theta(\frac{\partial U}{\partial x})_{i,j+1} +(1-\theta)(\frac{\partial U}{\partial x})_{i,j} = (\frac{\partial U}{\partial x})_{i,j+\theta} +O([\Delta t]^2) \qquad (2.11.11)$$

and similarly,

$$\theta(\frac{\partial^2 U}{\partial x^2})_{i,j+1} +(1-\theta)(\frac{\partial^2 U}{\partial x^2})_{i,j} = (\frac{\partial^2 U}{\partial x^2})_{i,j+\theta}+ O([\Delta t]^2) \ . \qquad (2.11.12)$$

Substituting equations (2.11.10), (2.11.11) and (2.11.12) into equation (2.11.8) we get,

$$T = \{-a(\frac{\partial U}{\partial x})_{i,j+\theta}-(\frac{\partial U}{\partial t})_{i,j+\theta}\}-a\Delta x(\tfrac{1}{2}-w)(\frac{\partial^2 U}{\partial x^2})_{i,j+\theta}+(\theta-\tfrac{1}{2})\Delta t(\frac{\partial^2 U}{\partial t^2})_{i,j+\theta}$$
$$+O([\Delta x]^2) + O([\Delta t]^2) \ .$$

From the differential equation (2.9.1) the first term in braces is zero and we also find by differentiating (2.9.1) with respect to t and x that,

$$(\frac{\partial^2 U}{\partial t^2})_{i,j+\theta} = a^2(\frac{\partial^2 U}{\partial x^2})_{i,j+\theta} \ .$$

Hence, we have the final form of the local truncation error as,

$$T = -a\Delta x(\tfrac{1}{2}-w)(\frac{\partial^2 U}{\partial x^2})_{i,j+\theta} + (\theta-\tfrac{1}{2})\Delta t a^2 (\frac{\partial^2 U}{\partial x^2})_{i,j+\theta} + O([\Delta x]^2) + O([\Delta t]^2) \tag{2.11.13}$$

i.e.,

$$T = a\Delta x[(w-\tfrac{1}{2})+\lambda(\theta-\tfrac{1}{2})] \ (\frac{\partial^2 U}{\partial x^2})_{i,j+\theta} + O([\Delta x]^2) + O([\Delta t]^2).$$

The principal part of the local truncation error is therefore

$$a\Delta x[(w-\tfrac{1}{2})+\lambda(\theta-\tfrac{1}{2})](\frac{\partial^2 U}{\partial x^2})_{i,j+\theta} \quad .$$

For the Lax-Wendroff formula (2.9.9), we find that with $\theta=0$ and $w=\frac{1}{2}(1+\lambda)$, $T=O([\Delta x]^2+O([\Delta t]^2)$ confirming that the method is second-order accurate in both space and time. The local truncation errors of the other formulae are described in Table 2.9.1.

2.12 OTHER APPROXIMATIONS FOR THE CONVECTION EQUATION

Other methods can be developed in much the same way as for the generalised schemes to approximate the convection equation (2.9.1).

(i) *Leapfrog method.* If we discretize $\frac{\partial U}{\partial x}$ and $\frac{\partial U}{\partial t}$ by their second order centred analogues as in equation (2.1.8) then we are led to the *three-time level formula*,

$$\frac{u_{i,j+1}-u_{i,j-1}}{2\Delta t} = -a\frac{(u_{i+1,j}-u_{i-1,j})}{2\Delta x}$$

or $$u_{i,j+1} = u_{i,j-1}-\lambda(u_{i+1,j}-u_{i-1,j}) \ , \qquad (2.12.1)$$

which is known as the explicit leapfrog scheme. We shall now establish the stability condition of this scheme. By substituting the error,

$$\varepsilon_{ij} = \xi^{j}\exp(i_c\beta i\Delta x)$$

into equation (2.12.1) results in

$$\xi^{j+1}\exp(i_c\beta i\Delta x) = \xi^{j-1}\exp(i_c\beta i\Delta x)-\lambda\xi^{j}(\exp(i_c\beta(i+1)\Delta x)- \exp(i_c\beta(i-1)\Delta x))$$

or $$\gamma = \gamma^{-1}-\lambda(\exp(i_c\beta\Delta x)-\exp(-i_c\beta\Delta x) \ . \qquad (2.12.2)$$

But we know from the identity,

$$\exp(i_c\beta\Delta x)-\exp(-i_c\beta\Delta x) = 2i_c\sin(\beta\Delta x) \ .$$

Hence equation (2.12.2) becomes,

$$\gamma^2+(2\lambda i_c\sin(\beta\Delta x))\gamma-1 = 0 \ .$$

The roots of this equation (the amplification factors) are

$$\gamma = i_c\lambda\sin(\beta\Delta x) \pm \sqrt{1-\lambda^2\sin^2(\beta\Delta x)}.$$

If $\lambda\sin(\beta\Delta x)>1$, then the absolute value of one of these roots exceeds unity; if $\lambda\sin(\beta\Delta x)\leq 1$ then both roots satisfy $|\gamma|=1$. Thus, assuming that the maximum value for $\sin(\beta\Delta x)$ must be considered possible, then the stability condition is $\lambda\leq 1$.

From equation (2.12.1) we observe that

$$\frac{1}{\Delta t}(u_{i,j+1}-u_{i,j-1}) = \frac{1}{\Delta x}\,a(u_{i+1,j}-u_{i-1,j}) \ .$$

An expression for the local truncation error is

$$\begin{aligned} T &= F(U_{ij}) \\ &= \frac{1}{\Delta t}(U_{i,j+1}-U_{i,j-1}) - \frac{1}{\Delta x}\,a(U_{i+1,j}-U_{i-1,j}) . \end{aligned}$$

On applying Taylor's theorem on this expression at the point $(i\Delta x, j\Delta t)$ we find that,

$$\begin{aligned} T &= \frac{1}{\Delta t}(2\Delta t(\frac{\partial U}{\partial t})_{ij} + O([\Delta t]^3)) \\ &\quad - \frac{1}{\Delta x}\,a(2\Delta x(\frac{\partial U}{\partial x})_{ij} + O([\Delta x]^3)) \\ &= 2(\frac{\partial U}{\partial t})_{ij} - 2a(\frac{\partial U}{\partial x})_{ij} + O([\Delta x]^2) + O([\Delta t]^2) \\ &= 2\{(\frac{\partial U}{\partial t})_{ij} - a(\frac{\partial U}{\partial x})_{ij}\} + O([\Delta x]^2) + O([\Delta t]^2) . \end{aligned}$$

Hence $T = O([\Delta x]^2)+O([\Delta t]^2)$ for the leapfrog scheme.

(ii) *Wendroff implicit method.* The algorithm due to Wendroff may be obtained by representing $\frac{\partial U}{\partial x}$ at the $(j+\frac{1}{2})$ level as the average of the derivatives at j and j+1 and by representing $\frac{\partial U}{\partial t}$ at the $(i+\frac{1}{2})$ level as the average of the derivatives at i and i+1. Thus,

$$(\frac{\partial U}{\partial t})_{i+\frac{1}{2},j+\frac{1}{2}} \approx \frac{1}{2}(\frac{u_{i+1,j+1}-u_{i+1,j}}{\Delta t} + \frac{u_{i,j+1}-u_{ij}}{\Delta t})$$

and

$$(\frac{\partial U}{\partial x})_{i+\frac{1}{2},j+\frac{1}{2}} \approx \frac{1}{2}(\frac{u_{i+1,j+1}-u_{i,j+1}}{\Delta x} + \frac{u_{i+1,j}-u_{ij}}{\Delta x}) \ .$$

The Wendroff algorithm then assumes the form,

$$(1+\lambda)u_{i+1,j+1}+(1-\lambda)u_{i,j+1} = (1+\lambda)u_{ij} + (1-\lambda)u_{i+1,j} \qquad (2.12.3)$$

It cannot be used for pure initial-value problems, that is, conditions on t=0 only because it would give an infinite number of simultaneous equations. If, however, initial values are known on the x-axis, $x \geq 0$, and boundary values on the t-axis, $t \gtrsim 0$, the equation can be used explicitly by writing it as

$$u_{i+1,j+1} = u_{ij} + \frac{(1-\lambda)}{(1+\lambda)}\{u_{i+1,j} - u_{i,j+1}\}.$$

To analyse its stability, we again insert the error $\varepsilon_{ij} = \xi^j \exp(i_c \beta i \Delta x)$ into (2.12.3). This provides us

$$(1+\lambda)\xi^{j+1}\exp(i_c\beta(i+1)\Delta x) + (1-\lambda)\xi^{j+1}\exp(i_c\beta i\Delta x) = (1+\lambda)\xi^{j}\exp(i_c\beta i\Delta x)$$
$$+(1-\lambda)\xi^{j}\exp(i_c\beta(i+1)\Delta x),$$

i.e.,

$$\gamma = \frac{(1+\lambda)+(1-\lambda)\exp(i_c\beta\Delta x)}{(1-\lambda)+(1+\lambda)\exp(i_c\beta\Delta x)}$$

the amplification factor.

This can be written as,

$$\gamma = \frac{[(1+\lambda)+(1-\lambda)\cos(\beta\Delta x)]+i_c(1-\lambda)\sin(\beta\Delta x)}{[(1-\lambda)+(1+\lambda)\cos(\beta\Delta x)]+i_c(1+\lambda)\sin(\beta\Delta x)}$$

Hence,

$$|\gamma|^2 = \frac{(1+\lambda)^2+(1-\lambda)^2+2(1-\lambda^2)\cos(\beta\Delta x)}{(1+\lambda)^2+(1-\lambda)^2+2(1-\lambda^2)\cos(\beta\Delta x)}$$

$$= 1.$$

Therefore, the Wendroff scheme is unconditionally stable for all values of λ.

We proceed now to investigate the truncation errors of the method. Using equation (2.12.3) we have,

$$T = (1+\lambda)\{U_{i+1,j+1} - U_{ij}\} + (1-\lambda)\{U_{i,j+1} - U_{i+1,j}\}.$$

U_{ij}, $U_{i,j+1}$, $U_{i+1,j}$ and $U_{i+1,j+1}$ are expanded about the point $((i+\frac{1}{2}\Delta x),(j+\frac{1}{2})\Delta t)$ by means of Taylor's series for multi-variables. After some extensive manipulations and cancellations of terms we obtain,

$$T = 2 \quad \{a(\frac{\partial U}{\partial x})_{i+\frac{1}{2},j+\frac{1}{2}} + (\frac{\partial U}{\partial t})_{i+\frac{1}{2},j+\frac{1}{2}}\}$$

$$+ \quad \{\frac{(\Delta x)^2}{12} \ (a(\frac{\partial^3 U}{\partial x^3})_{i+\frac{1}{2},j+\frac{1}{2}} + 3(\frac{\partial^3 U}{\partial x^2 \partial t})_{i+\frac{1}{2},j+\frac{1}{2}})$$

$$+ \frac{(\Delta t)^2}{12}((\frac{\partial^3 U}{\partial t^3})_{i+\frac{1}{2},j+\frac{1}{2}} + 3a(\frac{\partial^3 U}{\partial x \partial t^2})_{i+\frac{1}{2},j+\frac{1}{2}})+\ldots\}.$$

From the given differential equation, $a(\frac{\partial U}{\partial x})_{i+\frac{1}{2},j+\frac{1}{2}} + (\frac{\partial U}{\partial t})_{i+\frac{1}{2},j+\frac{1}{2}}=0$.

Therefore,

$$T = O([\Delta x]^2) + O([\Delta t]^2).$$

showing that the Wendroff scheme is second-order accurate in both space and time.

2.13 FINITE DIFFERENCE APPROXIMATIONS FOR SECOND ORDER HYPERBOLIC EQUATIONS

A natural extension to first-order hyperbolic equations is to employ finite difference procedures to second-order equations. In the process of developing these methods, we will, of course, bear in mind, as we did with first-order equations, the limitations imposed by the characteristics.

(a) *Explicit Methods.*

Let us consider the simplest of the second-order hyperbolic equations called the wave equation given by,

$$\frac{\partial^2 U}{\partial x^2} = \frac{\partial^2 U}{\partial t^2} \qquad (2.13.1)$$

and the Cauchy condition,

$$u(x,0) = f(x) \ , \ \frac{\partial U}{\partial t}(x,0) = g(x) \ . \qquad (2.13.2)$$

As before, we take a rectangular net with constant space and time intervals given by Δx and Δt respectively and we write $u_{ij}=u(i\Delta x,j\Delta t)$, $-\infty<i<\infty$, $0\leq j\leq\infty$. Both second partial derivatives are approximated by central difference expressions given by (2.1.9) whose truncation error is $O([\Delta x]^2)$. Thus equation (2.13.1) is approximated by the explicit formula

$$u_{i,j+1} = \lambda^2(u_{i-1,j}+u_{i+1,j})+2(1-\lambda^2)u_{ij}-u_{i,j-1} \qquad (2.13.3)$$

where $\lambda = \frac{\Delta t}{\Delta x}$. The first initial condition of (2.13.2) specifies $u_{i,0}$ on the line $t=0$. We can use the second condition to find values on the line $t=\Delta t$ by employing a 'false' boundary and the second-order central difference formula,

$$\left.\frac{\partial U}{\partial t}\right|_{i,0} = \frac{U_{i,1}-U_{i,-1}}{2\Delta t} + O([\Delta t]^2) \ . \qquad (2.13.4)$$

Writing $g(i\Delta x)=g_i$, we have the approximation,

$$u_{i,1}-u_{i,-1} = 2\Delta t g_i \ . \tag{2.13.5}$$

From equation (2.13.3) with $j=0$, we have

$$u_{i,1} = \lambda^2(u_{i-1,0}+u_{i+1,0})+2(1-\lambda^2)u_{i,0}-u_{i,-1} \ .$$

Upon replacing $u_{i,-1}$ with its value, from equation (2.13.5), and solving for $u_{i,1}$ we find

$$u_{i,1} = \tfrac{1}{2}\lambda^2(f_{i-1}+f_{i+1})+(1-\lambda^2)f_i + \Delta t g_i \ . \tag{2.13.6}$$

The computational molecule for equation (2.13.3) is shown in Figure (2.13.1). Superimposed on this figure are the characteristics of the wave equation, namely $t=\pm x+$ constant, whose slopes are ± 1, represented by the lines AC and BC. As we have seen from Section 2.6, the solution is uniquely determined in the triangle ACB, provided the solution up to AB is known. If the absolute value of λ (i.e. the slope) exceeds 1, then equation (2.13.3) would provide a 'solution' in a region 'not reached' by the continuous solution. In such a case we would expect the result to be quite incorrect.

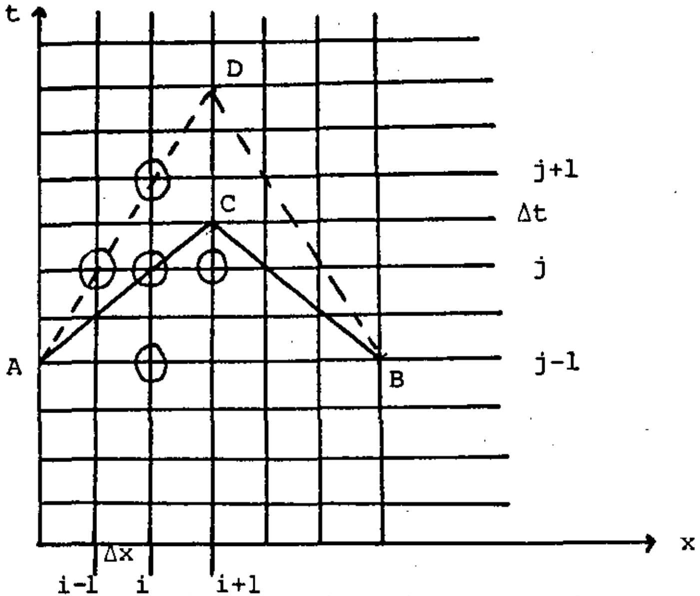


FIGURE 2.13.1: Comparison of the finite difference characteristics AD and DB with the true characteristics AC and BC.
]

If $|\lambda|\leq 1$, however, it can be shown that the method will converge under the usual assumption that certain higher derivatives exist. Convergence of the solution of equations (2.13.3) and (2.13.4) to that of the differential problem, equations (2.13.1) and (2.13.2) as $\Delta x\to 0$ and $\Delta t\to 0$, was first examined by Courant, Friedrichs, and Lewy (1928), by Lowan (1957) using operator methods, and by Collatz (1960).

The proof of convergence is somewhat complicated for $\lambda<1$ but is much simpler if the special mesh ratio condition $\lambda=1$ holds. We shall provide the proof for the latter when the derivative initial condition is approximated by forward differences and $\Delta t=\Delta x$.

Let U_{ij} be the exact solution of the differential (wave) equation at the point (x_i,t_j) and u_{ij} the exact solution of the explicit difference equation. From the initial conditions (2.13.2) we obtain

$$g_i = (u_{i,1}-u_{i,0})/\Delta x$$

giving

$$u_{i,1} = f_i + \Delta x g_i \ . \tag{2.13.7}$$

By the Taylor's expansion about the point $(i\Delta x, 0)$, we have

$$U_{i,1} = U_{i,0} + \Delta x\left(\frac{\partial U}{\partial t}\right)_{i,0} + \frac{1}{2!}(\Delta x)^2\left(\frac{\partial^2 U}{\partial t^2}\right)_{i,\theta} \ , \quad (0<\theta<1)$$

i.e.
$$U_{i,1} = f_i+\Delta x g_i + \frac{1}{2!}(\Delta x)^2 \left(\frac{\partial^2 U}{\partial t^2}\right)_{i,\theta} \ . \tag{2.13.8}$$

Hence, from equations (2.13.7) and (2.13.8) we get the discretization error,

$$\begin{aligned}|e_{i,1}| &= |U_{i,1}-u_{i,1}| \\ &\leq \tfrac{1}{2}(\Delta x)^2 M_2 \ ,\end{aligned} \tag{2.13.9}$$

where M_2 is the modulus of the largest value of $\frac{\partial^2 U}{\partial t^2}$ in the first time interval. Substitution of $u_{ij}=U_{ij}-e_{ij}$ into the finite difference equation (2.13.3) and expansion in terms of U_{ij} by the Taylor's theorem gives,

$$e_{i,j+1} = e_{i+1,j}+e_{i-1,j}-e_{i,j-1}+ \frac{1}{24}(\Delta x)^4((\frac{\partial^4 U}{\partial t^4})_{i,j+\theta_1}$$

$$+(\frac{\partial^4 U}{\partial t^4})_{i,j+\theta_2} - (\frac{\partial^4 U}{\partial x^4})_{i+\theta_3,j} -(\frac{\partial^4 U}{\partial x^4})_{i+\theta_4,j}$$

where $|\theta_s|<1$, $(s=1,2,3,4)$. Hence, if M_4 is the modulus of the largest of $\frac{\partial^4 U}{\partial t^4}$ and $\frac{\partial^4 U}{\partial x^4}$ throughout the solution domain and $|\eta|\leq 1$, then

$$e_{i,j+1} = e_{i+1,j}+e_{i-1,j}-e_{i,j-1}+ \frac{1}{6}(\Delta x)^4\eta M_4 \text{ .} \qquad (2.13.10)$$

Draw the straight line characteristics through the point $(i,j+1)$ at ±45° to the x-axis Ox until they meet the line $j=1$, and mark the points within this triangle contributing terms to $e_{i,j+1}$ when working backwards using the last equation so as to express $e_{i,j+1}$ entirely in terms of the errors along $j=1$. It will be seen that there is one point at $(i,j+1)$, two points along $t=j\Delta x$, three points along $t=(j-1)\Delta x$ and following the same manner $(j+1)$ points along $t=\Delta x$. As the $(j+1)$ points along $t=\Delta x$ each contribute an error bounded by $\frac{1}{2}(\Delta x)^2M_2$ as in (2.13.9) and $\sum_{k=1}^{j} k=\frac{1}{2}j(j+1)$ points between $t=2\Delta x$ and $t=(j+1)\Delta x$ each contribute an error bounded by $\frac{1}{6}(\Delta x)^4M_4$, it follows that

$$|e_{i,j+1}| \leq \tfrac{1}{2}(j+1)(\Delta x)^2M_2 + \frac{1}{12}\, j(j+1)(\Delta x)^4M_4 \qquad (2.13.11)$$

Changing j into $(j-1)$ completes the proof. Since $j\Delta x=t$, we have

$$|e_{ij}| \leq \tfrac{1}{2}t\Delta xM_2 + \frac{1}{12}t^2(\Delta x)^2M_4 \text{ .} \qquad (2.13.12)$$

As Δx tends to zero, this error tends to zero for finite values of t. Hence convergence is established. The proof for the case $\lambda<1$ is given by Forsythe and Wasow (1960) and Lowan (1957).

(b) *Implicit Methods*.

With the expectation of gaining stability advantages, we shall now attempt to derive implicit methods for second-order equations. For the wave equation (2.13.1), the simplest implicit system is obtained by approximating $\frac{\partial^2 U}{\partial t^2}$, as before, by a second central difference centred at (i,j) while $\frac{\partial^2 U}{\partial x^2}$ is approximated by the average of two second central differences, one centred at (i,j+1) and the other at (i,j-1). Thus, one simple implicit approximation takes the form,

$$u_{i,j+1}-2u_{ij}+u_{i,j-1} = \frac{\lambda^2}{2}\{(u_{i+1,j+1}-2u_{i,j+1}+u_{i-1,j+1})+(u_{i+1,j-1}-2u_{i,j-1}+u_{i-1,j-1})\}. \qquad (2.13.13)$$

The implicit nature of this formula is obvious by rewriting the expression to be solved on the (j+1) line in terms of the values on the two preceding lines. Thus one finds the equation,

$$\begin{aligned}&-\lambda^2 u_{i+1,j+1}+2(1+\lambda^2)u_{i,j+1}-\lambda^2 u_{i-1,j+1}\\ &= 4u_{ij}+\lambda^2 u_{i+1,j-1}-2(1+\lambda^2)u_{i,j-1}+\lambda^2 u_{i-1,j-1} \; . \qquad (2.13.14)\end{aligned}$$

Another implicit method discussed by Richtmyer and Morton (1967) to approximate (2.13.1) is given by,

$$\begin{aligned}-\frac{1}{4}\lambda^2 u_{i-1,j+1}+(1+\frac{\lambda^2}{2})u_{i,j+1} - \frac{\lambda^2}{4}u_{i+1,j+1} &= \tfrac{1}{2}\lambda^2 u_{i-1,j}+(2-\lambda^2)u_{ij} + \\ &\frac{\lambda^2}{2}u_{i+1,j} + \frac{\lambda^2}{4}u_{i-1,j-1} + (-1-\frac{\lambda^2}{2})u_{i,j-1}+\frac{\lambda^2}{4}u_{i+1,j-1}\end{aligned}$$

$$(2.13.15)$$

By assuming that there are m mesh values to be determined, then upon writing equations (2.13.14) and (2.13.15) for each i, i=1,2,...,m and inserting the discretized boundary conditions, *the tridiagonal nature* of the system becomes clear. Thus *the Thomas algorithm* of Section 1.9 may be applied to find a non-iterative solution.

Equations (2.13.3), (2.13.14) and (2.13.15) are special cases of a *general three-level implicit form* obtained by approximating $(\frac{\partial^2 U}{\partial t^2})_{i,j}$ by $\frac{1}{(\Delta t)^2}\delta_t^2 u_{ij}$ and approximating $(\frac{\partial^2 U}{\partial x^2})_{ij}$ with

$$\frac{1}{(\Delta x)^2}[\alpha\delta_x^2 u_{i,j+1}+(1-2\alpha)\delta_x^2 u_{ij}+\alpha\delta_x^2 u_{i,j-1}].$$

Hence,

$$\frac{1}{(\Delta t)^2}\delta_t^2 u_{ij} = \frac{1}{(\Delta x)^2}[\alpha\delta_x^2 u_{i,j+1}+(1-2\alpha)\delta_x^2 u_{ij}+\alpha\delta_x^2 u_{i,j-1}] \qquad (2.13.16)$$

where α is a weighting factor and δ^2 is the operator defined by

$$\delta_x^2 u_{ij} = u_{i+1,j}-2u_{ij}+u_{i-1,j} \ .$$

Note that α=0 gives the explicit method (2.13.3), $\alpha=\frac{1}{2}$ gives the implicit method (2.13.14) and for $\alpha=\frac{1}{4}$ we obtain the implicit equation (2.13.15).

2.14 STABILITY ANALYSIS OF THE GENERAL THREE-LEVEL FORMULA

To determine the conditions of stability of the general three level implicit method, we perform a Fourier series analysis on equation (2.13.16) by inserting the error,

$$\varepsilon_{ij} = \xi^j \exp(i_c \beta i \Delta x) \ .$$

This leads to,

$$\frac{1}{(\Delta t)^2}\delta_t^2\varepsilon_{ij} = \frac{1}{(\Delta x)^2}[\alpha\delta_x^2\varepsilon_{i,j+1}+(1-2\alpha)\delta_x^2\varepsilon_{ij}+\alpha\delta_x^2\varepsilon_{i,j-1}]$$

or

$$\xi^{j-1}\exp(i_c\beta i\Delta x)(1-2\gamma+\gamma^2) = \lambda^2[\alpha\xi^{j+1}\exp(i_c\beta i\Delta x)\{\exp(i_c\beta\Delta x)+$$

$$\exp(-i_c\beta\Delta x)-2\}+(1-2\alpha)\xi^j\exp(i_c\beta i\Delta x)\{\exp(i_c\beta\Delta x)+\{\exp(-i_c\beta\Delta x)-2\}$$

$$+\alpha\xi^{j-1}\exp(i_c\beta i\Delta x)\{\exp(i_c\beta\Delta x)+\exp(-i_c\beta\Delta x)-2\}] .$$

After the cancellation and grouping of terms, we get,

$$\begin{aligned}\gamma^2-2\gamma+1 &= \lambda^2(\exp(i_c\beta\Delta x)+\exp(-i_c\beta\Delta x)-2)(\alpha\gamma^2+(1-2\alpha)\gamma+\alpha)\\ &= \lambda^2[2(\cos(\beta\Delta x)-1)\{\alpha\gamma^2+(1-2\alpha)\gamma+\alpha\}] \ .\end{aligned} \qquad (2.14.1)$$

But,

$$\cos(\beta\Delta x)-1 = -2\sin^2(\frac{\beta\Delta x}{2}) \ .$$

Hence (2.14.1) reduces to the quadratic,

$$(1+4\alpha\lambda^2\sin^2(\frac{\beta\Delta x}{2}))\gamma^2+2(2(1-2\alpha)\lambda^2\sin^2(\frac{\beta\Delta x}{2})-1)\gamma+(1+4\alpha\lambda^2\sin^2(\frac{\beta\Delta x}{2})) = 0 .$$

(2.14.2)

The roots of this quadratic are given by

$$\gamma = \frac{-[2(1-2\alpha)\lambda^2\sin^2(\frac{\beta\Delta x}{2})-1]\pm 2\lambda\sin(\frac{\beta\Delta x}{2})\sqrt{\lambda^2\sin^2(\frac{\beta\Delta x}{2})[1-4\alpha]-1}}{(1+4\alpha\lambda^2\sin^2(\frac{\beta\Delta x}{2})} \qquad (2.14.3)$$

We now discuss the stability requirement of the explicit formula (2.13.3). Putting $\alpha=0$ into equations (2.14.2) and (2.14.3) we obtain

$$\gamma^2-2(1-2\lambda^2\sin^2(\frac{\beta\Delta x}{2}))\gamma+1 = 0 \ , \qquad (2.14.4)$$

and $\gamma = -[2\lambda^2\sin^2(\frac{\beta\Delta x}{2})-1]\pm 2\lambda\sin(\frac{\beta\Delta x}{2})\sqrt{\lambda^2\sin^2(\frac{\beta\Delta x}{2})-1}$

or $\gamma = (1-2\lambda^2\sin^2(\frac{\beta\Delta x}{2})\pm\sqrt{4\lambda^4\sin^4(\frac{\beta\Delta x}{2})-4\lambda^2\sin^2(\frac{\beta\Delta x}{2})}$. (2.14.5)

Letting $$A = 1-2\lambda^2\sin^2(\frac{\beta\Delta x}{2}) \ , \qquad (2.14.6)$$

equation (2.14.4) becomes,

$$\gamma^2-2A\gamma+1 = 0 \ , \qquad (2.14.7)$$

and using (2.14.5), the values of γ are

$$\gamma_1 = A+\sqrt{A^2-1} \qquad \text{and} \qquad \gamma_2 = A-\sqrt{A^2-1} \ .$$

As λ,β and Δx are real, then by (2.14.6), $A\leq 1$. When $A<-1$, $|\gamma_2| >1$ giving instability.

When, $$-1\leq A\leq 1,\ A^2\leq 1,\ \gamma_1=A+i_c\sqrt{1-A^2} \text{ and } \gamma_2=A-i_c\sqrt{1-A^2} \ .$$

It is clear that $|\gamma_1| =|\gamma_2|=1$ proving that the explicit method is stable for $-1\leq A\leq 1$. From equation (2.14.6), we then have

$$-1\leq 1-2\lambda^2\sin^2(\frac{\beta\Delta x}{2}) \leq 1.$$

The only useful inequality is $-1\leq 1-2\lambda^2\sin^2(\frac{\beta\Delta x}{2})$, giving $\lambda\leq 1$. Hence (2.13.3) is conditionally stable for $\lambda\leq 1$.

For $\alpha>0$, the stability equation is, from (2.14.2),

$$\gamma^2-2B\gamma +1 = 0 \ , \qquad (2.14.8)$$

where $B=1-[2A'/(1+4A'\alpha)]$ and $A'=\lambda^2\sin^2(\frac{\beta\Delta x}{2})$. We note that (2.14.8) is of the same form as (2.14.7) implying that $|\gamma_1|\leq 1$ and $|\gamma_2|\leq 1$ if and only if $-1\leq B\leq 1$. Thus, $-1\leq 1-2A'/(1+4A'\alpha)\leq 1$ or

$$1 \geq A'/(1+4A'\alpha) \geq 0.$$

Since A' and α are non-negative, the right inequality is trivial. The left inequality yields the two inequalities

$$A' \leq \frac{1}{1-4\alpha} \ ,\alpha \leq \frac{1}{4} \text{ and } A'> \frac{1}{1-4\alpha} \ ,\ \alpha>\frac{1}{4} \ .$$

The second inequality is trivial since $A'=\lambda^2\sin^2(\frac{\beta\Delta x}{2})\geq 0$ always. If we allow $\sin^2(\frac{\beta\Delta x}{2})$ to take on its largest possible value, we obtain from the first inequality,

$$\lambda \leq 1/\sqrt{1-4\alpha}, \quad \alpha<\frac{1}{4} \ .$$

For stability λ must satisfy the above condition. If $\alpha \geq \frac{1}{4}$, stability is obtained for all values of λ.

The implicit methods above clearly need some extra boundary conditions, for example along two lines x=constant, since otherwise we have more unknowns than equations along the new line. We can then use a matrix method for analysing stability which will automatically include the effects of the boundaries.

The application of the matrix method can, perhaps, be better illustrated by first considering a general two-level finite difference scheme approximating a given first-order hyperbolic differential equation with which the initial and boundary values are specified for example at $x=x_0$ and $x=x_m$. In particular if the boundary values are zero then the tridiagonal system of equations generated by the finite difference approximation can be written in the matrix form,

$$A\underline{u}_{j+1} = B\underline{u}_j \qquad (2.14.9)$$

where A and B are square matrices of order (m-1) and $\underline{u}_j$ is a column vector consisting of the u-values along the j-line, that is, $\underline{u}_j=(u_{1,j},u_{2,j},\ldots,u_{m-1,j})^T$. The equation governing stability is (2.14.9); with other than zero boundary conditions a vector will be added to equation (2.14.9) which can, at most, depend upon i, i.e., $A\underline{u}_{j+1} = B\underline{u}_j + \underline{c}$.

The non-singular nature of A allows us to rewrite equation (2.14.9) as,

$$\underline{u}_{j+1} = P\underline{u}_j , \quad P = A^{-1}B \ . \qquad (2.14.10)$$

Upon repeated application of equation (2.14.10) leads to

$$\underline{u}_{j+1} = P\underline{u}_j = P^2\underline{u}_{j-1}=\ldots=P^j\underline{u}_1 = P^{j+1}\underline{u}_0$$

where $\underline{u}_0$ is the vector of initial values. Now suppose we introduce errors at every mesh point along t=0 and start the computation with the vector of values $\underset{\sim}{u}_0^*$ instead of $\underset{\sim}{u}_0$. We shall then calculate

$$\underset{\sim}{u}_1^* = P\underset{\sim}{u}_0^*,\ \underset{\sim}{u}_2^* = P\underset{\sim}{u}_1^* = P^2\underset{\sim}{u}_0^*, \ldots, \underset{\sim}{u}_j^* = P^j\underset{\sim}{u}_0^*$$

where we assume that no further errors are introduced.

If we define the error vector $\underset{\sim}{\varepsilon}$ by

$$\underset{\sim}{\varepsilon} = \underset{\sim}{u} - \underset{\sim}{u}^* ,$$

then $\underset{\sim}{\varepsilon}_j = \underset{\sim}{u}_j - \underset{\sim}{u}_j^* = P^j(\underset{\sim}{u}_0 - \underset{\sim}{u}_0^*) = P^j\underset{\sim}{\varepsilon}_0$.

The finite difference scheme will be stable when $\underset{\sim}{\varepsilon}_j$ remains bounded as j increases indefinitely. This can always be investigated by expressing the initial error vector in terms of the eigenvectors of P. We assume that the matrix P has (m-1) linearly independent eigenvectors $\underset{\sim}{v}_s$, which will always be so if the eigenvalues μ_s of P are all distinct or P is real and symmetric (Hermitian). Then these eigenvectors can be used as a basis for our (m-1)-dimensional vector space and the error vector $\underset{\sim}{\varepsilon}_0$, with its (m-1) components, can be expressed uniquely as a linear combination of them, namely,

$$\underset{\sim}{\varepsilon}_0 = \sum_{s=1}^{m-1} c_s\underset{\sim}{v}_s \quad \text{where the } c_s,\ s=1,\ldots,m-1$$

are known scalars.

The errors along the time-level t=Δt, resulting from the initial perturbations $\underset{\sim}{\varepsilon}_0$, will be given by,

$$\underset{\sim}{\varepsilon}_1 = P\underset{\sim}{\varepsilon}_0 = P\sum_{s=1}^{m-1} c_s\underset{\sim}{v}_s = \sum_{s=1}^{m-1} c_s P\underset{\sim}{v}_s \ .$$

But $P\underset{\sim}{v}_s = \mu_s\underset{\sim}{v}_s$ by the definition of an eigenvalue. Therefore,

$$\underset{\sim}{\varepsilon}_1 = \sum_{s=1}^{m-1} c_s\mu_s\underset{\sim}{v}_s \ .$$

Similarly,

$$\underset{\sim}{\varepsilon}_j = \sum_{s=1}^{m-1} c_s \mu_s^j \underset{\sim}{v}_s .$$

This shows that the errors will not increase exponentially with j provided the eigenvalue with the largest modulus (the spectral radius of P) has a modulus less than or equal to unity. We call P the amplification matrix.

Before we proceed to establish stability, we state the following theorem which is useful for the analysis of three or more time-level difference equations.

Theorem 2.1 Stability of Three or More Time-Level Difference Equations

If the matrix P can be written as

$$P = \begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,m} \\ P_{2,1} & P_{2,2} & \cdots & P_{2,m} \\ \vdots & & & \vdots \\ P_{m,1} & P_{m,2} & \cdots & P_{m,m} \end{bmatrix}$$

where each $P_{i,j}$ is an n×n matrix, and all the P_{ij} have a common set of n linearly independent eigenvectors, then the eigenvalues of P are given by the eigenvalues of the matrices

$$\begin{bmatrix} \eta_{1,1}^{(k)} & \eta_{1,2}^{(k)} & \cdots & \eta_{1,m}^{(k)} \\ \eta_{2,1}^{(k)} & \eta_{2,2}^{(k)} & \cdots & \eta_{2,m}^{(k)} \\ \vdots & & & \\ \eta_{m,1}^{(k)} & \eta_{m,2}^{(k)} & \cdots & \eta_{m,m}^{(k)} \end{bmatrix} , k=1,\ldots,n.$$

where $\eta_{i,j}^{(k)}$ is the k^{th} eigenvalue of P_{ij} corresponding to the k^{th} eigenvector $\underset{\sim}{v}_k$ common to all the P_{ij}'s.

Proof

Let $\underline{v}_k$ be an eigenvector common to all the submatrices P_{ij}, $i,j=1,2,\ldots,m$ and denote the corresponding eigenvalues of $P_{1,1},\ldots,P_{2,1},\ldots$ by $\eta_{1,1}^{(k)},\eta_{2,1}^{(k)},\ldots$ respectively. For simplicity, consider $i,j=1,2$ and denote $\underline{v}_k$ by $\underline{v}$, $\eta_{i,j}^{(k)}$ by $\eta_{i,j}$. Then,

$$P_{1,1}\underline{v} = \eta_{1,1}\underline{v}\ , \quad P_{1,2}\underline{v} = \eta_{1,2}\underline{v}\ ,$$

$$P_{2,1}\underline{v} = \eta_{2,1}\underline{v}\ , \quad P_{2,2}\underline{v} = \eta_{2,2}\underline{v}\ .$$

We multiply these equations respectively by the non-zero constants α_1 and α_2 and write them as

$$\begin{bmatrix} P_{1,1} & P_{1,2} \\ P_{2,1} & P_{2,2} \end{bmatrix} \begin{bmatrix} \alpha_1\underline{v} \\ \alpha_2\underline{v} \end{bmatrix} = \begin{bmatrix} (\eta_{1,1}\alpha_1+\eta_{1,2}\alpha_2)\underline{v} \\ (\eta_{2,1}\alpha_1+\eta_{2,2}\alpha_2)\underline{v} \end{bmatrix} . \qquad (2.14.11)$$

Let us assume that

$$P = \begin{bmatrix} P_{1,1} & P_{1,2} \\ P_{2,1} & P_{2,2} \end{bmatrix}$$

has an eigenvalue μ corresponding to the eigenvector

$$\begin{bmatrix} \alpha_1\underline{v} \\ \alpha_2\underline{v} \end{bmatrix}$$

so that

$$\begin{bmatrix} P_{1,1} & P_{1,2} \\ P_{2,1} & P_{2,2} \end{bmatrix} \begin{bmatrix} \alpha_1\underline{v} \\ \alpha_2\underline{v} \end{bmatrix} = \mu \begin{bmatrix} \alpha_1\underline{v} \\ \alpha_2\underline{v} \end{bmatrix} . \qquad (2.14.12)$$

By the right-hand sides of equations (2.14.11) and (2.14.12),

$$(\eta_{1,1}-\mu)\alpha_1 + \eta_{1,2}\alpha_2 = 0$$

and $$\eta_{2,1}\alpha_1+(\eta_{2,2}-\mu)\alpha_2 = 0.$$

These two equations will have a non-trivial solution for α_1 and α_2 if and only if

$$\det \begin{bmatrix} (\eta_{1,1}-\mu) & \eta_{1,2} \\ \eta_{2,1} & (\eta_{2,2}-\mu) \end{bmatrix} = 0 ,$$

i.e. if and only if μ is an eigenvalue of the matrix

$$\begin{bmatrix} \eta_{1,1} & \eta_{1,2} \\ \eta_{2,1} & \eta_{2,2} \end{bmatrix}$$

We are now in a position to investigate the stability conditions that equation (2.13.16) has to fulfil. In matrix form, (2.13.16) may be written as,

$$A\underset{\sim}{u}_{j+1} = B\underset{\sim}{u}_j + C\underset{\sim}{u}_{j-1} + \underset{\sim}{b}_j , \qquad (2.14.13)$$

where,

$$A = (1+2\alpha\lambda^2)I-\alpha\lambda^2E ,$$

$$B = 2(1-(1-2\alpha)\lambda^2)I + (1-2\alpha)\lambda^2E \qquad (2.14.14)$$

and

$$C = (-1-2\alpha\lambda^2)I+\alpha\lambda^2E.$$

E is the matrix with 1's along each diagonal immediately above and below the main diagonal and zeros elsewhere, and $\underset{\sim}{b}_j$ is a column vector of known constants (boundary values). From (2.14.13) we have

$$\underset{\sim}{u}_{j+1} = A^{-1}B\underset{\sim}{u}_j + A^{-1}C\underset{\sim}{u}_{j-1} + A^{-1}\underset{\sim}{b}_j \qquad (2.14.15)$$

Therefore, a perturbation $\underset{\sim}{\varepsilon}_0$ of the initial values will satisfy

$$\underset{\sim}{\varepsilon}_{j+1} = A^{-1}B\underset{\sim}{\varepsilon}_j + A^{-1}C\underset{\sim}{\varepsilon}_{j-1} .$$

Hence,

$$\begin{bmatrix} \underset{\sim}{\varepsilon}_{j+1} \\ \underset{\sim}{\varepsilon}_j \end{bmatrix} = \begin{bmatrix} A^{-1}B & A^{-1}C \\ I & 0 \end{bmatrix} \begin{bmatrix} \underset{\sim}{\varepsilon}_j \\ \underset{\sim}{\varepsilon}_{j-1} \end{bmatrix}$$

i.e. $\underset{\sim}{v}_{j+1} = P\underset{\sim}{v}_j$. The matrices A,B and C have the same system of linearly independent eigenvectors as E. So have the matrices $A^{-1}B$ and $A^{-1}C$. Therefore, applying Theorem 2.1, the eigenvalues μ of P are given by,

$$\det \begin{bmatrix} (a_k^{-1}b_k-\mu) & a_k^{-1}c_k \\ 1 & -\mu \end{bmatrix} = 0, \; k=1,2,\ldots,(m-1), \qquad (2.14.16)$$

where a_k, b_k and c_k are the eigenvalues of A,B,C respectively. We note from (2.14.14) that each of the matrices A,B and C is of a common tridiagonal form and therefore from Section 1.6, we find that,

$$\begin{aligned} a_k &= 1+4\alpha\lambda^2\sin^2(\tfrac{k\pi}{2m}) \\ b_k &= 2-4(1-2\alpha)\lambda^2\sin^2(\tfrac{k\pi}{2m}) \;, \quad k=1,2,\ldots,m-1 \\ c_k &= -1-4\alpha\lambda^2\sin^2(\tfrac{k\pi}{2m}) \;. \end{aligned}$$

Using equation (2.14.16), we have,

$$a_k\mu^2-b_k\mu-c_k = 0 \text{ or}$$

$$(1+4\alpha\lambda^2\sin^2(\tfrac{k\pi}{2m}))\mu^2+2(2(1-2\alpha)\lambda^2\sin^2(\tfrac{k\pi}{2m})-1)\mu+(1+4\alpha\lambda^2\sin^2(\tfrac{k\pi}{2m})) = 0 \qquad (2.14.17)$$

which is exactly of the form (2.14.2). To avoid repetition of the mathematics involved to solve (2.14.17), we conclude that $\mu = 1$ if $\alpha \gtrsim 1/4$ and for all values of $\lambda = \Delta t/\Delta x$. Hence the implicit equation (2.13.16) is *always stable*. We observe that the stability criterion has not been changed by the incorporation of the boundary values, and in general different types of boundary conditions have only negligible effect.

2.15 TRUNCATION ERROR ANALYSIS OF THE GENERAL THREE-LEVEL FORMULA

As before, by virtue of equation (2.13.16), we have,

$$F(u_{ij}) = \delta_t^2 u_{ij} - \lambda^2 [\alpha\delta_x^2 u_{i,j+1} + (1-2\alpha)\delta_x^2 u_{ij} + \alpha\delta_x^2 u_{i,j-1}] \ ,$$

$$F(U_{ij}) = \delta_t^2 U_{ij} - \lambda^2 [\alpha\delta_x^2 U_{i,j+1} + (1-2\alpha)\delta_x^2 U_{ij} + \alpha\delta_x^2 U_{i,j-1}]$$

and $T = F(U_{ij})$.

If we expand $U_{i-1,j-1}, U_{i-1,j}, U_{i-1,j+1}, U_{i,j-1}, U_{i,j+1}, U_{i+1,j-1}$, $U_{i+1,j}$ and $U_{i+1,j+1}$ about the point (x_i, t_j), we find that

$$\text{(a)} \quad \delta_t^2 U_{ij} = (\Delta t)^2 \left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j} + \frac{2(\Delta t)^4}{4!} \left(\frac{\partial^4 U}{\partial t^4}\right)_{i,j} + \frac{2(\Delta t)^6}{6!} \left(\frac{\partial^6 U}{\partial t^6}\right) + \ldots$$

$$\text{(b)} \quad \delta_x^2 U_{i,j+1} = (\Delta x)^2 \left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} + (\Delta x)^2 (\Delta t) \left(\frac{\partial^3 U}{\partial x^2 \partial t}\right)_{i,j} + \frac{1}{12}(\Delta x)^4 \left(\frac{\partial^4 U}{\partial x^4}\right)_{i,j}$$
$$+ \tfrac{1}{2}(\Delta x)^2 (\Delta t)^2 \left(\frac{\partial^4 U}{\partial x^2 \partial t^2}\right)_{i,j} + \ldots$$

$$\text{(c)} \quad \delta_x^2 U_{ij} = (\Delta x)^2 \left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} + \frac{1}{12} (\Delta x)^4 \left(\frac{\partial^4 U}{\partial x^4}\right)_{i,j} + \ldots$$

$$\text{(d)} \quad \delta_x^2 U_{i,j-1} = (\Delta x)^2 \left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} - (\Delta x)^2 (\Delta t) \left(\frac{\partial^3 U}{\partial x^2 \partial t}\right)_{i,j} + \frac{1}{12}(\Delta x)^4 \left(\frac{\partial^4 U}{\partial x^4}\right)_{i,j}$$
$$+ \tfrac{1}{2}(\Delta x)^2 (\Delta t)^2 \left(\frac{\partial^4 U}{\partial x^2 \partial t^2}\right)_{i,j} + \ldots$$

Hence,

$$T = (\Delta t)^2 \left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j} + \frac{(\Delta t)^4}{12} \left(\frac{\partial^4 U}{\partial t^4}\right)_{i,j} + \ldots$$
$$-\lambda^2 \left\{ (\Delta x)^2 \left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} + \frac{(\Delta x)^4}{12} \left(\frac{\partial^4 U}{\partial x^4}\right)_{i,j} + \alpha(\Delta x)^2 (\Delta t)^2 \left(\frac{\partial^4 U}{\partial x^2 \partial t^2}\right)_{i,j} + \ldots \right\}$$
$$= (\Delta t)^2 \left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j} + \frac{(\Delta t)^4}{12} \left(\frac{\partial^4 U}{\partial t^4}\right)_{i,j} - (\Delta t)^2 \left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} - \frac{1}{12}(\Delta t)^2$$
$$(\Delta x)^2 \left(\frac{\partial^4 U}{\partial x^4}\right)_{i,j} - \alpha(\Delta t)^4 \left(\frac{\partial^4 U}{\partial x^2 \partial t^2}\right)_{i,j} + \ldots$$
$$= (\Delta t)^2 \left\{ \left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j} - \left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} \right\} - \frac{1}{12}(\Delta t)^2 (\Delta x)^2 \left(\frac{\partial^4 U}{\partial x^4}\right)_{i,j} - \alpha(\Delta t)^4$$
$$\left(\frac{\partial^4 U}{\partial x^2 \partial t^2}\right) + \frac{(\Delta t)^2}{12} \left(\frac{\partial^4 U}{\partial t^4}\right)_{i,j} + \ldots$$

From the given differential equation, $(\frac{\partial^2 U}{\partial t^2})_{i,j} - (\frac{\partial^2 U}{\partial x^2})_{i,j} = 0$ and $(\frac{\partial^4 U}{\partial t^4})_{i,j} = (\frac{\partial^4 U}{\partial x^2 \partial t^2})_{i,j} = (\frac{\partial^4 U}{\partial x^4})_{i,j}$. Therefore,

$$T = - (\Delta t)^2 \{\frac{(\Delta x)^2}{12} (1+(12\alpha-1)\lambda^2) (\frac{\partial^4 U}{\partial x^4})_{i,j} + \ldots\}$$

from which the principal part is $\frac{(\Delta t)^2}{12} (\Delta x)^2 (1+(12\alpha-1)\lambda^2) (\frac{\partial^4 U}{\partial x^4})_{ij}$.

We deduce that for $\alpha=0$, 1/4 and 1/2,

$$T = O([\Delta x]^2) + O([\Delta t]^2) \ .$$

In fact, for the explicit formula (with $\alpha=0$) the truncation error is

$$T = - \frac{(\Delta t)^2 (\Delta x)^2}{12} [(1-\lambda^2) (\frac{\partial^4 U}{\partial x^4})_{i,j} + \frac{1}{30}(\Delta x)^2 (1-\lambda^4) (\frac{\partial^6 U}{\partial x^6})_{i,j} + \ldots]$$

It vanishes completely when $\lambda=1$, and so the difference formula

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

is *an exact difference representation* of the wave equation.

2.16 OTHER APPROXIMATIONS FOR THE WAVE EQUATION

Mitchell (1969) uses the following method for deriving implicit approximations to the wave equation (2.13.1) that are accurate to fourth-order differences. If $U_{i,j+1}$ and $U_{i,j-1}$ are expanded about the point (x_i,t_j) by Taylor's series, we get

$$U_{i,j+1}-2U_{i,j}+U_{i,j-1} = (\Delta t)^2(\frac{\partial^2 U}{\partial t^2})_{i,j} + \frac{1}{12}(\Delta t)^4(\frac{\partial^4 U}{\partial t^4})_{i,j}+\ldots$$

If U is a solution of the wave equation, then

$$(\frac{\partial^4 U}{\partial x^4})_{i,j} = (\frac{\partial^4 U}{\partial t^4})_{i,j}\ ,\ (\frac{\partial^6 U}{\partial x^6})_{i,j} = (\frac{\partial^6 U}{\partial t^6})_{i,j}\ ,\ \ldots\ .$$

Hence,

$$U_{i,j+1}-2U_{ij}+U_{i,j-1} = (\Delta t)^2(\frac{\partial^2 U}{\partial x^2})_{i,j} + \frac{1}{12}\ (\Delta t)^4(\frac{\partial^4 U}{\partial x^4})_{i,j}+\ \ldots \tag{2.16.1}$$

From the central difference formula (cf. Hildebrand (1956)),

$$\frac{\partial^2 U}{\partial x^2} = \frac{1}{(\Delta x)^2}\ (\delta_x^2 - \frac{1}{12}\delta_x^4 + \frac{1}{90}\ \delta_x^6 + \ldots)U\ .$$

it follows that for fourth-order differences,

$$\frac{\partial^2 U}{\partial x^2} = \frac{1}{(\Delta x)^2}\ (\delta_x^2 - \frac{1}{12}\ \delta_x^4)U$$

and

$$\frac{\partial^4 U}{\partial x^4} = \frac{\partial^2}{\partial x^2}(\frac{\partial^2 U}{\partial x^2}) = \frac{1}{(\Delta x)^4}\ \delta_x^4\ U.$$

By substituting these approximations into equation (2.16.1), we find that to this order of accuracy,

$$U_{i,j+1}-2U_{ij}+U_{i,j-1} = \lambda^2\ \{1+ \frac{1}{12}(\lambda^2-1)\delta_x^2\}\delta_x^2 U_{i,j}, \tag{2.16.2}$$

where $\lambda = \frac{\Delta t}{\Delta x}$. Now, if we operate on both sides of this equation with $\{1+ \frac{1}{12}(\lambda^2-1)\delta_x^2\}^{-\frac{1}{2}}$ and expand each operator up to terms in δ_x^4 by the binomial expansion, we arrive at the following implicit difference approximation,

$$u_{i,j+1}-2u_{ij}+u_{i,j-1} = \frac{1}{24}(\lambda^2-1)(\delta_x^2 u_{i,j+1}+\delta_x^2 u_{i,j-1}) + \frac{1}{12}(11\lambda^2+1)\delta_x^2 u_{ij}$$

$$+ \frac{1}{192}(\lambda^2-1)(9\lambda^2-1)\delta_x^4 u_{ij} - \frac{1}{384}(\lambda^2-1)^2(\delta_x^4 u_{i,j+1}+\delta_x^4 u_{i,j-1}) . \tag{2.16.3}$$

Similarly, if both sides of (2.16.2) are operated on by $\{1+\frac{1}{12}(\lambda^2-1)\delta_x^2\}^{-1}$, the corresponding difference equation is,

$$u_{i,j+1}-2u_{ij}+u_{i,j-1} = \frac{1}{12}(\lambda^2-1)(\delta_x^2 u_{i,j+1}+\delta_x^2 u_{i,j-1}) + \frac{1}{6}(5\lambda^2+1)\delta_x^2 u_{ij}$$

$$+ \frac{1}{144}(\lambda^2-1)^2\{2\delta_x^4 u_{ij}-(\delta_x^4 u_{i,j+1}+\delta_x^4 u_{i,j-1})\} . \tag{2.16.4}$$

These high-order difference approximations would be difficult to implement in practice because of the problems associated with the boundary conditions.

von Neumann (cf. O'Brien *et al.* (1951)) introduced the difference equation,

$$\frac{\delta_t^2 u_{i,j}}{(\Delta t)^2} = \frac{\delta_x^2 u_{i,j}}{(\Delta x)^2} + \omega\left[\frac{1}{(\Delta x)^2(\Delta t)^2}\,\delta_t^2\delta_x^2 u_{i,j}\right] \tag{2.16.5}$$

to solve the wave equation (2.13.1). Except for $\omega=0$, (2.16.5) is an implicit equation whose solution at each time step is obtained by solving a tridiagonal system of linear equations. For $\omega=0$, the equation reduces to the classical explicit method (2.13.3). von Neumann proved that (2.16.5) is unconditionally stable if $\omega>1/4$ and is conditionally stable if $\omega\leq 1/4$, the stability condition in the latter case being $\lambda\leq\frac{1}{(1-4\omega)^{\frac{1}{2}}}$.

Friberg (1961) and Lees (1960) generalised this result to quasi-linear hyperbolic equations of the form,

$$\frac{\partial^2 U}{\partial t^2} = a(x,t)\,\frac{\partial^2 U}{\partial x^2} + F(x,t,U,\frac{\partial U}{\partial x}, \frac{\partial U}{\partial t}) . \tag{2.16.6}$$

The same result can even be extended to *certain linear multi-dimensional* systems. However, the linear equations that arise are no longer tridiagonal. To overcome this problem, Lees (1962) proposed modifications to equation (2.16.5) by applying *an alternating direction procedure* and employing *the energy method*, he showed that the modified schemes are unconditionally stable if $\omega > 1/4$.

2.17 SIMULTANEOUS FIRST-ORDER EQUATIONS

All partial differential equations of second and higher order in the independent variable, t, say, can be reduced to a system of simultaneous equations of first order in t which may then be approximated by stable and convergent difference schemes. Convergence is usually satisfied, as in the second-order problem, if the mesh ratio is chosen such that the region of finite difference determination lies completely within that of the differential equation. Stability can be examined by either the Fourier or the matrix method.

Let us consider the wave equation (2.13.1),

$$\frac{\partial^2 U}{\partial x^2} = \frac{\partial^2 U}{\partial t^2} .$$

If we put $p=\frac{\partial U}{\partial x}$ and $q=\frac{\partial U}{\partial t}$, then

$$\frac{\partial q}{\partial t} = \frac{\partial p}{\partial x} \text{ and } \frac{\partial p}{\partial t} = \frac{\partial q}{\partial x} . \qquad (2.17.1)$$

Suppose that the initial conditions specify p and q on the segment $0<x<1$ of the initial line $t=0$ and that we seek to find p and q at other points in the *region of determinacy*. The characteristics are $x-t=$ constant, $x+t=$constant and the region of determinacy in the region of positive t is bounded by the lines $t=0$, $x-t=0$, $x+t=1$ (Fig. 2.17.1).

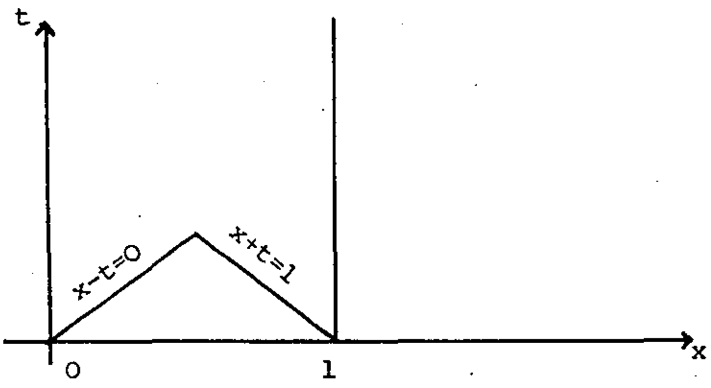


FIGURE 2.17.1: Region of determinacy of initial conditions

Now, to develop a simple explicit method, for example, we can approximate the t-derivatives by the first forward difference in the t-direction. In the x-direction, however, we have a choice between forward, backward and central differences. To remain within the region of determinacy, a backward difference must not be employed near the line x=t, nor a forward difference near x+t=1. On the other hand, central differences are *apparently* satisfactory everywhere inside the region of determinacy. To continue the solution, as t increases, boundary values of p or q have to be specified on the lines x=0 and x=1, for t>0.

One explicit method to approximate (2.17.1) is

$$\frac{1}{\Delta t}(q_{i,j+1}-q_{i,j}) = \frac{1}{2\Delta x}(p_{i+1,j}-p_{i-1,j})$$

and (2.17.2)

$$\frac{1}{\Delta t}(p_{i,j+1}-p_{i,j}) = \frac{1}{2\Delta x}(q_{i+1,j}-q_{i-1,j}).$$

Here the x-derivatives are approximated by central differences and the t-derivatives by forward differences. We now investigate the stability of this method using the von Neumann criterion. Let the initial perturbations in p and q along t=0 be $A\exp(i_c\beta x)$ and $B\exp(i_c\beta x)$ respectively, where A and B are different constants, $x=i\Delta x$ and $i_c=\sqrt{-1}$. Then we can assume that the perturbations ε_{ij} in $p_{i,j}$ and $\eta_{i,j}$ in $q_{i,j}$ are given by,

$$\varepsilon_{i,j} = A\exp(i_c\beta i\Delta x)\xi^j$$

and

$$\eta_{i,j} = B\exp(i_c\beta i\Delta x)\xi^j \ . \qquad (2.17.3)$$

If we substitute these into (2.17.2), we get,

$$2B(\gamma-1) = \lambda A(\exp(i_c\beta\Delta x)-\exp(-i_c\beta\Delta x)$$

and

$$2A(\gamma-1) = \lambda B(\exp(i_c\beta\Delta x)-\exp(-i_c\beta\Delta x)).$$

By eliminating A and B leads us to the quadratic,

$$(\gamma-1)^2 = -\lambda^2\sin^2(\beta\Delta x) \ .$$

Hence,
$$\gamma = 1\pm i_c\lambda\sin(\beta\Delta x)$$
and
$$|\gamma| = (1+\lambda^2\sin^2(\beta\Delta x))^{\frac{1}{2}}.$$

It is clear that we cannot have stability for any non-zero value of λ. This demonstrates that care must be taken in adopting apparently reasonable schemes for first-order equations.

A modification of (2.17.2) is the following explicit formula,

$$\frac{1}{\Delta t}[q_{i,j+1}-\tfrac{1}{2}(q_{i+1,j}+q_{i-1,j})] = \frac{1}{2\Delta x}(p_{i+1,j}-p_{i-1,j})$$

and
$$\frac{1}{\Delta t}[p_{i,j+1}-\tfrac{1}{2}(p_{i+1,j}+p_{i-1,j})] = \frac{1}{2\Delta x}(q_{i+1,j}-q_{i-1,j}) \qquad (2.17.4)$$

in which $q_{i,j}$ is replaced by the mean of $q_{i+1,j}$ and $q_{i-1,j}$. A similar replacement is made for p_{ij} in the second equation. To examine its stability, we substitute ε_{ij} and η_{ij} of (2.17.3) into (2.17.4). On eliminating A and B from these two equations results in

$$\gamma = \cos(\beta\Delta x)\pm i_c\lambda\sin(\beta\Delta x).$$

Hence,
$$|\gamma|^2 = \cos^2(\beta\Delta x)+\lambda^2\sin^2(\beta\Delta x)$$
$$\leqslant 1 \text{ for } \lambda\leqslant 1,$$

which implies that (2.17.4) is conditionally stable for $\lambda\leqslant 1$.

If we use central differences in the t-direction, then (2.17.2) can be replaced by,

$$\frac{q_{i,j+1}-q_{i,j-1}}{2\Delta t} = \frac{p_{i+1,j}-p_{i-1,j}}{2\Delta x}$$

and
$$\frac{p_{i,j+1}-p_{i,j-1}}{2\Delta t} = \frac{q_{i+1,j}-q_{i-1,j}}{2\Delta x} \ . \qquad (2.17.5)$$

The governing equation for stability is

$$(\gamma-\frac{1}{\gamma})^2 = -4\lambda^2\sin^2(\beta\Delta x)$$

and as above, we again have conditional stability for $\lambda \leq 1$.

To avoid difference quotients over the double interval $2\Delta x$, the use of midpoints of the interval is recommended for one of the functions. Thus, for example, *the scheme of Courant, Friedrichs and Lewy,*

$$\frac{1}{\Delta t}(q_{i,j+1}-q_{ij}) = \frac{1}{\Delta x}(p_{i+\frac{1}{2},j}-p_{i-\frac{1}{2},j})$$

and

$$\frac{1}{\Delta t}(p_{i-\frac{1}{2},j+1}-p_{i-\frac{1}{2},j}) = \frac{1}{\Delta x}(q_{i,j+1}-q_{i-1,j+1}) \tag{2.17.6}$$

is *equivalent* to the usual explicit scheme (2.13.3), if one identifies

$$p_{i-\frac{1}{2},j} = (u_{ij}-u_{i-1,j})/\Delta x \tag{2.17.7}$$

and

$$q_{ij} = (u_{ij}-u_{i,j-1})/\Delta t \ . \tag{2.17.8}$$

Equation (2.17.6) is recognised as a central divided difference approximation for $\frac{\partial u}{\partial x}$ at a midpoint while (2.17.8) is a backward divided difference for $\frac{\partial u}{\partial t}$ at a mesh point.

If we insert the perturbations $\varepsilon_{i,j}$ and η_{ij} of (2.17.3) into (2.17.6), we get the following set of equations,

$$B(\gamma-1) = \lambda A[e^{i}c^{(\frac{1}{2}\beta\Delta x)} - e^{-i}c^{(\frac{1}{2}\beta\Delta x)}] \tag{2.17.9}$$

and

$$A(\gamma-1) = \lambda\gamma B[e^{i}c^{(\frac{1}{2}\beta\Delta x)} - e^{-i}c^{(\frac{1}{2}\beta\Delta x)}] \ . \tag{2.17.10}$$

By eliminating A and B from these two equations, we finally obtain $(\gamma-1)^2=-4\lambda^2\sin^2(\frac{1}{2}\beta\Delta x)\gamma$, i.e.,

$$\gamma^2+(4\lambda^2\sin^2(\tfrac{1}{2}\beta\Delta x)-2)\gamma+1 = 0.$$

On solving this quadratic, we get,

$$\gamma = 1-2\lambda^2\sin^2(\tfrac{1}{2}\beta\Delta x)\pm 2\lambda\sin(\tfrac{1}{2}\beta\Delta x)\sqrt{\lambda^2\sin^2(\tfrac{1}{2}\beta\Delta x)-1} \ . \tag{2.17.11}$$

We note that (2.17.11) is of the same form as (2.14.4). Therefore, following the same line of argument as before, we get restricted stability for $\lambda \leq 1$ as we should expect due to the equivalence of (2.17.6) to (2.13.3).

Other explicit methods have been employed to produce stability for larger values of λ. In particular, we have

$$\begin{aligned} \frac{1}{\Delta t}(q_{i,j+1}-q_{ij}) &= \frac{1}{2\Delta x}(p_{i+1,j}-p_{i-1,j}) \\ \frac{1}{\Delta t}(p_{i,j+1}-p_{ij}) &= \frac{1}{2\Delta x}(q_{i+1,j+1}-q_{i-1,j+1}) \ . \end{aligned} \tag{2.17.12}$$

The stability equation is $(\gamma-1)^2+\lambda^2\sin^2(\beta\Delta x)\gamma=0$ which simplifies to, $\gamma^2+(\lambda^2\sin^2(\beta\Delta x)-2)\gamma+1=0$ whose roots are given by,

$$\gamma = \frac{-(\lambda^2\sin^2(\beta\Delta x)-2)\lambda\sin(\beta\Delta x)\sqrt{(\lambda\sin(\beta\Delta x)-2)(\lambda\sin(\beta\Delta x)+2)}}{2}$$

We note that $\lambda\sin(\beta\Delta x)+2\geq 0$ and if

$$\lambda\sin(\beta\Delta x) \leq 2 \ , \tag{2.17.12a}$$

then,

$$\gamma = \frac{-(\lambda^2\sin^2(\beta\Delta x)-2)\pm i_c\lambda\sin(\beta\Delta x)\sqrt{(2-\lambda\sin(\beta\Delta x)(2+\lambda\sin(\beta\Delta x))}}{2}$$

and $|\gamma|=1$ giving stability. From equation (2.17.12a), we deduce that the method is stable for $\lambda\leq 2$.

An implicit scheme to approximate the first-order equations of the wave equation is given by

$$\begin{aligned} q_{i,j+1}-q_{ij} &= \frac{\lambda}{2}(p_{i+\frac{1}{2},j-1}-p_{i-\frac{1}{2},j-1}+p_{i+\frac{1}{2},j+1}-p_{i-\frac{1}{2},j+1}) \\ p_{i-\frac{1}{2},j+1}-p_{i-\frac{1}{2},j-1} &= \lambda(q_{i,j+1}-q_{i-1,j+1}+q_{ij}-q_{i-1,j}) \end{aligned} \tag{2.17.13}$$

If we identify,

$$p_{i-\frac{1}{2},j} = \frac{(u_{ij}-u_{i-1,j})}{\Delta x} \quad \text{and} \quad q_{ij} = \frac{(u_{ij}-u_{i,j-1})}{\Delta t}$$

then (2.17.13) is equivalent to (2.13.14) and we therefore expect stability for all values of λ.

We might also use an implicit scheme such as

$$\begin{aligned} q_{i,j+1}-q_{ij} &= \frac{\lambda}{4}(p_{i+1,j+1}-p_{i-1,j+1}+p_{i+1,j}-p_{i-1,j}) \\ p_{i,j+1}-p_{ij} &= \frac{\lambda}{4}(q_{i+1,j+1}-q_{i-1,j+1}+q_{i+1,j}-q_{i-1,j}) \ . \end{aligned} \tag{2.17.14}$$

The equation governing stability is,

$$(\gamma-1)^2 = -\frac{1}{4}\lambda^2(\gamma+1)^2 \sin^2(\beta\Delta x)$$

or

$$(1+\frac{1}{4}\lambda^2\sin^2(\beta\Delta x))\gamma^2+(\tfrac{1}{2}\lambda^2\sin^2(\beta\Delta x)-2)\gamma+(1+\frac{1}{4}\lambda^2\sin^2(\beta\Delta x)) = 0.$$

Hence,

$$\gamma = \frac{-2(1+\frac{1}{4}\lambda^2\sin^2(\beta\Delta x))\pm i_c 2\lambda\sin(\beta\Delta x)}{2(1+\frac{1}{4}\lambda^2\sin^2(\beta\Delta x))}$$

and $|\gamma|^2=1$ giving stability for all λ.

A system of first-order equations can also be written in the more general form as

$$\frac{\partial \underline{U}}{\partial t} + A\frac{\partial \underline{U}}{\partial x} = 0 , \qquad (2.17.15)$$

where A is an $(m\times m)$ real matrix and $\underline{U}$ is an m–component column vector. Initially A is assumed to be constant. We consider the case where A has all real eigenvalues and m linearly independent eigenvectors, so that the system is hyperbolic. A is not necessarily a symmetric matrix.

The explicit and implicit finite difference approximations developed in the previous sections for the case of a *scalar coefficient* carry over in an obvious manner when A is a constant matrix. For example, the Lax-Wendroff method is now

$$\underline{u}_{i,j+1} = [I-\tfrac{1}{2}\lambda A(\Delta_x+\nabla_x) + \tfrac{1}{2}\lambda^2 A^2(\Delta_x-\nabla_x)]\underline{u}_{i,j} \qquad (2.17.16)$$

and the Wendroff's implicit formula is

$$[I+\tfrac{1}{2}(I+\lambda A)\Delta_x]\underline{u}_{i,j+1} = [I+\tfrac{1}{2}(I-\lambda A)\Delta_x]\underline{u}_{i,j} \qquad (2.17.17)$$

where

$$\Delta_x\underline{u}_{i,j} = \underline{u}_{i+1,j}-\underline{u}_{i,j}, \quad \nabla_x\underline{u}_{i,j} = \underline{u}_{i,j}-\underline{u}_{i-1,j}$$

and I is the unit matrix of order m.

If A, however, depends on x or x and t then equations (2.17.16) and (2.17.17) require modification to maintain second-order accuracy.

We shall consider two cases of variable coefficients:

(i) *A Depending on x.*

The Lax-Wendroff and Wendroff formulae are written in the forms,

$$\underline{u}_{i,j+1} = [I+\tfrac{1}{2}\lambda A_{i+\frac{1}{2}}(\Delta_x+\nabla_x)+\frac{1}{4}\lambda^2(A_{i+\frac{1}{2}}\Delta_x A_{i+\frac{1}{2}}\nabla_x+A_{i+\frac{1}{2}}\nabla_x A_{i+\frac{1}{2}})]\underline{u}_{i,j} \qquad (2.17.16a)$$

and

$$[I+\tfrac{1}{2}(I-\lambda A_{i+\frac{1}{2}})\Delta_x]\underline{u}_{i,j+1} = [I+\tfrac{1}{2}(I+\lambda A_{i+\frac{1}{2}})\Delta_x]\underline{u}_{i,j} \qquad (2.17.17a)$$

respectively, where $A_{i+\frac{1}{2}}$ implies that A is evaluated at $x=(i+\frac{1}{2})\Delta x$.

(ii) *A Depending on x and t.*

The required forms of the Lax-Wendroff and Wendroff formulae are

$$\underline{u}_{i,j+1} = [I+\tfrac{1}{2}\lambda A_{i,j+\frac{1}{2}}(\Delta_x+\nabla_x)+\frac{1}{4}\lambda^2(A_{i,j+\frac{1}{2}}\Delta_x A_{i,j+\frac{1}{2}}\nabla_x+A_{i,j+\frac{1}{2}}\nabla_x A_{i,j+\frac{1}{2}}\Delta_x)]\underline{u}_{i,j} \qquad (2.17.16b)$$

and

$$[I+\tfrac{1}{2}(I-\lambda A_{i+\frac{1}{2},j+\frac{1}{2}})\Delta_x]\underline{u}_{i,j+1} = [I+\tfrac{1}{2}(I+\lambda A_{i+\frac{1}{2},j+\frac{1}{2}})\Delta_x]\underline{u}_{i,j} \qquad (2.17.17b)$$

respectively, where $A_{i,j+\frac{1}{2}}$ denotes the evaluation of A at $x=i\Delta x$, $t=(j+\frac{1}{2})\Delta t$ and $A_{i+\frac{1}{2},j+\frac{1}{2}}$ implies A is evaluated at $x=(i+\frac{1}{2})\Delta x$, $t=(j+\frac{1}{2})\Delta t$. It is also interesting to see that the centred-in-distance and centred-in-time equation (the Crank-Nicolson type) takes the form,

$$[I-\frac{1}{4}\lambda A_{i,j+\frac{1}{2}}(\Delta_x+\nabla_x)]\underline{u}_{i,j+1} = [I+\frac{1}{4}\lambda A_{i,j+\frac{1}{2}}(\Delta_x+\nabla_x)]\underline{u}_{i,j}\ , \qquad (2.17.18)$$

which can be shown to possess second-order accuracy. Its local truncation error is given by,

$$\begin{aligned}\underline{T} &= F(\underline{U}_{i,j})\\ &= [I-\frac{1}{4}\lambda A_{i,j+\frac{1}{2}}(\Delta_x+\nabla_x)]\underline{U}_{i,j+1}-[I+\frac{1}{4}\lambda A_{i,j+\frac{1}{2}}(\Delta_x+\nabla_x)]\underline{U}_{i,j}\\ &= [\underline{U}_{i,j+1}-\underline{U}_{i,j}]-\frac{1}{4}\lambda A_{i,j+\frac{1}{2}}(\Delta_x+\nabla_x)[\underline{U}_{i,j+1}+\underline{U}_{i,j}]\ .\end{aligned}$$

By using the Taylor's expansion about the point $(i\Delta x, j\Delta t)$, we find that

$$(\Delta_x+\nabla_x)\underline{U}_{i,j+1} = 2(\Delta x)\left(\frac{\partial \underline{U}}{\partial x}\right)_{i,j+1}$$

$$= 2(\Delta x)\left[\left(\frac{\partial \underline{U}}{\partial x}\right)_{i,j} + \Delta t\left(\frac{\partial^2 \underline{U}}{\partial x\partial t}\right)_{i,j}\right] + \dots$$

and

$$A_{i,j+\frac{1}{2}} = A_{i,j} + \tfrac{1}{2}\Delta t\left(\frac{\partial A}{\partial t}\right)_{i,j} + \frac{1}{8}(\Delta t)^2\left(\frac{\partial^2 A}{\partial t^2}\right)_{i,j} + \dots$$

Therefore,

$$\frac{1}{4}\lambda A_{i,j+\frac{1}{2}}(\Delta_x+\nabla_x)\underline{U}_{i,j+1} = \tfrac{1}{2}\Delta t\left(A\frac{\partial \underline{U}}{\partial x} + \Delta t A\frac{\partial^2 \underline{U}}{\partial x\partial t} + \tfrac{1}{2}\Delta t\frac{\partial A}{\partial t}\frac{\partial \underline{U}}{\partial x}\right)_{i,j} + \dots$$

and

$$\frac{1}{4}\lambda A_{i,j+\frac{1}{2}}(\Delta_x+\nabla_x)\underline{U}_{i,j} = \tfrac{1}{2}\Delta t\left(A\frac{\partial \underline{U}}{\partial x} + \tfrac{1}{2}\Delta t\frac{\partial A}{\partial t}\frac{\partial \underline{U}}{\partial x}\right)_{i,j} + \dots$$

If we add the two equations, we get,

$$\frac{1}{4}\lambda A_{i,j+\frac{1}{2}}(\Delta_x+\nabla_x)[\underline{U}_{i,j+1}+\underline{U}_{i,j}] = \left(\Delta t A\frac{\partial \underline{U}}{\partial x} + \tfrac{1}{2}(\Delta t)^2 A\frac{\partial^2 \underline{U}}{\partial x\partial t} + \tfrac{1}{2}(\Delta t)^2\frac{\partial A}{\partial t}\frac{\partial \underline{U}}{\partial x}\right)_{i,j} + \dots$$

Also,

$$\underline{U}_{i,j+1}-\underline{U}_{i,j} = \left(\Delta t\frac{\partial \underline{U}}{\partial t} + \tfrac{1}{2}(\Delta t)^2\frac{\partial^2 \underline{U}}{\partial t^2}\right) + \dots$$

$$= \left((\Delta t)A\frac{\partial \underline{U}}{\partial x} + \tfrac{1}{2}(\Delta t)^2\frac{\partial A}{\partial t}\frac{\partial \underline{U}}{\partial x} + \tfrac{1}{2}(\Delta t)^2 A\frac{\partial^2 \underline{U}}{\partial x\partial t}\right)_{i,j} + \dots$$

Hence, using the last two equations, we find that $\underline{T}=\underline{O}$ to second-order accuracy. The accuracy of the Lax-Wendroff and the Wendroff schemes can also be established in the same manner.

The von Neumann criterion analogous to the scalar case can be used to examine the stability of the difference schemes approximating the given system of first-order equations.

If a typical Fourier term,

$$\underline{u} = \underline{u}_0\exp(i_c\beta x) \ ,$$

where $\underline{u}_o$ is a constant vector and $i_c=\sqrt{-1}$ is substituted into the difference equation for $\underline{u}_{i,j}$ it is found that $\underline{u}_{i,j+1}$ is of the same form but with $\Gamma\underline{u}_o$ replacing $\underline{u}_o$. The matrix Γ is the *amplification matrix*. For example, for the Lax-Wendroff method (2.17.16), the amplification matrix is,

$$\begin{aligned}\Gamma &= I-\tfrac{1}{2}\lambda A(\exp(i_c\beta\Delta x)-\exp(-i_c\beta\Delta x))+\tfrac{1}{2}\lambda^2A^2(\exp(i_c\beta\Delta x)+\exp(-i_c\beta\Delta x)-2)\\ &= [I-\lambda^2A^2(1-\cos(\theta))]-i_c\lambda A\sin(\theta)\ , \qquad (2.17.16c)\end{aligned}$$

where $\theta=\beta\Delta x$. For the Wendroff scheme (2.17.17), the amplification matrix is given implicitly by,

$$\begin{aligned}&[(I-\lambda A\tan^2(\tfrac{1}{2}\theta))+i_c(I+\lambda A)\tan(\tfrac{1}{2}\theta)]\Gamma\\ &\quad = [(I+\lambda A\tan^2(\tfrac{1}{2}\theta))+i_c(I-\lambda A)\tan(\tfrac{1}{2}\theta)]\ . \qquad (2.17.17c)\end{aligned}$$

The von Neumann necessary condition for the stability of a system is

$$\max_{1\leq j\leq m} |\mu_j| \leq 1 \qquad (2.17.19)$$

where μ_j (j=1,2,...,m) are the eigenvalues of Γ. It can be shown that condition (2.17.19) is satisfied for the Lax-Wendroff amplification matrix Γ if

$$\lambda|\eta_j| \leq 1 \ (j=1,2,\ldots,m),$$

where η_j are the eigenvalues of A. This is evident from (2.17.16c) in which Γ is a rational function of A and so has the same eigenvectors as A. Hence,

$$\mu_j = 1-\lambda^2\eta_j^2(1-\cos(\theta))-i_c\lambda\eta_j\sin(\theta)\ , \quad (j=1,2,\ldots,m).$$

As θ varies from 0 to 2π, μ_j describes an ellipse in the complex plane. This ellipse lies inside the unit circle of

$$\lambda|\eta_j| \leq 1\ ,$$

and the result follows. It can also be established that the Wendroff scheme is unconditionally stable.

2.18 NON-LINEAR HYPERBOLIC EQUATIONS OF FIRST ORDER

The general quasi-linear hyperbolic equation of first order is

$$a(U)\frac{\partial U}{\partial x} + b(U)\frac{\partial U}{\partial t} = c(U) \quad . \tag{2.18.1}$$

Finite difference methods which are centred midway between the previous and the present time levels are generally used to solve (2.18.1). One such method is the *centred-in-distance, centred-in-time formula (the Crank-Nicolson type).* In conjunction with this method, the non-linear coefficients are treated in a number of ways as we shall see later.

For this purpose, we consider the simplest form of equation (2.18.1) that is,

$$-a(U)\frac{\partial U}{\partial x} = \frac{\partial U}{\partial t} \tag{2.18.1a}$$

The *centred-in-distance, centred-in-time (henceforth, CD-CT)* analogues to the derivatives are centred about the time level $t_{j+\frac{1}{2}}$. An analogue to the non-linear coefficient a(U) is required at this time level. Obviously, if the resulting finite difference equations are to be linear then this analogue must not contain values of u at the time level t_{j+1}. The simplest way is to evaluate a(U) at the old time level and using $a(u_j)$ for $a(u_{j+\frac{1}{2}})$. These values can then be improved by next evaluating $a(u_{j+\frac{1}{2}})$ as $a[(u_j+u_{j+1}^{(1)})/2]$, where $u_{j+1}^{(1)}$ is the value obtained when $a(u_j)$ was used for $a(u_{j+\frac{1}{2}})$. By means of equation (2.9.12), our iterative procedure is described by the following equations,

$$-[a(\frac{u_{i,j}+u_{i,j+1}^{(p)}}{2})]\ \frac{1}{4\Delta x}\ [(\Delta_x+\nabla_x)(u_{i,j+1}^{(p+1)}+u_{ij})] = \frac{u_{i,j+1}^{(p+1)}-u_{ij}}{\Delta t} \tag{2.18.2}$$

and $u_{i,j+1}^{(0)}=u_{ij}$.

Convergence of the iterative process is achieved when $|u_{i,j+1}^{(p+1)}-u_{i,j+1}^{(p)}|\leq\varepsilon$

where ε is a preset tolerance or accuracy. The resulting finite difference equations are linear in $u_{i,j+1}^{(p+1)}$ with a tridiagonal coefficient matrix so that the *Thomas algorithm* can be used for the solution. This method, however, may suffer some restrictions on the time-step size for stability and the rate of convergence. The following *linearisation techniques* are employed to overcome these problems.

By using a truncated *forward Taylor series* expansion, we obtain

$$\begin{aligned} U_{i,j+\frac{1}{2}} &= U(x_i,t_j+\tfrac{1}{2}\Delta t) \\ &= U_{i,j} + \tfrac{1}{2}(\Delta t)\left(\frac{\partial U}{\partial t}\right)_{i,j} + \frac{1}{2!}\left(\frac{\Delta t}{2}\right)^2 \left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j} + \cdots \end{aligned} \tag{2.18.3}$$

The series in (2.18.3) is truncated after the second term to obtain a second-order-correct formula for $u_{i,j+\frac{1}{2}}$, that is,

$$u_{i,j+\frac{1}{2}} = u_{i,j} + \tfrac{1}{2}(\Delta t)\left(\frac{\partial u}{\partial t}\right)_{i,j} . \tag{2.18.4}$$

The time derivative $\left(\frac{\partial u}{\partial t}\right)_{i,j}$ is then obtained from (2.18.1a). Hence,

$$u_{i,j+\frac{1}{2}} = u_{i,j} + \frac{(\Delta t)}{4\Delta x}\,[-a(u_{i,j})]\,[(\Delta_x+\nabla_x)(u_{ij})]. \tag{2.18.5}$$

This value is then used in evaluating a(u) for use in the CD-CT approximation for (2.18.1a). Hence, we arrive at the resulting finite difference equation,

$$\{-a[u_{i,j} - \frac{\Delta t}{4\Delta x}\,a(u_{ij})(\Delta_x+\nabla_x)(u_{i,j})]\}\,\frac{1}{4\Delta x}\,[(\Delta_x+\nabla_x)(u_{i,j+1}+u_{ij})] = \frac{u_{i,j+1}-u_{i,j}}{\Delta t} . \tag{2.18.6}$$

We see that the actual application of (2.18.6) is performed in two steps. An iteration procedure similar to the above may be used, if desired, to improve the values of $u_{i,j+1}$ although this is usually found to be unnecessary for small Δt. In the first step, $u_{i,j+\frac{1}{2}}$ are determined from (2.18.5) and are then used to evaluate $a(u_{i,j+\frac{1}{2}})$. The elements of the coefficient matrix for (2.18.6) are then computed and the Thomas algorithm applied.

On the other hand, an analogue to $u_{i,j+\frac{1}{2}}$ for use in the non-linear coefficients can also be obtained from a truncated Taylor series expanded about the time level $t_{j+\frac{1}{2}}$. In this case, we have

$$U_{i,j} = U(x_i, t_{j+\frac{1}{2}} - \frac{\Delta t}{2}) = U_{i,j+\frac{1}{2}} - \frac{\Delta t}{2} (\frac{\partial U}{\partial t})_{i,j+\frac{1}{2}} + \frac{1}{2!}(\frac{\Delta t}{2})^2 (\frac{\partial^2 U}{\partial t^2})_{i,j+\frac{1}{2}} + \ldots \qquad (2.18.7)$$

Again, this *backward Taylor series* is truncated after the second term to obtain a second-order-correct analogue,

$$u_{ij} = u_{i,j+\frac{1}{2}} - \frac{\Delta t}{2} (\frac{\partial u}{\partial t})_{i,j+\frac{1}{2}} \qquad (2.18.8)$$

and as before, the time derivative is obtained from (2.18.1a). However, if the time derivative is evaluated at $t_{j+\frac{1}{2}}$ then this would lead to non-linear equations thus complicating the solution process. As an alternative, we evaluate the space derivative at $t_{j+\frac{1}{2}}$ and the non-linear coefficient is computed at t_j. We therefore obtain the implicit equation,

$$[-a(u_{i,j})] \frac{1}{2\Delta x}(\Delta_x + \nabla_x)(u_{i,j+\frac{1}{2}}) = \frac{u_{i,j+\frac{1}{2}} - u_{ij}}{\Delta t/2} . \qquad (2.18.9)$$

The resulting tridiagonal system is solved for $u_{i,j+\frac{1}{2}}$ which is then utilised in the CD-CT approximation to (2.18.1a) given by

$$[-a(u_{i,j+\frac{1}{2}})] \frac{1}{4\Delta x}(\Delta_x + \nabla_x)(u_{ij} + u_{i,j+1}) = \frac{u_{i,j+1} - u_{ij}}{\Delta t} . \qquad (2.18.10)$$

The values of $u_{i,j+1}$ may again be improved by using the same iteration technique as before if required.

Finally, the centred Taylor series expansion for the analogue of $u_{i,j+\frac{1}{2}}$ can be obtained from $U_{i,j+\frac{1}{2}}$ and U_{ij} which are written about $t_{j+1/4}$, that is,

$$U_{i,j+\frac{1}{2}} = U(x_i,t_{j+1/4}+\frac{\Delta t}{4}) = U_{i,j+1/4}+(\frac{\Delta t}{4})(\frac{\partial U}{\partial t})_{i,j+1/4}+\frac{1}{2!}(\frac{\Delta t}{4})^2(\frac{\partial^2 U}{\partial x^2})_{i,j+1/4} + \ldots$$

and

$$U_{ij} = U(x_i,t_{j+1/4}-\frac{\Delta t}{4}) = U_{i,j+1/4}-(\frac{\Delta t}{4})(\frac{\partial U}{\partial t})_{i,j+1/4}+\frac{1}{2!}(\frac{\Delta t}{4})^2(\frac{\partial^2 U}{\partial x^2})_{i,j+1/4} + \ldots$$

If we now subtract the two equations, we get the required second-order-correct analogue,

$$u_{i,j+\frac{1}{2}} = \frac{\Delta t}{2}(\frac{\partial u}{\partial t})_{i,j+1/4} + u_{i,j} \qquad (2.18.11)$$

By the same reasoning as above, $a(u_{i,j})$ is now used for $a(u_{i,j+\frac{1}{4}})$ and the resulting finite difference equation is

$$[-a(u_{i,j})]\ \frac{1}{4\Delta x}(\Delta_x+\nabla_x)(u_{i,j}+u_{i,j+\frac{1}{2}}) = \frac{u_{i,j+\frac{1}{2}}-u_{ij}}{\Delta t/2} \qquad (2.18.12)$$

The values of $u_{i,j+\frac{1}{2}}$ obtained from this implicit equation are then used in (2.18.10) to obtain the values of $u_{i,j+1}$. The basic difference between this method and that which employs the backward Taylor series expansion is that now $a(u_{ij})$ is used for $a(u_{i,j+1/4})$, in the latter, however, $a(u_{ij})$ replaces $a(u_{i,j+\frac{1}{2}})$. Both methods are in the class of the *predictor-corrector* schemes.

Non-linear first-order systems may be written in the more general form as,

$$\frac{\partial \underline{U}}{\partial t} + \frac{\partial \underline{f}}{\partial x} = 0 \ , \qquad (2.18.13)$$

where $\underline{f}$ is a function of $\underline{u}$ and

$$\underline{f} = (f_1, f_2, \ldots, f_m)^T ,$$

$$f_i = f_i(\underline{U}) \; , \; i=1,2,\ldots,m$$

and
$$\underline{U} = (U_1(x,t), U_2(x,t), \ldots, U_m(x,t))^T.$$

Equation (2.18.13) is said to be in *conservation form.* To derive difference analogues to (2.18.13), it is instructive to rewrite it in the form,

$$\frac{\partial \underline{U}}{\partial t} + A(\underline{U})\frac{\partial \underline{U}}{\partial x} = 0 \; , \qquad (2.18.14)$$

where $A(\underline{U}) = \frac{\partial \underline{f}}{\partial \underline{U}}$, the *Jacobian matrix* of $\underline{f}$ with respect to $\underline{U}$ and is defined by

$$A_{k,\ell} = \frac{\partial f_k}{\partial U_\ell} \; .$$

As an example, we may proceed with the formal development of the Lax-Wendroff approximation to (2.18.13) as follows. From the Taylor's series expansion of $\underline{U}_{i,j+1}$ about the point (x_i, t_j), we get

$$\underline{U}_{i,j+1} = \underline{U}_{i,j} + \Delta t\left(\frac{\partial \underline{U}}{\partial t}\right)_{i,j} + \tfrac{1}{2}(\Delta t)^2 \frac{\partial}{\partial t}\left(\frac{\partial \underline{U}}{\partial t}\right)_{i,j} + \ldots$$

By virtue of (2.8.13), it follows that,

$$\underline{U}_{i,j+1} = \underline{U}_{i,j} - \Delta t\left(\frac{\partial \underline{f}(\underline{U})}{\partial x}\right)_{i,j} - \tfrac{1}{2}(\Delta t)^2 \frac{\partial}{\partial t}\left(\frac{\partial \underline{f}(\underline{U})}{\partial x}\right)_{i,j} + \ldots \qquad (2.18.15)$$

But,

$$\begin{aligned} \frac{\partial}{\partial t}\left(\frac{\partial \underline{f}(\underline{U})}{\partial x}\right) &= \frac{\partial}{\partial x}\left(\frac{\partial \underline{f}(\underline{U})}{\partial t}\right) \\ &= \frac{\partial}{\partial x}\left\{\frac{\partial \underline{f}(\underline{U})}{\partial \underline{U}} \;\; \frac{\partial \underline{U}}{\partial t}\right\} \\ &= -\frac{\partial}{\partial x}\left\{\frac{\partial \underline{f}(\underline{U})}{\partial \underline{U}} \;\; \frac{\partial \underline{f}}{\partial x}\right\} \end{aligned}$$

where $\frac{\partial \underline{f}(\underline{U})}{\partial \underline{U}} = A(\underline{U})$.

Therefore, equation (2.18.15) can be written as,

$$\underline{U}_{i,j+1} = \underline{U}_{ij} - \Delta t\left(\frac{\partial \underline{f}(\underline{U})}{\partial x}\right)_{i,j} + \tfrac{1}{2}(\Delta t)^2 \frac{\partial}{\partial x}\left\{A(\underline{U}) \; \frac{\partial \underline{f}(\underline{U})}{\partial x}\right\}_{i,j} + \ldots$$

(2.18.16)

By using the central difference approximation, we obtain,

$$\left(\frac{\partial \underline{f}(\underline{U})}{\partial x}\right)_{i,j} \simeq \frac{1}{2\Delta x}\,[\underline{f}(\underline{u}_{i+1,j})-\underline{f}(\underline{u}_{i-1,j})]$$

and

$$\frac{\partial}{\partial x}\{A(\underline{U})\frac{\partial \underline{f}(\underline{U})}{\partial x}\}_{i,j} \simeq \frac{1}{\Delta x}\,\delta_x[A(\underline{u}_{i,j})\,\frac{1}{\Delta x}\,\delta_x\{\underline{f}(\underline{u}_{i,j})\}]\ ,$$

where,

$$\delta_x\underline{u}_{i,j} = \underline{u}_{i+\frac{1}{2},j}-\underline{u}_{i-\frac{1}{2},j}.$$

Hence,

$$\frac{\partial}{\partial x}\{A(\underline{U})\,\frac{\partial \underline{f}(\underline{U})}{\partial x}\}_{i,j} \simeq \frac{1}{(\Delta x)^2}\delta_x[A(\underline{u}_{i,j})\{\underline{f}(\underline{u}_{i+\frac{1}{2},j})-\underline{f}(\underline{u}_{i-\frac{1}{2},j})\}]$$

$$= \frac{1}{(\Delta x)^2}\{A(\underline{u}_{i+\frac{1}{2},j})\,[\underline{f}(\underline{u}_{i+1,j})-\underline{f}(\underline{u}_{i,j})]-A(\underline{u}_{i-\frac{1}{2},j})\,[\underline{f}(\underline{u}_{i,j})-\underline{f}(\underline{u}_{i-1,j})]\}.$$

If we denote $A_{i,j}=A(\underline{u}_{i,j})$ and $\underline{f}_{i,j}=\underline{f}(\underline{u}_{i,j})$ then the difference analogue for (2.18.16) is,

$$\underline{u}_{i,j+1} = \underline{u}_{i,j} - \tfrac{1}{2}\lambda(\underline{f}_{i+1,j}-\underline{f}_{i-1,j})+\tfrac{1}{2}\lambda^2[A_{i+\frac{1}{2},j}(\underline{f}_{i+1,j}-\underline{f}_{i,j})-A_{i-\frac{1}{2},j}$$
$$(\underline{f}_{i,j}-\underline{f}_{i-1,j})]\ . \qquad (2.18.17)$$

Evaluations at the mid-points can be avoided by replacing $A_{i+\frac{1}{2},j}$ by $\frac{1}{2}(A_{i,j}+A_{i+1,j})$ and $A_{i-\frac{1}{2},j}$ by $\frac{1}{2}(A_{i-1,j}+A_{ij})$. Equation (2.18.17) then becomes,

$$\underline{u}_{i,j+1} = \underline{u}_{i,j}-\tfrac{1}{2}\lambda(\underline{f}_{i+1,j}-\underline{f}_{i-1,j})+\tfrac{1}{2}\lambda^2(\underline{g}_{i,j}-\underline{g}_{i-1,j}) \qquad (2.18.18)$$

where,

$$\underline{g}_{i,j} = \tfrac{1}{2}(A_{i+1,j}+A_{i,j})(\underline{f}_{i+1,j}-\underline{f}_{i,j})$$

and

$$\underline{g}_{i-1,j} = \tfrac{1}{2}(A_{i,j}+A_{i-1,j})(\underline{f}_{i,j}-\underline{f}_{i-1,j})\ .$$

Finally, we present a procedure known as *the two-step Lax-Wendroff process*, which also has second-order accuracy and which is slightly simpler to use because of the absence of the coefficient matrix A.

This generalisation of the Lax-Wendroff scheme takes on a *Runge-Kutta form* with intermediate approximate values. For an $S_{\beta,\alpha}$ scheme, these intermediate values approximate $\underline{U}$ at the points $(x_{i+\beta},t_{j+\alpha})$. The following difference operators will be used,

$$\begin{aligned}
\Delta_x u_i &= u_{i+1}-u_i \\
\nabla_x u_i &= u_i-u_{i-1} \\
\delta_x u_i &= u_{i+\frac{1}{2}}-u_{i-\frac{1}{2}} \\
H_x u_i &= u_{i+1}-u_{i-1} \\
\mu_x u_i &= \tfrac{1}{2}(u_{i+\frac{1}{2}}+u_{i-\frac{1}{2}}) \ .
\end{aligned} \qquad (2.18.19)$$

The usual $S_{\frac{1}{2},\frac{1}{2}}$ scheme due to Ritchmyer (1963), consists of a Lax type step followed by a leap-frog step and is given by,

$$\begin{aligned}
\bar{\underline{u}}_{i+\frac{1}{2},j+\frac{1}{2}} &= \mu_x \underline{u}_{i+\frac{1}{2},j} - \tfrac{1}{2}\lambda\delta_x \underline{f}_{i+\frac{1}{2},j} \\
\underline{u}_{i,j+1} &= \underline{u}_{i,j} - \lambda\delta_x \bar{\underline{f}}_{i,j+\frac{1}{2}} \ .
\end{aligned} \qquad (2.18.20)$$

MacCormack (1971) introduced the very popular and successful $S_{1,1}$ and $S_{0,1}$ schemes given by,

$$\begin{aligned}
\bar{\underline{u}}_{i,j+1} &= \underline{u}_{i,j} - \lambda\nabla_x \underline{f}_{i,j} \\
\underline{u}_{i,j+1} &= \tfrac{1}{2}(\underline{u}_{i,j}+\bar{\underline{u}}_{i,j+1}) - \tfrac{1}{2}\lambda\Delta_x \bar{\underline{f}}_{i,j+1}
\end{aligned} \qquad (2.18.21)$$

and

$$\begin{aligned}
\bar{\underline{u}}_{i,j+1} &= \underline{u}_{i,j} - \lambda\Delta_x \underline{f}_{i,j} \\
\underline{u}_{i,j+1} &= \tfrac{1}{2}(\underline{u}_{i,j}+\bar{\underline{u}}_{i,j+1}) - \tfrac{1}{2}\lambda\nabla_x \bar{\underline{f}}_{i,j+1} \ .
\end{aligned} \qquad (2.18.22)$$

respectively. The advantage of the $S_{0,1}$ and $S_{1,1}$ schemes over the $S_{\frac{1}{2},\frac{1}{2}}$ approximations lies in the ease of adding a dissipative term $-b\dfrac{\partial^2 \underline{U}}{\partial x^2}$ to $\dfrac{\partial \underline{f}(\underline{U})}{\partial x}$ of (2.18.13) with suitable differences of $\underline{u}$ and $\bar{\underline{u}}_j$ being added to $\underline{f}$ and $\bar{\underline{f}}_j$ (Morton, 1977). If in schemes (2.18.20), (2.18.21) and (2.18.22) we put,

$$\underline{f} = A\underline{u} \text{ , A a constant matrix}$$

at the mesh points and eliminate the intermediate value $\bar{\underline{u}}$, we get in all cases the Lax-Wendroff formula (2.17.16).

McGuire and Morris (1974) generalised the $S_{\frac{1}{2},\frac{1}{2}}$ form of (2.18.20) to give the $S_{\frac{1}{2},\alpha}$ two-step procedure

$$\bar{\underline{u}}_{i+\frac{1}{2},j+\alpha} = \mu_x \underline{u}_{i+\frac{1}{2},j} - \alpha\lambda\delta_x \underline{f}_{i+\frac{1}{2},j}$$

$$\underline{u}_{i,j+1} = \underline{u}_{i,j} - \tfrac{1}{2}\lambda[(1-\frac{1}{2\alpha})H_x \underline{f}_{i,j} + \frac{1}{\alpha}\delta_x \bar{\underline{f}}_{i,j+\alpha}] \tag{2.18.23}$$

with $0<\alpha\leqslant1$. For $\alpha=\frac{1}{2}$ we get back the $S_{\frac{1}{2},\frac{1}{2}}$ form and when $\alpha=1$, (2.18.23) reduces to the scheme devised by Rubin and Burstein (1967). Again, as above, the procedure (2.18.23) can be reduced to the Lax-Wendroff formula (2.17.16).

Further work on the $S_{\beta,\alpha}$ form has been done by Shankar and Anderson (1975) and Lerat and Peyret (1975). The method has also been extended to third order by Warming, Kutler and Lomax (1973). The generalisation of the basic Lax-Wendroff method to two space dimensions leads to a wide variety of different methods which usually involve nine space points. For a survey of these schemes, the reader may refer to the works of Turkel (1974) and Gourlay and Morris (1968).

2.19 NON-LINEAR HYPERBOLIC EQUATIONS OF SECOND ORDER

Explicit methods have the disadvantage that they suffer from severe stability restrictions when employed to solve non-linear second-order hyperbolic equations. As an example, we consider the non-linear vibrations of a string fixed at both ends. The governing equations, due to Carrier (1945), for the simplified system are,

$$\frac{\partial}{\partial x}[T'\sin\theta] = \rho A\,\frac{\partial^2 U'}{\partial t^2}\ ,\ \theta = \tan^{-1}(\frac{\partial U'}{\partial x})$$

$$T' = T_O + EA\{(1+(\frac{\partial U'}{\partial x})^2)^{\frac{1}{2}}-1\}. \qquad (2.19.1)$$

By introducing the dimensionless variables,

$$U = \frac{U'}{L}\ ,\ X = \frac{x}{L}\ ,\ T = \frac{t}{L[\rho A/T_O]^{\frac{1}{2}}}$$

and by setting $B=EA/T_O$ we obtain the dimensionless equation,

$$\frac{1-B+B[1+(\frac{\partial U}{\partial x})^2]^{3/2}}{[1+(\frac{\partial U}{\partial x})^2]^{3/2}}\ \frac{\partial^2 U}{\partial X^2} = \frac{\partial^2 U}{\partial T^2} \qquad (2.19.2)$$

subject to the initial-boundary conditions,

$$U(0,T) = U(1,T) = 0,\ \frac{\partial U}{\partial T}(X,0) = 0\ ,\ U(X,0) = 4X(1-X)\ .$$

When (2.19.2) is approximated by the same second-order central differences employed to obtain the explicit equation (2.13.3) and a first-order forward difference is used for $\frac{\partial U}{\partial x}$, Ames (1965) discovered that the stability threshold is approximately $\lambda=0.55$, beyond which instability occurs.

In many cases, the algebraic equations which result from the use of implicit methods are non-linear. However, for a class of *time quasi-linear* problems, sets of linear equations are generated. The general form of second-order non-linear equations in this class of problems is given by,

$$\frac{\partial^2 U}{\partial t^2} + f(x,t,U,\frac{\partial U}{\partial x},\frac{\partial^2 U}{\partial x^2})\ \frac{\partial^2 U}{\partial x \partial t} + g(x,t,U,\frac{\partial U}{\partial x},\frac{\partial^2 U}{\partial x^2})\frac{\partial U}{\partial t} = P(x,t,U,\frac{\partial U}{\partial x},\frac{\partial^2 U}{\partial x^2}) \tag{2.19.3}$$

subject to the initial-boundary conditions,

$$U(0,t) = F(t),\ U(1,t) = G(t),\ U(x,0) = H(x),\ \frac{\partial U}{\partial t}(x,0) = J(x).$$

The finite-difference analogue to (2.19.3) is obtained in the following manner:

All first derivatives are approximated by the second-order formula (2.1.8). For example, we use,

$$(\frac{\partial U}{\partial t})_{i,j} = \frac{U_{i,j+1}-U_{i,j-1}}{2\Delta t} + O([\Delta t]^2)\ .$$

Second-derivatives $(\frac{\partial^2 U}{\partial x^2}$ and $\frac{\partial^2 U}{\partial t^2})$ are approximated by the central difference equation (2.1.9). Hence, we have,

$$(\frac{\partial^2 U}{\partial x^2})_{i,j} = \frac{U_{i+1,j}-2U_{ij}+U_{i-1,j}}{(\Delta x)^2} + O([\Delta x]^2)$$

and similarly for $(\frac{\partial^2 U}{\partial t^2})_{i,j}$.

Finally, $\frac{\partial^2 U}{\partial x \partial t}$ is approximated at the point (x_i,t_j) by

$$\frac{u_{i+1,j+1}-u_{i+1,j-1}-u_{i-1,j+1}+u_{i-1,j-1}}{4(\Delta x)(\Delta t)}\ .$$

The result of these approximations leads to the algebraic equation

$$a_i u_{i-1,j+1}+b_i u_{i,j+1}+c_i u_{i+1,j+1} = d_i\ , \tag{2.19.4}$$

where,

$$a_i = -\frac{f_{ij}}{4(\Delta x)(\Delta t)}\ ,\ b_i = \frac{1}{(\Delta t)^2} + \frac{g_{ij}}{2(\Delta t)}\ ,\ c_i = -a_i,$$

$$d_i = P_{i,j} + \frac{1}{4(\Delta t)(\Delta x)}(u_{i+1,j-1}-u_{i-1,j-1})f_{ij} + \frac{1}{2(\Delta t)}\ u_{i,j-1}g_{ij} + \frac{1}{(\Delta t)^2}\ (2u_{ij}-u_{i,j-1})$$

and the notation f_{ij} means

$$f_{i,j} = f\,[\,i\Delta x, j\Delta t, u_{ij}, (\frac{\partial u}{\partial x})_{i,j}, (\frac{\partial^2 u}{\partial x^2})_{i,j}]\ .$$

Equation (2.19.4) generates an implicit *linear* tridiagonal system which can be solved for values at the mesh points in the (j+1) time row by the *Thomas algorithm*

CHAPTER THREE

SURVEY OF CURRENT METHODS TO SOLVE PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

3.1 PARABOLIC EQUATIONS IN ONE SPACE DIMENSION

A linear parabolic partial differential equation takes the general form,

$$\sigma(x,t)\frac{\partial U}{\partial t} = \frac{\partial}{\partial x}(a(x,t)\frac{\partial U}{\partial x}) + b(x,t)\frac{\partial U}{\partial x} - c(x,t)U , \qquad (3.1.1)$$

which is defined within some prescribed domain D of the (x,t) space. Within this domain, the functions $\sigma(x,t)$,$a(x,t)$ are strictly positive and $c(x,t)$ is non-negative.

As in hyperbolic equations, we shall now focus our attention to developing finite difference methods for parabolic equations. A simplified form of (3.1.1), that is, the *diffusion (heat) equation with constant coefficients* $(\sigma(x,t)=a(x,t)=1,\ b(x,t)=c(x,t)=0)$,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} , \ 0<x<1, 0<t\leq T , \qquad (3.1.2)$$

subject to the initial-boundary conditions,

$$\begin{aligned} U(x,0) &= f(x) \ , \ 0<x<1 \ , \\ U(0,t) &= g(t) \ , \ 0<t\leq T \\ U(1,t) &= h(t) \ , \ 0<t\leq T \ , \end{aligned} \qquad (3.1.2a)$$

will be used. For this purpose, a uniformly-spaced network whose mesh points are $x_i=i\Delta x$, $t_j=j\Delta t$ for $i=0,1,2,\ldots,m-1,m$ and $j=0,1,\ldots,n-1,n$ with $\Delta x=\frac{1}{m}$ and $\Delta t=\frac{T}{n}$ being utilised.

3.2 EXPLICIT METHODS

If the space derivative is approximated by the central second difference (2.1.9) and the time derivative by the forward difference (2.1.6), then the resulting approximation is,

$$\frac{1}{\Delta t}(u_{i,j+1}-u_{i,j}) = \frac{1}{(\Delta x)^2}(u_{i+1,j}-2u_{ij}+u_{i-1,j}) \ . \qquad (3.2.1)$$

Upon solving this equation for $u_{i,j+1}$ we obtain the explicit formula,

$$\begin{aligned} u_{i,j+1} &= \lambda u_{i-1,j}+(1-2\lambda)u_{ij}+\lambda u_{i+1,j} \\ &= u_{ij}+\lambda(u_{i-1,j}-2u_{ij}+u_{i+1,j}) \qquad (3.2.1a) \end{aligned}$$

where $\lambda = \frac{\Delta t}{(\Delta x)^2}$ is the mesh ratio.

The *local truncation error* of (3.2.1) can be obtained as usual by the Taylor's series expansion about the point (x_i,t_j). Hence, we have,

$$T = U_{i,j+1}-\lambda U_{i-1,j}-(1-2\lambda)U_{ij}-\lambda U_{i+1,j} \ ,$$

which on expansion gives us,

$$T = \tfrac{1}{2}\{\Delta t(\frac{\partial^2 U}{\partial t^2})_{i,j} - \frac{(\Delta x)^2}{6}(\frac{\partial^4 U}{\partial x^4})_{i,j}\} + \frac{(\Delta t)^2}{6}(\frac{\partial^3 U}{\partial t^3})_{i,j} - \frac{(\Delta x)^4}{360}(\frac{\partial^6 U}{\partial x^6})_{i,j}+\ldots \qquad (3.2.2)$$

That is, $T = O([\Delta t])+O([\Delta x]^2)$. (3.2.3)

We now study the stability of equation (3.2.1) by means of the four methods mentioned in Section 2.3.

(a) The Maximum Principle

From (3.2.1a) and (3.2.2), we have

$$U_{i,j+1} = \lambda U_{i-1,j}+(1-2\lambda)U_{ij} + \lambda U_{i+1,j} + O(\Delta t) + O([\Delta x]^2). \qquad (3.2.4)$$

At the mesh point, we have

$$\varepsilon_{i,j} = U_{ij}-u_{ij} \ . \qquad (3.2.5)$$

Therefore, by subtracting equation (3.2.1a) from (3.2.4) leads to

$$\varepsilon_{i,j+1} = \lambda\varepsilon_{i-1,j}+(1-2\lambda)\varepsilon_{i,j}+\lambda\varepsilon_{i+1,j}+O(\Delta t) + O([\Delta x]^2) \ . \qquad i=1,2,\ldots,m-1. \qquad (3.2.6)$$

Since U agrees with u initially and on the boundary, then,

$$\varepsilon_{i,0} = 0 \ , \quad i=0,1,\ldots,m,$$

$$\varepsilon_{0,j} = \varepsilon_{m,j} = 0 \ , \quad j=0,1,\ldots,n.$$

We note that the sum of the coefficients in λ in (3.2.6) is unity and if,

$$\lambda \leqslant \tfrac{1}{2} \quad , \qquad (3.2.7)$$

then they are all non-negative. Let E_j denote the maximum value of $|\varepsilon_{i,j}|$ along the jth time row. From (3.2.6), we find that if (3.2.7) holds then,

$$\begin{aligned}|\varepsilon_{i,j+1}| &\leqslant \lambda|\varepsilon_{i-1,j}| +(1-2\lambda)|\varepsilon_{ij}|+\lambda|\varepsilon_{i+1,j}|+O(\Delta t+[\Delta x]^2)\\ &\leqslant \lambda E_j + (1-2\lambda)E_j + \lambda E_j + M(\Delta t+[\Delta x]^2)\\ &= E_j + M(\Delta t+[\Delta x]^2) \ .\end{aligned}$$

As this is true for all values of i, it is also true for $\max\limits_i|\varepsilon_{i,j+1}|$. Hence,

$$E_{j+1} \leqslant E_j + M(\Delta t+[\Delta x]^2) \ ,$$

i.e.

$$E_{j+1} \leqslant E_{j-1}+2M(\Delta t+[\Delta x]^2)$$

and so on, from which it follows that,

$$\begin{aligned}E_j &\leqslant E_0 + jM(\Delta t+[\Delta x]^2) \ ,\\ &\leqslant nM(\Delta t+(\Delta x)^2) \ ,\end{aligned}$$

i.e.

$$E_j \leqslant N(\Delta t+(\Delta x)^2) \qquad (3.2.8)$$

since $E_0 = 0$, $j \leqslant n$ and $N = nM$. But $\lambda = \dfrac{\Delta t}{(\Delta x)^2}$ and so (3.2.8) can be expressed as $E_j = O([\Delta x]^2)$.

The inequality (3.2.8) means that the finite difference analogue *converges* to the solution of the differential equation as Δx and Δt tend uniformly to zero. The above boundedness condition also implies stability and thus the stability condition is $\lambda \leqslant \tfrac{1}{2}$.

(b) The Energy Method

For convenience, we specify the boundary conditions in (3.1.2a) as,

$$U(0,t) = U(1,t) \ , \ 0<t\leq T \ . \qquad (3.1.2b)$$

We first observe that if we multiply the differential equation (3.1.2) by U and integrating with respect to x, we obtain

$$\int_0^1 U \frac{\partial U}{\partial t}\, dx = \int_0^1 U \frac{\partial^2 U}{\partial x^2}\, dx \ ,$$

which on integration by parts of the right hand side gives,

$$\frac{\partial}{\partial t} \int_0^1 U^2 dx = - \int_0^1 \left(\frac{\partial U}{\partial x}\right)^2 dx \leq 0 \ .$$

We deduce from this, that,

$$\int_0^1 U^2(x,t)dx \leq \int_0^1 U^2(x,0)dx = \int_0^1 f^2(x)dx \ ,$$

and therefore the quantity $\int_0^1 U^2(x,t)dx$ remains bounded as $t\to\infty$.

When the explicit scheme (3.2.1a) is used, the energy method may be employed to show that the analogous quantity $(\Delta x)\sum_{i=1}^{m-1}(\varepsilon_{ij})^2$ remains bounded as $j\to\infty$.

As before, the perturbations at the mesh points satisfy the difference equation

$$\varepsilon_{i,j+1}-\varepsilon_{ij} = \lambda(\varepsilon_{i+1,j}+\varepsilon_{i-1,j}-2\varepsilon_{ij}) \ , \qquad (3.2.9)$$

and the boundary conditions give,

$$\varepsilon_{0,j} = \varepsilon_{m,j} = 0 \ , \text{ for all } j.$$

Equation (3.2.9) is multiplied by $(\varepsilon_{i,j+1}+\varepsilon_{ij})$ and the result summed over $i=1,2,\ldots,m-1$ to give,

$$||\underline{\varepsilon}_{j+1}||^2-||\underline{\varepsilon}_j||^2 = \lambda \sum_{i=1}^{m-1}(\varepsilon_{i,j+1}+\varepsilon_{i,j})(\varepsilon_{i+1,j}+\varepsilon_{i-1,j}-2\varepsilon_{i,j}) \qquad (3.2.10)$$

where $||\underline{\varepsilon}_j||^2 = \sum_{i=1}^{m-1} (\varepsilon_{ij})^2$. We shall now use the following identities

$$\sum_{i=1}^{m-1} e_i(\varepsilon_{i-1}-\varepsilon_i) = \sum_{i=1}^{m-1} \varepsilon_i(e_{i+1}-e_i) \text{ if } \varepsilon_0=e_m=0 \text{ ,} \tag{3.2.11a}$$

and,

$$\sum_{i=1}^{m-1} e_i(\varepsilon_{i+1}-\varepsilon_i) = \sum_{i=1}^{m-1} \varepsilon_{i+1}(e_i-e_{i+1})-e_1\varepsilon_1 \text{ if } \varepsilon_m=0. \tag{3.2.11b}$$

The right hand side of (3.2.10) can be rearranged to read,

$$\lambda \sum_{i=1}^{m-1} e_i[(\varepsilon_{i+1,j}-\varepsilon_{i,j})-(\varepsilon_{ij}-\varepsilon_{i-1,j})] \tag{3.2.12}$$

with $e_i = \varepsilon_{i,j+1}+\varepsilon_{ij}$. Remembering that,

$$\varepsilon_{0,j} = \varepsilon_{0,j+1} = \varepsilon_{m,j} = \varepsilon_{m,j+1} = 0 \text{ ,}$$

then after substituting (3.2.11a) and (3.2.11b) into (3.2.12) we find that equation (3.2.10) becomes,

$$||\underline{\varepsilon}_{j+1}||^2-||\underline{\varepsilon}_j||^2 = -\lambda[\sum_{i=0}^{m-1}\{(\varepsilon_{i+1,j}-\varepsilon_{ij})^2+(\varepsilon_{i+1,j+1}-\varepsilon_{i,j+1})(\varepsilon_{i+1,j}-\varepsilon_{i,j})\}] \text{ .} \tag{3.2 13}$$

If we define,

$$E_j = ||\underline{\varepsilon}_j||^2 - \tfrac{1}{2}\lambda \sum_{i=0}^{m-1}(\varepsilon_{i+1,j}-\varepsilon_{ij})^2 \text{ , } j=0,1,2,\ldots$$

it follows that,

$$E_j \leqslant ||\underline{\varepsilon}_j||^2$$

and

$$E_{j+1}-E_j = ||\underline{\varepsilon}_{j+1}||^2-||\underline{\varepsilon}_j||^2-\tfrac{1}{2}\lambda\sum_{i=0}^{m-1}\{(\underline{\varepsilon}_{i+1,j+1}-\underline{\varepsilon}_{i,j+1})^2-(\underline{\varepsilon}_{i+1,j}-\underline{\varepsilon}_{i,j})^2\}.$$

By using (3.2.12), we get,

$$E_{j+1}-E_j = -\tfrac{1}{2}\lambda \sum_{i=0}^{m-1}(\varepsilon_{i+1,j+1}-\varepsilon_{i,j+1}+\varepsilon_{i+1,j}-\varepsilon_{i,j})^2 \leqslant 0 \text{ .}$$

We conclude therefore that E_j is a monotonic decreasing function of j. By summing the inequalities,

$$(\varepsilon_{i+1,j}-\varepsilon_{i,j})^2 \leq (\varepsilon_{i+1,j}-\varepsilon_{i,j})^2+(\varepsilon_{i+1,j}+\varepsilon_{i,j})^2 = 2(\varepsilon_{i+1,j})^2+2(\varepsilon_{i,j})^2$$

for $i=0,1,\ldots,m-1$, leads to,

$$\sum_{i=0}^{m-1}(\varepsilon_{i+1,j}-\varepsilon_{i,j})^2 \leq 2\sum_{i=0}^{m-1}[(\varepsilon_{i+1,j})^2+(\varepsilon_{i,j})^2] = 4\sum_{i=0}^{m-1}(\varepsilon_{i,j})^2 = 4||\underline{\varepsilon}_j||^2,$$

and so from the definition of E_j, we get,

$$E_j \geq (1-2\lambda)||\underline{\varepsilon}_j||^2 .$$

Now, if,

$$0 < \lambda < \tfrac{1}{2} , \qquad (3.2.14)$$

$$||\underline{\varepsilon}_j||^2 \leq \frac{1}{(1-2\lambda)} E_j ,$$

and in view of the following inequalities,

$$E_j \leq E_{j-1} \leq \cdots \leq E_0 \leq ||\underline{\varepsilon}_0||^2 ,$$

we have,

$$||\underline{\varepsilon}_j||^2 \leq \frac{1}{(1-2\lambda)} ||\underline{\varepsilon}_0||^2 .$$

The vectors $\underline{\varepsilon}_j$, $j=0,1,\ldots$ are therefore bounded provided $0<\lambda<\frac{1}{2}$ and stability results.

(c) The Matrix Analysis

By taking account of the boundary conditions (3.1.2b), the explicit scheme (3.2.1a) can be written in matrix form as,

$$\underline{u}_{j+1} = \Gamma\underline{u}_j \quad , \qquad (3.2.15)$$

where $\underline{u}_j = (u_{1,j},u_{2,j},\ldots,u_{m-1,j})^T$ and Γ is a square tridiagonal matrix of order $(m-1)$ given by,

$$\Gamma = \begin{bmatrix} (1-2\lambda) & \lambda & & & & \\ \lambda & (1-2\lambda) & \lambda & & 0 & \\ & \lambda & (1-2\lambda) & \lambda & & \\ & 0 & \ddots & \ddots & \ddots & \\ & & & \lambda & (1-2\lambda) & \lambda \\ & & & & \lambda & (1-2\lambda) \end{bmatrix}_{(m-1)\times(m-1)}$$

$$= \begin{bmatrix} 1 & 0 & 0 & & & \\ 0 & 1 & 0 & & O & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & O & & & 1 & 0 \\ & & & & 0 & 1 \end{bmatrix}_{(m-1)\times(m-1)} + \lambda \begin{bmatrix} -2 & 1 & & & & \\ 1 & -2 & 1 & & O & \\ & & \ddots & \ddots & \ddots & \\ & O & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{bmatrix}_{(m-1)\times(m-1)}$$

That is,

$$\Gamma = I + \lambda T \; . \tag{3.2.16}$$

Since the matrix T is of a *common tridiagonal form* then we know from Section 1.6 that the eigenvalues of T are given by,

$$\eta_i = -4\sin^2 \left(\frac{i\pi}{2m}\right) \;,\; i=1,2,\ldots,m-1.$$

Hence, from (3.2.16), the eigenvalues of the *amplification matrix* Γ are $\mu_i=1+\lambda\{-4 \sin^2(\frac{i\pi}{2m})\}$ and the condition for stability is $|1-4\lambda\sin^2(\frac{i\pi}{2m})|\leq 1$. The only useful inequality is $-1\leq 1-4\lambda\sin^2(\frac{i\pi}{2m})$ giving

$$\lambda \leq \frac{1}{2\sin^2(\frac{i\pi}{2m})}$$

which proves that the scheme is stable for $\lambda\leq\frac{1}{2}$.

(d) The von Neumann Criterion (Fourier Series Analysis)

The substitution of the perturbation (equation (2.10.4))

$$\varepsilon_{ij} = \xi^j \exp(i_c \beta i\Delta x) \text{ into equation (3.2.1a)}$$

results in the stability equation,

$$\begin{aligned}\gamma &= 1+\lambda(\exp(-i_c\beta\Delta x)+\exp(i_c\beta\Delta x)-2)\\ &= 1+\lambda(2\cos(\beta\Delta x)-2)\\ &= 1-2\lambda(1-\cos(\beta\Delta x))\\ &= 1-4\lambda\sin^2(\frac{\beta\Delta x}{2}) \;.\end{aligned}$$

The condition for stability, $|\gamma|\leq 1$, leads to,

$$|1-4\lambda\sin^2(\frac{\beta\Delta x}{2})| \leq 1.$$

By the same argument as above, we obtain conditional stability for $\lambda\leq\frac{1}{2}$.

3.3 IMPROVING THE ACCURACY OF EXPLICIT METHODS

Richardson (1910) proposed the following explicit method to solve (3.1.2),

$$\frac{(u_{i,j+1}-u_{i,j-1})}{2\Delta t} = \frac{1}{(\Delta x)^2}(u_{i+1,j}-2u_{ij}+u_{i-1,j}) \ . \qquad (3.3.1)$$

By using the familiar Taylor series expansion about the point (x_i,t_j), the local truncation error of (3.3.1) is found to be,

$$T = \{(\frac{\partial^2 U}{\partial x^2})_{i,j} - (\frac{\partial U}{\partial t})_{i,j}\} - \frac{1}{6}(\Delta t)^2(\frac{\partial^3 U}{\partial t^3})_{i,j} + \frac{(\Delta x)^2}{12}(\frac{\partial^4 U}{\partial x^4})_{i,j} + \ldots$$

i.e., $T = O([\Delta t]^2)+O([\Delta x]^2)$, (3.3.2)

since the terms in braces satisfy (3.1.2) at the mesh point. It is obvious that by comparing (3.3.2) and (3.2.3), equation (3.3.1) is of one order higher in time than equation (3.2.1). Nevertheless the difference equation (3.3.1) is completely unstable for all choices of the mesh ratio λ. By applying the von Neumann criterion, the stability equation is,

$$\gamma - \frac{1}{\gamma} = -8\lambda\sin^2(\frac{\beta\Delta x}{2}) \ , \qquad (3.3.3)$$

or

$$\gamma^2+8\lambda\sin^2(\frac{\beta\Delta x}{2})\gamma-1 = 0 \ . \qquad (3.3.4)$$

It is clear that the roots of this equation are negative reciprocals. The sum of the roots is $\gamma+(-\frac{1}{\gamma})$. From (3.3.3), we see that unless $\beta=0$, $\gamma-\frac{1}{\gamma}<0$ and either γ or $-\frac{1}{\gamma}$ is less than -1. Since γ must be assumed to be any real number, we conclude that the Richardson's representation is unconditionally unstable.

The accuracy of the explicit method (3.2.1) can be further improved and yet still maintaining its stability by considering the truncation error (3.2.2). Let us assume that U(x,t) possesses continuous bounded derivatives up to order six in x and order three in t, i.e. $U \in C^{6,3}$. Hence, from the given diffusion equation $\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2}$, we get,

$$\frac{\partial^2 U}{\partial t^2} = \frac{\partial^3 U}{\partial^2 x\partial t} = \frac{\partial^3 U}{\partial t\partial x^2} = \frac{\partial^4 U}{\partial x^4} , \tag{3.3.5}$$

and equation (3.2.2) becomes,

$$T = \tfrac{1}{2}[\Delta t - \frac{(\Delta x)^2}{6}](\frac{\partial^4 U}{\partial x^4})_{i,j} + \frac{(\Delta t)^2}{6}(\frac{\partial^3 U}{\partial t^3})_{i,j} - \frac{(\Delta x)^4}{360}(\frac{\partial^6 U}{\partial x^6})_{i,j} + \ldots$$

Therefore, if Δt and Δx go to zero in such a way that

$$\lambda = \frac{\Delta t}{(\Delta x)^2} = 1/6 ,$$

then the truncation error will be,

$$T = O([\Delta t]^2) , \tag{3.3.6a}$$

or

$$T = O([\Delta x]^4) , \tag{3.3.6b}$$

and the bound (3.2.8) will be replaced by,

$$E_j = \max|U_{i,j} - u_{i,j}| \leq N_1([\Delta t]^2 + [\Delta x]^4) . \tag{3.3.7}$$

We see that the difference equation converges to the solution of the differential equation at *a rate much faster* than that exhibited by (3.2.8).

Another attempt to gain better accuracy is to expand $U_{i,j+1}$ by the Taylor's series about the point (x_i,t_j), i.e.,

$$\begin{aligned} U_{i,j+1} &= U_{ij} + (\Delta t)(\frac{\partial U}{\partial t})_{i,j} + \frac{(\Delta t)^2}{2}(\frac{\partial^2 U}{\partial t^2})_{i,j} + \ldots \\ &= U_{ij} + (\Delta t)(\frac{\partial^2 U}{\partial x^2})_{i,j} + \frac{(\Delta t)^2}{2}(\frac{\partial^4 U}{\partial x^4})_{i,j} + \ldots \end{aligned} \tag{3.3.8}$$

by virtue of (3.3.5). If we approximate $\frac{\partial^2 U}{\partial x^2}$ and $\frac{\partial^4 U}{\partial x^4}$ at the mesh points by second and fourth differences respectively, i.e.,

$$\begin{aligned} (\frac{\partial^2 U}{\partial x^2})_{i,j} &\simeq \frac{1}{(\Delta x)^2}\delta_x^2 U_{i,j} \\ &= \frac{1}{(\Delta x)^2}[U_{i+1,j} - 2U_{ij} + U_{i-1,j}] , \end{aligned}$$

and $$\left(\frac{\partial^4 U}{\partial x^4}\right)_{i,j} \simeq \frac{1}{(\Delta x)^4}\delta_x^4 U_{ij}$$

$$= \frac{1}{(\Delta x)^4}[U_{i-2,j}-4U_{i-1,j}+6U_{i,j}-4U_{i+1,j}+U_{i+2,j}] ,$$

then, the difference analogue to (3.3.8) is

$$u_{i,j+1} = u_{i,j} + \frac{\Delta t}{(\Delta x)^2}\delta_x^2 u_{ij} + \frac{1}{2}\frac{(\Delta t)^2}{(\Delta x)^4}\delta_x^4 u_{i,j} \qquad (3.3.9)$$

with a local truncation error,

$$T = O([\Delta x]^2) + O([\Delta t]^2) . \qquad (3.3.10)$$

We also find that the stability requirement of (3.3.9) is the same as that of (3.2.1), i.e., $\lambda \leq \frac{1}{2}$ and the boundedness condition (3.2.8) is replaced by,

$$E_j = \max|U_{i,j}-u_{i,j}| \leq N_2([\Delta x]^2+[\Delta t]^2) . \qquad (3.3.11)$$

Since $\lambda = \frac{\Delta t}{(\Delta x)^2}$, (3.3.11) can also be written as $E_j = O([\Delta x]^2)$ which is the same as before. We observe from (3.3.10) or (3.3.11) that although the error in the t-direction is one order higher, the error in the x-direction remains the same as that of the original explicit method. In fact, no overall improvement is obtained because the rate of convergence of the difference equation to the solution of the differential equation is of the same magnitude as before.

This can be rectified if we utilise the fact that

$$\left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} \simeq \frac{1}{(\Delta x)^2}\delta_x^2 U_{i,j} - \frac{1}{12}(\Delta x)^2\left(\frac{\partial^4 U}{\partial x^4}\right)_{i,j}$$

$$= \frac{1}{(\Delta x)^2}\delta_x^2 U_{i,j} - \frac{1}{12}\frac{1}{(\Delta x)^2}\delta_x^4 U_{i,j} . \qquad (3.3.12)$$

When (3.3.12) is used in (3.3.8), we obtain the approximation,

$$u_{i,j+1} = u_{ij} + \frac{\Delta t}{(\Delta x)^2}\delta_x^2 u_{ij} + \frac{1}{(\Delta x)^4}[\tfrac{1}{2}(\Delta t)^2 - \tfrac{1}{12}(\Delta x)^2\Delta t]\delta_x^4 u_{i,j} \qquad (3.3.13)$$

with a local truncation error,

$$T = O([\Delta x]^4) + O([\Delta x]^2\Delta t) + O([\Delta t]^2) \ , \qquad (3.3.14)$$

and a stability ratio $\lambda \leqslant \frac{2}{3}$. If the maximum analysis is employed, it is found that,

$$E_j \leqslant N_3([\Delta x]^4+[\Delta x]^2\Delta t+[\Delta t]^2) \ , \qquad (3.3.15)$$

and for $\lambda = \frac{\Delta t}{(\Delta x)^2}$, $E_j = O((\Delta x)^4)$. Hence, we obtain improved accuracy and rate of convergence as well as a slightly better stability condition.

An adaptation of the Richardson's equation (3.3.1) is the formula of Dufort and Frankel (1953) which is both *explicit and unconditionally stable*. As we shall see, however, these two desirable properties are achieved at the expense of consistency. Dufort and Frankel replace the term $2u_{i,j}$ by $u_{i,j-1}+u_{i,j+1}$, thereby generating the *three-level formula*,

$$\frac{(u_{i,j+1}-u_{i,j-1})}{2\Delta t} = \frac{(u_{i+1,j}-(u_{i,j-1}+u_{i,j+1})+u_{i-1,j})}{(\Delta x)^2} \qquad (3.3.16)$$

which may be written as,

$$(1+2\lambda)u_{i,j+1} = 2\lambda(u_{i-1,j}+u_{i+1,j})+(1-2\lambda)u_{i,j-1} \ , \qquad (3.3.17)$$

where $\lambda = \frac{\Delta t}{(\Delta x)^2}$. For known boundary values with $m\Delta x=1$, these equations in matrix form are,

$$(1+2\lambda)\begin{bmatrix} u_{1,j+1} \\ u_{2,j+1} \\ u_{3,j+1} \\ \vdots \\ u_{m-1,j+1} \end{bmatrix} = 2\lambda\begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ & 1 & 0 & 1 & \bigcirc \\ & \bigcirc & & \ddots & \\ & & & 1 & 0 \end{bmatrix}\begin{bmatrix} u_{1j} \\ u_{2j} \\ u_{3j} \\ \vdots \\ u_{m-1,j} \end{bmatrix} + (1-2\lambda)\begin{bmatrix} u_{1,j-1} \\ u_{2,j-1} \\ u_{3,j-1} \\ \vdots \\ u_{m-1,j-1} \end{bmatrix} + 2\lambda\begin{bmatrix} u_{0,j} \\ 0 \\ 0 \\ \vdots \\ 0 \\ u_{m,j} \end{bmatrix}$$

giving

$$\underline{u}_{j+1} = \frac{2\lambda}{(1+2\lambda)} A\underline{u}_j + \frac{(1-2\lambda)}{(1+2\lambda)} \underline{u}_{j-1} + \underline{c}_j \ , \qquad (3.3.18)$$

where A is the matrix (of order (m-1)) as displayed and $\underline{c}_j$ is a vector of known values. If we put,

$$\underline{v}_j = (\underline{u}_j, \underline{u}_{j-1}) \ ,$$

then equation (3.3.18) can be written in the partitioned matrix form,

$$\left[\begin{array}{c} \underline{u}_{j+1} \\ \hline \underline{u}_j \end{array}\right] = \left[\begin{array}{c|c} \frac{2\lambda}{(1+2\lambda)}A & \frac{(1-2\lambda)}{(1+2\lambda)}I \\ \hline I & 0 \end{array}\right] \left[\begin{array}{c} \underline{u}_j \\ \hline \underline{u}_{j-1} \end{array}\right] + \left[\begin{array}{c} \underline{c}_j \\ \hline \underline{0} \end{array}\right] \ ,$$

i.e.,

$$\underline{v}_{j+1} = P\underline{v}_j + \underline{d}_j \ , \qquad (3.3.19)$$

where I is the unit matrix of order (m-1), P is the matrix shown and $\underline{d}_j$ a column vector of known constants. The reduced, two-level equations (3.3.19) will be stable when the eigenvalues of P are less than or equal to one in modulus. The matrix A has (m-1) different eigenvalues and so it has (m-1) linearly independent eigenvectors $\underline{v}_s$, s=1,2,...,m-1. Although I has (m-1) eigenvalues, each equal to 1, it has (m-1) linearly independent eigenvectors which may be taken as $\underline{v}_s$ since $I\underline{v}_s = 1.\underline{v}_s$. Hence, the eigenvalues μ of P are the eigenvalues of the matrix,

$$\begin{bmatrix} \frac{2\lambda\eta_k}{(1+2\lambda)} & \frac{(1-2\lambda)}{(1+2\lambda)} \\ 1 & 0 \end{bmatrix} \ ,$$

where η_k is the k^{th} eigenvalue of A given by,

$$\eta_k = 2\cos(\frac{k\pi}{m}) \quad , \ k=1,2,\ldots,m-1.$$

From,

$$\det\begin{bmatrix}\{\frac{2\lambda\eta_k}{1+2\lambda} - \mu\} & \frac{(1-2\lambda)}{(1+2\lambda)} \\ 1 & -\mu\end{bmatrix} = 0 ,$$

we get,

$$\mu^2 - \frac{2\lambda\eta_k}{(1+2\lambda)}\mu - \frac{(1-2\lambda)}{(1+2\lambda)} = 0 ,$$

the roots of which are obtained from,

$$\mu = \{2\lambda\cos(\frac{k\pi}{m}) \pm (1-4\lambda^2\sin^2(\frac{k\pi}{m}))^{\frac{1}{2}}\}/(1+2\lambda) . \qquad (3.3.20)$$

Case (i)

If $4\lambda^2\sin^2(\frac{k\pi}{m})\leqslant 1$, which corresponds to small values of Δt, then

$$|\mu| < \frac{(2\lambda+1)}{(1+2\lambda)} = 1 ,$$

Case (ii)

If $4\lambda^2\sin^2(\frac{k\pi}{m})>1$, which corresponds to large time steps (Δt),

then,

$$|\mu|^2 = \{(2\lambda\cos(\frac{k\pi}{m}))^2 + 4\lambda^2\sin^2(\frac{k\pi}{m})-1\}/(1+2\lambda)^2 ,$$

$$= \frac{(4\lambda^2-1)}{(4\lambda^2+4\lambda+1)} < 1 \text{ since } \lambda>0.$$

Therefore, the equations are unconditionally stable for all positive λ.

By the usual Taylor series expansion, we find that the local truncation error of the Dufort-Frankel scheme is

$$T = \frac{1}{6}(\Delta t)^2(\frac{\partial^3 U}{\partial t^3})_{i,j} - \frac{1}{12}(\Delta x)^2(\frac{\partial^4 U}{\partial x^4})_{i,j} + \frac{(\Delta t)^2}{(\Delta x)^2}(\frac{\partial^2 U}{\partial t^2})_{i,j} + O(\frac{[\Delta t]^4}{[\Delta x]^2}) \qquad (3.3.21)$$

Hence, $T = O([\Delta t]^2) + O([\Delta x]^2) + O(\frac{[\Delta t]^2}{[\Delta x]^2})$. (3.3.22)

Consistency requires that $\frac{\Delta t}{\Delta x} \rightarrow 0$ as $t \rightarrow 0$. This means that the Dufort-Frankel scheme (3.3.16) or (3.3.17) is *consistent* with the heat equation

(3.1.2) if and only if $\Delta t \to 0$ *faster* than $\Delta x \to 0$. If $\frac{\Delta t}{\Delta x}$ is kept fixed, say equal to β, then the third term of (3.3.21) would tend to $\beta^2 \frac{\partial^2 U}{\partial t^2}$, rather than to zero. Equation (3.3.16) or (3.3.17) would therefore be consistent with the hyperbolic equation

$$\frac{\partial U}{\partial t} - \frac{\partial^2 U}{\partial x^2} + \beta^2 \frac{\partial^2 U}{\partial t^2} = 0 \ ,$$

and not with the parabolic equation (3.1.2).

There are other explicit three-level schemes of high-order accuracy introduced by Russian mathematicians. An example is given by Miteladze and takes the form,

$$u_{i,j+1} = \frac{1}{40}[2u_{i,j-1}+32u_{i,j}+3(u_{i-1,j-1}+u_{i+1,j-1})] \ . \qquad (3.3.23)$$

Another formula is due to Yushkov and is given by,

$$u_{i,j+1} = \frac{1}{10}[3(u_{i-1,j}+u_{i+1,j})+2(u_{ij}+u_{i,j-1})] \qquad (3.3.24)$$

as well as,

$$u_{i,j+1} = \frac{1}{72}[25(u_{i-1,j}+u_{i+1,j})+4u_{i,j}+18u_{i,j-1}] \qquad (3.3.25)$$

Equations (3.3.23)-(3.3.25) must take on specific values of λ, i.e., $\lambda=\frac{1}{16}$, $\frac{1}{4}$ and $\frac{1}{36}$ respectively and all have local truncation errors of order $O([\Delta t]^2)+O([\Delta x]^4)$.

Following the argument of Jain (1984) *the general three-level explicit difference schemes* for (3.1.2) involves seven points and can be written as,

$$(1+\tau_1)u_{i,j+1} = [1+2\tau_1+\lambda(1-\gamma_1)\delta_x^2]u_{i,j} - (\tau_1-\lambda\gamma_1\delta_x^2)u_{i,j-1} \qquad (3.3.26)$$

where τ_1 and γ_1 are arbitrary parameters. The truncation error of (3.3.26) is of order:

(i) $(\Delta t+[\Delta x]^2)$ if γ_1 and τ_1 are arbitrary,

(ii) $([\Delta t]^2+[\Delta x]^2)$ if $\tau_1+\gamma_1+\frac{1}{2}=0$ and either γ_1 or τ_1 is arbitrary,

(iii) $([\Delta x]^4)$ if $\tau_1+\gamma_1+\frac{1}{2}-\frac{1}{12\lambda}=0$ and either γ_1 or τ_1 is arbitrary.

(3.3.27)

The necessary and sufficient conditions for (3.3.26) to be stable are:

(i) $1+2\tau_1-2(1-2\gamma_1)\lambda\geq 0$ and

(ii) $1-4\gamma_1\lambda\geq 0$,

(3.3.28)

We find that for,

(i) $\gamma_1\leq 0$, the conditions (3.3.28) are satisfied if,

$$0<\lambda\leq\frac{(1+2\tau_1)}{2(1-2\gamma_1)} \quad\text{and}\quad 1+2\tau_1>0 \ ,$$

(ii) $\gamma_1<\frac{1}{2}$, the stability condition is obtained as $0<\lambda<\lambda_{min}$, where $1+2\tau_1>0$ and

$$\lambda_{min}=\min\left[\frac{1}{4\gamma_1},\frac{(1+2\tau_1)}{2(1-2\gamma_1)}\right] \ . \qquad (3.3.29)$$

The above conditions are shown in Figure 3.3.1 below.

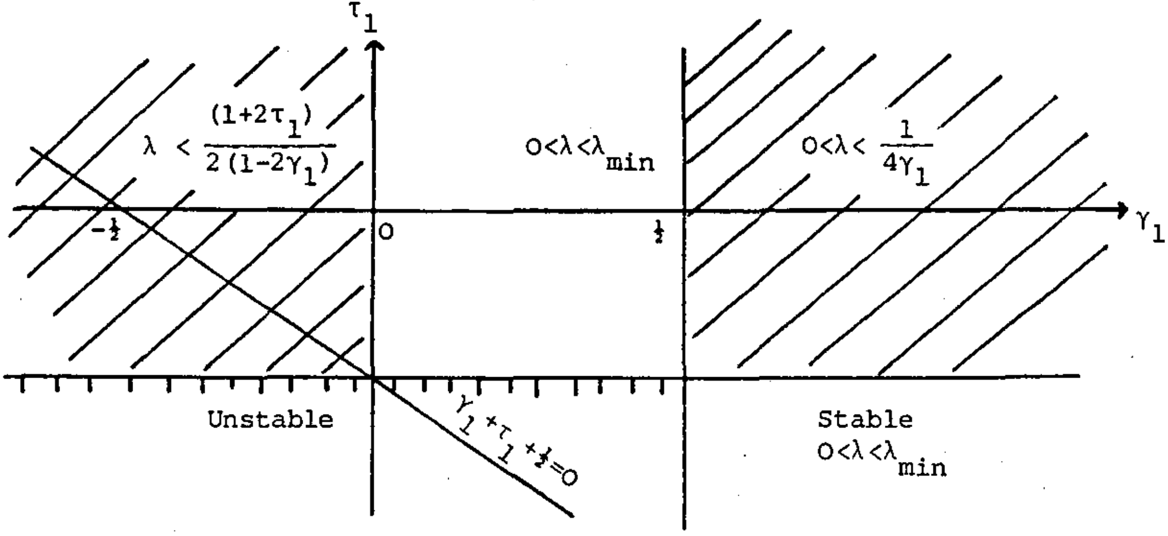


FIGURE 3.3.1

We now demonstrate the derivation of some well-known methods from equation (3.3.26). For $\gamma_1=0$ and $\tau_1=-\frac{1}{2}+\lambda$, we are led to the Dufort-Frankel formula $(\nabla_t-\frac{1}{2}\nabla_t^2+\lambda\nabla_t^2)u_{i,j+1}=\lambda\delta_x^2u_{i,j}$.

For $\gamma_1=0$ and $\tau_1=-\frac{1}{2}$ the unstable Richardson formula is obtained. If we prescribe γ_1 and τ_1 to satisfy (3.3.27) (iii), such as $\gamma_1=0$, $\tau_1=-\frac{1}{2}+\frac{1}{12\lambda}$, we get,

$$\tfrac{1}{2}(1+\frac{1}{6\lambda})u_{i,j+1} = \lambda(u_{i-1,j}+u_{i+1,j})-(2\lambda-\frac{1}{6\lambda})u_{i,j}+\tfrac{1}{2}(1-\frac{1}{6\lambda})u_{i,j-1}. \tag{3.3.30}$$

On the other hand, if $\tau_1=0$ and $\gamma_1=-\frac{1}{2}+\frac{1}{12\lambda}$, equation (3.3.26) yields,

$$\begin{aligned} u_{i,j+1} &= (\frac{7}{6}-3\lambda)u_{i,j} + \tfrac{1}{2}(3\lambda-\frac{1}{6})(u_{i-1,j}+u_{i+1,j}) \\ &\quad -\tfrac{1}{2}(\lambda-\frac{1}{6})(u_{i-1,j-1}-2u_{i,j-1}+u_{i+1,j-1}) , \end{aligned} \tag{3.3.31}$$

which is stable for $\lambda\leq\frac{1}{3}$. The local truncation error is $T=O([\Delta x]^6)$ if $\lambda=\frac{1}{10}$.

3.4 IMPLICIT METHODS

The simplest implicit method is that which was first suggested by O'Brien *et al* (1951). Upon approximating $\frac{\partial^2 U}{\partial x^2}$ of equations (3.1.2) in the (j+1) row instead of in the j row, we obtain,

$$\frac{(u_{i,j+1}-u_{i,j})}{\Delta t} = \frac{1}{(\Delta x)^2}(u_{i+1,j+1}-2u_{i,j+1}+u_{i-1,j+1}) \text{ ,} \qquad (3.4.1)$$

or
$$-\lambda u_{i-1,j+1}+(1+2\lambda)u_{i,j+1}-\lambda u_{i+1,j+1} = u_{i,j} \text{ .} \qquad (3.4.2)$$

Crank and Nicolson (1947) used an average of the approximations in the j and (j+1) row. More generally, one can introduce a *weighting parameter* θ and replace equation (3.4.1) by

$$u_{i,j+1}-u_{i,j} = \lambda\{\theta[u_{i-1,j+1}-2u_{i,j+1}+u_{i+1,j+1}]+(1-\theta)[u_{i-1,j}- 2u_{i,j}+u_{i+1,j}]\} \qquad (3.4.3)$$

with $0\leq\theta\leq 1$. A single application of equation (3.4.3) equates a linear combination of three unknowns in the (j+1) row to a linear combination of the three known values in the j row and can be solved by the usual Thomas algorithm. Thus, we obtain,

$$-\lambda\theta u_{i-1,j+1}+(1+2\lambda\theta)u_{i,j+1}-\lambda\theta u_{i+1,j+1}$$
$$= \lambda(1-\theta)u_{i-1,j}+[1-2\lambda(1-\theta)]u_{i,j}+\lambda(1-\theta)u_{i+1,j} \text{ .} \qquad (3.4.4)$$

If θ=1, equation (3.4.4) becomes the fully implicit O'Brien *et al* form of (3.4.2). If θ=½, we obtain the Crank-Nicolson formula. On the other hand, if θ=0 the explicit relation (3.2.1a) is recovered. We now examine the stability of (3.4.4) by the matrix method. We assume that the boundary values are zero and the initial values are known. With $u_{0,j}=u_{m,j}=0$, equation (3.4.4), can be written in the matrix form as,

$$\begin{bmatrix} (1+2\lambda\theta) & -\lambda\theta & & & \\ -\lambda\theta & (1+2\lambda\theta) & -\lambda\theta & & \bigcirc \\ & \ddots & \ddots & \ddots & \\ \bigcirc & & -\lambda\theta & (1+2\lambda\theta) & -\lambda\theta \\ & & & -\lambda\theta & (1+2\lambda\theta) \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j+1}$$

$$= \begin{bmatrix} (1-2\lambda(1-\theta)) & \lambda(1-\theta) & & & \\ \lambda(1-\theta) & (1-2\lambda(1-\theta)) & \lambda(1-\theta) & & \bigcirc \\ & \ddots & \ddots & \ddots & \\ \bigcirc & & \lambda(1-\theta) & (1-2\lambda(1-\theta)) & \lambda(1-\theta) \\ & & & \lambda(1-\theta) & (1-2\lambda(1-\theta)) \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j},$$

i.e.
$$(I-\lambda\theta T)\underline{u}_{j+1} = (I+\lambda(1-\theta)T)\underline{u}_j \,, \tag{3.4.5}$$

where I is the identity matrix of order (m-1) and T is as given in (3.2.16). From (3.4.5), we have,

$$\underline{u}_{j+1} = \Gamma\underline{u}_j \,, \tag{3.4.6}$$

where
$$\Gamma = (I-\lambda\theta T)^{-1}(I+\lambda(1-\theta)T) \,, \tag{3.4.7}$$

the amplification matrix. As before, the eigenvalues of T are given by,

$$\eta_i = -4\sin^2(\frac{i\pi}{2m}), \quad i=1,2,\ldots,m-1.$$

Hence, from equation (3.4.7), the eigenvalues of Γ are

$$\mu_i = \frac{(1-4\lambda(1-\theta)\sin^2(\frac{i\pi}{2m}))}{(1+4\lambda\theta\sin^2(\frac{i\pi}{2m}))} \,, \tag{3.4.8}$$

and the condition for stability is,

$$-1 \leqslant \frac{1-4\lambda(1-\theta)\sin^2(\frac{i\pi}{2m})}{1+4\lambda\theta\sin^2(\frac{i\pi}{2m})} \leqslant 1 . \qquad (3.4.9)$$

The upper inequality is automatically satisfied for $\lambda>0$. The lower inequality gives,

$$(8\lambda\theta-4\lambda)\sin^2(\frac{i\pi}{2m}) \geqslant -2.$$

Hence, $2\lambda(1-2\theta)\leqslant 1$. We conclude that the equation (3.4.4) is conditionally stable for $\lambda\leqslant \frac{1}{2(1-2\theta)}$ when $0\leqslant\theta<\frac{1}{2}$ (3.4.10)

and is

unconditionally stable for all $\lambda>0$ when $\frac{1}{2}\leqslant\theta\leqslant 1$. (3.4.11)

To analyse the local truncation error of the general method, we have from (3.4.4),

$$T = -\lambda\theta(U_{i-1,j+1}+U_{i+1,j+1})-\lambda(1-\theta)(U_{i-1,j}+U_{i+1,j})+(1+2\lambda\theta)U_{i,j+1}$$
$$-(1-2\lambda(1-\theta))U_{i,j} . \qquad (3.4.12)$$

On expanding the terms $U_{i,j+1}$, $U_{i-1,j}$, $U_{i+1,j}$, $U_{i-1,j+1}$ and $U_{i+1,j+1}$ about the mesh points (x_i,t_j) by the Taylor's series and on recognising that,

$$(\frac{\partial U}{\partial t})_{i,j} = (\frac{\partial^2 U}{\partial x^2})_{i,j}$$

$$(\frac{\partial^2 U}{\partial t^2})_{i,j} = (\frac{\partial^3 U}{\partial x^2 \partial t})_{i,j} = (\frac{\partial^4 U}{\partial x^4})_{i,j} \qquad (3.4.13)$$

$$(\frac{\partial^3 U}{\partial t^3})_{i,j} = (\frac{\partial^4 U}{\partial x^2 \partial t^2})_{i,j} = (\frac{\partial^5 U}{\partial x^4 \partial t})_{i,j}$$

and

$$(\frac{\partial^4 U}{\partial t^4})_{i,j} = (\frac{\partial^5 U}{\partial x^2 \partial t^3})_{i,j} \quad ,$$

we get,

$$T = (\frac{(\Delta t)^2}{2} -\lambda\theta(\Delta x)^2\Delta t - \frac{\lambda}{12}(\Delta x)^4)(\frac{\partial^2 U}{\partial t^2})_{i,j}+(\frac{(\Delta t)^2}{6} - \frac{\lambda\theta}{2}(\Delta x)^2(\Delta t)^2 - \frac{\lambda\theta}{12}$$
$$(\Delta x)^4\Delta t)(\frac{\partial^3 U}{\partial t^3})_{i,j} +$$

$$+ (\frac{(\Delta t)^4}{24} - \frac{1}{6}\lambda\theta(\Delta x)^2(\Delta t)^3)(\frac{\partial^4 U}{\partial t^4})_{i,j} + \frac{(\Delta t)^5}{120}(\frac{\partial^5 U}{\partial t^5})_{i,j} + \ldots \quad (3.4.14)$$

But $\lambda = \frac{\Delta t}{(\Delta x)^2}$. Therefore, we obtain,

$$T = (\frac{(\Delta t)^2}{2} - (\Delta t)^2\theta - \frac{\Delta t}{12}(\Delta x)^2)(\frac{\partial^2 U}{\partial t^2})_{i,j} + (\frac{(\Delta t)^3}{6} - \frac{(\Delta t)^3}{2}\theta - \frac{(\Delta t)^2}{12}(\Delta x)^2\theta)$$

$$(\frac{\partial^3 U}{\partial t^3})_{i,j} + (\frac{(\Delta t)^4}{24} - \frac{1}{6}(\Delta t)^4\theta)(\frac{\partial^4 U}{\partial t^4})_{i,j} + \frac{(\Delta t)^5}{120}(\frac{\partial^5 U}{\partial t^5})_{i,j} + \ldots \quad (3.4.15)$$

i.e.,

$$T = \Delta t\{[\Delta t(\tfrac{1}{2}-\theta) - \frac{(\Delta x)^2}{12}](\frac{\partial^2 U}{\partial t^2})_{i,j} + [\frac{(\Delta t)^2}{2}(\frac{1}{3}-\theta) - \frac{(\Delta t)}{12}(\Delta x)^2\theta](\frac{\partial^3 U}{\partial t^3})_{i,j}$$

$$+ \frac{(\Delta t)^3}{6}(\frac{1}{4}-\theta)(\frac{\partial^4 U}{\partial t^4})_{i,j} + \frac{(\Delta t)^4}{120}(\frac{\partial^5 U}{\partial t^5})_{i,j} + \ldots\} = 0.$$

Hence, an expression for the local truncation error is,

$$T = [\Delta t(\tfrac{1}{2}-\theta) - \frac{(\Delta x)^2}{12}](\frac{\partial^2 U}{\partial t^2})_{i,j} + [\frac{(\Delta t)^2}{2}(\frac{1}{3}-\theta) - \frac{(\Delta t)}{12}(\Delta x)^2\theta](\frac{\partial^3 U}{\partial t^3})_{i,j}$$

$$+ \frac{(\Delta t)^3}{6}(\frac{1}{4}-\theta)(\frac{\partial^4 U}{\partial t^4})_{i,j} + \frac{(\Delta t)^4}{120}(\frac{\partial^5 U}{\partial t^5})_{i,j} + \ldots \quad (3.4.16)$$

We find that for $\theta=0$ (the explicit scheme (3.2.1)),

$$T_1 = O(\Delta t)+O([\Delta x]^2) , \quad (3.4.17)$$

for $\theta=\frac{1}{2}$ (the Crank-Nicolson scheme),

$$T_2 = O([\Delta x]^2)+O([\Delta t]^2) , \quad (3.4.18)$$

and for $\theta=1$ (the fully implicit scheme (3.4.4)),

$$T_3 = O(\Delta t)+([\Delta x]^2) . \quad (3.4.19)$$

3.5 PARABOLIC PROBLEMS WITH DERIVATIVE BOUNDARY CONDITIONS

Boundary conditions which are expressed in terms of derivatives occur very frequently in practice. To solve (3.1.2), the boundary conditions (3.1.2a) are replaced by,

$$\begin{aligned} \frac{\partial U}{\partial x} &= h_1(U-v_1) \text{ at } x=0,\ 0<t\leq T \\ \frac{\partial U}{\partial x} &=-h_2(U-v_2) \text{ at } x=1,\ 0<t\leq T\ , \end{aligned} \qquad (3.5.1)$$

where h_1,h_2,v_1 and v_2 are constants with $h_1\geq 0$, $h_2\geq 0$. The boundary conditions may be approximated in a number of ways. For example, they may be approximated by central differences and an explicit scheme is used to approximate the differential equation. Alternatively, they can be replaced by forward differences at x=0 and backward differences at x=1 and an explicit scheme is employed for the solution.

When the boundary conditions are approximated by the central difference equations,

$$\frac{(u_{1,j}-u_{-1,j})}{2\Delta x} = h_1(u_{0,j}-v_1)$$

$$\frac{(u_{m+1,j}-u_{m-1,j})}{2\Delta x} = -h_2(u_{m,j}-v_2)$$

with $m\Delta x=1$, and the differential equation by the Crank-Nicolson scheme (3.4.3) with $\theta=\frac{1}{2}$, i.e.,

$$-\frac{\lambda}{2}u_{i-1,j+1}+(1+\lambda)u_{i,j+1}-\frac{\lambda}{2}u_{i+1,j+1}=\frac{\lambda}{2}u_{i-1,j}+(1-\lambda)u_{i,j}+\frac{\lambda}{2}u_{i+1,j}\ , \qquad (3.5.2)$$

the elimination of $u_{-1,j}$ and $u_{m+1,j}$ leads to the equation,

$$A\underline{u}_{j+1} = B\underline{u}_j + \underline{b}_j\ , \qquad (3.5.3)$$

where $\underline{u}_j = (u_{0,j},u_{1,j},\ldots,u_{m,j})^T$,

$$A = I-\tfrac{1}{2}\lambda Q\ , \qquad (3.5.4)$$

and

$$B = I+\tfrac{1}{2}\lambda Q\ , \qquad (3.5.5)$$

where Q is a matrix of order (m+1) given by,

$$Q = \begin{bmatrix} -2(1+\Delta x h_1) & 2 & & & \\ 1 & -2 & 1 & & \bigcirc \\ & \ddots & \ddots & \ddots & \\ \bigcirc & & 1 & -2 & 1 \\ & & & 2 & -2(1+\Delta x h_2) \end{bmatrix} \qquad (3.5.6)$$

and $\quad \underline{b}_j^T = (2\lambda v_1 \Delta x h_1, 0, \ldots, 0, 2\lambda v_2 \Delta x h_2)^T$.

For the analysis of stability we write (3.5.3) as

$$\underline{u}_{j+1} = A^{-1}B\underline{u}_j + A^{-1}\underline{b}_j \ ,$$

i.e. $\quad \underline{u}_{j+1} = (I-\frac{1}{2}\lambda Q)^{-1}(I+\frac{1}{2}\lambda Q)\underline{u}_j + \hat{\underline{b}}_j \ , \qquad (3.5.7)$

where $\quad \hat{\underline{b}}_j = (I-\frac{1}{2}\lambda Q)^{-1}\underline{b}_j$ and it is assumed that $\det(I-\frac{1}{2}\lambda Q) \neq 0$.

We note that Q is unsymmetric. Therefore, we introduce the diagonal matrix,

$$D = \begin{bmatrix} \sqrt{2} & & & & \\ & 1 & & \bigcirc & \\ & & \ddots & & \\ & \bigcirc & & 1 & \\ & & & & \sqrt{2} \end{bmatrix}_{(m+1)\times(m+1)} \qquad (3.5.8)$$

of order (m+1) such that Q is similar to the symmetric matrix,

$$\tilde{Q} = D^{-1}QD \ . \qquad (3.5.9)$$

Then,

$$\begin{aligned} (\widetilde{A^{-1}B}) &= D^{-1}(I-\tfrac{1}{2}\lambda Q)^{-1}(I+\tfrac{1}{2}\lambda Q)D \\ &= [D^{-1}(I-\tfrac{1}{2}\lambda Q)^{-1}D]\,[D^{-1}(I+\tfrac{1}{2}\lambda Q)D] \\ &= [D^{-1}(I-\tfrac{1}{2}\lambda Q)D]^{-1}[D^{-1}(I+\tfrac{1}{2}\lambda Q)D] \\ &= [I-\tfrac{1}{2}\lambda\tilde{Q}]\,[I+\tfrac{1}{2}\lambda\tilde{Q}] \ . \qquad (3.5.10) \end{aligned}$$

However, the matrices $(I-\frac{1}{2}\lambda\tilde{Q})$ and $(I+\frac{1}{2}\lambda\tilde{Q})$ are symmetric and commute, and so $(\widetilde{A^{-1}B})$ is symmetric. Therefore, $A^{-1}B$ is similar to the symmetric matrix $(\widetilde{A^{-1}B})$ and

$$\rho(A^{-1}B) = \rho(\widetilde{A^{-1}B}) \leqslant 1 \ , \qquad (3.5.11)$$

is a necessary and sufficient condition for stability, where ρ denotes the spectral radius.

The eigenvalues η_j $(j=0,1,\ldots,m)$ of $A^{-1}B$ are given by,

$$\eta_j = \frac{(1+\frac{1}{2}\lambda\mu_j)}{(1-\frac{1}{2}\lambda\mu_j)} \qquad (3.5.12)$$

where μ_j are the eigenvalues of the matrix Q. Since η and ρ are related by,

$$\rho(A^{-1}B) = \max_j |\eta_j| ,$$

the condition for stability (3.5.11) together with (3.5.12) gives,

$$\mu_j \leqslant 0 \text{ for all } j.$$

By means of the Gerschgorin's Circle Theorem, it is easily seen that μ_j lies on the negative line for all j. Hence, equation (3.5.3) is unconditionally stable.

3.6 IMPROVING THE ACCURACY OF IMPLICIT METHODS

(a) *Reduction of the Local Truncation Error*

As we have seen from Section 2.16 on hyperbolic equations, all derivatives of U(x,t) can be expressed exactly in terms of the *infinite series* of forward, backward or central differences. For example, in our previous analysis, we used

$$\frac{\partial^2 U}{\partial x^2} = \frac{1}{(\Delta x)^2}(\delta_x^2 U - \frac{1}{12}\delta_x^4 U + \frac{1}{90}\delta_x^6 U + \ldots) . \qquad (3.6.1)$$

When the diffusion equation (3.1.2) is approximated at the point $(x_i, t_{j+\frac{1}{2}})$ by

$$\begin{aligned}\frac{1}{\Delta t}(u_{i,j+1}-u_{ij}) &= \tfrac{1}{2}\{(\frac{\partial^2 u}{\partial x^2})_{i,j+1} + (\frac{\partial^2 u}{\partial x^2})_{i,j}\} \\ &= \frac{1}{2(\Delta x)^2}(\delta_x^2 - \frac{1}{12}\delta_x^4 + \frac{1}{90}\delta_x^6 + \ldots)(u_{i,j+1}+u_{ij}) ,\end{aligned} \qquad (3.6.2)$$

then the terms involving δ_x^4 can be eliminated by operating on both sides with $(1+\frac{1}{12}\delta_x^2)$. This gives us,

$$(1+\frac{1}{12}\delta_x^2)(u_{i,j+1}-u_{ij}) = \frac{1}{2}\lambda(\delta_x^2 u_{i,j+1}+\delta_x^2 u_{ij}) , \qquad (3.6.3)$$

where we have neglected the terms of order δ_x^6.

Equation (3.6.3) can be rearranged as,

$$[1+(\frac{1}{12} - \tfrac{1}{2}\lambda)\delta_x^2]u_{i,j+1} = [1+(\frac{1}{12} + \tfrac{1}{2}\lambda)\delta_x^2]u_{ij} ,$$

which on expanding leads to the *Douglas equation*,

$$\begin{aligned}&(1-6\lambda)u_{i-1,j+1}+(10+12\lambda)u_{i,j+1}+(1-6\lambda)u_{i+1,j+1} \\ &\quad = (1+6\lambda)u_{i-1,j}+(10-12\lambda)u_{i,j}+(1+6\lambda)u_{i+1,j} .\end{aligned} \qquad (3.6.4)$$

We note that the Douglas equation (3.6.4) can be considered as a special case of the weighted equation (3.4.3) by putting $\theta=\frac{1}{2} - \frac{1}{12\lambda}$.

By substituting this value of θ into (3.4.8), we find that

$$\mu_i = \frac{(1-\frac{1}{3}(6\lambda+1)\sin^2(\frac{i\pi}{2m}))}{(1+\frac{1}{3}(6\lambda+1)\sin^2(\frac{i\pi}{2m}))} \quad , \; i=1,2,\ldots,m-1,$$

and it is clear that $|\mu_i| \leqslant 1$ for all values of $\lambda>0$. Hence, the Douglas equation has unrestricted stability. Furthermore, with the same value of θ, the expression for the truncation error, i.e. equation (3.4.16) becomes,

$$T = \frac{1}{12}[-(\Delta t)^2 + \frac{1}{12}(\Delta x)^4](\frac{\partial^3 U}{\partial t^3})_{i,j} + \frac{1}{72}[\Delta t]^2((\Delta x)^2-3(\Delta t))(\frac{\partial^4 U}{\partial t^4})_{i,j}$$

$$+ \frac{(\Delta t)^4}{120}(\frac{\partial^5 U}{\partial t^5})_{i,j} + \ldots$$

i.e. $T = O([\Delta t]^2)+O([\Delta x]^4)$ as opposed to $T=O([\Delta t]^2)+O([\Delta x]^2)$ for the Crank-Nicolson method. We note that the resulting tridiagonal system of equations for the Douglas scheme can be solved by the Thomas algorithm and it requires the same amount of arithmetic as the Crank-Nicolson method.

(b) *Use of Three-Level Difference Equations*

In the construction of difference equations of high accuracy, or, occasionally, improved stability, one often uses more time levels than the minimum number required by the given differential equations. As an example, a fully implicit two-level approximation to the heat equation (3.1.2) is, from (3.4.1),

$$\frac{(u_{i,j+1}-u_{i,j})}{\Delta t} = \frac{(u_{i+1,j+1}-2u_{i,j+1}+u_{i-1,j+1})}{(\Delta x)^2}$$

and from (3.4.19), this has a local truncation error,

$$T = O(\Delta t) + O([\Delta x]^2) \; .$$

When the analogue for $\frac{\partial U}{\partial t}$ is replaced by $\frac{(\frac{3}{2}u_{i,j+1}-2u_{i,j}+\frac{1}{2}u_{i,j-1})}{\Delta t}$,

we obtain a three-level equation given by,

$$\frac{1}{\Delta t}(\frac{3}{2}u_{i,j+1}-2u_{ij}+\frac{1}{2}u_{i,j-1}) = \frac{1}{(\Delta x)^2}(u_{i+1,j+1}-2u_{i,j+1}+u_{i-1,j+1}) \tag{3.6.5}$$

which has a local truncation error of the order,

$$T = O([\Delta t]^2 + O([\Delta x]^2) \ . \tag{3.6.6}$$

As with the wave equation in Chapter 2, initial data are required at t=0 and t=Δt to start the calculation. The data at t=0 are given whilst the data at t=Δt are calculated from a two-level difference formula or by using a power series expansion. The data calculated at t=Δt should be of an accuracy comparable with that of the three-level scheme.

Equation (3.6.5) can be written as,

$$-2\lambda u_{i-1,j+1}+(3+4\lambda)u_{i,j+1}-2\lambda u_{i+1,j+1} = 4u_{i,j}-u_{i,j-1} \ .$$

In matrix form, this becomes,

$$\underline{u}_{j+1} = 4A^{-1}\underline{u}_j-A^{-1}\underline{u}_{j-1}+A^{-1}\underline{c}_{j+1} \ , \tag{3.6.7}$$

where $\underline{c}_{j+1}$ is a vector of known boundary values and

$$A = \begin{bmatrix} (3+4\lambda) & -2\lambda & & & \\ -2\lambda & (3+4\lambda) & -2\lambda & & O \\ & \ddots & \ddots & \ddots & \\ O & & & & \\ & & & -2\lambda & (3+4\lambda) \end{bmatrix} \quad \text{of order } (m-1).$$

Equation (3.6.7) can be written as,

$$\begin{bmatrix} \underline{u}_{j+1} \\ \hline \underline{u}_j \end{bmatrix} = \left[\begin{array}{c|c} 4A^{-1} & -A^{-1} \\ \hline I & O \end{array}\right] \begin{bmatrix} \underline{u}_j \\ \hline \underline{u}_{j-1} \end{bmatrix} + \begin{bmatrix} A^{-1}\underline{c}_{j+1} \\ \hline \underline{O} \end{bmatrix} ,$$

i.e. $\underline{v}_{j+1} = P\underline{v}_j + \underline{c}'$. As before, the eigenvalues μ of P are the eigenvalues of,

$$\begin{bmatrix} \frac{4}{\eta_k} & -\frac{1}{\eta_k} \\ 1 & 0 \end{bmatrix}$$

where η_k is the k^{th} eigenvalue of A given by,

$$\eta_k = 3+8\lambda\sin^2(\frac{k\pi}{2m}) \quad , \quad k=1,2,\ldots,m-1.$$

As $\det(P-\mu I)=0$

$$= \mu^2-(\frac{4}{\eta_k})\mu + (\frac{1}{\eta_k})$$

then
$$\mu = [2\pm\sqrt{(4-\eta_k)}]/\eta_k$$

$$= [2\pm\sqrt{(1-8\lambda\sin^2(\frac{k\pi}{2m})}]/(3+8\lambda\sin^2(\frac{k\pi}{2m})) \ .$$

When the roots are real, $|\mu|<(2+1)/(3+\delta)$, $\delta>0$, Hence $|\mu|<1$.
When the roots are complex, $|\mu|=1/\sqrt{\eta_k}<1$. Therefore, the approximation (3.6.7) is unconditionally stable.

We note that the truncation error of (3.6.5) is of the same order as the Crank-Nicolson method and is often used when the initial data are discontinuous or varies rapidly with x since it damps the short-wavelength components (components with $\beta\Delta x\simeq\pi$) more rapidly.

The Crank-Nicolson approximation should be used when the initial data and its derivatives are continuous. This is because the coefficient of $(\Delta t)^2$ in its truncation error is smaller. A three-level variation of the Douglas equation is,

$$\frac{1}{12(\Delta t)}\{\tfrac{3}{2}(u_{i+1,j+1}-u_{i+1,j})\}-\tfrac{1}{2}(u_{i+1,j}-u_{i+1,j-1})\} + \frac{5}{6(\Delta t)}$$
$$\{\tfrac{3}{2}(u_{i,j+1}-u_{ij}) - \tfrac{1}{2}(u_{ij}-u_{i,j-1})\} + \frac{1}{12(\Delta t)}\{\tfrac{3}{2}(u_{i-1,j+1}-u_{i-1,j})$$
$$-\tfrac{1}{2}(u_{u-1,j}-u_{i-1,j-1})\}$$
$$= \frac{1}{(\Delta x)^2}(u_{i+1,j+1}-2u_{i,j+1}+u_{i-1,j+1}) \ , \qquad (3.6.8)$$

and like the Douglas equation, is stable and has a truncation error $T=O([\Delta t]^2)+O([\Delta x]^4)$.

Let us now consider another method which is second-order accurate in both space and time. If we replace $\frac{\partial^2 u}{\partial x^2}$ by the average of the second differences at t_{j-1}, t_j and t_{j+1} and $\frac{\partial u}{\partial t}$ by a central first difference, we obtain

$$\frac{(u_{i,j+1}-u_{i,j-1})}{2(\Delta t)} = \frac{1}{3}\frac{\delta_x^2}{(\Delta x)^2}(u_{i,j+1}+u_{ij}+u_{i,j-1}) \qquad (3.6.9)$$

where $\delta_x^2 u_{ij}=u_{i+1,j}-2u_{ij}+u_{i-1,j}$ and $T = O([\Delta t]^2)+O([\Delta x]^2)$. The method (3.6.9) can now be improved to attain a higher order accuracy. As

$$(\frac{\partial^2 U}{\partial x^2})_{i,j} = \frac{1}{3}\frac{\delta_x^2}{(\Delta x)^2}(U_{i,j+1}+U_{ij}+U_{i,j-1}) - \frac{(\Delta x)^2}{12}(\frac{\partial^4 U}{\partial x^4})_{i,j}+O([\Delta t]^2)+O([\Delta x]^4), \qquad (3.6.10)$$

then,

$$(\frac{\partial^2 U}{\partial x^2})_{i,j} = \frac{1}{3}\frac{\delta_x^2}{(\Delta x)^2}(U_{i,j+1}+U_{i,j}+U_{i,j-1}) - \frac{(\Delta x)^2}{12}(\frac{\partial^2 U}{\partial t^2})_{i,j} +O([\Delta t]^2)+O([\Delta x]^4) \qquad (3.6.11)$$

$$= \frac{1}{3}\frac{\delta_x^2}{(\Delta x)^2}(U_{i,j+1}+U_{i,j}+U_{i,j-1}) - \frac{(\Delta x)^2}{12}\frac{\delta_t^2}{(\Delta t)^2}U_{i,j}+ O([\Delta t]^2)+O([\Delta x]^4) \ .$$

Thus, we obtain the difference equation,

$$\frac{(u_{i,j+1}-u_{i,j-1})}{2\Delta t} = \frac{1}{3}\frac{\delta_x^2}{(\Delta x)^2}(u_{i,j+1}+u_{i,j}+u_{i,j-1}) - \frac{(\Delta x)^2}{12}\frac{\delta^2}{(\Delta t)^2}u_{i,j} \qquad (3.6.12)$$

with a local truncation error $T = O([\Delta t]^2)+O([\Delta x]^4)$.

An alternative substitute for $\frac{\partial^4 U}{\partial x^4}$ is $\frac{\partial^3 U}{\partial x^2 \partial t}$. This substitution leads to the difference equation,

$$\frac{(u_{i,j+1}-u_{i,j-1})}{2\Delta t} = \frac{1}{3}\frac{\delta_x^2}{(\Delta x)^2}(u_{i,j+1}+u_{i,j}+u_{i,j-1}) - \frac{1}{24(\Delta t)}\delta_x^2 (u_{i,j+1}-u_{i,j-1}) \qquad (3.6.13)$$

which is also fourth-order correct in space and second-order in time for a fixed λ.

The *general three-level implicit formulae* based upon nine points for approximating (3.1.2) is given by (see Jain (1984)),

$$[(1+\tau_1)+\{\alpha(1+\tau_1)-\lambda(1-\gamma_1+\gamma_2)\}\delta_x^2]u_{i,j+1}$$

$$= [(1+2\tau_1)+\{\alpha(1+2\tau_1)+\lambda(\gamma_1-2\gamma_2)\}\delta_x^2]u_{i,j}-[\tau_1+(\alpha\tau_1-\lambda\gamma_2)\delta_x^2]u_{i,j-1} \tag{3.6.14}$$

where τ_1,γ_1,γ_2 and α are arbitrary constants. The truncation error for this class of methods is given by,

$$T = \Delta t[(\alpha-\frac{1}{12})+\lambda(\tau_1+\gamma_1-\tfrac{1}{2})](\Delta x)^2(\frac{\partial^4 U}{\partial x^4})_{i,j} +\Delta t[\tfrac{1}{2}\gamma_1-\gamma_2-\frac{1}{3})\lambda^2$$

$$+(\alpha(\tau_1+\tfrac{1}{2})+\frac{1}{12}(\gamma_1-1))\lambda+\frac{1}{12}(\alpha-\frac{1}{30})](\Delta x)^4(\frac{\partial^6 U}{\partial x^6})_{i,j}+\ldots \tag{3.6.15}$$

and the necessary and sufficient conditions for stability, as given by Jain (1979), are,

$$(\tau_1+\tfrac{1}{2})(1-4\alpha\sin^2(\frac{\beta\Delta x}{2}))+\lambda\sin^2(\frac{\beta\Delta x}{2})(1-2\gamma_1+4\gamma_2) > 0$$

and $$\frac{1}{4}(1-4\alpha\sin^2(\frac{\beta\Delta x}{2}))+\lambda\sin^2(\frac{\beta\Delta x}{2})(1-\gamma_1) > 0 \tag{3.6.16}$$

where $\alpha<\frac{1}{4}$ is taken.

Finally, it is worth noting that more than three time levels can also be used. For instance, the difference equation,

$$\frac{\delta_x^2}{(\Delta x)^2}[\tfrac{1}{2}u_{i,j+1}+\frac{1}{3}(u_{i,j}+u_{i,j-1}+u_{i,j-2})] = \frac{(u_{i,j+1}-u_{i,j-1})}{2\Delta t} \tag{3.6.17}$$

is a stable, second-order analogue of the Heat Equation. The determination of stability for more than three levels, however, can be very complicated. The works of Jain (1984), Ritchmyer and Morton (1967) and Saulev (1964) may be consulted for further discussion on multi-level schemes.

(c) *The Deferred Correction Method*

In this method, the finite difference equations approximating the given differential equation are solved as usual and their solution is then used to calculate a *correction term* at each mesh point, of the solution domain. This term is added to the approximating difference equation which in turn is solved again and the process is repeated if necessary. The correction terms are numbers obtained by differencing the numerical solution in either the x- or the t-direction or both. To derive the correction term, we will be using the following operators (c.f. Hildebrand (1956)),

operator μ defined by $\mu f_{j+\frac{1}{2}} = \frac{1}{2}(f_j+f_{j+1})$, (3.6.18a)

operator δ defined by $\delta f_{j+\frac{1}{2}} = f_{j+1}-f_j$, (3.6.18b)

and the following results,

$$\Delta t \frac{\partial}{\partial t} \equiv 2\sinh^{-1}(\tfrac{1}{2}\delta_t) \tag{3.6.19}$$

and

$$\mu_t \equiv (1+\tfrac{1}{4}\delta_t^2)^{\frac{1}{2}} \quad . \tag{3.6.20}$$

Now,

$$\tfrac{1}{2}\Delta t\{(\frac{\partial U}{\partial t})_{i,j}+(\frac{\partial U}{\partial t})_{i,j+1}\} = \Delta t\mu_t(\frac{\partial U}{\partial t})_{i,j+\frac{1}{2}}$$

from (3.6.18a).

Then by using (3.6.19) and (3.6.20), this becomes $\{(1+\frac{1}{4}\delta_t^2)^{\frac{1}{2}}2\sinh^{-1}(\frac{1}{2}\delta_t)\}U_{i,j+\frac{1}{2}}$. We find that when this expression is expanded into positive powers of δ_t,

$$\tfrac{1}{2}\Delta t\{(\frac{\partial U}{\partial t})_{i,j}+(\frac{\partial U}{\partial t})_{i,j+1}\} = (\delta_t+\frac{1}{12}\delta_t^3-\frac{1}{120}\delta_t^5+\ldots)U_{i,j+\frac{1}{2}}$$

which can be rearranged as,

$$\delta_t U_{i,j+\frac{1}{2}} = \tfrac{1}{2}\Delta t\{(\frac{\partial U}{\partial t})_{i,j} + (\frac{\partial U}{\partial t})_{i,j+1}\}+ C_t U_{i,j+\frac{1}{2}} \ .$$

This gives,

$$U_{i,j+1}-U_{i,j} = \tfrac{1}{2}\Delta t \frac{\partial}{\partial t}(U_{i,j}+U_{i,j+1}) + C_t U_{i,j+\frac{1}{2}} \ , \tag{3.6.21}$$

where, $$C_t = -\frac{1}{12}\delta_t^3 + \frac{1}{120}\delta_t^5 + \dots \qquad (3.6.22)$$

From the differential equation,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} \quad ,$$

it follows that, $$\frac{\partial}{\partial t} \equiv \frac{\partial^2}{\partial x^2} \quad .$$

Hence, equation (3.6.21) can be written as,

$$U_{i,j+1}-U_{ij} = \tfrac{1}{2}\Delta t \frac{\partial^2}{\partial x^2}(U_{ij}+U_{i,j+1})+C_tU_{i,j+\frac{1}{2}} \; .$$

By virtue of (3.6.1), this becomes,

$$U_{i,j+1}-U_{ij} = \tfrac{1}{2}\frac{\Delta t}{(\Delta x)^2}(\delta_x^2 - \frac{1}{12}\delta_x^4 + \frac{1}{90}\delta_x^6+\dots)(U_{i,j+1}+U_{ij})+C_tU_{i,j+\frac{1}{2}}$$

$$= \tfrac{1}{2}\lambda(\delta_x^2U_{i,j+1}+\delta_x^2U_{i,j})+C \qquad (3.6.23)$$

where,

$$C = \tfrac{1}{2}\lambda\{(-\frac{1}{12}\delta_x^4U_{i,j+1}+\frac{1}{90}\delta_x^6U_{i,j+1}+\dots)$$

$$+(-\frac{1}{12}\delta_x^4U_{i,j}+\frac{1}{90}\delta_x^6U_{i,j}+\dots)\}+(-\frac{1}{12}\delta_t^3U_{i,j+\frac{1}{2}}+\frac{1}{120}\delta_t^5U_{i,j+\frac{1}{2}})$$

(3.6.24)

In our first approximation, the correction term C is obviously 0 and the resulting Crank-Nicolson equation,

$$u_{i,j+1}-u_{ij} = \tfrac{1}{2}\lambda(\delta_x^2u_{i,j+1}+\delta_x^2u_{i,j}) \; ,$$

is solved in the usual manner as before. In the next step, the correction term must now be calculated from a truncated approximation to C of equation (3.6.24). An example of such a truncation at the mesh point is,

$$C = -\frac{\lambda}{24}(\delta_x^4U_{i,j+1}+\delta_x^4u_{i,j}) - \frac{1}{12}\delta_t^3u_{i,j+\frac{1}{2}} \qquad (3.6.25)$$

and equation (3.6.23) is solved again. We observe that this method, effectively, includes *higher-order* difference terms in the approximations to the derivatives. It is therefore essential that during computation, double precision arithmetic is used in evaluating these terms. The tridiagonal nature of the resulting system is retained and solved using the Thomas elimination process.

(d) *Richardson's Extrapolation to the Limit*

It is well known (see, for example, Henrici (1962)), that if,

$$y_1(x+\Delta x) = y_1(x) + f(x,y_1(x))(\Delta x) \quad , \quad x=k\Delta x \ , \ k=0,1,\ldots$$

$$y_2(x+\frac{\Delta x}{2}) = y_2(x)+f(x,y_2(x))\,(\frac{\Delta x}{2}) \quad , \quad x=\tfrac{1}{2}k\Delta x, \ k=0,1,\ldots$$

then the linear combination,

$$Y(x) = 2y_2(x)-y_1(x) \quad , \quad x=k\Delta x, \ k=0,1,\ldots$$

is a second-order approximation to the solution of

$$\frac{dU}{dx} = f(x,U) \ ,$$

even though $y_1(x)$ and $y_2(x)$ are only first-order correct. This is an example of a technique known as the Richardson's extrapolation to the limit which can be extended to difference analogues of parabolic equations. Let us consider the *backward difference equation* as an example. Then,

$$\frac{1}{(\Delta x)^2}\,\delta_x^2 U_{i,j+1} = \frac{(U_{i,j+1}-U_{ij})}{\Delta t} - \tfrac{1}{2}(\frac{\partial^2 U}{\partial t^2})_{i,j+1}(\Delta t) + T_{ij} \ , \qquad (3.6.26)$$

where $\quad T_{ij} = O([\Delta x]^2)+O([\Delta t]^2)$

for a sufficiently smooth U. It is clear that the difference analogue for equation (3.6.26) to approximate (3.1.2) is,

$$\frac{1}{(\Delta x)^2}\,\delta_x^2 u_{i,j+1} = \frac{(u_{i,j+1}-u_{i,j})}{\Delta t} \qquad (3.6.27)$$

with a truncation error $O([\Delta x]^2)+O[\Delta t]$.

As previously, if we let ε_{ij} be the discretisation error at the mesh point, i.e. $\varepsilon_{ij}=U_{ij}-u_{ij}$, then from equations (3.6.26) and (3.6.27) we have,

$$\frac{1}{(\Delta x)^2}\,\delta_x^2\varepsilon_{i,j+1} = \frac{(\varepsilon_{i,j+1}-\varepsilon_{ij})}{\Delta t} - \tfrac{1}{2}\left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j+1}(\Delta t) + T_{ij} \qquad (3.6.28)$$

with $\varepsilon=0$ on B (the boundary).

Now, we consider the effects of the two error terms separately as follows. Let $E(x,t)$ be the solution of

$$\frac{\partial^2 E}{\partial x^2} = \frac{\partial E}{\partial t} - \tfrac{1}{2}\,\frac{\partial^2 U}{\partial t^2}$$

with $\quad E(x,t) = 0$ on B . (3.6.29)

Then,

$$E(x,t) = O(1)\ , \qquad (3.6.30)$$

and,

$$\frac{1}{(\Delta x)^2}\delta_x^2 E_{i,j+1} = \frac{(E_{i,j+1}-E_{ij})}{\Delta t} - \tfrac{1}{2}\left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j+1}+O([\Delta x]^2)+O(\Delta t)\ . \qquad (3.6.31)$$

If we multiply equation (3.6.31) with Δt, we get,

$$\frac{1}{(\Delta x)^2}\delta_x^2 E_{i,j+1}\Delta t = (E_{i,j+1}-E_{i,j})-\tfrac{1}{2}\left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j}\Delta t+O((\Delta t)[\Delta x]^2)+O([\Delta t]^2)\ . \qquad (3.6.32)$$

By subtracting equation (3.6.32) from (3.6.28) yields,

$$\frac{1}{(\Delta x)^2}\delta_x^2(\varepsilon_{i,j+1}-E_{i,j+1}\Delta t) = \frac{(\varepsilon_{i,j+1}-E_{i,j+1}\Delta t)-(\varepsilon_{ij}-E_{ij}\Delta t)}{\Delta t} + O([\Delta t]^2)+O([\Delta x]^2)\ ,$$

i.e.,

$$\frac{1}{(\Delta x)^2}\delta_x^2\, e_{i,j+1} = \frac{(e_{i,j+1}-e_{ij})}{\Delta t} + O([\Delta t]^2)+O([\Delta x]^2)$$

with $\quad e_{ij} = 0$ on B , (3.6.33)

where $\quad e_{ij} = \varepsilon_{ij}-E_{ij}\Delta t$. (3.6.34)

It is clear that,

$$||\underline{e}_j|| = \max_i |e_{ij}| = O([\Delta t]^2)+O[\Delta x]^2)$$

and $$\varepsilon_{ij} = E_{ij}\Delta t + O([\Delta x]^2)+O([\Delta t]^2) \quad . \qquad (3.6.35)$$

With this relation, we are now in a position to improve the convergence rate. Let $u_1(x,t)$ denote the solution of the difference equation corresponding to $(\Delta x,\Delta t)$ and similarly for $u_2(x,t)$ corresponding to $(\Delta x, \frac{\Delta t}{2})$. Both $u_1(x,t)$ and $u_2(x,t)$, as we know, have an order of accuracy of $O([\Delta x]^2)+O(\Delta t)$. Let us now take the linear combination $v(x,t) = 2u_2(x,t)-u_1(x,t)$, $x=i\Delta x$, $t=j\Delta t$. If we denote the discretization errors for $u_1(x,t)$, $u_2(x,t)$ and $v(x,t)$ by $\varepsilon_1(x,t)$, $\varepsilon_2(x,t)$ and $\varepsilon(x,t)$ respectively, then,

$$\begin{aligned}\varepsilon(x,t) &= 2\varepsilon_2(x,t)-\varepsilon_1(x,t) \quad \text{and from (3.6.35)},\\ &= 2[E(x,t)\frac{\Delta t}{2} + O([\Delta x]^2)+O([\Delta t]^2)]\\ &\quad -[E(x,t)\Delta t+O([\Delta x]^2)+O([\Delta t]^2)]\\ &= O([\Delta x]^2)+O([\Delta t]^2) \quad .\end{aligned}$$

Hence, we see that the extrapolation strategy leads to an increase in the order of accuracy. The above argument remains valid for the general linear differential equation of the form,

$$\frac{\partial U}{\partial t} = a(x)\ \frac{\partial^2 U}{\partial x^2} + b(x)\ \frac{\partial U}{\partial x} + c(x)U + d(x,t) \quad .$$

A proper linear combination of several solutions should lead to the elimination of several leading error terms and therefore increase the accuracy in both the space and time directions. The ideas can also be extended to several space variables.

3.7 OTHER METHODS TO SOLVE THE DIFFUSION EQUATION

Some of the well-known finite-difference methods that are available for the treatment of parabolic equations are summarised in Table 3.7.1 below. The truncation error and the stability requirement of each method are also included.

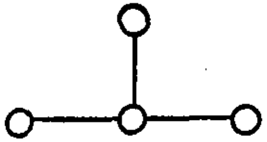
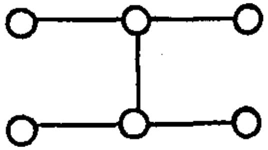
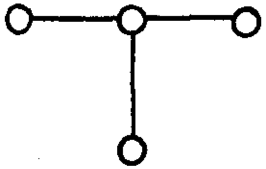
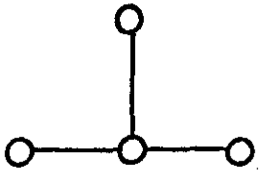
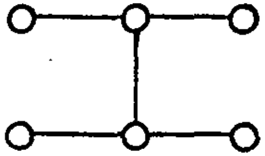
Finite Difference Form	Equation	Stability Condition	Truncation Error	Computational Molecule
1. $u_{i,j+1}=u_{ij}+\lambda(u_{i+1,j}-2u_{ij}+u_{i-1,j})$ (Classical explicit method)	(3.2.1) Explicit	$\lambda\leqslant\frac{1}{2}$	$T=O(\Delta t)+O([\Delta x]^2)$	
2. $\frac{\lambda}{2}u_{i+1,j+1}-(1+\lambda)u_{i,j+1}+\frac{\lambda}{2}u_{i-1,j+1}=-\frac{\lambda}{2}u_{i+1,j}-(1-\lambda)u_{ij}-\frac{\lambda}{2}u_{i-1,j}$ (Crank-Nicolson method)	(3.4.4) with $\theta=\frac{1}{2}$ Implicit	Always stable	$T=O([\Delta t]^2)+O([\Delta x]^2)$	
3. $-\lambda u_{i+1,j+1}+(1+2\lambda)u_{i,j+1}-\lambda u_{i-1,j+1}=u_{ij}$ (Fully implicit method)	(3.4.4) with $\theta=1$ Implicit	Always stable	$T=O(\Delta t)+O([\Delta x]^2)$	
4. $u_{i,j+1}=u_{ij}+\frac{1}{6}(u_{i+1,j}-2u_{ij}+u_{i-1,j})$ (Special explicit method)	(3.2.1) with $\lambda=\frac{1}{6}$ Explicit	Stable	$T=O([\Delta t]^2)=O([\Delta x]^4)$	
5. $\lambda\theta u_{i+1,j+1}-(1+2\lambda\theta)u_{i,j+1}+\lambda\theta u_{i-1,j+1}$ $=-\lambda(1-\theta)u_{i+1,j}-(1-2\lambda(1-\theta))u_{ij}-\lambda(1-\theta)u_{i-1,j}$ (Weighted formula $(0\leqslant\theta\leqslant 1)$)	(3.4.4) Implicit	$\lambda\leqslant\frac{1}{(2-4\theta)}$ for $0\leqslant\theta\leqslant 1$ Always stable for $\frac{1}{2}\leqslant\theta\leqslant 1$	$T=\begin{cases}O((\Delta t)^2)+O((\Delta x)^2) & \text{for } \theta=\frac{1}{2}\\ O(\Delta t)+O((\Delta x)^2) & \text{for } \theta\neq\frac{1}{2}\end{cases}$	

TABLE 3.7.1: Finite-difference approximations to $\frac{\partial U}{\partial t}=\frac{\partial^2 U}{\partial x^2}$

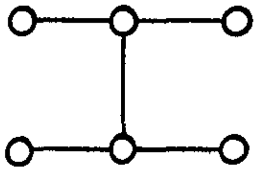
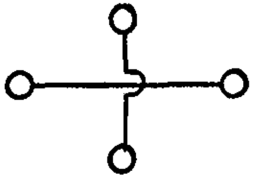
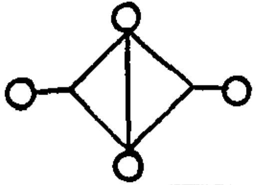
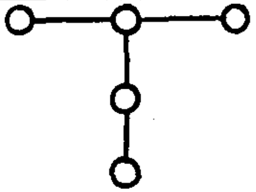
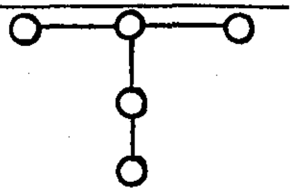
Finite Difference Form	Equation	Stability Condition	Truncation Error	Computational Molecule
6. As in 5. with $\theta=\frac{1}{2}-\frac{1}{12\lambda}$ (Douglas Method)	(3.6.3) Implicit	Always stable	$T=O([\Delta t]^2)+O([\Delta x]^4)$	
7. $u_{i,j+1}=u_{i,j-1}+2\lambda(u_{i+1,j}-2u_{ij}+u_{i-1,j})$ (Richardson method)	(3.3.1) Explicit, 3-Level	Always unstable	$T=O([\Delta t]^2)+O([\Delta x]^2)$	
8. $u_{i,j+1}=\frac{1}{(1+2\lambda)}\{2\lambda(u_{i+1,j}+u_{i-1,j})+(1-2\lambda)u_{i,j-1}\}$ (Dufort-Frankel method)	(3.3.16) Explicit, 3-Level	Always stable	$T=O([\Delta t]^2)+O([\Delta x]^2)+$ $O([\frac{\Delta t}{\Delta x}]^2)$ with $\frac{\Delta t}{\Delta x}\to 0$ as $\Delta t,\Delta x\to 0$	
9. $\lambda u_{i+1,j+1}-\frac{1}{2}(3+4\lambda)u_{i,j+1}+\lambda u_{i-1,j+1}=-2u_{ij}+\frac{1}{2}u_{i,j-1}$	(3.6.5) Implicit, 3-Level	Always stable	$T=O((\Delta t)^2)+O([\Delta x]^2)$	
10. $\lambda u_{i+1,j+1}-(1+\theta+2\lambda)u_{i,j+1}+\lambda u_{i-1,j}=-(1+2\theta)u_{ij}+u_{i,j-1}$ where $\theta\geq 0$	Implicit, 3-Level	Always stable	$T=\begin{cases}O([\Delta t]^2)+O([\Delta x]^2) & \text{for } \theta=\frac{1}{2}\\ O(\Delta t)+O([\Delta x]^2)\end{cases}$	

TABLE 3.7.1: (continued).

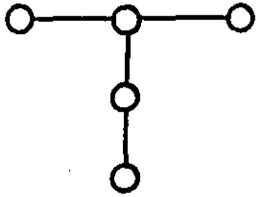
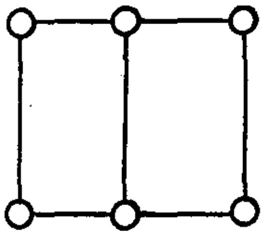
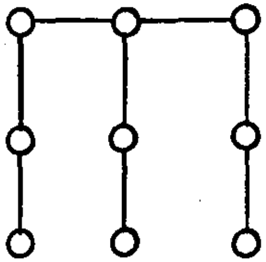
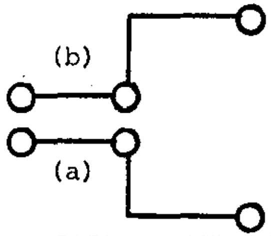
Finite Difference Form	Equation	Stability Condition	Truncation Error	Computational Molecule
11. As in 10. with $\theta=\frac{1}{2}+\frac{1}{12\lambda}$	Implicit, 3-Level	Always stable	$T=O([\Delta t]^2)=O([\Delta x]^4)$	
12. $(\frac{1}{12}-\frac{\lambda}{2})u_{i+1,j+1}+(\frac{5}{6}+\lambda)u_{i,j+1}+(\frac{1}{12}-\frac{\lambda}{2})u_{i-1,j+1}$ $=(\frac{1}{12}+\frac{\lambda}{2})u_{i+1,j}+(\frac{5}{6}-\lambda)u_{ij}+(\frac{1}{12}+\frac{\lambda}{2})u_{i-1,j}$ (Douglas method)	(3.6.4) Implicit	Always stable	$T=O([\Delta t]^2)+O([\Delta x]^4)$	
13. $(\frac{1}{8}-\lambda)u_{i+1,j+1}+(\frac{5}{4}+2\lambda)u_{i,j+1}+(\frac{1}{8}-\lambda)u_{i-1,j+1}$ $=\frac{1}{6}u_{i+1,j}+\frac{5}{3}u_{ij}+\frac{1}{6}u_{i-1,j}-\frac{1}{24}u_{i+1,j-1}-\frac{5}{12}u_{i,j-1}-\frac{1}{24}u_{i-1,j-1}$ (Variation of Douglas Equation)	(3.6.8) Implicit, 3-Level	Always stable	$T=O([\Delta t]^2)+O([\Delta x]^4)$	
14 (a) $u_{i,j+1}-u_{ij}=\lambda(u_{i+1,j}-u_{ij}-u_{i,j+1}+u_{i-1,j+1})$ (b) $u_{i,j+2}-u_{i,j+1}=\lambda(u_{i+1,j+2}-u_{i,j+2}-u_{i,j+1}+u_{i-1,j+1})$ (Saulev's alternating method)	Semi- implicit	Always stable	$T=O(\Delta t)=O([\Delta x]^2)$ for fixed λ	

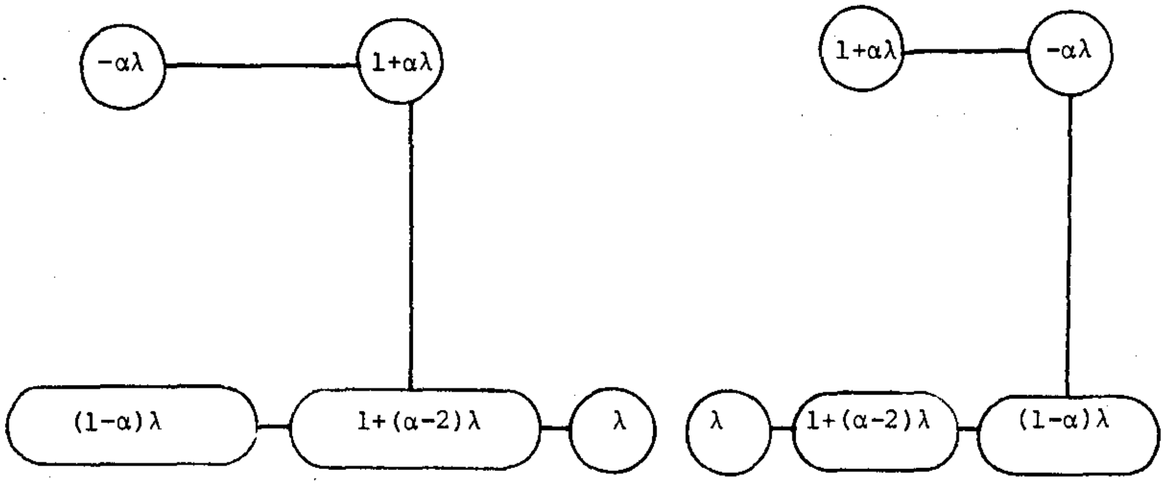
TABLE 3.7.1: continued

We note that each of the equations in entry 14. of Table 3.7.1 belongs to a class of *asymmetric finite difference equations* of order $O(\Delta x)$ introduced by Saulev (1964) of the form,

$$(1+\alpha\lambda)u_{i,j+1}-\alpha\lambda u_{i-1,j+1} = (1-\alpha)\lambda u_{i-1,j}+[1+(\alpha-2)\lambda]u_{ij}+\lambda u_{i+1,j}, \tag{3.7.1}$$

and $$(1+\alpha\lambda)u_{i,j+1}-\alpha\lambda u_{i+1,j+1} = \lambda u_{i-1,j}+[1+(\alpha-2)\lambda]u_{i,j}+(1-\alpha)\lambda u_{i+1,j}. \tag{3.7.2}$$

where $0\leq\alpha\leq 1$. The computational molecules of these equations are given by Figure 3.7.1 below.



Equation (3.7.1) Equation (3.7.2)

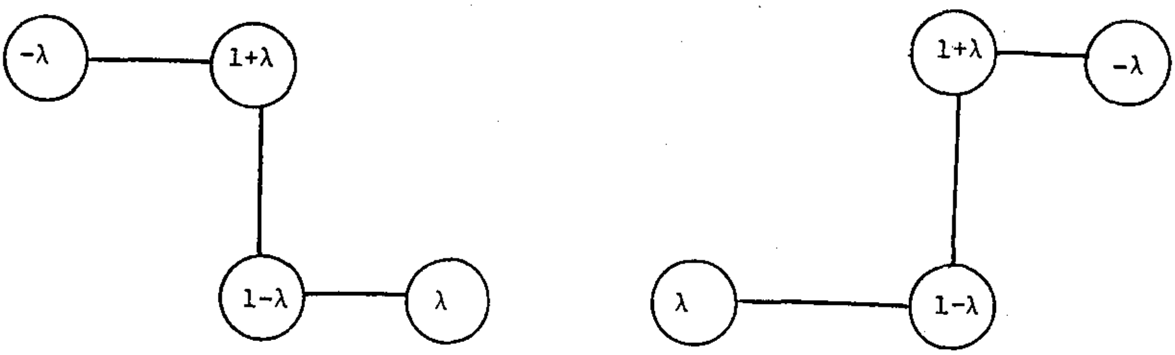
FIGURE 3.7.1

For $\alpha=0$, both formulae reduce to the classical explicit form of (3.2.1a). When $\alpha=1$, we obtain,

$$(1+\lambda)u_{i,j+1}-\lambda u_{i-1,j+1} = (1-\lambda)u_{i,j}+\lambda u_{i+1,j}, \tag{3.7.3}$$

and $$(1+\lambda)u_{i,j+1}-\lambda u_{i+1,j+1} = \lambda u_{i-1,j}+(1-\lambda)u_{i,j} \tag{3.7.4}$$

whose computational molecules are given in Figure 3.7.2.



Equation (3.7.3)

Equation (3.7.4)

FIGURE 3.7.2

For an initial-boundary value problem, equations (3.7.3) and (3.7.4) are both implicit-explicit if the computations proceed in the positive (LR) and negative (RL) directions respectively. Hence, equations (3.7.1) and (3.7.2) may be considered as generalisations of the classical explicit equation.

The stability of the asymmetric formulae (3.7.1) and (3.7.2) for the case $\alpha=1$ is now investigated by using the von Neumann criterion. The equation governing the stability of equation (3.7.1) is found to be

$$-\alpha\lambda\xi^{j+1}\exp(i_c\beta(i-1)\Delta x)+(1+\alpha\lambda)\xi^{j+1}\exp(i_c\beta i\Delta x) = (1-\alpha)\lambda\xi^{j}\exp(i_c\beta(i-1)\Delta x)$$
$$+(1+(\alpha-2)\lambda)\xi^{j}\exp(i_c\beta i\Delta x)+\lambda\xi^{j}\exp(i_c\beta(i+1)\Delta x)$$

which reduces to,

$$\gamma = \frac{(1-\alpha)\lambda\exp(-i_c\beta\Delta x)+\lambda\exp(i_c\beta\Delta x)+[1+(\alpha-2)\lambda]}{(1+\alpha\lambda)-\alpha\lambda\exp(-i_c\beta\Delta x)}$$

where $\gamma=\xi^{j+1}/\xi^{j}$. For $\alpha=1$, we get,

$$\gamma = \frac{1-\lambda[1-\exp(i_c\beta\Delta x)}{1+\lambda[1-\exp(-i_c\beta\Delta x)]} ,$$

i.e.,
$$\gamma = \frac{(1+2\lambda^2\cos(\beta\Delta x)-2\lambda^2\cos^2(\beta\Delta x))+i_c\lambda\sin(\beta\Delta x)[2\lambda-2\lambda\cos(\beta\Delta x)]}{1+2\lambda(1-\cos(\beta\Delta x))+2\lambda^2(1-\cos(\beta\Delta x))}$$

$$\therefore|\gamma| = \frac{\{1+8\lambda^2\sin^2(\frac{\beta\Delta x}{2})-16\lambda^2\sin^4(\frac{\beta\Delta x}{2})+16\lambda^4\sin^4(\frac{\beta\Delta x}{2})\}^{\frac{1}{2}}}{\{1+4\lambda\sin^2(\frac{\beta\Delta x}{2})+4\lambda^2\sin^2(\frac{\beta\Delta x}{2})\}}$$

$$= \frac{\{[1+4\lambda^2\sin^2(\frac{\beta\Delta x}{2})]^2-[4\lambda\sin^2(\frac{\beta\Delta x}{2})]^2\}^{\frac{1}{2}}}{\{1+4\lambda\sin^2(\frac{\beta\Delta x}{2})+4\lambda^2\sin^2(\frac{\beta\Delta x}{2})\}}$$

$$= \frac{[1+4\lambda^2\sin^2(\frac{\beta\Delta x}{2})+4\lambda\sin^2(\frac{\beta\Delta x}{2})]^{\frac{1}{2}}[1+4\lambda^2\sin^2(\frac{\beta\Delta x}{2})-4\lambda\sin^2(\frac{\beta\Delta x}{2})]^{\frac{1}{2}}}{[1+4\lambda\sin^2(\frac{\beta\Delta x}{2})+4\lambda^2\sin^2(\frac{\beta\Delta x}{2})]}$$

$$= \frac{[1+4\lambda^2\sin^2(\frac{\beta\Delta x}{2})-4\lambda\sin^2(\frac{\beta\Delta x}{2})]^{\frac{1}{2}}}{[1+4\lambda^2\sin^2(\frac{\beta\Delta x}{2})+4\lambda\sin^2(\frac{\beta\Delta x}{2})]^{\frac{1}{2}}}$$

Therefore, $|\gamma| \leqslant 1$ for all $\lambda>0$.

In the same manner, it can also be shown that $|\gamma|\leqslant 1$ for equation (3.7.2) with $\alpha=1$. We conclude that both asymmetric equations (3.7.3) and (3.7.4) are always stable. More generally, it was mentioned by A.F. Filipov (Saulev (1964)) that in the case of the Cauchy's problem, a necessary and sufficient condition for the stability of the equations (3.7.1) and (3.7.2) is,

$$\lambda \leqslant \frac{1}{2(1-\alpha)} \quad , 0\leqslant\alpha\leqslant 1. \tag{3.7.5}$$

Due to their low accuracy, the asymmetric formulae (3.7.1) and (3.7.2) are not highly recommended for use in the numerical integration of the heat equation. In the method of Saulev, however, different equations are used on *alternate time steps*, that is,

$$u_{i,j+1} = \frac{1}{(1+\alpha\lambda)}\{\alpha\lambda u_{i-1,j+1}+(1-\alpha)\lambda u_{i-1,j}+[1+(\alpha-2)\lambda]u_{ij}+\lambda u_{i+1,j}\} \tag{3.7.6}$$

and
$$u_{i,j+2} = \frac{1}{(1+\alpha\lambda)}\{\alpha\lambda u_{i+1,j+2}+\lambda u_{i-1,j+1}+[1+(\alpha-2)\lambda]u_{i,j+1}+(1-\alpha)\lambda u_{i+1,j+1}\}, \qquad (3.7.7)$$

The first system of equations (3.7.6) can be solved in the order $i=1,2,\ldots,m-1$ and the second system of equations (3.7.7) takes the order $i=m-1,m-2,\ldots,2,1$ in turn for $j=0,2,4,\ldots$. In particular, for $\alpha=1$, equations (3.7.6) and (3.7.7) reduce to the equations given in entry 14. of Table 3.7.1, that is,

$$u_{i,j+1}-u_{ij} = \lambda(u_{i+1,j}-u_{ij}-u_{i,j+1}+u_{i-1,j+1}), \qquad (3.7.8)$$

and
$$u_{i,j+2}-u_{i,j+1} = \lambda(u_{i+1,j+2}-u_{i,j+2}-u_{i,j+1}+u_{i-1,j+1}). \qquad (3.7.9)$$

A stability analysis and an estimate of the accuracy can be obtained by eliminating the $u_{i,j+1}$ from the two equations. From the equations (3.7.8) and (3.7.9), we have, respectively,

$$(1+\lambda)u_{i,j+1}-\lambda u_{i-1,j+1} = (1-\lambda)u_{i,j}+\lambda u_{i+1,j}, \qquad (3.7.8a)$$

and
$$(1-\lambda)u_{i,j+1}+\lambda u_{i-1,j+1} = (1+\lambda)u_{i,j+2}-\lambda u_{i+1,j+2}. \qquad (3.7.9a)$$

If we add the two equations, we get

$$u_{i,j+1} = \tfrac{1}{2}\{(1-\lambda)u_{ij}+(1+\lambda)u_{i,j+2}+\lambda(u_{i+1,j}-u_{i+1,j+2})\}$$

and similarly,

$$u_{i-1,j+1} = \tfrac{1}{2}\{(1-\lambda)u_{i-1,j}+(1+\lambda)u_{i-1,j+2}+\lambda(u_{i,j}-u_{i,j+2})\}.$$

By substituting these equations into equation (3.7.9) and after some manipulation of terms, leads to the *implicit scheme*,

$$u_{i,j+2}-(\lambda+\lambda^2)\delta_x^2u_{i,j+2} = u_{i,j}+(\lambda-\lambda^2)\delta_x^2u_{i,j}. \qquad (3.7.10)$$

If the stability analysis is carried out by means of the von Neumann criterion, we arrive at the overall stability polynomial,

$$\gamma^2\{1-(\lambda+\lambda^2)[\exp(i_c\beta\Delta x)+\exp(-i_c\beta\Delta x)-2]\} = 1+(\lambda-\lambda^2)[\exp(i_c\beta\Delta x)+\exp(-i_c\beta\Delta x)-2].$$

Therefore,

$$\gamma^2 = \frac{\{1+(\lambda-\lambda^2)[\exp(i_c\beta\Delta x)+\exp(-i_c\beta\Delta x)-2]\}}{\{1-(\lambda+\lambda^2)[\exp(i_c\beta\Delta x)+\exp(-i_c\beta\Delta x)-2]\}}$$

which, upon simplification, becomes,

$$\gamma^2 = \frac{1+2(\lambda-\lambda^2)[\cos(\beta\Delta x)-1]}{1-2(\lambda+\lambda^2)[\cos(\beta\Delta x)-1]} \ .$$

Hence,

$$|\gamma| = \frac{\{[1-\lambda(1-\cos(\beta\Delta x))]^2+\lambda^2\sin^2(\beta\Delta x)\}^{\frac{1}{2}}}{\{[1+\lambda(1-\cos(\beta\Delta x))]^2+\lambda^2\sin^2(\beta\Delta x)\}^{\frac{1}{2}}}$$

It is clear that since $1-\cos(\beta\Delta x)\geqslant 0$, $|\gamma|\leqslant 1$ which implies that the *Saulev's alternating method* has unrestricted stability.

Let us now consider equation (3.7.10), i.e.,

$$u_{i,j+2}-(\lambda+\lambda^2)\delta_x^2 u_{i,j+2} = u_{i,j}+(\lambda-\lambda^2)\delta_x^2 u_{i,j} \ .$$

If the terms λ^2 were absent then this equation would be the Crank-Nicolson scheme whose truncation error is $O([\Delta t]^2)+O([\Delta x]^2)$. With the presence of the λ^2 terms, however, their contribution to $(u_{i,j+2}-u_{i,j})/2\Delta t$ is,

$$O([\lambda\Delta x]^2\ \frac{\partial^3 u}{\partial x^2\partial t}) = O(\frac{[\Delta t]^2}{[\Delta x]^2}\ \frac{\partial^3 u}{\partial x^2\partial t}) \ .$$

For any fixed λ, the truncation error is $O(\Delta t)=O([\Delta x]^2)$. If Δt and Δx are regarded as being independent, then, just as we have seen for the Dufort-Frankel scheme, the Saulev's scheme will be consistent with the heat flow equation only if $\frac{\Delta t}{\Delta x}\rightarrow 0$ as the net is refined.

3.8 PARABOLIC EQUATIONS WITH VARIABLE COEFFICIENTS

The general, linear parabolic equation (3.1.1) may be written in the form,

$$\frac{\partial U}{\partial t} = L(x,t,D_x,D_x^2)U \tag{3.8.1}$$

or

$$D_t U = L(x,t,D_x,D_x^2)U \tag{3.8.1a}$$

where the operator L is linear, $D_x = \frac{\partial}{\partial x}$ and $D_t = \frac{\partial}{\partial t}$. We shall continue to use the Taylor series expansion,

$$\begin{aligned} U(x,t+\Delta t) &= (1+(\Delta t)\frac{\partial}{\partial t} + \tfrac{1}{2}(\Delta t)^2\frac{\partial^2}{\partial t^2} + \ldots)U(x,t) \\ &= \exp(\Delta t\,\frac{\partial}{\partial t})U(x,t) \;, \end{aligned} \tag{3.8.2}$$

to derive two-level schemes approximating (3.8.1). At the point (x_i,t_j), equation (3.8.2) becomes,

$$U_{i,j+1} = \exp((\Delta t)L)U_{i,j} \;, \tag{3.8.3}$$

where L is assumed to be independent of t. We recall that,

$$D_x = \frac{2}{\Delta x}\sinh^{-1}(\frac{\delta_x}{2}) = \frac{1}{\Delta x}(\delta_x - \frac{1^2}{2^2.3!}\delta_x^3 + \frac{1^2.3^2}{2^4.5!}\delta_x^5 + \ldots). \tag{3.8.4}$$

If equation (3.8.4) is used to eliminate D_x in terms of δ_x in equation (3.8.3), the exact difference replacement,

$$u_{i,j+1} = \exp\{(\Delta t)L[i\Delta x,j\Delta t,\frac{2}{\Delta x}\sinh^{-1}(\frac{\delta_x}{2}),(\frac{2}{\Delta x}\sinh^{-1}(\frac{\delta_x}{2}))]\}u_{i,j}, \tag{3.8.5}$$

is obtained.

Differential equations with various cases of *variable coefficients* will be treated and both explicit and implicit finite difference methods will be derived.

3.9 EXPLICIT METHODS

(a) *Coefficients Depending On x.*

The simplest differential equation in this case is

$$\frac{\partial U}{\partial t} = a(x)\,\frac{\partial^2 U}{\partial x^2} \qquad (3.9.1)$$

where $a(x)\neq 0$ for all x. Here $L\equiv aD_x^2$ and equation (3.8.3) becomes,

$$\begin{aligned} U_{i,j+1} &= \exp((\Delta t)aD_x^2)U_{i,j} \\ &= [1+(\Delta t)aD_x^2+\tfrac{1}{2}(\Delta t)^2 aD_x^2(aD_x^2)+\ldots]U_{i,j} \\ &= [1+(\Delta t)aD_x^2+\tfrac{1}{2}(\Delta t)^2 a(a''D_x^2+2a'D_x^3+aD_x^4)+\ldots]U_{i,j} \end{aligned}$$

where the prime(') notation means differentiation with respect to x. The differential operators $D_x^2, D_x^3, \ldots$ are now replaced by difference operators using equation (3.8.4). An explicit difference formula that is commonly used is,

$$u_{i,j+1} = (1-2a_i\lambda)u_{i,j}+a_i\lambda(u_{i+1,j}+u_{i-1,j}) \;. \qquad (3.9.2)$$

To find its local truncation error, we have, on expanding the U terms by means of the Taylor series about the point (x_i,t_j) yields,

$$\begin{aligned} T &= \{(\frac{\partial U}{\partial t})_{i,j}-a_i(\frac{\partial^2 U}{\partial x^2})_{i,j}\}+\tfrac{1}{2}(\Delta t)(\frac{\partial^2 U}{\partial t^2})_{i,j} - \frac{a_i}{12}[\Delta x]^2(\frac{\partial^4 U}{\partial x^4})_{i,j}+\ldots \\ &= \tfrac{1}{2}(\Delta t)(\frac{\partial^2 U}{\partial t^2})_{i,j} - \frac{a_i}{12}[\Delta x]^2(\frac{\partial^4 U}{\partial x^4})_{i,j} + \ldots \end{aligned}$$

Hence, the local truncation error of the formula (3.9.2) is given by

$$T = O(\Delta t)+O([\Delta x]^2) \;. \qquad (3.9.3)$$

(b) *The Self-Adjoint Form.*

We consider the case,

$$\frac{\partial U}{\partial t} = \frac{\partial}{\partial x}(a(x)\frac{\partial U}{\partial x}) \;, \qquad (3.9.4)$$

where $a(x)\neq 0$ for all x. Differentiation of the right-hand side term gives,

$$L \equiv aD_x^2+a'D_x$$

and so,

$$U_{i,j+1} = \exp(\Delta t(aD_x^2+a'D_x))U_{i,j}$$
$$= [1+\Delta t(aD_x^2+a'D_x)+\ldots]U_{i,j} \ .$$

If all the terms of order $(\Delta t)^2$ and above are neglected and $D_x^2 u_{i,j}$ and $D_x u_{i,j}$ are replaced by $(\frac{1}{(\Delta x)^2})\delta^2 u_{i,j}$ and $(\frac{1}{2\Delta x})(u_{i+1,j}-u_{i-1,j})$ respectively, the difference analogue,

$$u_{i,j+1} = (1-2a_i\lambda)u_{i,j}+(a_i+\tfrac{1}{4}(\Delta x)a_i')\lambda u_{i+1,j}+(a_i-\tfrac{1}{4}(\Delta x)a_i')\lambda u_{i-1,j} \tag{3.9.5}$$

is obtained. An expression for the local truncation error of this explicit scheme is,

$$T = [\frac{\partial U}{\partial t} - \frac{\partial}{\partial x}(a_i \frac{\partial U}{\partial x})]_{i,j}+\tfrac{1}{2}(\Delta t)\{(\frac{\partial^2 U}{\partial t^2})_{i,j}- \frac{1}{6\lambda}[\tfrac{1}{2}a_i((\frac{\partial^4 U}{\partial x^4})_{i,j}+(\frac{\partial^4 U}{\partial x^4})_{i,j})$$
$$+2a_i'(\frac{\partial^3 U}{\partial x^3})_{i,j}] - \frac{(\Delta x)}{24\lambda} a_i'[(\frac{\partial^4 U}{\partial x^4})_{i,j}-(\frac{\partial^4 U}{\partial x^4})_{i,j}]\}+ \ldots$$

Hence, we find that,

$$T = O(\Delta t) + O([\Delta x]^2) \ . \tag{3.9.6}$$

Alternatively, difference approximations to equation (3.9.4) can be obtained without differentiating the right-hand side term and destroying the self-adjoint nature of the operator. The following method is due to Tikhonnov and Smarskii (1961). Equation (3.9.4) can be written in conservation form as,

$$\frac{\partial U}{\partial t} + \frac{\partial W}{\partial x} = 0 \ , \tag{3.9.7}$$

where

$$W = -a(x) \frac{\partial U}{\partial x}$$

or

$$\frac{W}{a(x)} = - \frac{\partial U}{\partial x} \ .$$

By integrating this equation with respect to x over the interval $[(i-1)\Delta x, i\Delta x]$ of the x-axis and assuming that $W=W_{i-\frac{1}{2}}$, we obtain,

$$W_{i-\frac{1}{2}} \int_{(i-1)\Delta x}^{i\Delta x} \frac{dx}{a(x)} = U_{i-1}-U_i \ .$$

Similarly, if the interval $[i\Delta x,(i+1)\Delta x]$ is considered, we get

$$W_{i+\frac{1}{2}}\int_{i\Delta x}^{(i+1)\Delta x}\frac{dx}{a(x)} = U_i-U_{i+1} .$$

Now from equation (3.9.7) we have,

$$\begin{aligned}\frac{\partial U}{\partial t} &= -\frac{\partial W}{\partial x}\\ &= -\frac{1}{\Delta x}\delta_x W+O([\Delta x]^2) \qquad \text{(from equation (3.8.4))}\\ &= \frac{1}{\Delta x}(W_{i-\frac{1}{2}}-W_{i+\frac{1}{2}})+O([\Delta x]^2)\\ &= B_i(U_{i+1}-U_i)-A_i(U_i-U_{i-1})+O([\Delta x]^2) \qquad (3.9.8)\end{aligned}$$

where

$$B_i = \frac{1}{\Delta x}[\int_{i\Delta x}^{(i+1)\Delta x}\frac{dx}{a(x)}]^{-1}, \quad A_i = \frac{1}{\Delta x}[\int_{(i-1)\Delta x}^{i\Delta x}\frac{dx}{a(x)}]^{-1} .$$

Since $B_i=A_{i+1}$, equation (3.9.8) can be written as,

$$(\frac{\partial U}{\partial t})_{i,j} = A_{i+1}U_{i+1}-(A_{i+1}+A_i)U_i+A_iU_{i-1}+O([\Delta x]^2), \qquad (3.9.9)$$

and using the formula,

$$(\frac{\partial U}{\partial t})_{i,j} = \frac{1}{\Delta t}(U_{i,j+1}-U_{i,j})+O(\Delta t) ,$$

we obtain,

$$u_{i,j+1} = [1-(\Delta t)(A_{i+1}+A_i)]u_{i,j}+\Delta t[A_{i+1}u_{i+1,j}+A_iu_{i-1,j}]. \quad (3.9.10)$$

A more standard method of approximating equation (3.9.4) is to consider the central difference approximation to the self-adjoint operator, i.e.,

$$\begin{aligned}\frac{1}{\Delta x}\delta_x(a_i\delta_x)U_i &= \frac{1}{(\Delta x)^2}\delta_x[a_i(U_{i+\frac{1}{2}}-U_{i-\frac{1}{2}})]\\ &= \frac{1}{(\Delta x)^2}[a_{i+\frac{1}{2}}U_{i+1}-a_{i-\frac{1}{2}}U_i-a_{i+\frac{1}{2}}U_i+a_{i-\frac{1}{2}}U_{i-1}]\end{aligned}$$

$$= \frac{1}{(\Delta x)^2}[a_{i+\frac{1}{2}}(U_{i+1}-U_i)-a_{i-\frac{1}{2}}(U_i-U_{i-1})].$$

This leads to the explicit formula,

$$u_{i,j+1} = [1-\lambda(a_{i+\frac{1}{2}}+a_{i-\frac{1}{2}})]u_{i,j}+\lambda[a_{i+\frac{1}{2}}u_{i+1,j}+a_{i-\frac{1}{2}}u_{i-1,j}]. \quad (3.9.11)$$

We note that if a(x) is replaced by a(x,t), formulae (3.9.2),(3.9.5), (3.9.10) and (3.9.11) still hold with the same degree of accuracy.

3.10 IMPLICIT METHODS

Implicit formulae can often be obtained from equation (3.8.3) written in the central form:

$$\exp(-\tfrac{1}{2}(\Delta t)L)U_{i,j+1} = \exp(\tfrac{1}{2}(\Delta t)L)U_{i,j} \quad . \qquad (3.10.1)$$

(a) *Coefficients Depending on x and t.*

We consider equations of the form

$$\frac{\partial U}{\partial t} = a(x,t)\,\frac{\partial^2 U}{\partial x^2} \quad . \qquad (3.10.2)$$

When the Crank-Nicolson averaging concept is applied to this equation, we get,

$$\frac{(u_{i,j+1}-u_{ij})}{\Delta t} = \tfrac{1}{2}a(x_i,t_{j+\frac{1}{2}})\,\delta_x^2(u_{i,j+1}+u_{i,j})/(\Delta x)^2$$

$$= \tfrac{1}{2}a_{i,j+\frac{1}{2}}\delta_x^2(u_{i,j+1}+u_{ij})/(\Delta x)^2 , \qquad (3.10.3)$$

with a local truncation error equal to $T=O([\Delta x]^2)+O([\Delta t]^2)$.

A higher-order accuracy method equivalent to the Douglas formula can be derived in the same way as before. By means of the Taylor series expansions, we have,

$$\frac{1}{2(\Delta x)^2}\,\delta_x^2(U_{i,j+1}+U_{ij}) = \left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j+\frac{1}{2}} + \frac{1}{12}(\Delta x)^2\left(\frac{\partial^4 U}{\partial x^4}\right)_{i,j+\frac{1}{2}} + O([\Delta x]^4) + O([\Delta t]^2)$$

$$= \left(\frac{1}{a}\,\frac{\partial U}{\partial t}\right)_{i,j+\frac{1}{2}} + \frac{1}{12}(\Delta x)^2\,\frac{\partial^2}{\partial x^2}\left(\frac{1}{a}\,\frac{\partial U}{\partial t}\right)_{i,j+\frac{1}{2}} + O([\Delta x]^4)+O([\Delta t]^2) \qquad (3.10.4)$$

where we have used equation (3.10.2).

It can also be shown that,

$$\frac{\partial^2}{\partial x^2}\left(\frac{1}{a}\,\frac{\partial U}{\partial t}\right)_{i,j+\frac{1}{2}} = \frac{1}{(\Delta x)^2\Delta t}\,\delta_x^2\left[\frac{1}{a_{i,j+\frac{1}{2}}}(U_{i,j+1}-U_{ij})\right]+O([\Delta x]^2)+O([\Delta t]^2) .$$

The substitution of this into equation (3.10.4) gives,

$$\frac{1}{2(\Delta x)^2}\,\delta_x^2(U_{i,j+1}+U_{i,j}) = \frac{1}{a_{i,j+\frac{1}{2}}(\Delta t)}(U_{i,j+1}-U_{i,j})+\frac{1}{12\lambda(\Delta x)^2}\,\delta_x^2$$

$$[\frac{1}{a_{i,j+\frac{1}{2}}}(U_{i,j+1}-U_{i,j})]+O([\Delta x]^4)+O([\Delta t]^2)$$

which leads to the difference equation,

$$\frac{1}{\lambda a_{i,j+\frac{1}{2}}}(u_{i,j+1}-u_{i,j}) = \tfrac{1}{2}\delta_x^2[1-\frac{1}{6\lambda a_{i,j+\frac{1}{2}}}]u_{i,j+1}+\tfrac{1}{2}\delta_x^2[1+\frac{1}{6\lambda a_{i,j+\frac{1}{2}}}]u_{i,j}, \qquad (3.10.5)$$

with a local truncation error $T=O([\Delta x]^4)=O([\Delta t]^2)$ when λ is kept fixed. Formula (3.10.5) can be rewritten in the more convenient form,

$$[1+\frac{1}{12}\,a_{i,j+\frac{1}{2}}\delta_x^2(a_{i,j+\frac{1}{2}})^{-1}-\tfrac{1}{2}\lambda a_{i,j+\frac{1}{2}}\delta_x^2]u_{i,j+1}$$

$$= [1+\frac{1}{12}\,a_{i,j+\frac{1}{2}}\delta_x^2(a_{i,j+\frac{1}{2}})^{-1}+\tfrac{1}{2}\lambda a_{i,j+\frac{1}{2}}\delta_x^2]u_{i,j}\ . \qquad (3.10.6)$$

A *three-level scheme* similar to the one introduced in (3.6.13) can also be developed to approximate $\frac{\partial^2 U}{\partial x^2} = a(x,t)\frac{\partial U}{\partial t}$. The term $\frac{\partial^4 U}{\partial x^4}$ can be replaced as follows:

$\frac{\partial^4 U}{\partial x^4} = \frac{\partial^2}{\partial x^2}(a(x,t)\frac{\partial U}{\partial t})$. Then (3.6.13) would become,

$$a_{i,j}\,\frac{(u_{i,j+1}-u_{i,j-1})}{2\Delta t} = \frac{1}{3}\,\frac{\delta_x^2}{(\Delta x)^2}(u_{i,j+1}+u_{ij}+u_{i,j-1})-\frac{1}{24\lambda}\,\frac{\delta_x^2}{(\Delta x)^2}$$

$$[a_{ij}(u_{i,j+1}-u_{i,j-1})]\ , \qquad (3.10.7)$$

which is again fourth-order correct in space and second-order in time for a fixed λ.

(b) *The More General Form*

We consider the equation,

$$\frac{\partial U}{\partial t} = \frac{\partial}{\partial x}(a(x,t)\frac{\partial U}{\partial x}) + b(x,t)\frac{\partial U}{\partial x} + c(x,t)U\ . \qquad (3.10.8)$$

At the point $x_i=i\Delta x$ and $t_{j+\frac{1}{2}}=(j+\frac{1}{2})\Delta t$, we have,

$$\frac{1}{\Delta t}(U_{i,j+1}-U_{i,j}) = \tfrac{1}{2}[\frac{\partial}{\partial x}(a\frac{\partial U}{\partial x})+b\frac{\partial U}{\partial x}+cU]_{i,j+1}$$

$$+\tfrac{1}{2}[\frac{\partial}{\partial x}(a\frac{\partial U}{\partial x})+b\frac{\partial U}{\partial x}+cU]_{i,j} + O([\Delta t]^2) \ . \qquad (3.10.9)$$

Now, by using equation (3.9.9) we have, *at any fixed time*,

$$\frac{\partial}{\partial x}(a\frac{\partial U}{\partial x}) = A_{i+1}U_{i+1}-(A_{i+1}+A_i)U_i+A_iU_{i-1}+O([\Delta x]^2)$$

where,

$$A_i = \frac{1}{\Delta x}[\int_{(i-1)\Delta x}^{i\Delta x} \frac{dx}{a(x)}]^{-1} ,$$

and
$$b\frac{\partial U}{\partial x} = b_i \frac{1}{2\Delta x}(U_{i+1}-U_{i-1})+O([\Delta x]^2) \ .$$

The substitution of these values into equation (3.10.9) leads to the implicit difference scheme,

$$[1+\tfrac{1}{2}(\Delta t)(A_{i+1,j+1}+A_{i,j+1})-\tfrac{1}{2}(\Delta t)c_{i,j+1}]u_{i,j+1}$$

$$-\tfrac{1}{2}(\Delta t)[(A_{i+1,j+1}+\frac{1}{2(\Delta x)}b_{i,j+1})u_{i+1,j+1}+(A_{i,j+1}-\frac{1}{2(\Delta x)}b_{i,j+1})u_{i-1,j+1}]$$

$$= \tfrac{1}{2}(\Delta t)[(A_{i+1,j}+\frac{1}{2(\Delta x)}b_{ij})u_{i+1,j}+(A_{i,j}-\frac{1}{2(\Delta x)}b_{ij})u_{i-1,j}]$$

$$+[1-\tfrac{1}{2}(\Delta t)(A_{i+1,j}+A_{i,j})+\tfrac{1}{2}(\Delta t)c_{i,j}]u_{i,j} \ . \qquad (3.10.10)$$

If b is large in modulus, we put,

$$b\frac{\partial U}{\partial x} = b_i \frac{1}{(\Delta x)}(U_i-U_{i-1}) + O(\Delta x) \ , \quad \text{if } b<0 \ ,$$

$$= b_i \frac{1}{(\Delta x)}(U_{i+1}-U_i)+O(\Delta x) \ , \quad \text{if } b>0 \ ,$$

and equation (3.10.10) is changed accordingly.

Another class of implicit formulae approximating particular cases of equation (3.1.1) are the backward difference schemes. As we have already seen, equation (3.6.27) is an example in this class. Most of these formulae can be derived from (3.8.3) when written in the form,

$$\exp(-(\Delta t)L)U_{i,j+1} = U_{ij} \ . \qquad (3.10.11)$$

For example, when

$$L \equiv D_x^2 = \frac{1}{(\Delta x)^2}\delta_x^2 + O([\Delta x]^2) \ ,$$

then expansion of (3.10.11) leads to the formula,

$$(1-\lambda\delta_x^2)u_{i,j+1} = u_{i,j} \ .$$

which is just (3.6.27). Formulae (3.8.3) and (3.10.11) are the special cases θ=0 and 1 respectively of the general θ formula,

$$\exp(-\theta(\Delta t)L)U_{i,j+1} = \exp((1-\theta)(\Delta t)L)U_{i,j} \ , \ 0\leq\theta\leq 1. \qquad (3.10.12)$$

The central form with $\theta=\frac{1}{2}$ (Crank Nicolson) is very widely used.

3.11 THE DIFFUSION-CONVECTION EQUATION

The general form of the diffusion-convection equation in one space dimension is given by,

$$\frac{\partial U}{\partial t} = K\frac{\partial^2 U}{\partial x^2} - V\frac{\partial U}{\partial x}\ , \tag{3.11.1}$$

where U describes the concentration of a suspension being convected with a velocity V and is diffused according to the diffusion coefficient K. If we put a=K and a' =-V in equation (3.9.5), we get the explicit approximation,

$$u_{i,j+1} = (1-2\lambda K)u_{i,j}+\lambda(K-W)u_{i+1,j}+\lambda(K+W)u_{i-1,j}\ , \tag{3.11.2}$$

where $W = \frac{1}{2}(\Delta x)V$. By employing the von Neumann criterion, the equation governing stability is,

$$\gamma = [1+2\lambda K(\cos(\beta\Delta x)-1)]-i_{c}2\lambda W\sin(\beta\Delta x)\ .$$

Hence,

$$|\gamma|^2 = [1+2\lambda K(\cos(\beta\Delta x)-1)]^2+4\lambda^2W^2\sin^2(\beta\Delta x)$$

$$= [1+\frac{2K\Delta t}{(\Delta x)^2}(\cos(\beta\Delta x)-1)]^2+(\frac{V\Delta t}{\Delta x})^2(1-\cos^2(\beta\Delta x)) \tag{3.11.3}$$

and for stability, we must have the conditions,

$$\frac{V\Delta t}{\Delta x} \leqslant 1\ , \tag{3.11.4a}$$

$$\frac{K\Delta t}{(\Delta x)^2} \leqslant \frac{1}{2}\ . \tag{3.11.4b}$$

Inequality (3.11.4a) is the requirement associated with *convective stability* whilst (3.11.4b) contributes to the condition for *diffusive stability*.

The solution to equation (3.11.1), which is given by (3.11.2) *oscillates* for W>K (Siemieniuch and Gladwell (1978)) and so for large values of V, it is customary to replace central differences for the first space derivative by backward differences. In this way, the oscillations are minimised although the local accuracy of the difference scheme is reduced from second to first order. The difference formula where D_x, the differential operator in (3.8.1), is replaced by $\frac{1}{\Delta x}\nabla_x$ with,

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

is given by,

$$u_{i,j+1} = [1-2\lambda(K+W)]u_{ij} + \lambda K u_{i+1,j} + \lambda(K+2W)u_{i-1,j} \text{ .} \qquad (3.11.5)$$

As we have seen the explicit equations (3.11.2) and (3.11.5) result from the approximations,

$$\frac{1}{\Delta t}\Delta_t u_{i,j} = K \frac{\delta_x^2}{(\Delta x)^2} u_{i,j} - \frac{V}{2\Delta x}(\Delta_x + \nabla_x)u_{i,j} \text{ ,} \qquad (3.11.2a)$$

and,

$$\frac{1}{\Delta t}\Delta_t u_{i,j} = K \frac{\delta_x^2}{(\Delta x)^2} u_{i,j} - \frac{V}{\Delta x}\nabla_x u_{i,j} \text{ ,} \qquad (3.11.5a)$$

respectively. Formula (3.11.5a) is known as the *explicit method with upwinding* approximating the diffusion-convection equation (3.11.1). A more general method to (3.11.5a) can be developed by considering an approximation involving a θ-weighted type of upwinding for the first spatial derivative in (3.11.1) at two time levels, i.e.,

$$\left(\frac{\partial U}{\partial x}\right)_{i,j} \simeq \frac{\theta}{\Delta x}\nabla_x u_{i,j+1} + \frac{(1-\theta)}{\Delta x}\nabla_x u_{ij} \text{ ,}$$

where $0 \leq \theta \leq 1$. Therefore, equation (3.11.5a) is replaced by the more *general, exact explicit type of method with upwinding*,

$$(1+2\lambda\theta W)U_{i,j+1} - 2\lambda\theta W U_{i-1,j+1} = \lambda K U_{i+1,j} + \{1-2\lambda(K+(1-\theta)W)\}U_{i,j}$$
$$+\lambda\{K+2(1-\theta)W\}U_{i-1,j} + \Delta t T_{i,j} \text{ ,} \qquad (3.11.6)$$

and the finite-difference analogue to (3.11.6) is,

$$(1+2\lambda\theta W)u_{i,j+1}-2\lambda\theta Wu_{i-1,j+1} = \lambda Ku_{i+1,j}+\{1-2\lambda(K+(1-\theta)W\}u_{i,j}$$
$$+\lambda\{K+2(1-\theta)W\}u_{i-1,j}\ , \qquad (3.11.7)$$

with a local truncation error,

$$T_{ij} = -\tfrac{1}{2}(\Delta t)\{-(\frac{\partial^2 U}{\partial t^2})_{i,j}-2V\theta(\frac{\partial^2 U}{\partial x\partial t})_{i,j} + V\theta(\frac{\partial^3 U}{\partial x^2\partial t})_{i,j}\}$$
$$-\tfrac{1}{2}V(\Delta x)(\frac{\partial^2 U}{\partial x^2})_{i,j} + \ldots \qquad (3.11.8)$$

implying that $T=O(\Delta t)+O(\Delta x)$. It is evident from (3.11.8) that 'numerical diffusion' gradually builds up when V becomes significantly large. As for stability, the following conditions are obtained,

(i) $$0<\lambda\leqslant\frac{(1+W)}{2(1+2W(1-\theta)+W^2(1-2\theta))} \quad \text{for } 0\leqslant\theta\leqslant\tfrac{1}{2}\ , \qquad (3.11.9)$$

(ii) $$0<\lambda\leqslant\frac{(1+W)}{2(1+2W(1-\theta)-W^2(2\theta-1))} \quad \text{for } \tfrac{1}{2}\leqslant\theta\leqslant 1,\ 0\leqslant W\leqslant\frac{1}{2\theta-1}\ , \qquad (3.11.10)$$

and λ is unrestricted for $W\geqslant\frac{1}{2\theta-1}$. The effect of the numerical diffusion may be reduced by rewriting equation (3.11.6) as

$$(1+2\lambda\theta W)U_{i,j+1}-2\lambda\theta WU_{i-1,j+1} = \lambda KU_{i+1,j}+\{1-2\lambda(K+(1-\theta)W\}U_{i,j}$$
$$+ \lambda\{K+2(1-\theta)W\}U_{i-1,j}-\tfrac{1}{2}(\Delta t)(\Delta x)V(\frac{\partial^2 U}{\partial x^2})_{i,j} +\Delta tT'_{i,j}.$$
$$(3.11.11)$$

The second spatial derivative $\frac{\partial^2 U}{\partial x^2}$ is then discretised at the mesh point by means of the second central difference to give,

$$(\frac{\partial^2 U}{\partial x^2})_{i,j} = \frac{(U_{i+1,j}-2U_{i,j}+U_{i-1,j})}{(\Delta x)^2} \qquad (3.11.12)$$

By substituting this expression into (3.11.11) yields,

$$(1+2\lambda\theta W)U_{i,j+1}-2\lambda\theta WU_{i-1,j+1} = \lambda(K-W)U_{i+1,j}+\{1-2\lambda(K-\theta)W\}U_{i,j}$$
$$+\lambda\{K+(1-\theta)W\}U_{i-1,j}+\Delta tT'_{i,j}\ , \qquad (3.11.13)$$

and the finite-difference analogue to (3.11.13) is,

$$(1+2\lambda\theta W)u_{i,j+1}-2\lambda\theta Wu_{i-1,j+1}=\lambda(K-W)u_{i+1,j}+\{1-2\lambda(K-\theta)W\}u_{i,j}$$

$$+\lambda\{K+(1-\theta)W\}u_{i-1,j}\ , \qquad (3.11.14)$$

with a local truncation error,

$$T'_{ij}=-\tfrac{1}{2}(\Delta t)\{-(\frac{\partial^2 U}{\partial t^2})_{i,j}-2V\theta(\frac{\partial^2 U}{\partial x\partial t})_{i,j}+V\theta(\frac{\partial^3 U}{\partial x^2\partial t})_{i,j}\}+\ldots$$

$$(3.11.15)$$

Similarly, we can get rid of the effect of the mixed-derivatives such as $\frac{\partial^2 U}{\partial x\partial t}$ by discretising it as,

$$(\frac{\partial^2 U}{\partial x\partial t})_{i,j}=(U_{i,j+1}+U_{i-1,j}-U_{i-1,j+1}-U_{ij})/(\Delta x\Delta t)$$

and this gives us,

$$U_{i,j+1}=\{1-2\lambda(K+W)\}U_{ij}+\lambda\{KU_{i+1,j}+(K+2W)U_{i-1,j}\}+\Delta tT''_{ij}$$

whose finite-difference analogue is,

$$u_{i,j+1}=\{1-2\lambda(K+W)\}u_{ij}+\lambda\{Ku_{i+1,j}+(K+2W)u_{i-1,j}\}. \qquad (3.11.16)$$

Another explicit scheme can also be derived by first considering the hyperbolic part of the equation (3.11.1), i.e.,

$$\frac{\partial U}{\partial t}=-V\frac{\partial U}{\partial x}\ .$$

This equation is of the form (2.17.15) with A=V. Hence the Lax-Wendroff formula of (2.17.16) gives us,

$$u_{i,j+1}=u_{ij}+\tfrac{1}{2}V(\frac{\Delta t}{\Delta x})(\Delta_x+\nabla_x)u_{ij}+\tfrac{1}{2}V^2(\frac{\Delta t}{\Delta x})^2(\Delta_x-\nabla_x)u_{ij}$$

$$=u_{ij}+\tfrac{1}{2}V(\frac{\Delta t}{\Delta x})(u_{i+1,j}-u_{i-1,j})+\tfrac{1}{2}V^2(\frac{\Delta t}{\Delta x})^2(u_{i+1,j}-2u_{ij}+u_{i-1,j}).$$

$$(3.11.17)$$

After having established the difference analogue for the 'convective part' of (3.11.1), we are now left with approximating the 'diffusive part' $\frac{\partial^2 U}{\partial x^2}$ by the usual central difference formula,

$$\left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} \simeq \frac{1}{(\Delta x)^2}\,\delta_x^2 u_{i,j}$$

$$= \frac{1}{(\Delta x)^2}\,(u_{i+1,j}-2u_{ij}+u_{i-1,j}) \;. \qquad (3.11.18)$$

By combining equations (3.11.17) and (3.11.18), we obtain the final approximation to the diffusion-convection equation (3.11.1), i.e.,

$$u_{i,j+1} = \{1-2\lambda(K+2W^2\lambda)\}u_{i,j}+\lambda\{[K+W(1+2W\lambda)]u_{i-1,j}+[K-W(1-2W\lambda)]u_{i+1,j} \qquad (3.11.19)$$

with a local truncation error,

$$T_{ij} = -2W^2\lambda\left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} - \frac{\Delta t}{2}\left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j} - \frac{(\Delta t)^2}{6}\left(\frac{\partial^3 U}{\partial t^3}\right)_{i,j} - \frac{1}{3}W(\Delta x)\left(\frac{\partial^3 U}{\partial x^3}\right)_{i,j}$$

$$+ \frac{1}{12}(K+2W^2\lambda)(\Delta x)^2\left(\frac{\partial^4 U}{\partial x^4}\right)_{i,j} \;. \qquad (3.11.20)$$

The stability condition for (3.11.19) is found to be (Siemieniuch and Gladwell (1976)),

$$0 < \lambda \leq 1/(1+\sqrt{(1+4W^2)}) \;. \qquad (3.11.21)$$

A *fully implicit scheme with upwinding* to solve (3.11.1) takes the form,

$$U_{i,j+1}-U_{ij} = \lambda K\{U_{i+1,j+1}-2U_{i,j+1}+U_{i-1,j+1}\}- V\left(\frac{\Delta t}{\Delta x}\right)\{U_{i,j+1}-U_{i-1,j+1}\}$$

$$+\Delta t T_{ij} \;, \qquad (3.11.22)$$

and its analogue is,

$$u_{i,j+1}-u_{ij} = \lambda K\{u_{i+1,j+1}-2u_{i,j+1}+u_{i-1,j+1}\}- V\left(\frac{\Delta t}{\Delta x}\right)\{u_{i,j+1}-u_{i-1,j+1}\}, \qquad (3.11.23)$$

with a local truncation error given by,

$$T_{ij} = -\tfrac{1}{2}V(\Delta x)\left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} -\tfrac{1}{2}(\Delta t)\left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j} + \ldots \qquad (3.11.24)$$

As before, the effect of numerical diffusion is minimised by replacing the second space derivative in (3.11.24) by the central difference (3.11.12). By substituting this into equation (3.11.24), formula (3.11.22) is changed to

$$U_{i,j+1}-U_{i,j} = \lambda K\{U_{i+1,j+1}-2U_{i,j+1}+U_{i-1,j+1}\} - V\frac{\Delta t}{\Delta x}\{U_{i,j+1}-U_{i-1,j+1}$$

$$+\tfrac{1}{2}(U_{i+1,j}-2U_{i,j}+U_{i-1,j})\}+ \Delta t\ T'_{ij}\ . \qquad (3.11.25)$$

While the original equation (3.11.22) is first-order accurate in both space and time, equation (3.11.25) is accurate to second order spatially.

The *Crank-Nicolson scheme* can be derived in exactly the same manner as we did by employing the 'averaging' concept for the heat equation (3.1.2). We shall need the following equations:

$$\left(\frac{\partial U}{\partial t}\right)_{i,j+\frac{1}{2}} = \frac{(U_{i,j+1}-U_{i,j})}{\Delta t} + O([\Delta t]^2) \qquad (3.11.26)$$

$$\left(\frac{\partial U}{\partial x}\right)_{i,j+\frac{1}{2}} = \tfrac{1}{2}\left[\frac{(U_{i+1,j+1}-U_{i-1,j+1})}{2\Delta x} + \frac{(U_{i+1,j}-U_{i-1,j})}{2\Delta x}\right] +O([\Delta x]^2)$$

$$(3.11.27)$$

$$\left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j+\frac{1}{2}} = \tfrac{1}{2}\left[\frac{(U_{i+1,j+1}-2U_{i,j+1}+U_{i-1,j+1})}{(\Delta x)^2} + \frac{(U_{i+1,j}-2U_{i,j}+U_{i-1,j})}{(\Delta x)^2}\right]$$

$$+O([\Delta x]^2)\ , \qquad (3.11.28)$$

By substituting these derivatives into equation (3.11.1) yields the approximation,

$$-\alpha_1\lambda u_{i-1,j+1}+(1+\lambda K)u_{i,j+1}-\alpha_2\lambda u_{i+1,j+1} = \alpha_1\lambda u_{i-1,j}+(1-\lambda K)u_{i,j}+\alpha_2\lambda u_{i+1,j},$$

$$(3.11.29)$$

where $\alpha_1=\frac{1}{2}(K+W)$ and $\alpha_2=\frac{1}{2}(K-W)$. The truncation error is $T=O([\Delta t]^2)+O([\Delta x]^2)$. A more general equation than (3.11.29) is given by Peaceman and Rachford (1962) and takes the form,

$$\frac{(u_{i,j+1}-u_{ij})}{\Delta t} = \frac{K}{2(\Delta x)^2}[(u_{i-1,j}-2u_{ij}+u_{i+1,j})+(u_{i-1,j+1}-2u_{i,j+1}+u_{i+1,j+1})] - \frac{V}{2(\Delta x)}[u^*_{i+\frac{1}{2},j}+u^*_{i+\frac{1}{2},j+1}-u^*_{i-\frac{1}{2},j}-u^*_{i-\frac{1}{2},j+1}] \tag{3.11.30}$$

If we choose $u^*_{i+\frac{1}{2}}=\frac{1}{2}(u_{i+1}+u_i)$, we retrieve the Crank-Nicolson scheme (3.11.29). If, on the other hand, we take $u^*_{i+\frac{1}{2}}=u_i$, then (3.11.30) reduces to the 'upwinding' or 'backward-in-distance' formula. The upwinding scheme has a truncation error of first order in space. However, the method is capable of suppressing oscillations of the solution for small Δx.

Another general formula due to Stone and Brian (1963) is given by,

$$\frac{1}{\Delta t}[g(u_{i,j+1}-u_{ij})+\frac{\theta}{2}(u_{i-1,j+1}-u_{i-1,j})+m(u_{i+1,j+1}-u_{i+1,j})]$$
$$=K[\tfrac{1}{2}\Delta_x^2(u_{i,j}+u_{i,j+1})] - \frac{V}{\Delta x}[a(u_{i+1,j}-u_{ij})+\frac{\varepsilon}{2}(u_{i,j}-u_{i-1,j})$$
$$+ c(u_{i+1,j+1}-u_{i,j+1})+d(u_{i,j+1}-u_{i-1,j+1})] \ . \tag{3.11.31}$$

The weighting coefficients $a,\varepsilon,c,d,g,\theta$ and m have to satisfy the following restrictions:

$$\begin{aligned} a + \frac{\varepsilon}{2} + c + d &= 1 \ , \\ g + \frac{\theta}{2} + m &= 1 \ . \end{aligned} \tag{3.11.32}$$

In approximating (3.11.1), the method devised by Price, Varga and Warren (1966) maintains equations (3.11.26) and (3.11.28) but replaces (3.11.27) by,

$$(\frac{\partial U}{\partial x})_{i,j+\frac{1}{2}} \simeq \tfrac{1}{2}[\frac{(3u_{i,j+1}-4u_{i-1,j+1}+u_{i-2,j+1})}{2\Delta x} + \frac{(3u_{i,j}-4u_{i-1,j}+u_{i-2,j})}{2\Delta x}] \tag{3.11.33}$$

This method greatly suppresses unwarranted oscillations of the solution and permits the use of a coarser spatial grid. However, it requires the solution of a pentadiagonal system which can be time consuming.

3.12 THE DIFFUSION EQUATION IN CYLINDRICAL AND SPHERICAL POLAR CO-ORDINATES

The non-dimensional form of the equation for heat conduction in three dimensions is $\frac{\partial U}{\partial t} = \nabla^2 U$, which in *cylindrical polar co-ordinates (r,θ,z)* is expressed as,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial r^2} + \frac{1}{r}\frac{\partial U}{\partial r} + \frac{1}{r^2}\frac{\partial^2 U}{\partial \theta^2} + \frac{\partial^2 U}{\partial z^2} \quad . \tag{3.12.1}$$

If, for simplicity, we assume that U is independent of z, then (3.12.1) reduces to the two-dimensional equation,

$$\begin{aligned}\frac{\partial U}{\partial t} &= \frac{\partial^2 U}{\partial r^2} + \frac{1}{r}\frac{\partial U}{\partial r} + \frac{1}{r^2}\frac{\partial^2 U}{\partial \theta^2} \\ &= \frac{1}{r}\left\{\frac{\partial}{\partial r}\left(r\frac{\partial U}{\partial r}\right) + \frac{\partial}{\partial \theta}\left(\frac{1}{r}\frac{\partial U}{\partial \theta}\right)\right\}\end{aligned} \tag{3.12.2}$$

with $U=U(r,\theta,t)$. Finite-difference approximations in polar co-ordinates analogous to the familiar Cartesian system may be used to solve (3.12.2) on the cylinder $R\times[0\leq t\leq T]$, where t denotes time and the region R consists either of *the unit circle* $(0\leq r\leq 1)$ or *the ring* $(a\leq r\leq 1)$. The solution to (3.12.2) satisfies the initial condition,

$$U(r,\theta,0) = f(r,\theta) \quad , \tag{3.12.3a}$$

where $f(r,\theta)$ is given for all $(r,\theta) \in \bar{R}$, and the boundary condition

$$U(1,\theta,t) = h_1(\theta,t)\,,\ 0\leq t\leq T \quad , \tag{3.12.3b}$$

and it also fulfils the additional condition,

$$U(a,\theta,t) = h_a(\theta,t)\,,\ 0\leq t\leq T \quad , \tag{3.12.3c}$$

in the case of the ring $a\leq r\leq 1$. The functions $h_1(\theta,t)$ and $h_a(\theta,t)$ are prescribed for all $0\leq t\leq T$, $0\leq\theta\leq 2\pi$.

We shall first consider the solution of (3.12.2) when R is the ring. A *polar grid* in the rθ plane is defined by the *concentric* *circles* $r=i\Delta r$ and *the radial lines* $\theta=j\Delta\theta$. The peripheral ordering of the mesh points is illustrated in Figure 3.12.1.

We shall call the peripheral of unknown grid points adjacent to the boundary $r=a$, as *the first peripheral*, and number the remaining peripherals in the direction of increasing r. The mesh points are the points of intersection of the circles (peripherals) $r_i=i\Delta r$ for $i=i_{min}, i_{min}+1,\ldots,i_{min}+(m-1)$ $(=i_{max})$, where

$$(i_{min}-1)\Delta r = a \ ,$$

and

$$(i_{max}+1)\Delta r = 1 \ , \qquad (3.12.4)$$

(i.e. the radius of the first and last peripherals are $i_{min}\Delta r$ and $i_{max}\Delta r$, respectively) and the straight lines $\theta_j=j\Delta\theta$, $j=0,1,\ldots,n-1$ where,

$$\Delta\theta = \frac{2\pi}{n} \quad . \qquad (3.12.5)$$

Hence for Figure 3.12.1, we have $m=3$, $n=8$ and $\Delta r=(1-a)/(m+1)$. In the same manner as before, we can derive *an explicit finite difference analogue* of (3.12.2) by using central difference approximations for the spatial derivatives and a forward difference formula for $\frac{\partial U}{\partial t}$ at the mesh point $(r_i,\theta_j,t_k)=(i\Delta r,j\Delta\theta,k\Delta t)$ to give,

$$\frac{(u_{i,j,k+1}-u_{i,j,k})}{\Delta t} = \left\{\frac{(1-\frac{1}{2i})u_{i-1,j,k}-2u_{i,j,k}+(1+\frac{1}{2i})u_{i+1,j,k}}{(\Delta r)^2}\right\}$$

$$+ \left\{\frac{u_{i,j-1,k}-2u_{i,j,k}+u_{i,j+1,k}}{(r_i\Delta\theta)^2}\right\} \qquad (3.12.6)$$

If we put $\lambda = \frac{\Delta t}{(\Delta r)^2}$ and $\lambda' = \frac{\Delta t}{(r_i\Delta\theta)^2}$ then equation (3.12.6) becomes,

$$u_{i,j,k+1} = [1-2\lambda+2\lambda']u_{i,j,k}+\lambda(1-\frac{1}{2i})u_{i-1,j,k}+\lambda(1+\frac{1}{2i})u_{i+1,j,k}$$

$$+\lambda' u_{i,j-1,k}+\lambda' u_{i,j+1,k};\ i_{min}\leq i\leq i_{max},\ 0\leq j\leq n-1. \quad (3.12.7)$$

This can be written in the more compact form as,

$$\underset{\sim}{u}_{k+1} = (I-A)\underset{\sim}{u}_k + \underset{\sim}{b}_k \ , \qquad (3.12.8)$$

where $\underset{\sim}{b}_k$ is independent of the $\underset{\sim}{u}$'s, $\underset{\sim}{u}_k$ (k>0) is the vector consisting of the values $u_{i,j,k}$ and,

$$A = \begin{bmatrix} D_{i_{min}} & E_{i_{min}} & & & \\ C_{i_{min+1}} & D_{i_{min+1}} & E_{i_{min+1}} & & O \\ & \ddots & \ddots & \ddots & \\ & O & \ddots & \ddots & E_{i_{max-1}} \\ & & & C_{i_{max}} & D_{i_{max}} \end{bmatrix}_{(nm\times nm)} ,$$

$$D_i = \begin{bmatrix} 2(\lambda+\lambda') & -\lambda' & & & -\lambda' \\ -\lambda' & \ddots & \ddots & O & \\ & \ddots & \ddots & \ddots & \\ & O & \ddots & \ddots & -\lambda' \\ -\lambda' & & & -\lambda' & 2(\lambda+\lambda') \end{bmatrix} , \ i=i_{min},i_{min}+1,\ldots,i_{max} \ (=i_{min}+m-1);$$

$$E_i = \begin{bmatrix} -\lambda(1+\frac{1}{2i}) & & O \\ & \ddots & \\ O & & -\lambda(1+\frac{1}{2i}) \end{bmatrix}_{n\times n} , \ i=i_{min},i_{min}+1,\ldots,i_{max}-1$$

and

$$C_i = \begin{bmatrix} -\lambda(1-\frac{1}{2i}) & & O \\ & \ddots & \\ O & & -\lambda(1-\frac{1}{2i}) \end{bmatrix} , \ i=i_{min}+1,\ldots,i_{max}. \qquad (3.12.8a)$$

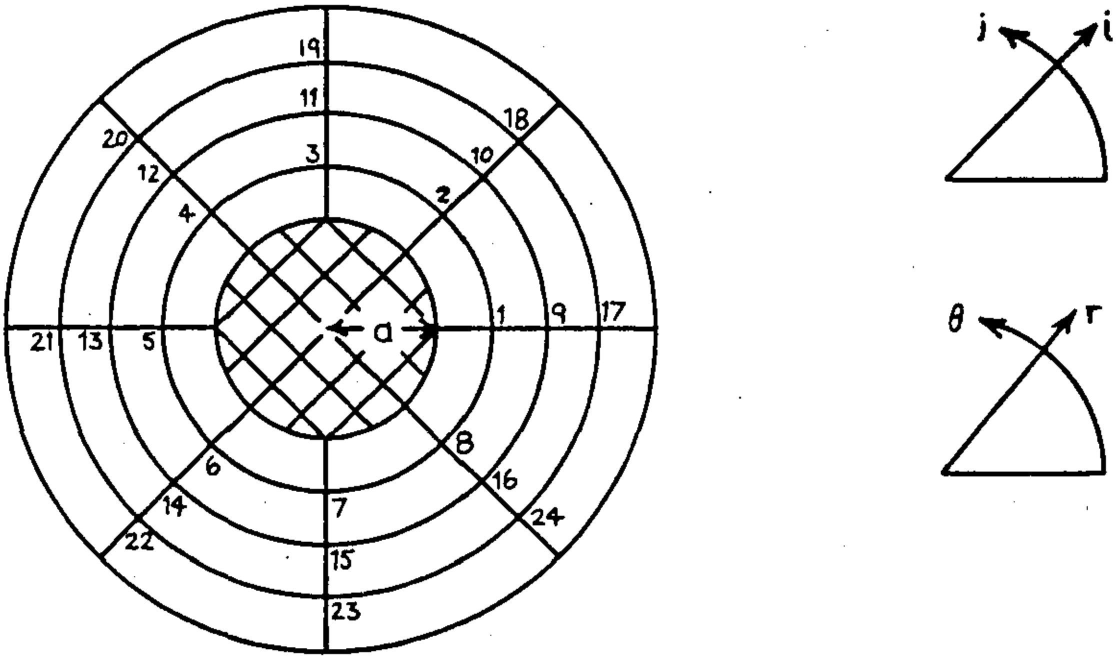


FIGURE 3.12.1

Let us now consider the case when R is the unit circle. As can be seen from Figure 3.12.2, the "Δr mesh lengths", are arranged along each diameter such that the centre point of the circle lies at the midpoint of a "Δr mesh length". This means that if m is the number of internal peripherals, then $\Delta r = \frac{2}{(2m+1)}$. For this region, the first peripheral is of radius $\frac{\Delta r}{2}$, so that since $r_{i_{min}} = i_{min}\Delta r$ we get $i_{min}=\frac{1}{2}$. Hence, when the explicit formula (3.12.7) is applied on this peripheral, we find that the coefficient of $u_{i-1,j,k}$, i.e., $(1 - \frac{1}{2i_{min}})=0$. If the grid points (in Figure 3.12.2) are ordered in the same way as in Figure 3.12.1, then we obtain an analogous system of equations to (3.12.8) where in this case, $i_{min}=\frac{1}{2}$ (which from (3.12.4) is equivalent to defining $a = -\frac{\Delta r}{2}$). Hence for the circle, we have $\lambda = \frac{\Delta t}{(\Delta r)^2}$ and $\lambda' = \frac{\Delta t}{(r_{i_{min}}\Delta\theta)^2} = \frac{\Delta t}{(\frac{1}{2}(\Delta r)(\Delta\theta))^2}$.

For the stability of the explicit formula (3.12.7), we must have from (3.12.8), the condition,

$$||(I-A)|| \leqslant 1 , \qquad (3.12.9)$$

where I-A is *the amplification matrix*. In the L_∞ norm, this becomes,

$$||I-A|| = \max_i\{|1-2(\lambda+\lambda')|+2(\lambda+\lambda')\} \leqslant 1$$

if $$\max_i\{4(\lambda+\lambda')\} \leqslant 2.$$

Now, $$\max_i\lambda' = \max_i \{\frac{\Delta t}{(r_i\Delta\theta)^2}\} = \lambda''$$

where $\lambda'' = \frac{\Delta t}{(r_{i_{min}}\Delta\theta)^2}$.

Hence, the stability condition is,

$$\lambda + \lambda'' \leqslant \tfrac{1}{2} . \qquad (3.12.10)$$

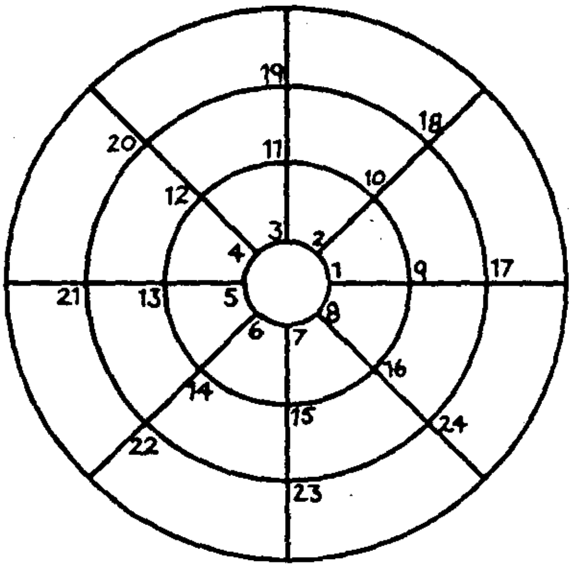


FIGURE 3.12.2

When R is the ring $(a \leq r \leq 1)$, condition (3.12.10) is equivalent to

$$\frac{\Delta t}{(\Delta t)^2} + \frac{\Delta t}{[(a+\Delta r)\Delta\theta]^2} \leq 0.5 \ , \tag{3.12.11}$$

and when R is the circle $(0 \leq r \leq 1)$, (3.12.10) becomes,

$$\frac{\Delta t}{(\Delta r)^2} + \frac{4\Delta t}{[\Delta r \Delta\theta]^2} \leq 0.5 \ . \tag{3.12.12}$$

By means of the Taylor series approximation, the local truncation error of the explicit approximation (3.12.7) is given by

$$T_{i,j,k} = \tfrac{1}{2}\Delta t \left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j,k} - \frac{(\Delta r)^2}{12}\left\{\left(\frac{\partial^4 U}{\partial r^4}\right)_{i,j,k} + \frac{2}{r_i}\left(\frac{\partial^3 U}{\partial r^3}\right)_{i,j,k}\right\}$$

$$- \frac{(\Delta\theta)^2}{12 r_i^2}\left(\frac{\partial^4 U}{\partial \theta^4}\right)_{i,j,k} + \ldots \tag{3.12.13}$$

Hence, $T = O(\Delta t) + O([\Delta r]^2) + O([\Delta\theta]^2)$.

Due to the poor stability of the above method, a natural recourse is to develop implicit schemes for two dimensional problems. The reader may consult, for example, the work of Gane (1974), in which *the two-step Peaceman-Rachford process* and *the hopscotch scheme* were applied to solve the heat conduction equation (3.12.2).

The heat conduction equation $\frac{\partial U}{\partial t} = \nabla^2 U$ in *spherical polar co-ordinates* (r,ϕ,θ) is given by,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial r^2} + \frac{2}{r}\frac{\partial U}{\partial r} + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 U}{\partial \phi^2} + \frac{1}{r^2}\frac{\partial^2 U}{\partial \theta^2} + \frac{\cot\theta}{r^2}\frac{\partial U}{\partial \theta} \tag{3.12.14}$$

$$= \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2 \frac{\partial U}{\partial r}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 U}{\partial \phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial U}{\partial \theta}\right)\right]. \tag{3.12.15}$$

It is of interest to note that when the cylindrical problem (3.12.2) is symmetrical with respect to the origin, then the equation reduces to

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial r^2} + \frac{1}{r}\frac{\partial U}{\partial r} \ , \tag{3.12.16}$$

with $$\frac{\partial U}{\partial r} = 0 \quad \text{at } r=0 \; . \qquad (3.12.16a)$$

Similarly, the spherical symmetry of the heat problem (3.12.14) leads to,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial r^2} + \frac{2}{r}\frac{\partial U}{\partial r} \; , \qquad (3.12.17)$$

and $$\frac{\partial U}{\partial r} = 0 \text{ at } r=0 \; . \qquad (3.12.17a)$$

By an appropriate change of variable that *excludes* $r=0$, the above problems can be solved by simpler equations. The change of the independent variable defined by $R=\log r$, transforms the cylindrical equation (3.12.16) to $e^{2R}\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial R^2}$. By the same token, the change of the dependent variable given by $U = \frac{W}{r}$, transforms the spherical equation (3.12.17) to $\frac{\partial W}{\partial t} = \frac{\partial^2 W}{\partial r^2}$.

3.13 EXPLICIT METHODS FOR SPECIAL GEOMETRIES

We now consider the following one-dimensional problem

$$\frac{\partial U}{\partial t} = \frac{1}{r^{\alpha}} \frac{\partial}{\partial r}\left(r^{\alpha}\frac{\partial U}{\partial r}\right),\ 0<r<1,\ 0<t\leq T\ , \qquad (3.13.18)$$

$$= \frac{\partial^2 U}{\partial r^2} + \frac{\alpha}{r}\frac{\partial U}{\partial r}\ , \qquad (3.13.19)$$

subject to,

$$U(r,0) = f(r),\ 0\leq r\leq 1\ , \qquad (3.13.19a)$$

and $\frac{\partial U(0,t)}{\partial r} = 0,\ U(1,t) = 0,\quad 0\leq t\leq T$.

For $\alpha=0$ (the Cartesian plane case), equation (3.13.18) reduces to the aforementioned equation (3.1.2). For $\alpha=1$, we get equation (3.12 .16) whilst $\alpha=2$ gives us (3.12.17). We note that, if the variation of the function U along the cylinder ($\alpha=1$) cannot be neglected (this happens when the height of the cylinder is not very large in comparison with the diameter), then instead of the one-dimensional equation (3.13.18), the two-dimensional equation,

$$\frac{\partial U}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial U}{\partial r}\right) + \frac{\partial^2 U}{\partial z^2} \qquad (3.13.20)$$

$$= \frac{\partial^2 U}{\partial r^2} + \frac{1}{r}\frac{\partial U}{\partial r} + \frac{\partial^2 U}{\partial z^2} \qquad (3.13.21)$$

is used, with the condition,

$$\frac{\partial U}{\partial r}(0,z,t) = 0\ . \qquad (3.13.22)$$

An explicit finite-difference approximation to (3.13.19) is given by

$$\frac{1}{\Delta t}\Delta_t u_{ij} = \frac{1}{(\Delta r)^2}\delta_r^2\, u_{i,j} + \frac{\alpha}{r_i}\ \frac{1}{(2\Delta r)}(\Delta_r+\nabla_r)u_{ij}\ ,$$

$$\text{for } i=1,2,\ldots,m-1 \qquad (3.13.23)$$

i.e.,

$$\frac{(u_{i,j+1}-u_{ij})}{\Delta t} = \frac{(u_{i-1,j}-2u_{ij}+u_{i+1,j})}{(\Delta r)^2} + \frac{\alpha}{r_i}\ \frac{(u_{i+1,j}-u_{i-1,j})}{2\Delta r}$$

$$\text{for } i=1,2,\ldots,m-1. \qquad (3.13.24)$$

After rearrangement of terms, this becomes,

$$u_{i,j+1} = \lambda[(1 - \frac{\alpha\Delta r}{2r_i})u_{i-1,j} + (1+ \frac{\alpha\Delta r}{2r_i})u_{i+1,j}]+(1-2\lambda)u_{i,j},$$

$$\text{for } i=1,2,\ldots,m-1. \qquad (3.13.25)$$

Another explicit approximation takes the form,

$$\frac{1}{\Delta t}\Delta_t u_{ij} = \frac{(r_i - \frac{\Delta r}{2})^\alpha u_{i-1,j} - [(r_i - \frac{\Delta r}{2})^\alpha + (r_i + \frac{\Delta r}{2})^\alpha]u_{ij} + (r_i + \frac{\Delta r}{2})^\alpha u_{i+1,j}}{r_i^\alpha(\Delta r)^2}$$

$$\text{for } i=1,2,\ldots,m-1, \qquad (3.13.26)$$

which on simplification reduces to,

$$u_{i,j+1} = \frac{\lambda}{r_i^\alpha}[(r_i - \frac{\Delta r}{2})^\alpha u_{i-1,j} + (r_i + \frac{\Delta r}{2})^\alpha u_{i+1,j}] + [1 - \frac{\lambda}{r_i^\alpha}\{(r_i - \frac{\Delta r}{2})^\alpha +$$

$$(r_i + \frac{\Delta r}{2})^\alpha\}]u_{ij}, \text{ for } i=1,2,\ldots m-1.$$

$$(3.13.27)$$

Now, in view of the condition $\frac{\partial U}{\partial r}(0,t)=0$ at *the axis* $r=0$, we have,

$$\left.\frac{\partial U}{\partial t}\right|_{r=0} = \lim_{r\to 0}\left(\frac{\partial^2 U}{\partial r^2} + \frac{\alpha}{r}\frac{\partial U}{\partial r}\right) = \left.\frac{\partial^2 U}{\partial r^2}\right|_{r=0} + \alpha \lim_{r\to 0}\frac{\partial U/\partial r}{r} = (1+\alpha)\left.\frac{\partial^2 U}{\partial r^2}\right|_{r=0}$$

$$(3.13.28)$$

and accordingly, for $i=0$, we use the approximation,

$$\frac{(u_{0,j+1}-u_{0,j})}{\Delta t} = (1+\alpha)\frac{(u_{-1,j}-2u_{0,j}+u_{1,j})}{(\Delta r)^2} \qquad (3.13.29)$$

The fictituous value $u_{-1,j}$ is eliminated by using the approximation at the axis $r=0$, i.e.,

$$\left.\frac{\partial U}{\partial r}\right|_{r=0} \simeq \frac{1}{2\Delta r}(\Delta_r+\nabla_r)u_{0j} = 0 .$$

This gives us $u_{-1,j} = u_{1j}$ and with this value, equation (3.13.28) is simplified to

$$u_{0,j+1} = u_{0j} + 2(1+\alpha)\lambda(u_{1j}-u_{0j}) . \qquad (3.13.30)$$

Equations (3.13.30) and (3.13.25) can be combined and written in the matrix form as,

$$\underset{\sim}{u}_{j+1} = \Gamma \underset{\sim}{u}_j \ , \tag{3.13.31}$$

where $\underset{\sim}{u}_j=(u_{0j},u_{1j},u_{2j},\ldots,u_{m-1,j})^T$ and Γ is the $(m\times m)$ amplification matrix given by,

$$\Gamma = \begin{bmatrix} 1-2(1+\alpha)\lambda & 2(1+\alpha)\lambda & & & & \\ \lambda(1-\frac{\alpha}{2}) & (1-2\lambda) & \lambda(1+\frac{\alpha}{2}) & & O & \\ & \lambda(1-\frac{\alpha}{4}) & (1-2\lambda) & \lambda(1+\frac{\alpha}{4}) & & \\ & & \ddots & \ddots & \ddots & \\ & O & & & & \\ & & & & \lambda(1-\frac{\alpha}{2(m-1)}) & (1-2\lambda) \end{bmatrix}$$

When,

$$\lambda \leqslant \frac{1}{2(1+\alpha)} \ , \tag{3.13.32}$$

the sum of the moduli of the terms along each row of Γ is less than or equal to 1. Therefore, by *Gerschgorin's first theorem*, the spectral radius, $\rho(\Gamma)\leqslant 1$, giving stability. We note that the condition (3.13.32) is in fact the stability requirement of the equation (3.13.30) for the left boundary whilst the stability condition for (3.13.25) turns out to be

$$\lambda \leqslant \tfrac{1}{2} \tag{3.13.32a}$$

for $\alpha=0,1,2$. It is clear that it is possible to achieve a better overall stability bound by improving the stability range of (3.13.30). Instead of the explicit equation (3.13.30), we shall use the following implicit approximation for the left boundary, i.e.,

$$\frac{(u_{0,j+1}-u_{0j})}{\Delta t} = 2(1+\alpha)\frac{(u_{1,j+1}-u_{0,j+1})}{(\Delta x)^2} \ . \tag{3.13.33}$$

This formula may be used explicitly. To do so, it is only necessary to arrive first at the value of $u_{1,j+1}$ by the formula (3.13.25). Equation (3.13.33) can then be readily rewritten as

$$u_{0,j+1} = \frac{(u_{0j}+2(1+\alpha)\lambda u_{1,j+1})}{(1+2(1+\alpha)\lambda)} \tag{3.13.34}$$

which is stable for all values of λ.

From Taylor's series expansion, it can be easily ascertained that the truncation error of (3.13.33) or (3.13.34) is,

$$T_{0,j} = \frac{\Delta t}{2}\left(\frac{\partial^2 U}{\partial t^2}\right)_{0,j} + \frac{(1+\alpha)}{12}(\Delta r)^2\left(\frac{\partial^4 U}{\partial r^4}\right)_{0,j} + \cdots \tag{3.13.35}$$

It can be similarly derived that the local truncation error of (3.13.25) is given by,

$$T_{ij} = \tfrac{1}{2}\Delta t\left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j} - (\Delta r)^2\left\{\frac{\alpha}{24}\left(\frac{\partial^4 U}{\partial r^4}\right)_{i,j} + \frac{\alpha}{6r_i}\left(\frac{\partial^3 U}{\partial r^3}\right)_{i,j}\right\} + \ldots \tag{3.13.36}$$

Hence, from (3.13.35) and (3.13.36), we find that the local truncation error of the explicit method (3.13.25) when applied to *the cartesian plane* (α=0), *cylindrical* (α=1) and *spherical* (α=2) problem is,

$$T = O(\Delta t) + O([\Delta r]^2) .$$

Let us now return to the second explicit formula (3.13.27). As before, the left boundary at the axis is determined either by equation (3.13.30) or equation (3.13.33). The latter, however, is the best choice because of stability. It can also be shown by the application of Gerschgorin's theorem that the conditions for stability of (3.13.27) are,

$$\lambda \leq \begin{cases} \frac{1}{2} & \text{if } \alpha=0 \text{ or } 1, \\ \frac{2}{5} & \text{if } \alpha=2 . \end{cases} \tag{3.13.37}$$

An expression for the local truncation error at the mesh point (r_i,t_j) is,

$$T_{ij} = \tfrac{1}{2}\Delta t\left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j} - (\Delta r)^2\left\{\frac{\alpha}{24}\left(\frac{\partial^4 U}{\partial r^4}\right)_{i,j} + \frac{\alpha}{6r_i}\left(\frac{\partial^3 U}{\partial r^3}\right)_{i,j} + \frac{1}{4r_i^2}\left(\frac{\partial^2 U}{\partial r^2}\right)_{i,j}\right\} + \ldots \qquad (3.13.38)$$

Again, we find that $T=O(\Delta t)+O([\Delta r]^2)$.

We now proceed to discuss briefly the numerical solution of the cylindrical heat conduction equation in two dimensions given by (3.13.21). Instead of equation (3.13.27), for example, we have the equation (with $\alpha=1$ and for simplicity, we assume $\Delta r=\Delta z$),

$$\frac{(u_{i,j,k+1}-u_{i,j,k})}{\Delta t} = \frac{[(1-\frac{\Delta r}{2r_i})u_{i-1,j,k}-2u_{i,j,k}+(1+\frac{\Delta r}{2r_i})u_{i+1,j,k}]}{(\Delta r)^2} + \frac{[u_{i,j-1,k}-2u_{i,j,k}+u_{i,j+1,k}]}{(\Delta r)^2} \qquad (3.13.39)$$

This can be rearranged as,

$$u_{i,j,k+1} = (1-4\lambda)u_{i,j,k}+\lambda\{(1-\frac{\Delta r}{2r_i})u_{i-1,j,k}+u_{i,j-1,k}+(1+\frac{\Delta r}{2r_i})u_{i+1,j,k} + u_{i,j+1,k}\}, \qquad (3.13.40)$$

with the stability ratio $\lambda\leq 1/4$. At the axis $r=0$, equation (3.13.21) transforms to the equation,

$$\frac{\partial U}{\partial t} = 2\frac{\partial^2 U}{\partial r^2} + \frac{\partial^2 U}{\partial z^2},$$

and accordingly, equation (3.13.39) is replaced by,

$$\frac{(u_{0,j,k+1}-u_{0,j,k})}{\Delta t} = 2\frac{(u_{-1,j,k}-2u_{0,j,k}+u_{1,j,k})}{(\Delta r)^2} + \frac{(u_{0,j-1,k}-2u_{0,j,k}+u_{0,j+1,k})}{(\Delta r)^2} \qquad (3.13.41)$$

In view of the condition (3.13.22), we have $u_{-1,j,k}=u_{1,j,k}$.

Thus, equation (3.13.41) becomes,

$$u_{0,j,k+1} = (1-6\lambda)u_{0,j,k}+\lambda(u_{0,j-1,k}+u_{0,j+1,k}+4\lambda u_{1,j,k}) \tag{3.13.42}$$

which is stable only for $\lambda \leq 1/6$. This, inevitably, affects the overall stability of the explicit scheme which approximates the heat conduction equation. An implicit approximation is therefore employed by "shifting" the first term on the right-hand side in (3.13.41) from the kth level to the (k+1)th level to give,

$$\frac{(u_{0,j,k+1}-u_{0,j,k})}{\Delta t} = \frac{4(u_{1,j,k+1}-u_{0,j,k+1})}{(\Delta r)^2} + \frac{(u_{0,j-1,k}-2u_{0,j,k}+u_{0,j+1,k})}{(\Delta r)^2} \tag{3.13.43}$$

or,

$$u_{0,j,k+1} = \frac{1}{(1+4\lambda)}\{4\lambda u_{1,j,k+1}+(1-2\lambda)u_{0,j,k}+\lambda(u_{0,j-1,k}+u_{0,j+1,k})\} \tag{3.13.44}$$

The values of $u_{1,j,k+1}$ are first obtained by the explicit formula (3.13.40). Formula (3.13.44) is stable even for $\lambda \leq \frac{1}{2}$ and the stability of the fundamental scheme (3.13.40) is therefore not dominated by this ratio. We observe that rather than approximating equation (3.13.21) at the axis, we could have simply used $u_{0,j,k}=u_{1,j,k}$ by virtue of the condition (3.13.22). In doing so, however, the local truncation error of the approximation at the axis (r=0) is $T=O(\Delta r)$ whereas the local truncation error of the difference analogue (3.13.40) at mesh points away from the axis (r>0) is $T=O(\Delta t)+O([\Delta r]^2)$.

3.14 IMPLICIT METHODS FOR SPECIAL GEOMETRIES

A fully implicit scheme to approximate (3.13.19) at the mesh points not on the axis (r>0) can be derived by using central difference analogues for the spatial derivatives and the usual forward difference for $\frac{\partial U}{\partial t}$. Our approximation, therefore, takes the form,

$$\frac{(u_{i,j+1}-u_{ij})}{\Delta t} = \frac{1}{(\Delta r)^2}(\delta_r u_{i+\frac{1}{2},j+1}-\delta_r u_{i-\frac{1}{2},j+1}) + \frac{\alpha}{(i\Delta r)}\,\frac{1}{(2\Delta r)}(\Delta_r u_{i,j+1}+\nabla_r u_{i,j+1}) \qquad (3.14.1)$$

which on expansion becomes,

$$u_{i,j+1} = u_{ij}+\lambda(u_{i+1,j+1}-2u_{i,j+1}+u_{i-1,j+1}) + \frac{\alpha}{2i}\lambda(u_{i+1,j+1}-u_{i-1,j+1})$$

or,

$$-p_i u_{i-1,j+1}+(1+2\lambda)u_{i,j+1}-q_i u_{i+1,j+1} = u_{ij}\ , \qquad (3.14.1a)$$

$$\text{for } i=1,2,\ldots,m-1,$$

where $p_i = (1-\frac{\alpha}{2i})\lambda,\ q_i = (1+\frac{\alpha}{2i})\lambda.$ (3.14.2)

At the axis, the equation yielding the boundary values on the left is given by (3.13.34), i.e.,

$$[1+2(1+\alpha)\lambda]u_{0,j+1}-2(1+\alpha)\lambda u_{1,j+1} = u_{0j}\ . \qquad (3.14.3)$$

Bearing in mind that $u_{m,j}=0$ from (3.13.19a), equations (3.14.3) and (3.14.1) can be combined in matrix form as,

$$\begin{bmatrix} 1+2(1+\alpha)\lambda & -2(1+\alpha)\lambda & & & & \\ -p_1 & (1+2\lambda) & -q_1 & & O & \\ & -p_2 & (1+2\lambda) & -q_2 & & \\ & & \ddots & \ddots & \ddots & \\ & O & & -p_{m-1} & (1+2\lambda) & -q_{m-2} \\ & & & & -p_{m-1} & (1+2\lambda) \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j+1} = \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j}$$

i.e. $$A\underline{u}_{j+1} = \underline{u}_j \tag{3.14.4}$$

or $$\underline{u}_{j+1} = \Gamma\underline{u}_j \tag{3.14.5}$$

where $\Gamma=A^{-1}$, the amplification matrix of order (m×m).

Now, the matrix A of order (m×m) can be split as

$$A = I+\lambda C \tag{3.14.6}$$

where,

$$C = \begin{bmatrix} 2(\alpha+1) & -2(\alpha+1) & & & & \\ -c_1 & 2 & -d_1 & & & \\ & -c_2 & 2 & -d_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -c_{m-2} & 2 & -d_{m-2} \\ & & & & -c_{m-1} & 2 \end{bmatrix}_{(m\times m)}$$

and $c_i = 1 - \frac{\alpha}{2i}$ and $d_i = 1 + \frac{\alpha}{2i}$, for i=1,2,...,m-1.

For our special geometrical problem, $\alpha=1$ or 2. It is clear that for these values of α, C is positive definite. If we let η_i to be the eigenvalues of C, then this implies that η_i are all real and positive. Hence, from equation (3.14.6),

$$\mu_i = \frac{1}{(1+\lambda\eta_i)} \quad , \; i=0,1,\ldots,m-1$$

where μ_i are the eigenvalues of A^{-1}. It is obvious that $|\mu_i|<1$ for all positive λ. Thus, *the fully implicit scheme* is unconditionally stable.

The local truncation error of (3.14.3) at points on the axis is given by (3.13.35), i.e.,

$$T_{0,j} = \tfrac{1}{2}\Delta t\left(\frac{\partial^2 U}{\partial t^2}\right)_{0,j} + \frac{(1+\alpha)}{12}\,(\Delta r)^2\,\left(\frac{\partial^4 U}{\partial r^4}\right)_{0,j} + \ldots \tag{3.14.7}$$

The Taylor series expansion for (3.14.1a) at the mesh point (r_i,t_j)

for $r_i>0$ also yields,

$$T_{ij} = (2\lambda-(p_i+q_i))U_{i,j}+\Delta r(p_i-q_i)(\frac{\partial U}{\partial r})_{i,j} +\Delta t(1+(2\lambda-(p_i+q_i)))(\frac{\partial U}{\partial t})_{i,j}$$

$$-\tfrac{1}{2}(\Delta r)^2(p_i+q_i)(\frac{\partial^2 U}{\partial r^2})_{i,j}+(\Delta r)(\Delta t)(p_i-q_i)(\frac{\partial^2 U}{\partial r\partial t})_{i,j}$$

$$+\tfrac{1}{2}(\Delta t)^2(1+(2\lambda-(p_i+q_i))) + \frac{1}{6}(\Delta r)^3(p_i-q_i)(\frac{\partial^3 U}{\partial r^3})_{i,j}$$

$$-\tfrac{1}{2}(\Delta r)^2\Delta t(p_i+q_i)(\frac{\partial^3 U}{\partial r^2\partial t})_{i,j} + \tfrac{1}{2}\Delta r(\Delta t)^2(p_i-q_i)(\frac{\partial^3 U}{\partial r\partial t^2})_{i,j}$$

$$+ \frac{1}{6}(\Delta t)^3(1+(2\lambda-(p_i+q_i)))(\frac{\partial^3 U}{\partial t^3})_{i,j} + \ldots$$

$$= \Delta t\{(\frac{\partial U}{\partial t})_{i,j} - (\frac{\partial^2 U}{\partial r^2})_{i,j} + \frac{\alpha}{r_i}(\frac{\partial U}{\partial r})_{i,j}\} + \frac{\alpha}{r_i}(\Delta t)^2(\frac{\partial^2 U}{\partial r\partial t})_{i,j} + \frac{(\Delta t)^2}{2}$$

$$(\frac{\partial^2 U}{\partial t^2})_{i,j} + \frac{\alpha}{6r_i}(\Delta r)^2(\frac{\partial^3 U}{\partial r^3})_{i,j} - (\Delta t)^2(\frac{\partial^3 U}{\partial r^2\partial t})_{i,j} + \frac{\alpha}{2r_i}(\Delta t)^2$$

$$(\frac{\partial^3 U}{\partial r\partial t^2})_{i,j} + \frac{1}{6}(\Delta t)^3 (\frac{\partial^3 U}{\partial t^3})_{i,j} .$$

Therefore,

$$T_{ij} = \Delta t(\frac{2\alpha}{r_i}(\frac{\partial U}{\partial r})_{i,j}) + (\Delta r)^2 \frac{\alpha}{6r_i}(\frac{\partial^3 U}{\partial r^3})_{i,j} + (\Delta t)^2 \{\frac{\alpha}{r_i}(\frac{\partial^2 U}{\partial r\partial t})_{i,j} + \tfrac{1}{2}(\frac{\partial^2 U}{\partial t^2})_{i,j}$$

$$-(\frac{\partial^3 U}{\partial r^2\partial t})_{i,j} + \frac{\alpha}{2r_i}(\frac{\partial^3 U}{\partial r\partial t^2})_{i,j} + \frac{\Delta t}{6}(\frac{\partial^3 U}{\partial t^3})_{i,j}\} + \ldots \qquad (3.14.8)$$

We see from (3.14.7) and (3.14.8) that $T=O(t)+O([\Delta r]^2)$ for the fully implicit scheme.

By averaging the central difference approximations for the spatial derivatives at the jth time level and the (j+1)th level, we obtain the following *Crank-Nicolson analogue* of the differential equation (3.3.19),

$$\frac{1}{\Delta t}\Delta_t u_{ij} = \tfrac{1}{2}\frac{1}{(\Delta r)^2}\{\delta_r^2 u_{i,j+1} + \delta_r^2 u_{ij}\} + (\frac{\alpha}{i\Delta r})\ \tfrac{1}{2}\{\frac{(\Delta_r u_{i,j+1} + \nabla_r u_{i,j+1}) + (\Delta_r u_{ij} + \nabla_r u_{ij})}{2\Delta r}$$

$$(3.14.9)$$

$$= \frac{1}{2(\Delta r)^2}\{(u_{i+1,j+1}-2u_{i,j+1}+u_{i-1,j+1})+(u_{i+1,j}-2u_{ij}+u_{i-1,j})\}$$

$$+ \frac{\alpha}{4i(\Delta r)^2}\{(u_{i+1,j+1}-u_{i-1,j+1})+(u_{i+1,j}-u_{i-1,j})\},$$

$$\text{for } i=1,2,\ldots,m-1. \qquad (3.14.9a)$$

This leads to the tridiagonal system,

$$\tfrac{1}{2}p_i u_{i-1,j+1}-(1+\lambda)u_{i,j+1}+\tfrac{1}{2}q_i u_{i+1,j+1}$$

$$= -\tfrac{1}{2}p_i u_{i-1,j}-(1-\lambda)u_{ij}-\tfrac{1}{2}q_i u_{i+1,j}\ ,\ i=1,2,\ldots,m-1 \qquad (3.14.10)$$

where, as before,

$$p_i = (1-\frac{\alpha}{2i})\lambda,\quad q_i = (1+\frac{\alpha}{2i})\lambda\ ,\ i=1,2,\ldots,m-1, \qquad (3.14.11)$$

By virtue of equation (3.13.28), the approximations for points on the axis are given by,

$$\frac{1}{\Delta t}\Delta_t u_{0,j} = (1+\alpha)\ \tfrac{1}{2}\ \frac{1}{(\Delta r)^2}\{\delta_r^2 u_{0,j+1}+\delta_r^2 u_{0,j}\}\ ,\ \text{i.e.,}$$

$$\frac{(u_{0,j+1}-u_{0j})}{\Delta t} = \frac{(1+\alpha)}{2(\Delta r)^2}\{u_{1,j+1}-2u_{0,j+1}+u_{-1,j+1}+u_{1j}-2u_{0j}+u_{-1,j}\}$$

which leads to (with $u_{-1,j}=u_{1,j}$),

$$[1+(1+\alpha)\lambda]u_{0,j+1}-(1+\alpha)\lambda u_{1,j+1} = [1-(1+\alpha)\lambda]u_{0j}+(1+\alpha)\lambda u_{1j}\ .$$

$$(3.14.12)$$

In matrix notation, equations (3.14.12) and (3.14.10) can be written as,

$$A\underline{u}_{j+1} = B\underline{u}_j \qquad (3.14.13)$$

where,

$$A = \begin{bmatrix} -1-(1+\alpha)\lambda & (1+\alpha)\lambda & & & & \\ \frac{1}{2}p_1 & -(1+\lambda) & \frac{1}{2}q_1 & & & \\ & \frac{1}{2}p_2 & -(1+\lambda) & \frac{1}{2}q_2 & \bigcirc & \\ & & \ddots & \ddots & \ddots & \\ & \bigcirc & & \frac{1}{2}p_{m-2} & -(1+\lambda) & \frac{1}{2}q_{m-2} \\ & & & & \frac{1}{2}p_{m-1} & -(1+\lambda) \end{bmatrix}_{(m\times m)} \tag{3.14.14}$$

$$B = \begin{bmatrix} -1+(1+\alpha)\lambda & -(1+\alpha)\lambda & & & & \\ -\frac{1}{2}p_1 & -(1-\lambda) & -\frac{1}{2}q_1 & & & \\ & -\frac{1}{2}p_2 & -(1-\lambda) & -\frac{1}{2}q_2 & \bigcirc & \\ & & \ddots & \ddots & \ddots & \\ & \bigcirc & & -\frac{1}{2}p_{m-2} & -(1-\lambda) & -\frac{1}{2}q_{m-2} \\ & & & & -\frac{1}{2}p_{m-1} & -(1-\lambda) \end{bmatrix}_{(m\times m)}$$

and $\underline{u}_j = (u_{0,j}, u_{1,j}, \ldots, u_{m-1,j})^T$.

If we let,

$$C = \begin{bmatrix} 2(1+\alpha) & -2(1+\alpha) & & & & \\ -c_1 & 2 & -d_1 & & & \\ & -c_2 & 2 & -d_2 & \bigcirc & \\ & & \ddots & \ddots & \ddots & \\ & \bigcirc & & -c_{m-2} & 2 & -d_{m-2} \\ & & & & -c_{m-1} & 2 \end{bmatrix}_{(m\times m)} \tag{3.14.15}$$

then $$A = -(I+\frac{\lambda}{2}C), \quad B = \frac{\lambda}{2}C-I \ . \tag{3.14.16}$$

From equation (3.14.13), we have

$$\underset{\sim}{u}_{j+1} = A^{-1}B\underset{\sim}{u}_j \ , \tag{3.14.17}$$

$$= \Gamma\underset{\sim}{u}_j \tag{3.14.18}$$

where Γ is the amplification matrix of the Crank-Nicolson scheme. Now, if the μ_i are the eigenvalues of $A^{-1}B$ corresponding to the eigenvectors $\underset{\sim}{v}_i$, then the equation defining μ_i and $\underset{\sim}{v}_i$ is

$$(A^{-1}B-\mu_i I)\underset{\sim}{v}_i = \underset{\sim}{O} \ , \tag{3.14.19}$$

or equivalently,

$$(B-\mu_i A)\underset{\sim}{v}_i = \underset{\sim}{O} \ . \tag{3.14.20}$$

On using the relation (3.14.16), we obtain,

$$\{\frac{\lambda}{2}(1+\mu_i)C+(\mu_i-1)I\}\underset{\sim}{v}_i = 0 \ ,$$

or $$\{C - \frac{2(1-\mu_i)}{\lambda(1+\mu_i)} I\}\underset{\sim}{v}_i = 0$$

which shows that $\frac{2(1-\mu_i)}{\lambda(1+\mu_i)}$ are the eigenvalues of C. Hence, if we denote η_i as the eigenvalues of C, then

$$\eta_i = \frac{2(1-\mu_i)}{\lambda(1+\mu_i)}$$

or $$\mu_i = \frac{(2-\lambda\eta_i)}{(2+\lambda\eta_i)} \tag{3.14.21}$$

We observe from (3.14.15) that the matrix C is positive definite for the values of α equal to 1 or 2. This implies that η_i, the eigenvalues of C are all real and positive and we deduce from (3.14.21) that $|\mu_i|<1$ for all values of λ. Therefore, the Crank-Nicolson formula possesses unrestricted stability.

To determine the order of accuracy of this scheme, we again resort to Taylor's series expansion for equation (3.14.10). Hence,

$$T_{ij} = [\frac{\alpha}{r_i}(\frac{\partial U}{\partial r})_{i,j} - (\frac{\partial U}{\partial t})_{i,j} + (\frac{\partial^2 U}{\partial r^2})_{i,j}] + \frac{\alpha}{2r_i}\Delta t(\frac{\partial^2 U}{\partial r \partial t})_{i,j}$$

$$-\tfrac{1}{2}\Delta t(\frac{\partial^3 U}{\partial t \partial r^2})_{i,j} - \frac{\alpha \Delta t}{2r_i}(\frac{\partial^2 U}{\partial t \partial r})_{i,j} + \tfrac{1}{2}\Delta t(\frac{\partial^3 U}{\partial r^2 \partial t})_{i,j} + \frac{\alpha}{8r_i}(\Delta t)^2(\frac{\partial^3 U}{\partial r \partial t^2})_{i,j}$$

$$+ \frac{\alpha}{6r_i}(\Delta r)^2(\frac{\partial^3 U}{\partial t^3})_{i,j} - \frac{1}{24}(\Delta t)^2(\frac{\partial^3 U}{\partial t^3})_{i,j} + \dots ,$$

which reduces to,

$$T_{ij} = \frac{\alpha}{6r_i}(\Delta r)^2(\frac{\partial^3 U}{\partial r^3})_{i,j} + \frac{\alpha}{8r_i}(\Delta t)^2(\frac{\partial^3 U}{\partial r \partial t^2})_{i,j} - \frac{1}{24}(\Delta t)^2(\frac{\partial^3 U}{\partial t^3})_{i,j} + \dots \tag{3.14.22}$$

Therefore, $T_{ij}=O([\Delta r]^2)+O([\Delta t]^2)$. In the same manner, we also find from (3.14.12) that $T_{0j} = O([\Delta r]^2)+O([\Delta t]^2)$. We conclude that the Crank-Nicolson approximation to solve our special geometrical parabolic problem (α=1 and α=2) is second-order accurate in both radial space and time.

3.15 IMPROVING THE ACCURACY OF FINITE DIFFERENCE SCHEMES FOR DIFFUSION EQUATIONS IN CYLINDRICAL COORDINATES

We shall restrict ourselves to solving the cylindrical problem,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial r^2} + \frac{1}{r}\frac{\partial U}{\partial r} \qquad (3.15.1)$$

and develop methods (due to Polak (1974)) to improve the accuracy of the solution.

In conformity with the convention of Section 3.8, equation (3.15.1) can also be written in the form,

$$D_t U = L(D_r^2, D_r)U \ . \qquad (3.15.2)$$

The θ-weighted schemes of (3.10.12) may be employed to approximate the solution of (3.15.2), i.e.,

$$\exp((\theta-1)(\Delta t)L(D_r^2,D_r))U_{i,j+1} = \exp(\theta(\Delta t)L(D_r^2,D_r)U_{i,j} \ . \qquad (3.15.3)$$

As we have seen in the last section (cf. equation (3.14.1) *central difference approximations* are again used for the derivatives in (3.15.3). Hence D_r is replaced by $\frac{1}{2\Delta r}\delta_{2r}$ and D_r^2 by $\frac{1}{(\Delta r)^2}\delta_r^2$, where we note that,

$$\delta_{2r}U_{i,j} = (\Delta_r+\nabla_r)U_{i,j}$$

and

$$\frac{1}{(\Delta r)^2}\delta_r^2 U_{i,j} = \frac{1}{(\Delta r)^2}(\delta_r U_{i+\frac{1}{2},j} - \delta_r U_{i-\frac{1}{2},j}) \qquad (3.15.4)$$

If we neglect the higher order terms of the Taylor's development of the exponential of equation (3.15.3), we obtain the following approximation to (3.15.2),

$$(1+(\theta-1)(\Delta t)L(\frac{1}{(\Delta r)^2}\delta_r^2, \frac{1}{2\Delta r}\delta_{2r}))u_{i,j+1} = (1+\theta(\Delta t)L(\frac{1}{(\Delta r)^2}\delta_r^2, \frac{1}{2\Delta r}\delta_{2r}))u_{i,j} \ . \qquad (3.15.5)$$

Equation (3.15.5) gives us (with $\alpha=1$),

the classical explicit scheme (3.13.23) for $\theta=0$,

the Crank-Nicolson scheme (3.14.9) for $\theta=\frac{1}{2}$,

and the classical (fully) implicit scheme (3.14.1) for $\theta=1$.

We shall now demonstrate the strategy of reducing the truncation error of (3.15.3) and hence improve its accuracy by first considering the parabolic equation on the plane,

$$D_t U = D_r^2 U \ . \qquad (3.15.6)$$

As stated above, by setting $\theta=\frac{1}{2}$ in (3.15.5) we obtain *the Crank-Nicolson scheme* whose local truncation error is $T_{CN}=O([\Delta t]^2)+O([\Delta r]^2)$. This can be further reduced if $\frac{1}{(\Delta r)^2}$ on the left hand side of (3.15.5) is replaced by $\frac{1}{\Delta t}(\frac{1}{6}+\frac{\Delta t}{(\Delta r)^2})$ (i.e. $\frac{1}{\Delta t}(\frac{1}{6}+\lambda)$) and by $\frac{1}{\Delta t}(\frac{1}{6}-\lambda)$ on the right. The result is,

$$(1+\tfrac{1}{2}(\tfrac{1}{6}-\lambda)\delta_r^2)u_{i,j+1} = (1+\tfrac{1}{2}(\tfrac{1}{6}+\lambda)\delta_r^2)u_{ij} \ , \qquad (3.15.7)$$

which, in fact, corresponds with *the Douglas equation* (3.6.4) whose truncation error is known to be $T_D=O([\Delta t]^2)+O([\Delta r]^4)$. By the same token, to investigate the local truncation error of the general equation (3.15.5), let us set,

$$\alpha_2 = \frac{1}{(\Delta r)^2} \text{ and } \beta_2 = \frac{1}{2\Delta r} \ , \qquad (3.15.8)$$

in the left hand side of (3.15.5) and similarly,

$$\alpha_1 = \frac{1}{(\Delta r)^2} \text{ and } \beta_1 = \frac{1}{2\Delta r} \qquad (3.15.8a)$$

in the right hand side of the same equation. This yields,

$$(1+(\theta-1)\mathcal{L}(\Delta t)(\alpha_2\delta_r^2,\beta_2\delta_{2r}))u_{i,j+1} = (1+\theta(\Delta t)\mathcal{L}(\alpha_1\delta_r^2,\beta_1\delta_{2r}))u_{i,j} \ . \qquad (3.15.9)$$

A judicious choice of the parameters $\alpha_1,\beta_1,\alpha_2$ and β_2 will then minimise the truncation error.

Let us return to the formula (3.15.7) to determine its truncation error. From the relation,

$$\delta_r = E_r^{\frac{1}{2}} - E_r^{-\frac{1}{2}} \quad , \qquad (3.15.10)$$

where E is the shifting operator, defined by,

$$\begin{aligned} E_r^s U_{ij} &= E_r^s U(r_i,t_j) \\ &= U(r_i+s\Delta r,t_j) \\ &= U(r_i,t_j)+s\Delta r(\frac{\partial U}{\partial r})_{i,j} + \frac{(s\Delta r)^2}{2!}(\frac{\partial^2 U}{\partial r^2})_{i,j} + \frac{(s\Delta r)^3}{3!}(\frac{\partial^3 U}{\partial r^3})_{i,j} + \ldots \\ &= (1+s\Delta r\, D_r + \frac{1}{2!}(s\Delta r)^2 D_r^2 + \frac{1}{3!}(s\Delta r)^3 D_r^3 + \ldots)U_{i,j} \end{aligned}$$

i.e. $$E_r^s = \exp(s\Delta r D_r) \quad , \qquad (3.15.11)$$

we have,

$$\begin{aligned} \delta_r^2 &= E_r + E_r^{-1} - 2 \\ &= \exp(\Delta r D_r) + \exp(-\Delta r D_r) - 2 \quad \text{(from (3.15.11))}, \end{aligned}$$

i.e. $$\delta_r^2 = (\Delta r)^2 D_r^2 + \frac{1}{12}(\Delta r)^4 D_r^4 + \frac{1}{360}(\Delta r)^6 D_r^6 + \ldots \qquad (3.15.12)$$

If we put $\theta=1$, in (3.15.3), we find that,

$$\begin{aligned} U_{i,j+1} &= \exp((\Delta t)L(D_r^2,D_r))U_{i,j} \\ &= \exp((\Delta t)D_r^2)U_{ij} \quad \text{(from (3.15.6))} \qquad (3.15.13) \\ &= (1+(\Delta t)D_r^2 + \frac{1}{2!}(\Delta t)^2 D_r^4 + \frac{1}{3!}(\Delta t)^3 D_r^6 + \ldots)U_{ij} \qquad (3.15.13a) \end{aligned}$$

The substitution of the truncated series of (3.15.12) and (3.15.13a) into (3.15.7) yields,

$$(1+\tfrac{1}{2}(\frac{1}{6}+\lambda)[(\Delta r)^2 D_r^2 + \frac{1}{12}(\Delta r)^4 D_r^4])U_{i,j} = (1+\tfrac{1}{2}(\frac{1}{6}-\lambda)[(\Delta r)^2 D_r^2 + \frac{1}{12}(\Delta r)^4 D_r^4])$$

$$(1+(\Delta t)D_r^2 + \tfrac{1}{2}(\Delta t)^2 D_r^4)U_{i,j} + T_D .$$

Hence the local truncation error is given by,

$$T_D = \{\tfrac{1}{2}(\frac{1}{6}+\lambda)(\Delta r)^2 - [\tfrac{1}{2}(\frac{1}{6}-\lambda)+\Delta t]\}D_r^2(U_{i,j})$$

$$+\{\frac{1}{12}(\Delta r)^4(\tfrac{1}{2}(\frac{1}{6}+\lambda))-[\frac{1}{12}(\Delta r)^4(\tfrac{1}{2}(\frac{1}{6}-\lambda))+\tfrac{1}{2}(\Delta t)^2+\tfrac{1}{2}(\frac{1}{6}-\lambda)\Delta t(\Delta r)^2]\}D_r^4(U_{i,j})$$

$$+\Delta t\{[-\frac{1}{4}(\Delta t)^2 + \frac{(\Delta r)^4}{144}]D_r^6 + \frac{\Delta t}{48}(\frac{1}{6} - \frac{\Delta t}{(\Delta r)^2})(\Delta r)^4 D_r^8+\ldots\}U_{i,j}.$$

The coefficients of the differential operators D_r^2 and D_r^4 are all equal to zero, i.e.,

for D_r^2,

$$\tfrac{1}{2}(\frac{1}{6}+\lambda)(\Delta r)^2 = \tfrac{1}{2}(\frac{1}{6}-\lambda)+\Delta t$$

and for D_r^4,

$$\frac{1}{12}(\Delta r)^4(\tfrac{1}{2}(\frac{1}{6}+\lambda)) = \frac{1}{12}(\Delta r)^4(\tfrac{1}{2}(\frac{1}{6}-\lambda))+\tfrac{1}{2}(\Delta t)^2+\tfrac{1}{2}(\frac{1}{6}-\lambda)\Delta t(\Delta r)^2.$$

Hence, the truncation error is $O([\Delta t]^2)+O([\Delta r]^4)$.

We are now in a position to construct an analogous higher order approximation to the differential equation (3.15.1). As a solution to (3.15.1), equation (3.15.5) can be written as,

$$(1+(\theta-1)(\Delta t)(\frac{\delta_r^2}{(\Delta r)^2} + \frac{\delta_{2r}}{2r_i\Delta r}))u_{i,j+1} = (1+\theta(\Delta t)(\frac{\delta_r^2}{(\Delta r)^2} + \frac{\delta_{2r}}{2r_i\Delta r}))u_{i,j}. \tag{3.15.14}$$

In particular, with $\theta=\tfrac{1}{2}$, the Crank-Nicolson formula is,

$$(1-\tfrac{1}{2}\Delta t(\frac{1}{2r_i(\Delta r)}\delta_{2r} + \frac{\delta_r^2}{(\Delta r)^2}))u_{i,j+1} = (1+\tfrac{1}{2}\Delta t(\frac{1}{2r_i(\Delta r)}\delta_{2r}+ \frac{\delta_r^2}{(\Delta r)^2}))u_{i,j} \tag{3.15.15}$$

which has a truncation error $O([\Delta t]^2)+O([\Delta r]^2)$. The corresponding form of (3.15.9) for equation (3.15.1) is

$$(1+(\theta-1)\Delta t[\alpha_2\delta_r^2+ \frac{1}{r_i}\beta_2\delta_{2r}])u_{i,j+1} = (1+\theta\Delta t[\alpha_1\delta_r^2+ \frac{1}{r_i}\beta_1\delta_{2r}])u_{i,j}. \tag{3.15.16}$$

If we set,

$$a_{1i} = \frac{\theta(\Delta t)\beta_1}{r_i},\ b_{1i} = \theta(\Delta t)\alpha_1,\ a_{2i} = \frac{(\theta-1)(\Delta t)\beta_2}{r_i}$$

and $b_{2i} = (\theta-1)(\Delta t)\alpha_2$, equation (3.15.16) becomes,

$$(1+a_{2i}\delta_{2r}+b_{2i}\delta_r^2)u_{i,j+1} = (1+a_{1i}\delta_{2r}+b_{1i}\delta_r^2)u_{i,j} \qquad (3.15.17)$$

If we put θ=1 in (3.15.3), we get,

$$U_{i,j+1} = \exp(\Delta t D_t)U_{i,j} = \exp(\Delta t(\frac{1}{r_i} D_r+D_r^2))U_{i,j} , \qquad (3.15.18)$$

$$= (1+ \frac{\Delta t}{r_i} D_r +(\Delta t+ \frac{(\Delta t)^2}{2r_i^2})D_r^2+ \frac{(\Delta t)^2}{r_i} D_r^3+\tfrac{1}{2}(\Delta t)^2D_r^4)U_{i,j}. \qquad (3.15.18a)$$

We also have,

$$\delta_{2r} = E_{2r}^{\frac{1}{2}} - E_{2r}^{-\frac{1}{2}}$$

$$= e^{\Delta rD_r} - e^{-\Delta rD_r}$$

i.e. $$\delta_{2r} = 2(\Delta r)D_r+\frac{1}{3}(\Delta r)^3D_r^3 + \frac{1}{120} (\Delta r)^5D_r^5 + \ldots \qquad (3.15.19)$$

The substitution of the truncated series of (3.15.12), (3.15.18a) and (3.15.19) into (3.15.17) leads to,

$$[1+2a_{1i}(\Delta r)D_r+b_{1i}(\Delta r)^2D_r^2+ \frac{1}{3}a_{1i}(\Delta r)^3D_r^3+ \frac{1}{12}b_{1i}(\Delta r)^4D_r^4]U_{i,j}$$

$$= [(1+2a_{2i}(\Delta r)D_r+b_{2i}(\Delta r)^2D_r^2+ \frac{1}{3}a_{2i}(\Delta r)^3D_r^3 + \frac{1}{12}b_{2i}(\Delta r)^4D_r^4)$$

$$\times (1+ \frac{\Delta t}{\Delta r} D_r +(\Delta t+ \frac{(\Delta t)^2}{2r_i^2})D_r^2+ \frac{(\Delta t)^2}{r_i} D_r^3 + \frac{1}{2}(\Delta t)^2D_r^4)]U_{i,j}+T, \qquad (3.15.20)$$

where T is the truncation error. By multiplying the terms of the right hand side in (3.15.20), we obtain,

$$[1+2a_{1i}(\Delta r)D_r+b_{1i}(\Delta r)^2D_r^2+ \frac{1}{3}a_{1i}(\Delta r)^3D_r^3 + \frac{1}{12}b_{1i}(\Delta r)^4D_r^4]U_{i,j}$$

$$= [1+(2a_{2i}\Delta r + \frac{\Delta t}{r_i})D_r+(b_{2i}(\Delta r)^2+\Delta t+ \frac{(\Delta t)^2}{2(\Delta r)^2} + 2a_{2i} \frac{(\Delta t)(\Delta r)}{r_i})D_r^2$$

$$+ (\frac{1}{3}a_{2i}(\Delta r)^3 + \frac{(\Delta t)^2}{r_i} + 2a_{2i}\Delta r(\Delta t + \frac{(\Delta t)^2}{2r_i^2}) + \frac{\Delta t b_{2i}(\Delta r)^2}{r_i})D_r^3$$

$$+(\frac{1}{12}b_{2i}(\Delta r)^4+2a_{2i}\frac{(\Delta r)(\Delta t)^2}{r_i} + \frac{1}{3}a_{2i}\ \frac{(\Delta r)^3\Delta t}{r_i} + b_{2i}(\Delta r)^2(\Delta t+\frac{(\Delta t)^2}{2(\Delta r)^2})$$

$$+\frac{1}{2}(\Delta t)^2)D_r^4 +(\frac{b_{2i}}{12r_i}(\Delta r)^4\Delta t +a_{2i}\Delta r(\Delta t)^2+ \frac{1}{3}a_{2i}(\Delta r)^3(\Delta t+\frac{(\Delta t)^2}{2r_i})+$$

$$+\frac{b_2(\Delta r)^2(\Delta t)^2}{r_i})D_r^5+(\frac{1}{2}b_{2i}(\Delta r)^2(\Delta t)^2+ \frac{1}{3r_i}\ a_{2i}(\Delta r)^3(\Delta t)^2+ \frac{1}{12}b_{2i}(\Delta r)^4$$

$$(\Delta t+ \frac{(\Delta t)^2}{2r_i^2}))D_r^6+(\frac{1}{6}a_{2i}(\Delta r)^3(\Delta t)^2+ \frac{1}{12r_i}\ b_{2i}(\Delta r)^4(\Delta t)^2)D_r^7 +$$

$$\frac{1}{24}b_{2i}(\Delta r)^4(\Delta t)^2D_r^8]U_{i,j}+T. \qquad (3.15.21)$$

By equating the coefficients of the differential operators of each side, we find that,

for D_r,

$$2a_{1i}\Delta r = 2a_{2i}\Delta r + \frac{\Delta t}{r_i}\ , \qquad (3.15.22)$$

for D_r^2,

$$b_{1i}(\Delta r)^2 = b_{2i}(\Delta r)^2 +\Delta t+ \frac{(\Delta t)^2}{2r_i^2} + 2a_{2i}\ \frac{(\Delta t)(\Delta r)}{r_i}\ , \qquad (3.15.23)$$

for D_r^3,

$$\frac{1}{3}a_{1i}(\Delta r)^3 = \frac{1}{3}a_{2i}(\Delta r)^3+\frac{(\Delta t)^2}{r_i} +2a_{2i}(\Delta r)(\Delta t + \frac{(\Delta t)^2}{2r_i^2})+ \frac{\Delta t b_{2i}(\Delta r)^2}{r_i}$$

$$(3.15.24)$$

and for D_r^4,

$$\frac{1}{12}b_{1i}(\Delta r)^4 = \frac{1}{12}b_{2i}(\Delta r)^4+ \frac{2a_{2i}(\Delta r)(\Delta t)^2}{r_i} + \frac{1}{3}a_{2i}\ \frac{(\Delta r)^3(\Delta t)}{r_i} +$$

$$+b_{2i}(\Delta r)^2(\Delta t+ \frac{(\Delta t)^2}{2r_i^2})+ \frac{1}{2}(\Delta t)^2. \qquad (3.15.25)$$

From equation (3.15.22), we introduce,

$$a_{1i} = F + \frac{\Delta t}{4(\Delta r)r_i}\ ,\quad a_{2i} = F - \frac{\Delta t}{4(\Delta r)r_i} \qquad (3.15.26)$$

and from (3.15.23),

$$b_{1i} = B + \frac{\Delta t}{2(\Delta r)^2} + F\,\frac{\Delta t}{(\Delta r)r_i},\ b_{2i} = B - \frac{\Delta t}{2(\Delta r)^2} - F\,\frac{\Delta t}{(\Delta r)r_i}\,, \tag{3.15.26a}$$

for some F and B. If we substitute (3.15.26) and (3.15.26a) into (3.15.24) and neglect terms of order $O([\Delta t]^3)$, we find that,

$$\frac{1}{2}(\frac{1}{6} - B)\frac{\Delta r}{r_i} = F\ . \tag{3.15.27}$$

A further substitution of (3.15.27) into (3.15.25) results in

$$B = \frac{1}{12} \text{ and } F = \frac{\Delta r}{24r_i}\ , \tag{3.15.28}$$

where, here, we have ignored terms of order $O([\Delta t]^3+(\Delta t)[\Delta r]^4)$. Thus, since $r_i = i\Delta r$, we obtain,

$$a_{1i} = \frac{1}{4i}(\frac{1}{6} + \lambda),\qquad a_{2i} = \frac{1}{4i}(\frac{1}{6} - \lambda)\ , \tag{3.15.29}$$

$$b_{1i} = \frac{1}{2}(\frac{1}{6} + (1 + \frac{1}{12i^2})\lambda)\ ,\quad b_{2i} = \frac{1}{2}(\frac{1}{6} - (1 + \frac{1}{12i^2})\lambda)\ . \tag{3.15.29a}$$

With these values our truncation error in (3.15.21) will then be,

$$T = \{\frac{1}{72}(\Delta r)^4(1+\frac{1}{2r_i}) - \frac{3}{4r_i}(\Delta t)^2 + \frac{(\Delta t)(\Delta r)^4}{144r_i}(1 - \frac{1}{2r_i^2}) - \frac{(\Delta t)^2(\Delta r)^2}{24r_i}(\frac{1}{r_i^2} + 1)\}D_r^5(U_{i,j}) + \ldots$$

Hence, $T=O([\Delta t]^2)+O([\Delta r]^4)$.

Polak (1974) also showed that the aforementioned scheme (3.15.17) with the values in (3.15.29) and (3.15.29a) is always stable. The *amplification factor* is given by,

$$|\gamma| = \frac{(1+2b_{1i}(x-1))^2+4a_{1i}^2(1-x^2)}{(1+2b_{2i}(x-1))^2+4a_{2i}^2(1-x^2)}\ ,\quad |x|\leq 1. \tag{3.15.30}$$

To prove that $|\gamma|\leq 1$ for all positive integer i and stability ratio λ, we write $a_{2i}=a_{1i}-(\frac{1}{2i})\lambda$ and $b_{2i}=b_{1i}-(1+(\frac{1}{12i^2}))\lambda$. Then,

$$|\gamma| = \frac{(1+2b_{1i}(x-1))^2+4a_{1i}^2(1-x^2)}{(1+2b_{1i}(x-1))^2+4a_{1i}^2(1-x^2)+h(x)} \qquad (3.15.31)$$

where $h(x) = -4(x-1)\lambda f(x)$ (3.15.32)

and

$$f(x) = (1+2b_{1i}(x-1))(1+\frac{1}{12i^2})-(1+\frac{1}{12i^2})^2(x-1)\lambda-\frac{a_{1i}}{i}(1+x)+\frac{1}{4i^2}(1+x)\lambda.$$

(3.15.33)

It is clear that $f(x)$ is *a linear function* of x. Hence, for $|x|\leq 1$ or $-1\leq x\leq 1$, the values of $f(x)$ must lie between $f(1)$ and $f(-1)$ because of the linearity. Furthermore, for the same values of x (i.e. for $-1\leq x\leq 1$), the factor $-4(x-1)\lambda$ in (3.15.32) is non-negative and from (3.15.33), $f(1)=1$ and $f(-1)=\frac{2}{3}(1+\frac{i^2}{12})$ which implies that $h(x)$ is *non-negative* for $|x|\leq 1$. Therefore, from (3.15.31), $|\gamma|\leq 1$ proving that the scheme is *unconditionally stable.*

It is observed that when the cylindrical equation (3.15.1) is approximated by its finite difference analogues as we did in the previous sections, the truncation error contains D_rU and D_r^2U. This can be seen by obtaining the time derivatives D_t^sU from (3.15.1) in the form,

$$\begin{aligned}
D_tU &= \frac{1}{r}D_rU+D_r^2U\\
D_t^2U &= \frac{1}{r^3}D_rU - \frac{1}{r^2}D_r^2U + \frac{2}{r}D_r^3U + D_r^4U\\
D_t^3U &= \frac{9}{r^5}D_rU - \frac{9}{r^4}D_r^2U + \frac{4}{r^3}D_r^3U - \frac{3}{r^2}D_r^4U + \frac{3}{r}D_r^5U+D_r^6U\\
&\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots
\end{aligned} \qquad (3.15.34)$$

and noticing that the terms D_rU and D_r^2U occur in all *the time derivatives.*

Thus, no matter how accurately D_tU is replaced by finite differences, the error always contains *the low order derivatives of U with respect to r.*

To avoid the occurrence of this type of derivatives, Mitchell and Pearce (1963) suggested the following transformation,

$$r = 2\sqrt{x} \ , \tag{3.15.35}$$

which converts (3.15.1) to

$$D_tU = D_xU + xD_x^2U \ , \ 0 \leq x \leq 1/4 \ , \tag{3.15.36}$$

The higher time derivatives of U are given by,

$$\begin{aligned} D_t^2U &= 2D_x^2U + 4xD_x^3U + x^2D_x^4U \\ D_t^3U &= 6D_x^3U + 18xD_x^4U + 9x^2D_x^5U + x^3D_x^6U \\ \vdots \quad & \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \end{aligned} \tag{3.15.37}$$

and we see that these expressions no longer contain the low order derivatives of u with respect to x.

It is also of significance to note that when finite differences are employed to solve (3.15.1) on *a rectangular net equally spaced* in the r-direction, the solution is less accurate in the vicinity of the cylinder axis than in the remainder of the field. This is presumably due to the term $\frac{1}{r}D_rU$ in (3.15.1), which for adequate representation by finite differences, requires points closer together in the r-direction when r is small. As we have seen, on the axis where r=0, the term becomes D_r^2U. This difficulty in the neighbourhood of the axis may be avoided to some extent by considering a rectangular net which is *uniformly spaced* in the t-direction given by $t=j\Delta t$ $(j=1,2,\ldots)$ and *unequally spaced* in the x-direction indicated by $x=i^2(\Delta x)$ $(i=1,2,\ldots)$. The latter is consistent with equal spacing in the original r-direction. Mitchell and Pearce (1963) proposed the following procedure to derive four and five point explicit and implicit replacements of (3.15.36).

(a) *Explicit Formulae*

We consider initially the points P,Q,R and S in Figure 3.15.1, where RP=Δt, QR=$(2i-1)\Delta x$ and RS=$(2i+1)\Delta x$.

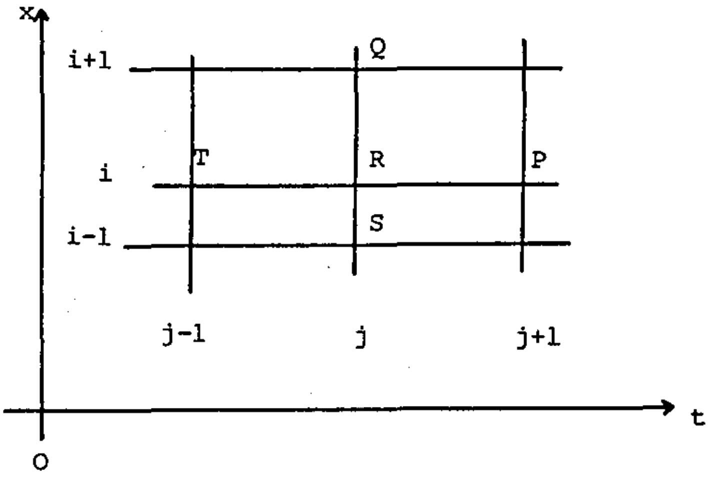


FIGURE 3.15.1

If equations (3.15.36) and (3.15.37) are used to eliminate the time derivatives of U(x,t), the Taylor's expansions of U at P,Q and S are given by,

$$U_{i,j+1} = U_{i,j}+\lambda_T A+\lambda_T(i^2+\lambda_T)B+\lambda_T^2(2i^2+\lambda_T)C + \ldots$$

$$U_{i+1,j} = U_{i,j}+(2i+1)A+\tfrac{1}{2}(2i+1)^2B + \frac{1}{6}(2i+1)^3C + \ldots$$

$$U_{i-1,j} = U_{i,j}-(2i-1)A+\tfrac{1}{2}(2i-1)^2B - \frac{1}{6}(2i-1)^3C + \ldots$$

where $U_{i,j}$,A,B,C... are the values of U(x,t),$\Delta x D_x U$, $(\Delta x)^2 D_x^2 U$, $(\Delta x)^3 D_x^3 U$,... at the point $R(x_i,t_j)$ and $\lambda_T = \frac{\Delta t}{\Delta x}$. The linear combination of the values of U at the points P,Q,R and S which eliminates A and B results in *the four point explicit analogue*

$$u_{i,j+1} = [1 - \frac{2\lambda_T(i^2+\lambda_T-1)}{4i^2-1}]u_{i,j} + \frac{\lambda_T(2i^2+2i+2\lambda_T-1)}{4i(2i-1)} u_{i+1,j}$$

$$+ \frac{\lambda_T(2i^2-2i+2\lambda_T-1)}{4i(2i-1)} u_{i-1,j}, \quad i=1,2,\ldots \qquad (3.15.38)$$

and the principal part of the local truncation error is,

$$-\lambda_T[(2i^2+\lambda_T)(\lambda_T-\frac{2}{3}) + \frac{1}{6}]C \; . \qquad (3.15.38a)$$

The error reduces to $-\frac{C}{9}$ at all points in the field when $\lambda_T=2/3$. By the Gerschgorin's theorem, it can also be shown that the above scheme is stable for $\lambda_T \leq \sqrt{\frac{3}{2}}$. A modified formula is used to deal with points on the axis. It is

$$u_{0,j+1} = \frac{1}{4}(4-5\lambda_T+2\lambda_T^2)u_{0,j} - \frac{2}{3}\lambda_T(\lambda_T-2)u_{1,j} + \frac{\lambda_T}{12}(2\lambda_T-1)u_{2,j} \qquad (3.15.39)$$

with a truncation error of $-\frac{\lambda_T}{3}(3\lambda_T^2-5\lambda_T+2)C$. This is derived by expanding $U_{0,j+1}, U_{1,j}, U_{2,j}$ in terms of U and its derivatives at the point (0,j) and eliminating the time derivatives of U by using (3.15.36) and (3.15.37).

If greater accuracy is required, a fifth point T, where TR=Δt can be introduced. By proceeding as before, *the five point, three-level explicit formula* is

$$u_{i,j+1} = \frac{\lambda_T^2}{i} \frac{[12i^4+12i^3-2i^2-4i+1-6\lambda_T^2]}{(2i+1)X} u_{i+1,j}$$

$$+ \frac{\lambda_T^2}{i} \frac{[12i^4-12i^3-2i^2+4i+1-6\lambda_T^2]}{(2i-1)X} u_{i-1,j}$$

$$+ \frac{[16(4-3\lambda_T^2)i^4-8(2\lambda_T^2+3)i^2+2+24\lambda_T^4]}{(4i^2-1)X} u_{i,j}$$

$$+ \frac{[4(3\lambda_T-2)i^2+(1-4\lambda_T+6\lambda_T^2)]}{X} u_{i,j-1}, \quad i=1,2,\ldots \qquad (3.15.40)$$

where $$X = 4(2+3\lambda_T)i^2 - (1+4\lambda_T+6\lambda_T^2) . \tag{3.15.41}$$

The coefficients sum to unity and are positive for all values of i if $\frac{2}{3} \leq \lambda_T \leq \sqrt{\frac{1}{2}}$. Therefore, by Gerschgorin's theorem, this also constitutes the stability condition. The principal part of the truncation error in the scheme (3.15.40) is

$$\frac{\lambda_T^2}{X}[(\frac{53}{3} -66\lambda_T^2)i^4-(\frac{10}{3} -4\lambda_T^2)i^2-(\frac{1}{6} -\lambda_T^2+12\lambda_T^4)]D,$$

where,

$$D = (\Delta x)^4(D_r^4U)_{i,j} , \tag{3.15.42}$$

and X is as given in (3.15.41). The modified formula used to deal with points on the axis is

$$\begin{aligned} u_{0,j+1} &= \frac{-(2-5\lambda_T+3\lambda_T^2)}{(2+5\lambda_T+3\lambda_T^2)} u_{0,j-1} + \frac{(8-9\lambda_T^2+6\lambda_T^4)}{2(2+5\lambda_T+3\lambda_T^2)} u_{0,j} \\ &+ \frac{4\lambda_T^2(8-3\lambda_T^2)}{3(2+5\lambda_T+3\lambda_T^2)} u_{1,j} - \frac{\lambda_T^2(1-6\lambda_T^2)}{6(2+5\lambda_T+3\lambda_T^2)} u_{2,j} \end{aligned} \tag{3.15.43}$$

with a truncation error of

$$\frac{\lambda_T^2}{18(2+5\lambda_T+3\lambda_T^2)}(24-117\lambda_T^2+108\lambda_T^4)D .$$

This is derived in a similar manner to (3.15.39), with the incorporation of the additional mesh point (0,j-1).

(b) *Implicit formulae*

From Figure 3.15.1, the optimum *four point implicit formula* involving the values of u at the points Q,R,S, and T is

$$\begin{aligned} [1+ \frac{2\lambda_T(i^2-\lambda_T-1)}{4i^2-1}]u_{i,j} &+ \frac{\lambda_T(-2i^2-2i+2\lambda_T+1)}{4i(2i+1)} u_{i+1,j} + \\ &+ \frac{\lambda_T(-2i^2+2i+2\lambda_T+1)}{4i(2i-1)} u_{i-1,j} = u_{i,j-1} . \end{aligned} \tag{3.15.44}$$

This is obtained in a similar manner to (3.15.38) and the principal part of the truncation error is,

$$\lambda_T[(2i^2-\lambda_T)(\lambda_T+\frac{2}{3}) - \frac{1}{6}]C \ . \qquad (3.15.44a)$$

The values of u at points on the axis are obtained from (3.15.39) with the index (j-1) replacing j. The scheme (3.15.44) generates a tridiagonal system of equations and is stable if *the eigenvalue of minimum modulus* of the matrix,

$$\begin{bmatrix} (1-\frac{2\lambda_T^2}{3}) & \frac{\lambda_T(2\lambda_T-3)}{12} & & & \\ \frac{\lambda_T(2\lambda_T-3)}{24} & 1+\frac{2\lambda_T(3-\lambda_T)}{15} & \frac{\lambda_T(2\lambda_T-11)}{40} & & \\ & \ddots & \ddots & \ddots & \\ & \frac{\lambda_T(-2i^2+2i+2\lambda_T+1)}{4i(2i-1)} & 1+\frac{2\lambda_T(i^2-\lambda_T-1)}{4i^2-1} & \frac{\lambda_T(-2i^2-2i+2\lambda_T+1)}{4i(2i+1)} & \\ & & & \frac{\lambda_T(-2m^2+2m+1+2\lambda_T)}{4m(2m-1)} & 1+\frac{2\lambda_T(m^2-1-\lambda_T)}{(4m^2-1)} \end{bmatrix}$$

exceeds unity, where $m^2 = \frac{1}{4\Delta x}$ (from equations (3.13.18), (3.13.19a) and (3.15.35)). The above implicit scheme has the same order of accuracy as the four point explicit formula (3.15.38) but has the advantage that if the tridiagonal system is solved by means of the simple Thomas algorithm, it is always stable for all positive values of λ_T.

3.16 MULTI-DIMENSIONAL SPHERICALLY SYMMETRIC DIFFUSION EQUATION

Eisen (1966) extended the special geometrical problem of (3.13.18/19) to the *n-dimensional spherically symmetric* diffusion equation,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial r^2} + \frac{(n-1)}{r}\frac{\partial U}{\partial r}, \quad 0 \leq t \leq T, \quad 0 \leq r \leq R, \tag{3.16.1}$$

subject to the initial and boundary conditions of the form,

$$U(r,0) = U_0, \ \frac{\partial U}{\partial r}(0,t) = 0, \ \beta_1 U(R,t) + \beta_2 \frac{\partial U}{\partial r}(R,t) = 0 , \tag{3.16.1a}$$

where the domain is an n-sphere of radius R and β_1, β_2 are constants. By using a similar explicit approach to derive (3.13.24) but with α replaced by (n-1), the corresponding formula to solve (3.16.1) is

$$u_{i,j+1} = u_{i,j} + \lambda[u_{i+1,j} - 2u_{ij} + u_{i-1,j}] + \frac{\lambda(n-1)}{2i}[u_{i+1,j} - u_{i-1,j}] . \tag{3.16.2}$$

This time, however, we avoid having to approximate U at the axis by defining $u_{-\frac{1}{2},j} = u_{\frac{1}{2},j}$, setting $i = \frac{1}{2}, \frac{3}{2}, \ldots, m + \frac{1}{2}$ and taking $u_{m+\frac{1}{2},j} = 0$ as the boundary condition. When written in matrix form, equation (3.16.2) takes the form

$$\underline{u}_{j+1} = \Gamma \underline{u}_j \tag{3.16.3}$$

where $\underline{u}_j = (u_{\frac{1}{2},j}, u_{\frac{3}{2},j}, \ldots, u_{m-\frac{1}{2},j})^T$

and Γ, the amplification matrix of order (m×m) is given by,

$$\Gamma = \begin{bmatrix} (1-n\lambda) & n\lambda & & & & \\ (1-\frac{n-1}{3})\lambda & (1-2\lambda) & (1+\frac{n-1}{3})\lambda & & & \bigcirc \\ & (1-\frac{n-1}{5})\lambda & (1-2\lambda) & (1+\frac{n-1}{5})\lambda & & \\ & & \ddots & \ddots & \ddots & \\ & \bigcirc & & \ddots & \ddots & \ddots \\ & & & & (1-\frac{n-1}{2m-1})\lambda & (1-2\lambda) \end{bmatrix}_{(m\times m)} \tag{3.16.4}$$

We note immediately that, for

$$2\leq n\leq 4 \text{ and } \lambda\leq\frac{1}{n}\ , \tag{3.16.5}$$

in the L_∞ norm, $||\Gamma||_\infty = \max\limits_{1\leq j\leq m} \sum\limits_{k=1}^{m} |\Gamma_{jk}| = 1$

which implies stability of the scheme for the indicated range of λ. Eisen (1966) was also able to derive the stability condition of equation (3.16.2) for the case n even and n>2. By assuming that Γ of order (m×m) can be *partitioned* to

$$\Gamma = \left[\begin{array}{c|c} \Gamma_{11} & \Gamma_{12} \\ \hline 0 & 22 \end{array}\right]_{(m\times m)}$$

then,

$$\left[\begin{array}{c} \underline{u}_{1,j+1} \\ \hline \underline{u}_{2,j+1} \end{array}\right] = \left[\begin{array}{c|c} \Gamma_{11} & \Gamma_{12} \\ \hline 0 & \Gamma_{22} \end{array}\right] \left[\begin{array}{c} \underline{u}_{1,j} \\ \hline \underline{u}_{2,j} \end{array}\right] \qquad \text{i.e.,}$$

$$\underline{u}_{1,j+1} = \Gamma_{11}\underline{u}_{1,j} + \Gamma_{12}\underline{u}_{2,j} \tag{3.16.6}$$

and

$$\underline{u}_{2,j+1} = \Gamma_{22}\underline{u}_{2,j}\ . \tag{3.16.7}$$

where Γ_{11} is a p×p submatrix, Γ_{12} is a submatrix of order p×(m-p) and Γ_{22} a submatrix of order (m-p)×(m-p) and

$$\underline{u}_{1,j} = (u_{\frac{1}{2},j}, u_{\frac{3}{2},j}, \ldots, u_{\frac{2p-1}{2},j})^T \text{ and } \underline{u}_{2,j} = (u_{\frac{2p+1}{2},j}, u_{\frac{2p+3}{2},j}, \ldots, u_{m-\frac{1}{2},j})^T\ .$$

Eisen then established (the reader is referred to his paper for details) that with respect to the L_∞ norm, the difference equation (3.16.3) is stable for some range of λ if the following conditions are satisfied:

(i) Equation (3.16.7) is stable for a certain range of λ,

(ii) $\{\Gamma_{12}\}$ is uniformly bounded in the corresponding norm (3.16.8)

and

(iii) the spectral radius, $\rho(\Gamma_{11})<1$.

We shall now attempt to derive the condition of stability for equation (3.16.3) for the case $n>2$. Since n is even, we set $n=2(p+1)$ and note that the subdiagonal element in the (p+1)st row of (3.16.4) must vanish (we actually have $(1-\frac{n-1}{2p+1})\lambda$ and with $n=2(p+1)$, this reduces to zero) and that the superdiagonal element in the pth row is $(1+\frac{2p+1}{2p-1})\lambda$ which, in fact, is the only entry of Γ_{12}. Hence $||\Gamma_{12}||_\infty=(1+\frac{2p+1}{2p-1})\lambda$, which readily satisfies condition (ii) of (3.16.8). We also have,

$$\Gamma_{22} = \begin{bmatrix} (1-2\lambda) & 2\lambda & & & & \\ \hat{p}_1 & (1-2\lambda) & \hat{q}_1 & & O & \\ & \hat{p}_2 & (1-2\lambda) & \hat{q}_2 & & \\ & & \ddots & \ddots & \ddots & \\ & O & & & & \\ & & & & \hat{p}_{m-p-1} & (1-2\lambda) \end{bmatrix}_{(m-p)\times(m-p)} , \qquad (3.16.9)$$

where $\hat{p}_i = [1-\frac{(2p+1)}{2(p+i)+1}]\lambda$ and $\hat{q}_i = [1+\frac{(2p+1)}{2(p+i)+1}]\lambda$, $i=1,2,\ldots,(m-p-1)$.

Now,

$$||\Gamma_{22}||_\infty = \max_{1\leq j\leq m} \sum_{k=1}^{m} |(\Gamma_{22})_{jk}| = 1$$

if,

$$\lambda \leq \tfrac{1}{2} . \qquad (3.16.10)$$

This implies that the difference formula (3.16.7) is stable for $\lambda\leq\frac{1}{2}$ thus fulfilling condition (i). Finally, the matrix Γ_{11} is of the form,

$$\Gamma_{11} = \begin{bmatrix} 1-2(p+1)\lambda & 2(p+1)\lambda & & & \\ \tilde{p}_1 & (1-2\lambda) & \tilde{q}_1 & & \\ & \tilde{p}_2 & (1-2\lambda) & \tilde{q}_2 & \bigcirc \\ & \bigcirc & \ddots & \ddots & \ddots \\ & & & \tilde{p}_{p-1} & (1-2\lambda) \end{bmatrix}_{(p\times p)} \tag{3.16.11}$$

where $\tilde{p}_i=(1-\frac{(2p+1)}{(2i+1)})\lambda$ and $\tilde{q}_i=(1+\frac{(2p+1)}{(2i+1)})\lambda$, $i=1,2,\ldots,p-1$.

It is found that Γ_{11} has an eigenvalue $1-4\lambda$ of *multiplicity* p.

Therefore $\rho(\Gamma_{11}) = |1-4\lambda| < 1$ if

$$\lambda < \tfrac{1}{2} \quad . \tag{3.16.12}$$

From (3.16.10) and (3.16.12), we conclude that the explicit method (3.16.2) to solve the differential equation (3.16.1) is conditionally stable when n is even and $\lambda<\frac{1}{2}$.

3.17 NON-LINEAR PARABOLIC EQUATIONS

The general case of the non-linear parabolic equation is given by

$$F(x,t,U,\frac{\partial U}{\partial t},\frac{\partial U}{\partial x},\frac{\partial^2 U}{\partial x^2}) = 0 \ , \qquad (3.17.1)$$

which is defined on $D=\{(x,t)|0<x<1,\ 0<t\leqslant T\}$ and with the conditions,

$$\begin{aligned} U(x,0) &= f(x),\ 0\leqslant x\leqslant 1 \ , \\ U(0,t) &= g(t),\ 0<t\leqslant T \\ U(1,t) &= h(t),\ 0<t\leqslant T \ . \end{aligned} \qquad (3.17.1a)$$

Equation (3.17.1) becomes $F(x,t,z,p,q,s)=0$ if we write $z=U$, $p=\frac{\partial U}{\partial t}$, $q=\frac{\partial U}{\partial x}$ and $s=\frac{\partial^2 U}{\partial x^2}$. F is assumed continuous for $(x,t) \in \bar{D}$ and the derivatives $\frac{\partial F}{\partial z}$, $\frac{\partial F}{\partial p}$, $\frac{\partial F}{\partial q}$ and $\frac{\partial F}{\partial s}$ are bounded in absolute magnitude. The following inequalities also hold,

$$\frac{\partial F}{\partial p} \geqslant a > 0 \ , \qquad (3.17.2)$$

and

$$\frac{\partial F}{\partial s} \leqslant b < 0 \ . \qquad (3.17.3)$$

We shall investigate briefly *the classical explicit difference equation* for (3.17.1). The corresponding analogue is,

$$F(x_i,t_j,u_{i,j},\frac{\Delta_t u_{i,j}}{\Delta t}\ ,\frac{\delta_{2x}u_{i,j}}{2\Delta x},\frac{\delta_x^2 u_{i,j}}{(\Delta x)^2}) = 0 \ , \qquad (3.17.4)$$

and since $\frac{\partial F}{\partial p} \neq 0$, this can be written as,

$$\frac{(u_{i,j+1}-u_{i,j})}{\Delta t} = \phi(x_i,t_j,u_{i,j},\frac{\delta_{2x}u_{i,j}}{2\Delta x}\ ,\frac{\delta_x^2 u_{i,j}}{(\Delta x)^2}) \ . \qquad (3.17.5)$$

To derive the local truncation error of (3.17.5), we have, for a sufficiently smooth function U,

$$F(x_i,t_j,U_{i,j},\frac{\partial U_{i,j}}{\partial t},\frac{\partial U_{i,j}}{\partial x},\frac{\partial^2 U_{i,j}}{\partial x^2})$$

$$= F(x_i,t_j,U_{i,j},\frac{\Delta_t U_{i,j}}{\Delta t} + \frac{\Delta t}{2}\frac{\widetilde{\partial^2 U}_{i,j}}{\partial t^2},\frac{\delta_{2x}U_{i,j}}{2\Delta x} - \frac{(\Delta x)^2}{6}\frac{\widetilde{\partial^3 U}_{i,j}}{\partial x^3},\frac{\delta_x^2 U_{i,j}}{(\Delta x)^2} - \frac{(\Delta x)^2}{12}\frac{\widetilde{\partial^4 U}_{i,j}}{\partial x^4})$$

$$= F(x_i,t_j,U_{i,j}, \frac{\Delta_t U_{i,j}}{\Delta t}, \frac{\delta_{2x} U_{i,j}}{2\Delta x}, \frac{\delta_x^2 U_{i,j}}{(\Delta x)^2}) + \widetilde{\frac{\partial F}{\partial p}}(\frac{\Delta t}{2}) \widetilde{\frac{\partial^2 U_{i,j}}{\partial t^2}} - \widetilde{\frac{\partial F}{\partial q}} \frac{(\Delta x)^2}{6} \widetilde{\frac{\partial^3 U_{i,j}}{\partial x^3}} - \widetilde{\frac{\partial F}{\partial s}} \frac{(\Delta x)^2}{12} \widetilde{\frac{\partial^4 U_{i,j}}{\partial x^4}} .$$

Hence, the local truncation error is,

$$T_{i,j} = (\frac{\Delta t}{2}) \widetilde{\frac{\partial F}{\partial p}} \widetilde{\frac{\partial^2 U_{i,j}}{\partial t^2}} - \frac{(\Delta x)^2}{6} \widetilde{\frac{\partial F}{\partial q}} \widetilde{\frac{\partial^3 U_{i,j}}{\partial x^3}} - \frac{(\Delta x)^2}{12} \widetilde{\frac{\partial F}{\partial s}} \widetilde{\frac{\partial^4 U_{i,j}}{\partial x^4}} , \tag{3.17.6}$$

or $T=O(\Delta t)+O((\Delta x)^2)$ where the tilde sign indicates that the given quantity is evaluated at an appropriate intermediate point. The condition of stability for the explicit equation (3.17.4/5) is (Saul'yev (1964)),

$$\lambda \leq \frac{\min(\frac{\partial F}{\partial p})}{2\max|\frac{\partial F}{\partial s}| + (\Delta x)^2 \max|\frac{\partial F}{\partial z}|} . \tag{3.17.7}$$

In the case of the linear parabolic equation, in which $\frac{\partial F}{\partial p} \equiv 1$, $\frac{\partial F}{\partial s} \equiv -1$ and $\frac{\partial F}{\partial z} \equiv 0$, (3.17.7) reduces to $\lambda \leq \frac{1}{2}$ which corresponds to the stability condition of the classical explicit equation (3.2.1).

There are other variants arising from the general equation (3.17.1) and extensive research has been carried out to construct implicit methods to solve them. A full discussion on some of these will be deferred to a later chapter as the author feels that it is more appropriate to treat them in conjunction with the development of a new iterative method.

3.18 MULTI-DIMENSIONAL PROBLEMS: PARABOLIC EQUATIONS IN SEVERAL SPACE VARIABLES

Finite-difference methods for parabolic equations in several space variables can be separated into two categories, namely, as extensions of the difference schemes which have been described before and as *the alternating direction implicit (A.D.I.) methods* which have no analogue in one space dimension at the present time. We will confine our attention to discuss methods in two and three space variables as they typify the general multi-dimensional problem.

Let us first consider *the two-dimensional heat flow equation*,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} \quad , \tag{3.18.1}$$

in the semi-infinite cylinder $0 \leq x \leq N_1$, $0 \leq y \leq N_2$, $t > 0$ with the initial condition,

$$U(x,y,0) = F(x,y) \quad , \tag{3.18.1a}$$

and the boundary conditions,

$$\left.\begin{aligned} U(0,y,t) &= f_1(y,t) \quad , \\ U(N_1,y,t) &= f_2(y,t) \quad , \\ U(x,0,t) &= f_3(x,t) \quad , \\ \text{and} \qquad U(x,N_2,t) &= f_4(x,t) . \end{aligned}\right. \tag{3.18.1b}$$

It is obvious that the region R of the xy-plane is a rectangle. The value of $U(x,y,t)$ at the point $P(x_i,y_j,t_k)$ in the solution domain is denoted by $U_{i,j,k}$ where $x_i = i\Delta x$, $y_j = j\Delta y$ for $0 \leq i \leq (m+1)$, $0 \leq j \leq (m_1+1)$ and $\Delta x = \frac{N_1}{(m+1)}$, $\Delta y = \frac{N_2}{(m_1+1)}$. The increment in t, Δt is chosen such that $t_k = k\Delta t$ for $k=0,1,2,\ldots$

The classical explicit equation (3.2.1) generalises to

$$u_{i,j,k+1} = u_{i,j,k} + \frac{\Delta t}{(\Delta x)^2}\delta_x^2 u_{i,j,k} + \frac{\Delta t}{(\Delta y)^2}\delta_y^2 u_{i,j,k} , \tag{3.18.2}$$

where, as before, we use,

$$\delta_x^2 u_{i,j,k} = (u_{i+1,j,k} - 2u_{i,j,k} + u_{i-1,j,k})$$

and

$$\delta_y^2 u_{i,j,k} = (u_{i,j+1,k} - 2u_{i,j,k} + u_{i,j-1,k}) .$$

To analyse the stability of (3.18.2), we employ the von Neumann criterion by assuming an error of the form,

$$\varepsilon_{i,j,k} = \xi^k \exp(i_c \beta_1 i\Delta x) \exp(i_c \beta_2 j\Delta y) \ . \tag{3.18.3}$$

The substitution of (3.18.3) into (3.18.2) and after the cancellation of common factors, leads to the amplification factor,

$$\begin{aligned} \gamma &= \xi^{k+1}/\xi^k \\ &= 1-4\left(\left(\frac{\Delta t}{[\Delta x]^2}\right)\sin^2(\beta_1 \Delta x/2) - \left(\frac{\Delta t}{[\Delta y]^2}\right)\sin^2(\beta_2 \Delta y/2)\right) . \end{aligned} \tag{3.18.4}$$

For stability, we require $-1 \leq \gamma \leq 1$. By using (3.18.4), the only useful inequality is for the left-hand side which gives,

$$-2 \leq -4((\Delta t/[\Delta x]^2)\sin^2(\beta_1 \Delta x/2) - (\Delta t/[\Delta y]^2)\sin^2(\beta_2 \Delta y/2))$$

and this leads to the condition,

$$\Delta t\left(\frac{1}{[\Delta x]^2} + \frac{1}{[\Delta y]^2}\right) \leq \tfrac{1}{2} \ . \tag{3.18.5}$$

In particular, for a square region R where $\Delta x = \Delta y = h$ and $N_1 = N_2$, this condition becomes

$$\lambda = \frac{\Delta t}{h^2} \leq 1/4 \tag{3.18.6}$$

which is even more restrictive than that for the one-dimensional problem ($\lambda \leq \frac{1}{2}$). The local truncation error can be easily derived using the Taylor's series expansion. From equation (3.18.2),

$$\begin{aligned} T_{i,j,k} &= U_{i,j,k+1} - U_{i,j,k} - \Delta t\left[\frac{\delta_x^2}{(\Delta x)^2} U_{i,j,k} + \frac{\delta_y^2}{(\Delta y)^2} U_{i,j,k}\right] \\ &= \left(\frac{\partial U}{\partial t}\right)_{i,j,k} - \left[\left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j,k} + \left(\frac{\partial^2 U}{\partial y^2}\right)_{i,j,k}\right] \end{aligned}$$

$$+ \tfrac{1}{2}(\Delta t)\left(\frac{\partial^2 U}{\partial t^2}\right)_{i,j,k} - \frac{(\Delta x)^2}{12}\left(\frac{\partial^4 U}{\partial x^4}\right)_{i,j,k} - \frac{(\Delta y)^2}{12}\left(\frac{\partial^4 U}{\partial y^4}\right)_{i,j,k} + \ldots$$

Hence,

$$T_E = O(\Delta t) + O([\Delta x]^2) + O([\Delta y]^2) \ . \qquad (3.18.7)$$

To achieve unconditional stability, we may use the following *backward-difference fully implicit formula* which is a generalisation of equation (3.4.1), i.e.,

$$\frac{(u_{i,j,k+1} - u_{i,j,k})}{\Delta t} = \frac{1}{(\Delta x)^2}\,\delta_x^2 u_{i,j,k+1} + \frac{1}{(\Delta y)^2}\,\delta_y^2 u_{i,j,k+1} \, . \qquad (3.18.8)$$

As above, the amplification factor is found to be,

$$\gamma = \frac{1}{1+4(\Delta t/[\Delta x]^2)\sin^2(\beta_1 \Delta x/2) + 4(\Delta t/[\Delta y]^2)\sin^2(\beta_2 \Delta y/2)} \ , \qquad (3.18.9)$$

which is always less than one in magnitude. The local truncation error is

$$T_I = O(\Delta t) + O([\Delta x]^2) + O([\Delta y]^2) \ . \qquad (3.18.10)$$

The Crank-Nicolson analogue of (3.18.1) is

$$\frac{(u_{i,j,k+1} - u_{i,j,k})}{\Delta t} = \tfrac{1}{2}\left[\frac{1}{(\Delta x)^2}\,\delta_x^2 u_{i,j,k+1} + \frac{\delta_y^2}{(\Delta y)^2}\, u_{i,j,k+1} + \frac{\delta_x^2}{(\Delta x)^2} u_{i,j,k} + \frac{\delta_y^2}{(\Delta y)^2} u_{i,j,k}\right] \ . \qquad (3.18.11)$$

The amplification factor works out to be,

$$\gamma = \frac{1-[2(\Delta t/[\Delta x]^2)\sin^2(\beta_1 \Delta x/2) + 2(\Delta t/[\Delta y]^2)\sin^2(\beta_2 \Delta y/2)]}{1+[2(\Delta t/[\Delta x]^2)\sin^2(\beta_1 \Delta x/2) + 2(\Delta t/[\Delta y]^2)\sin^2(\beta_2 \Delta y/2)]} \qquad (3.18.12)$$

which is clearly less than one in magnitude. Therefore, the Crank-Nicolson formula is always stable with a local truncation error given by,

$$T_{CN} = O([\Delta t]^2) + O([\Delta x]^2) + O([\Delta y]^2) \ . \qquad (3.18.13)$$

Although the implicit equations (3.18.8) and (3.18.11) are unconditionally stable, they are much more difficult to solve than their one-dimensional equivalents. The Crank-Nicolson equation, for example, requires the solution of an [mn×mn] linear system of equations to advance to level $t=t_{k+1}$ from level $t=t_k$. Each equation of (3.18.11) involves five unknowns $(u_{i-1,j,k+1}, u_{i,j-1,k}, u_{i,j,k+1}, u_{i+1,j,k+1}, u_{i,j+1,k+1})$, and hence the system is *no longer tridiagonal* and can lead to laborious computation. This drawback is countered by introducing the A.D.I. methods which are intended to simplify the solution of the algebraic equations and to preserve unconditional stability and reasonable accuracy.

Let us first consider the heat equation (3.18.1) on the unit square, with $\Delta x=\Delta y$. For arithmetic simplicity, we wish to retain the tridiagonal nature of the resulting equations of our approximation to (3.18.1). This can be achieved by allowing only *one of the spatial derivatives* to be evaluated at t_{k+1}. This leads to the difference equation,

$$\frac{1}{(\Delta x)^2}\,\delta_x^2 u_{i,j,k+1} + \frac{1}{(\Delta y)^2}\,\delta_y^2 u_{i,j,k} = \frac{(u_{i,j,k+1}-u_{i,j,k})}{\Delta t}\,. \qquad (3.18.14)$$

Unfortunately, this does not result in unlimited stability. A von Neumann analysis shows that the equation governing stability is

$$\xi^{k+1}[1+4\lambda\sin^2(\beta_1\Delta x/2)] = \xi^k[1-4\lambda\sin^2(\beta_2\Delta y/2)] \qquad (3.18.15)$$

or
$$\gamma = \frac{\xi^{k+1}}{\xi^k} = \frac{1-4\lambda\sin^2(\beta_2\Delta y/2)}{1+4\lambda\sin^2(\beta_1\Delta x/2)}\,, \qquad (3.18.16)$$

where $\lambda=\frac{\Delta t}{(\Delta x)^2}=\frac{\Delta t}{(\Delta y)^2}$. If $\lambda>\frac{1}{2}$, $\beta_1=1$ and $\beta_2=J-1$, then $|\gamma|>1$, so that the scheme (3.18.14) is unstable, at least for $\lambda>\frac{1}{2}$.

On the other hand, if $\frac{\delta_y^2}{(\Delta y)^2}u$ instead of $\frac{1}{(\Delta x)^2}\,\delta_x^2 u$ is evaluated at the advanced time level, then the scheme,

$$\frac{1}{(\Delta x)^2}\delta_x^2 u_{i,j,k} + \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k+1} = \frac{(u_{i,j,k+1}-u_{i,j,k})}{\Delta t} \tag{3.18.17}$$

would lead to a stability ratio in which the positions of β_1 and β_2 are interchanged. Again, this scheme does not possess unrestricted stability.

The A.D.I. method of Peaceman and Rachford (P.R.) (1955) amounts to taking one time step using the backward-difference approximation (3.18.14) that is implicit in only the x-direction. The governing equation for stability at this step remains the same as (3.18.16). The next step of the A.D.I. counters the bias introduced above by using (3.18.17) that is implicit in only the y-direction. This time, the governing equation for stability is,

$$\frac{\xi^{k+2}}{\xi^{k+1}} = \frac{1-4\lambda\sin^2(\beta_1\Delta x/2)}{1+4\lambda\sin^2(\beta_2\Delta x/2)} \quad . \tag{3.18.18}$$

The overall stability ratio for *the double step*, is therefore, given by,

$$\frac{\xi^{k+2}}{\xi^{k}} = \frac{1-4\lambda\sin^2(\beta_1\Delta x/2)}{1+4\lambda\sin^2(\beta_1\Delta x/2)} \cdot \frac{1-4\lambda\sin^2(\beta_2\Delta y/2)}{1+4\lambda\sin^2(\beta_2\Delta y/2)} \quad , \tag{3.18.19}$$

which is bounded in magnitude by unity for any size time step. Thus, the effect of using two possibly unstable difference equations *alternately* is to produce a stable equation. Since we are interested in the solution only after the double step, let us alter Δt to be the double step and introduce an intermediate value notation $(u_{i,j,k+\frac{1}{2}})$ for the solution at the end of one time step. Then, *the P.R. procedure* becomes,

$$\frac{1}{(\Delta x)^2}\delta_x^2 u_{i,j,k+\frac{1}{2}} + \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k} = \frac{(u_{i,j,k+\frac{1}{2}}-u_{i,j,k})}{\Delta t/2}$$

$$\frac{1}{(\Delta x)^2}\delta_x^2 u_{i,j,k+\frac{1}{2}} + \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k+1} = \frac{(u_{i,j,k+1}-u_{i,j,k+\frac{1}{2}})}{\Delta t/2} . \qquad (3.18.20)$$

To investigate the accuracy of the method, an overall equation can be obtained by eliminating $u_{i,j,k+\frac{1}{2}}$. The elimination process leads to,

$$\tfrac{1}{2}\left(\frac{1}{(\Delta x)^2}\delta_x^2 + \frac{1}{(\Delta y)^2}\delta_y^2\right)(u_{i,j,k}+u_{i,j,k+1}) = \frac{(u_{i,j,k+1}-u_{i,j,k})}{\Delta t} + \left(\frac{\Delta t}{4}\right) \frac{\delta_x^2}{(\Delta x)^2}\,\frac{\delta_y^2}{(\Delta y)^2}(u_{i,j,k+1}-u_{i,j,k}) . \qquad (3.18.21)$$

which appears to be very similar with the Crank-Nicolson equation (3.18.11) with the 'perturbation term' $-(\frac{\Delta t}{4})\,\frac{\delta_x^2}{(\Delta x)^2}\,\frac{\delta_y^2}{(\Delta y)^2}(u_{i,j,k+1}-u_{i,j,k})$. It can be shown, as before, that

$$T_{PR} = O([\Delta t]^2) + O([\Delta x]^2) + O([\Delta y]^2) . \qquad (3.18.22)$$

Although the P.R. method has second-order accuracy, it is, however, *not unconditionally stable* when extended to three dimensions. *The three-dimensional parabolic equation* is given by,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2}, \quad (x,y,z,t) \in R\times(0,T], \qquad (3.18.23)$$

with the initial condition,

$$U(x,y,z,0) = f(x,y,z), \quad (x,y,z,t) \in R \times \{0\}, \qquad (3.18.23a)$$

and the boundary conditions,

$$U(x,y,z,t) = g(x,y,z,t), \quad (x,y,z,t) \in \partial R\times(0,T] , \qquad (3.18.23b)$$

where R is the cube $0<x,y,z<1$ and ∂R its boundary. If we let i,j,k and N be the indices in the x,y,z and t-direction respectively with increments $\Delta x,\Delta y,\Delta z$ and Δt (for a cube, $0\leq i,j,k\leq(m+1)$, $N=0,1,\ldots$ and $\Delta x=\Delta y=\Delta z=\frac{1}{(m+1)}$), and if we denote $U_{i,j,k}^{N}$ be the value of $U(x_i,y_j,z_k)$ at the time level N, then *the P.R. analogue* of (3.18.23) is given by,

$$\frac{1}{(\Delta x)^2}\delta_x^2 u_{i,j,k}^{N+1/3} + \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k}^{N} + \frac{1}{(\Delta z)^2}\delta_z^2 u_{i,j,k}^{N} = \frac{(u_{i,j,k}^{N+1/3}-u_{i,j,k}^{N})}{\Delta t/3}$$

$$\frac{1}{(\Delta x)^2}\delta_x^2 u_{i,j,k}^{N+1/3} + \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k}^{N+2/3} + \frac{1}{(\Delta z)^2}\delta_z^2 u_{i,j,k}^{N} = \frac{(u_{i,j,k}^{N+2/3}-u_{i,j,k}^{N+1/3})}{\Delta t/3}$$

$$(3.18.24)$$

$$\frac{1}{(\Delta x)^2}\delta_x^2 u_{i,j,k}^{N+1/3} + \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k}^{N+2/3} + \frac{1}{(\Delta z)^2}\delta_z^2 u_{i,j,k}^{N+1} = \frac{(u_{i,j,k}^{N+1}-u_{i,j,k}^{N+2/3})}{\Delta t/3} .$$

After some extensive manipulation, a von Neumann stability analysis yields,

$$\gamma = \frac{\xi^{N+1}}{\xi^{N}} = 1 + \frac{3(Y-9)(X+Y+Z)}{(3+X)(3+Y)(3+Z)} ,$$

where,

$$X = 4(\Delta t/(\Delta x)^2)\sin^2(\beta_1 \Delta x/2) ,$$

$$Y = 4(\Delta t/(\Delta y)^2)\sin^2(\beta_2 \Delta y/2) , \qquad (3.18.24a)$$

and

$$Z = 4(\Delta t/(\Delta z)^2)\sin^2(\beta_3 \Delta z/2) .$$

If $X>6$ and $Y=Z=0$, then $\gamma<-1$. Thus, $\lambda = \frac{\Delta t}{(\Delta x)^2} > \frac{3}{2}$ implies instability.

Another unconditionally stable alternating-direction scheme that has been developed to solve the two-dimensional problem (3.18.1) results from a modification of the Crank-Nicolson method due to Douglas (1962). It is given by,

$$\tfrac{1}{2}\frac{1}{(\Delta x)^2}\delta_x^2(u_{i,j,k+\frac{1}{2}}+u_{i,j,k}) + \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k} = \frac{(u_{i,j,k+\frac{1}{2}}-u_{i,j,k})}{\Delta t}$$

and

$$\tfrac{1}{2}\frac{1}{(\Delta x)^2}\delta_x^2(u_{i,j,k+\frac{1}{2}}+u_{i,j,k}) + \tfrac{1}{2}\frac{1}{(\Delta y)^2}\delta_y^2(u_{i,j,k+1}+u_{i,j,k}) = \frac{(u_{i,j,k+1}-u_{i,j,k})}{\Delta t} \qquad (3.18.25)$$

We note that if the intermediate solution $u_{i,j,k+\frac{1}{2}}$ is eliminated, then again the overall equation (3.18.21) is satisfied for a rectangular region. Hence, the order of accuracy of *the Crank-Nicolson modified A.D.I. scheme* (3.18.25) due to Douglas (1962) is given by,

$$T_{CND} = O([\Delta t]^2)+O([\Delta x]^2)+O([\Delta y]^2) \ . \qquad (3.18.26)$$

The generalisation of the scheme to three space variables takes the form,

$$\tfrac{1}{2}\frac{1}{(\Delta x)^2}\delta_x^2(u_{i,j,k}^{N+1/3}+u_{i,j,k}^{N}) + \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k}^{N} + \frac{1}{(\Delta z)^2}\delta_z^2 u_{i,j,k}^{N} = \frac{(u_{i,j,k}^{N+1/3}-u_{i,j,k}^{N})}{\Delta t}$$

$$\tfrac{1}{2}\frac{1}{(\Delta x)^2}\delta_x^2(u_{i,j,k}^{N+1/3}+u_{i,j,k}^{N}) + \tfrac{1}{2}\frac{1}{(\Delta y)^2}\delta_y^2(u_{i,j,k}^{N+2/3}+u_{i,j,k}^{N}) + \frac{1}{(\Delta z)^2}\delta_z^2 u_{i,j,k}^{N} = \frac{(u_{i,j,k}^{N+2/3}-u_{i,j,k}^{N})}{\Delta t} \qquad (3.18.27)$$

$$\tfrac{1}{2}\frac{1}{(\Delta x)^2}\delta_x^2(u_{i,j,k}^{N+1/3}+u_{i,j,k}^{N}) + \tfrac{1}{2}\frac{1}{(\Delta y)^2}\delta_y^2(u_{i,j,k}^{N+2/3}+u_{i,j,k}^{N}) + \tfrac{1}{2}\frac{1}{(\Delta z)^2}(u_{i,j,k}^{N+1}+u_{i,j,k}^{N}) = \frac{(u_{i,j,k}^{N+1}-u_{i,j,k}^{N})}{\Delta t} \ ,$$

By referring to equation (3.18.27), if we subtract the first equation from the second and the second equation from the third separately, we obtain a more convenient form of the scheme given by,

$$\left(\frac{1}{(\Delta x)^2}\delta_x^2 - \frac{2}{\Delta t}\right)u_{i,j,k}^{N+1/3} = -\left(\frac{1}{(\Delta x)^2}\delta_x^2 + \frac{2}{(\Delta y)^2}\delta_y^2 + \frac{2}{(\Delta z)^2}\delta_z^2 + \frac{2}{\Delta t}\right)u_{i,j,k}^{N}$$

$$\left(\frac{1}{(\Delta y)^2}\delta_y^2 - \frac{2}{\Delta t}\right)u_{i,j,k}^{N+2/3} = \frac{1}{(\Delta y)^2}\delta_y^2\, u_{i,j,k}^{N} - \frac{2}{\Delta t}\, u_{i,j,k}^{N+1/3} \qquad (3.18.28)$$

$$\left(\frac{1}{(\Delta z)^2}\delta_z^2 - \frac{2}{\Delta t}\right)u_{i,j,k}^{N+1} = \frac{1}{(\Delta z)^2}\delta_z^2 u_{i,j,k}^{N} - \frac{2}{\Delta t}\, u_{i,j,k}^{N+2/3} \ .$$

When the region is a cube, the intermediate solutions $u_{i,j,k}^{N+1/3}$ and $u_{i,j,k}^{N+2/3}$ can be eliminated using the last two relations of (3.18.28). The resulting difference equation is,

$$\left(\frac{1}{(\Delta x)^2}\delta_x^2+\frac{1}{(\Delta y)^2}\delta_y^2+\frac{1}{(\Delta z)^2}\delta_z^2\right)(u_{i,j,k}^{N+1}+u_{i,j,k}^{N}) = 2\,\frac{(u_{i,j,k}^{N+1}-u_{i,j,k}^{N})}{\Delta t}$$

$$+\quad \frac{\Delta t}{2}\left(\frac{1}{(\Delta x)^2(\Delta y)^2}\delta_x^2\delta_y^2+\frac{1}{(\Delta y)^2(\Delta z)^2}\delta_y^2\delta_z^2+\frac{1}{(\Delta z)^2(\Delta x)^2}\delta_z^2\delta_x^2\right)$$

$$(u_{i,j,k}^{N+1}-u_{i,j,k}^{N})\underset{\uparrow}{-}\left(\frac{\Delta t}{2}\right)^2\frac{1}{(\Delta x)^2(\Delta y)^2(\Delta z)^2}\delta_x^2\delta_y^2\delta_z^2(u_{i,j,k}^{N+1}-u_{i,j,k}^{N}) . \tag{3.18.29}$$

As expected, this formula is a perturbation of the Crank-Nicolson difference equation in three dimensions and the truncation error is,

$$T_{CND} = O([\Delta t]^2)+O([\Delta x]^2)+O([\Delta y]^2)+O([\Delta z]^2) \ , \tag{3.18.30}$$

or with $\Delta x=\Delta y=\Delta z=h$ on the cube,

$$T_{CND} = O([\Delta t]^2)+O([h]^2) \ . \tag{3.18.30a}$$

From (3.18.24a), if we let,

$$X_1 = \tfrac{1}{2}X \ ,$$

$$Y_1 = \tfrac{1}{2}Y \ ,$$

$$Z_1 = \tfrac{1}{2}Z \ ,$$

then an application of the von Neumann analysis yields the following amplification factor,

$$\gamma = \frac{1-(X_1+Y_1+Z_1)+(X_1Y_1+Y_1Z_1+Z_1X_1)+X_1Y_1Z_1}{1+(X_1+Y_1+Z_1)+(X_1Y_1+Y_1Z_1+Z_1X_1)+X_1Y_1Z_1}$$

and thus for any $\Delta t>0, |\gamma|<1$ indicating unconditional stability.

The following method, discovered by Douglas and Rachford (1956) has also received widespread application. In the two-dimensional problem, *the Douglas-Rachford procedure* is given by,

$$\frac{1}{(\Delta x)^2}\delta_x^2 u_{i,j,k+\frac{1}{2}} + \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k} = \frac{(u_{i,j,k+\frac{1}{2}}-u_{i,j,k})}{\Delta t}$$

$$(3.18.31)$$

and

$$\frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k+\frac{1}{2}} = \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k} + \frac{(u_{i,j,k+1}-u_{i,j,k+\frac{1}{2}})}{\Delta t}$$

The elimination of the intermediate values $u_{i,j,k+\frac{1}{2}}$ results in the overall equation,

$$\frac{1}{(\Delta x)^2}\delta_x^2 u_{i,j,k+1} + \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k+1} = \frac{(u_{i,j,k+1}-u_{i,j,k})}{\Delta t} + (\Delta t)\,\frac{\delta_x^2}{(\Delta x)^2}\,\frac{\delta_y^2}{(\Delta y)^2}(u_{i,j,k+1}-u_{i,j,k}) \quad (3.18.32)$$

which, this time, is a perturbation of the backward-difference fully implicit formula (3.18.8). As expected, it can be shown by the usual Taylor's expansion in several variables that,

$$T_{DR} = O(\Delta t)+O([\Delta x]^2)+O([\Delta y]^2)\ , \quad (3.18.33)$$

or on the square region,

$$T_{DR} = O(\Delta t)+O([h]^2)\ . \quad (3.18.33a)$$

The stability condition of the scheme (on the square) can be found by employing the von Neumann criterion. The resulting amplification factor is,

$$\gamma = \frac{1+X^2Y^2}{1+X+Y+Y^2}$$

where X,Y (and Z) are given by (3.18.24a). Now, for $\Delta t>0, |\gamma|<1$ giving absolute stability of the method.

The corresponding difference equation for the three-dimensional problem is,

$$\frac{1}{(\Delta x)^2}\delta_x^2 u_{i,j,k+1/3} + \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k} + \frac{1}{(\Delta z)^2}\delta_z^2 u_{i,j,k} = \frac{(u_{i,j,k+1/3}-u_{i,j,k})}{\Delta t}$$

$$\frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k+2/3} = \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k} + \frac{(u_{i,j,k+2/3}-u_{i,j,k+1/3})}{\Delta t}$$

(3.18.34)

$$\frac{1}{(\Delta z)^2}\delta_z^2 u_{i,j,k+1} = \frac{1}{(\Delta z)^2}\delta_z^2 u_{i,j,k} + \frac{(u_{i,j,k+1}-u_{i,j,k+2/3})}{\Delta t}\ .$$

The standard von Neumann analysis shows that (on the cube),

$$\gamma = \frac{1+XY+XZ+YZ+XYZ}{1+XY+XZ+YZ+XYZ+X+Y+Z} < 1\ ,$$

giving unconditional stability. The overall equation for (3.18.34) can be obtained by eliminating $u_{i,j,k+1/3}$ and $u_{i,j,k+2/3}$. This yields,

$$\begin{aligned}
&\frac{1}{(\Delta x)^2}\delta_x^2 u_{i,j,k+1} + \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k+1} + \frac{1}{(\Delta z)^2}\delta_z^2 u_{i,j,k+1} = \\
&\frac{(u_{i,j,k+1}-u_{i,j,k})}{\Delta t} - \frac{(\Delta t)^2}{(\Delta x)^2(\Delta y)^2(\Delta z)^2}\delta_x^2\delta_y^2\delta_z^2(u_{i,j,k+1}-u_{i,j,k}) \\
&\Delta t\left(\frac{1}{(\Delta x)^2}\,\frac{1}{(\Delta y)^2}\delta_x^2\delta_y^2 + \frac{1}{(\Delta x)^2}\,\frac{1}{(\Delta z)^2}\delta_x^2\delta_z^2 + \frac{1}{(\Delta y)^2}\,\frac{1}{(\Delta z)^2}\delta_y^2\delta_z^2\right) \\
&\qquad\qquad (u_{i,j,k+1}-u_{i,j,k})
\end{aligned}$$

(3.18.35)

which is again a perturbation of the fully implicit formula in three dimensions. Like the latter, it has an error term of,

$$T_{DG} = O(\Delta t)+O([\Delta x]^2)+O([\Delta y]^2)+O([\Delta z]^2) \qquad (3.18.36)$$

or $$T_{DG} = O(\Delta t)+O([h]^2) \qquad (3.18.36a)$$

on the cube.

Mention must also be made on the following method which was introduced by Brian (1961) to solve a three-dimensional problem. The *Brian procedure* takes the form,

$$\frac{1}{(\Delta x)^2}\delta_x^2 u_{i,j,k}^{N+1/3} + \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k}^{N} + \frac{1}{(\Delta z)^2}\delta_z^2 u_{i,j,k}^{N} = \frac{(u_{i,j,k}^{N+1/3} - u_{i,j,k}^{N})}{\Delta t}$$

$$\frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k}^{N+2/3} = \frac{1}{(\Delta y)^2}\delta_y^2 u_{i,j,k}^{N} + \frac{(u_{i,j,k}^{N+2/3} - u_{i,j,k}^{N+1/3})}{\Delta t} \tag{3.18.37}$$

$$\frac{1}{(\Delta z)^2}\delta_z^2 u_{i,j,k}^{N+1} = \frac{1}{(\Delta z)^2}\delta_z^2 u_{i,j,k}^{N} + \frac{(u_{i,j,k}^{N+1} - u_{i,j,k}^{N+2/3})}{\Delta t} .$$

When the intermediate solutions are eliminated, the overall equation reduces to equation (3.18.29). Hence the local truncation error takes the form (3.18.30) and (3.18.30a) which is second-order correct in both space and time. It can also be shown in the same manner as before that the Brian procedure is always stable.

3.19 ITERATIVE PROCEDURE FOR THE A.D.I. METHODS

As we have seen above, the A.D.I. methods enable us *to solve directly* the resulting tridiagonal system of linear equations *along or parallel to each direction* of the Cartesian axes. The simple, efficient and stable *Thomas algorithm* is often applied for this purpose. We now recall from the last section that the overall formula of the Peaceman-Rachford method to solve the two-dimensional heat flow equation is given by,

$$\tfrac{1}{2}\left(\frac{\delta_x^2}{(\Delta x)^2} + \frac{\delta_y^2}{(\Delta y)^2}\right)(u_{i,j,k}+u_{i,j,k+1}) = \frac{(u_{i,j,k+1}-u_{i,j,k})}{\Delta t} + \left(\frac{\Delta t}{4}\right)\frac{\delta_x^2\delta_y^2}{(\Delta x)^2(\Delta y)^2}\,(u_{i,j,k+1}-u_{i,j,k}) . \qquad (3.19.1)$$

It is interesting to note that equation (3.19.1) can be taken to represent *an iterative procedure* which converges if

$$u_{i,j,k} = u_{i,j,k+1} = u_{i,j} \qquad (3.19.2)$$

for k sufficiently large (i.e. as the temperature (solution) reaches a steady state). The substitution of the values of (3.19.2) into (3.19.1) leads to

$$\left(\frac{1}{(\Delta x)^2}\delta_x^2 + \frac{1}{(\Delta y)^2}\delta_y^2\right)u_{i,j} = 0 , \qquad (3.19.3)$$

which is the standard *five-point* difference replacement of *the Laplace's elliptic equation*. Thus, the PR method (3.18.20) applied to a heat conduction problem *where the boundary conditions are independent of the time*, represents an iterative procedure for solving Laplace's equation on a square with Dirichlet boundary conditions. Equation (3.19.3) can be written as,

$$\{u_{i+1,j}-2u_{i,j}+u_{i-1,j}\}+\{u_{i,j+1}-2u_{i,j}+u_{i,j-1}\} = 0, \text{ for } i,j=1,2,\ldots,m \qquad (3.19.4)$$

Let the unknowns be ordered along the horizontal mesh lines and let,

$$\underline{u}_j = (u_{1j}, u_{2j}, \ldots, u_{mj})^T , \quad j=1,2,\ldots,m .$$

With the boundary conditions,

$$u_{i,j} = g_{i,j} = \begin{cases} i=0,m+1; & j=1,2,\ldots,m \\ \\ j=0,m+1; & i=1,2,\ldots,m \end{cases} \tag{3.19.5}$$

(3.19.4) can be written in the familiar *block tridiagonal form*

$$\begin{bmatrix} -D & I_m & & & \\ I_m & -D & I_m & & O \\ & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \\ O & & I_m & -D & I_m \\ & & & I_m & -D \end{bmatrix} \begin{bmatrix} \underline{u}_1 \\ \underline{u}_2 \\ \vdots \\ \vdots \\ \underline{u}_{m-1} \\ \underline{u}_m \end{bmatrix} = \underline{b} , \tag{3.19.6}$$

where the right-hand side is obtained from the data, I_m is the m×m unit matrix and $D=4I_m-(L_m+L_m^T)$ is tridiagonal with

$$L_m = \begin{bmatrix} 0 & & & & & \\ 1 & 0 & & & & \\ & 1 & 0 & & O & \\ & & 1 & 0 & & \\ & & & \ddots & \ddots & \\ & O & & & \ddots & \ddots \\ & & & & 1 & 0 \end{bmatrix}_{(m\times m)} \quad \text{of order } (m\times m) \quad .$$

Therefore, (3.19.6) can be written as,

$$A\underline{u} = \underline{b}, \quad \underline{u} = (\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_m)^T \quad , \tag{3.19.7}$$

with A=H+V, where H and V arise from the representation of the respective bracketed terms in (3.19.4). Thus, the component of the vector $H\underline{u}$ corresponding to the mesh point (x_i,y_j) is given by

$$u_{i+1,j}-2u_{i,j}+u_{i-1,j}$$

which is a central difference approximation to the *one-dimensional* operator $\frac{\partial^2 U}{\partial x^2}$ along different *horizontal lines*. Similarly, the components of $V\underline{u}$ correspond to the discretisation of $\frac{\partial^2 U}{\partial y^2}$ along different *vertical lines*. Hence, when viewed as an approximation to the operator $\frac{1}{(\Delta x)^2}\delta_x^2 + \frac{1}{(\Delta y)^2}\delta_y^2$ (cf. equations (3.19.3) and (3.19.7)), A=H+V is the *natural splitting* of A into its one-dimensional components.

If we define,

$$L = \begin{bmatrix} L_m & & & \\ & L_m & & O \\ & & \ddots & \\ O & & & L_m \end{bmatrix}_{(m^2\times m^2)} \qquad \text{and } B = \begin{bmatrix} 0 & & & & O \\ I_m & 0 & & & \\ & I_m & \ddots & & \\ O & & \ddots & \ddots & \\ & & & I_m & 0 \end{bmatrix}_{(m^2\times m^2)}$$

then $H=2I-(L+L^T)$ and $V=2I-(B+B^T)$. Various iterative schemes can now be devised based on the splitting of A.

From equation (3.19.7) we have,

$$(H+V)\underline{u} = \underline{b}\ ,$$

which is equivalent to,

$$(H+rI)\underline{u} = (rI-V)\underline{u} + \underline{b}$$

where r is any positive scalar. This suggests the simple *block iterative method*,

$$(H+rI)\underline{u}^{(p+1)} = (rI-V)\underline{u}^{(p)} + \underline{b}\ ,\quad p=0,1,2,\ldots \tag{3.19.8}$$

with $\underline{u}^{(0)}$ given. Now, H+rI is *block diagonal* where *each* block is given by $(2+r)I_m-(L_m+L_m^T)$. Hence, to obtain the new approximation $\underline{u}_j^{(p+1)}$, we only need to sweep through the mesh on line j by solving a system of m linear equations with a symmetric diagonally dominant tridiagonal matrix. Similarly, we also have the scheme,

$$(V+rI)\underline{u}^{(p+1)} = (rI-H)\underline{u}^{(p)} + \underline{b}, \quad p=0,1,2,\ldots \tag{3.19.9}$$

with $\underline{u}^{(0)}$ given. However, V+rI is *block tridiagonal* and the left-hand side of (3.19.9) takes the form (along rows),

$$(V+rI)\underline{u}^{(p+1)} = \left[\begin{array}{c} (2+r)\underline{u}_1^{(p+1)} - \underline{u}_2^{(p+1)} \\ \hdashline -\underline{u}_1^{(p+1)} + (2+r)\underline{u}_2^{(p+1)} - \underline{u}_3^{(p+1)} \\ \hdashline \vdots \\ \hdashline -\underline{u}_{m-2}^{(p+1)} + (2+r)\underline{u}_{m-1}^{(p+1)} - \underline{u}_m^{(p+1)} \\ \hdashline -\underline{u}_{m-1}^{(p+1)} + (2+r)\underline{u}_m^{(p+1)} \end{array}\right] . \tag{3.19.10}$$

The equations (3.19.9) can be simplified as follows. We write down the first component of each of (3.19.10). This yields, for the left-hand side of (3.19.9),

$$\{(2+r)I_m - (L_m+L_m^T)\}\hat{\underline{u}}_1^{(p+1)}$$

where $\hat{\underline{u}}_1 = (u_{11}, u_{12}, \ldots, u_{1m})^T$ which are the unknowns along the first vertical mesh line (column). Similarly, the jth component of each of (3.19.10) yields equations with the left-hand side,

$$\{(2+r)I_m - (L_m+L_m^T)\}\hat{\underline{u}}_j^{(p+1)} \quad .$$

Hence, a *reordering* which is column-wise in this way, produces as before m systems each with a tridiagonal matrix. The iterative process (3.19.9) is therefore easily implemented and improves the approximations one vertical line (column) at a time.

The iterative process can be extended to more general differential equations than the one (the simple Laplace equation) that we have considered above. In general, we will be solving the finite difference equations $A\underline{u}=\underline{f}$ where $A=H+V+\Sigma$ is *block tridiagonal* and Σ is *a positive diagonal matrix*. Furthermore, H and V are *symmetric and positive definite*. H is *block diagonal* where each block is a *tridiagonal symmetric and positive definite matrix*, and similarly for the *reordered* form of V. We shall assume in the subsequent discussion that the matrix is of order $(n \times n)$.

The A.D.I. iteration combines the features of the schemes (3.19.8) and (3.19.9) by writing $(H+V+\Sigma)\underline{u}=\underline{f}$ as a pair of equations,

$$(H+\tfrac{1}{2}\Sigma+rI)\underline{u} = (rI-V-\tfrac{1}{2}\Sigma)\underline{u} + \underline{f} \;,$$

and

$$(V+\tfrac{1}{2}\Sigma+rI)\underline{u} = (rI-H-\tfrac{1}{2}\Sigma)\underline{u} + \underline{f} \;, \qquad (3.19.11)$$

for any constant $r>0$. With $H_1=H+\tfrac{1}{2}\Sigma$ and $V_1=V+\tfrac{1}{2}\Sigma$, *the A.D.I. method of Peaceman and Rachford* is given by,

$$(H_1+r_{p+1}I)\underline{u}^{(p+\frac{1}{2})} = (r_{p+1}I-V_1)\underline{u}^{(p)} + \underline{f}$$

and

$$(V_1+r_{p+1}I)\underline{u}^{(p+1)} = (r_{p+1}I-H_1)\underline{u}^{(p+\frac{1}{2})} + \underline{f} \;,\; p\geq 0 \;, \qquad (3.19.12)$$

where $\underline{u}^{(0)}$ is a starting approximation and the r_p are positive constants called *acceleration parameters* whose values are chosen to maximise the rate of convergence. It is clear that the first stage of (3.19.12) corresponds to the process of iterating horizontally along rows and the second stage to iterating vertically along columns. From (3.19.12), we can write,

$$\underline{u}^{(p+1)} = M(r_{p+1})\underline{u}^{(p)} + \underline{q}(r_{p+1}) \,,\; p\geq 0 \;,$$

where the ADI *iteration matrix* (cf. Goult *et al* (1974)) is given by (setting $r_{p+1}=r$),

$$M(r) = (V_1+rI)^{-1}(rI-H_1)(H_1+rI)^{-1}(rI-V_1) \;. \qquad (3.19.13)$$

If we denote the error vector by $\underline{e}$ then $\underline{e}^{(p)}=\underline{u}^{(p)}-\underline{U}$ and $\underline{e}^{(p+1)}=M(r_{p+1})\underline{e}^{(p)}$. Hence, we have,

$$\underset{\sim}{e}^{(p)} = \left(\prod_{i=1}^{p} M(r_i) \right) \underset{\sim}{e}^{(O)} \quad , \; p \geq 1 \; . \qquad (3.19.14)$$

The *convergence* of the ADI scheme can be easily shown for *the stationary case* with constant parameters $r_i=r$. If we let $\overline{M}(r)=(V_1+rI)M(r)(V_1+rI)^{-1}$, then by similarity, $M(r)$ and $\overline{M}(r)$ have the same eigenvalues. Hence, from (3.19.13), we obtain,

$$\begin{aligned} \rho(M(r)) &= \rho(\overline{M}(r)) \leq ||\overline{M}(r)|| \\ &\leq ||(rI-H_1)(H_1+rI)^{-1}|| \; ||(rI-V_1)(V_1+rI)^{-1}|| \end{aligned}$$

where $\rho(M(r))$ is the spectral radius of $M(r)$. Since H_1 and V_1 are symmetric and positive definite, then in the L_2 norm, we find that

$$||(rI-H_1)(H_1+rI)^{-1}||_2 = \max_{1 \leq i \leq n} \left| \frac{r-\mu_i}{r+\mu_i} \right| < 1,$$

where μ_i, $1 \leq i \leq n$ are the eigenvalues (positive) of H_1. A similar argument applies to the norm involving V_1. Hence, $\rho(M(r))<1$ for all $r>0$ and therefore the P.R. iteration (3.9.12) converges.

For our approximation (3.19.3) with a uniform spacing, it is possible to determine the optimum parameter r_b such that

$$\rho(M(r_b)) \leq \rho(M(r)) \quad \text{for all } r>0 \; .$$

It turns out that $\rho(M(r_b))=\rho(\hat{M}_{\omega_b})$, the optimised spectral radius for *the point SOR iteration* (see, for example, Goult *et al* (1974)). Hence, the two schemes have an identical asymptotic rate of convergence. The A.D.I. iteration does, however, involve far more computation and it is therefore essential to vary the acceleration parameters r_p - *the non-stationary case*.

For our approximation (3.19.3), it can be shown that H_1 and V_1 *commute*, i.e. $H_1V_1=V_1H_1$. Thus, if in general, H_1 and V_1 commute, H_1 and V_1 have a common set of (orthonormal) eigenvectors. Now, let

$(\mu_\ell,\underline{v}_\ell)_{\ell=1}^{n}$ and $(\eta_\ell,\underline{v}_\ell)_{\ell=1}^{n}$ be the eigensystems of H_1 and V_1 respectively For p iterations of (3.19.12), the relation (3.19.13) yields, for $1\leq\ell\leq n$,

$$\left(\prod_{i=1}^{p} M(r_i)\right)\underline{v}_\ell = \left\{\prod_{i=1}^{p}\left(\frac{r_i-\mu_\ell}{r_i+\mu_\ell}\right)\left(\frac{r_i-\eta_\ell}{r_i+\eta_\ell}\right)\right\}\underline{v}_\ell \quad .$$

It follows that $\prod_{i=1}^{p} M(r_i)$ is symmetric and therefore > why it follows.

$$\left\|\prod_{i=1}^{p} M(r_i)\right\|_2 = \rho\left(\prod_{i=1}^{p} M(r_i)\right)$$

$$= \max_{1\leq\ell\leq n}\prod_{i=1}^{p}\left|\frac{r_i-\mu_\ell}{r_i+\mu_\ell}\right|\left|\frac{r_i-\eta_\ell}{r_i+\eta_\ell}\right| < 1 \qquad (3.19.15)$$

which establishes the convergence of the iteration. Generally, μ_ℓ and η_ℓ are unknown. However, the estimates for α and β where

$$0<\alpha \leq \mu_\ell,\eta_\ell \leq \beta$$

may be found by using variants of *the power method*. Clearly,

$$\max_{1\leq\ell\leq n}\prod_{i=1}^{p}\left|\frac{r_i-\mu_\ell}{r_i+\mu_\ell}\right|\left|\frac{r_i-\eta_\ell}{r_i+\eta_\ell}\right| \leq \max_{\ell}\prod_{i=1}^{p}\left|\frac{r_i-\mu_\ell}{r_i+\mu_\ell}\right| \max_{\ell}\prod_{i=1}^{p}\left|\frac{r_i-\eta_\ell}{r_i+\eta_\ell}\right|$$

$$\leq \left\{\max_{\alpha\leq z\leq\beta}\prod_{i=1}^{p}\left|\frac{r_i-z}{r_i+z}\right|\right\}^2$$

so that $\rho\left(\prod_{i=1}^{p} M(r_i)\right) \leq \max_{\alpha\leq z\leq\beta}|R_p(z)|^2$

where $R_p(z) \equiv \prod_{i=1}^{p}\frac{(r_i-z)}{(r_i+z)}$. Thus, the problem of minimising the spectral radius in (3.19.15) is now one of minimising the uniform norm of the rational function R. This can be done by using Chebyshev polynomials. The reader may consult Varga (1962) for a detailed discussion on this subject.

3.20 VARIANTS OF THE A.D.I. SCHEME AND THREE DIMENSIONAL PROBLEMS

Many variants of the basic Peaceman-Rachford scheme have been proposed. For example, we have, on modifying the second stage of (3.19.12),

$$(H_1+r_{p+1}I)\underline{u}^{(p+\frac{1}{2})} = (r_{p+1}I-V_1)\underline{u}^{(p)} + \underline{f} \qquad (3.20.1)$$

and $$(V_1+r_{p+1}I)\underline{u}^{(p+1)} = (V_1-(1-\omega)r_{p+1}I)\underline{u}^{(p)}+(2-\omega)r_{p+1}\underline{u}^{(p+\frac{1}{2})}$$

where ω is a parameter. For $\omega=0$, we have the Peaceman-Rachford scheme (3.19.12) and for $\omega=1$, we obtain the scheme due to Douglas and Rachford (1956). For H_1 and V_1 symmetric and positive definite and with a fixed acceleration parameter $r>0$, the resulting *generalised A.D.I. scheme* is convergent for any $0\leqslant\omega\leqslant 2$.

An important feature of the Douglas-Rachford scheme is that it generalises to equations with multi-space variables. To solve the three-dimensional Laplace equation, for example, the approximating seven-point finite-difference formula is given by,

$$(X+Y+Z)\underline{u} = \underline{f} \qquad (3.20.2)$$

where X,Y and Z are *symmetric positive definite* and which after appropriate *reordering* are *block diagonal (each diagonal block being tridiagonal)*. The Douglas-Rachford scheme takes the form,

$$\begin{aligned}(X+r_{p+1}I)\underline{u}^{(p+1/3)} &= (r_{p+1}I-X-2Y-2Z)\underline{u}^{(p)} + 2\underline{f}\\ (Y+r_{p+1}I)\underline{u}^{(p+2/3)} &= Y\underline{u}^{(p)}+r_{p+1}\underline{u}^{(p+1/3)} \qquad (3.20.3)\end{aligned}$$

and $$(Z+r_{p+1}I)\underline{u}^{(p+1)} = Z\underline{u}^{(p)}+r_{p+1}\underline{u}^{(p+2/3)}$$

and corresponds to sweeping through the mesh parallel to the three coordinate axes in turn; each stage consisting of solving tridiagonal systems. When X,Y and Z commute, the scheme is convergent for any fixed iteration parameter $r>0$ (this is equivalent to the stability of the scheme (3.18.34)).

3.21 OTHER FORMS OF APPROXIMATION FOR MULTI-DIMENSIONAL PARABOLIC EQUATIONS

Since the A.D.I. methods were first introduced by Peaceman and Rachford (1955), several other alternating direction techniques have been considered for parabolic problems. As we have seen, many splitting schemes can be constructed by "perturbing" in a suitable way an implicit formula. Douglas and Gunn (1964) extended this basic idea to develop schemes of the form

$$(I+A)\underline{u}^{N+1} = B\underline{u}^{N} , \qquad (3.21.1)$$

where $A = \sum_{\ell=1}^{q} A_\ell$ and $\{A_\ell\}_{\ell=1}^{q}$ are easily "inverted" (for example, by solving sets of tridiagonal equations). In the two dimensional case, they used

$$(I+A_1)\underline{u}_{(1)}^{N+1} = B\underline{u}^{N} - \sum_{i=2}^{q} A_i\underline{u}^{N}$$

and $$(I+A_i)\underline{u}_{(\ell)}^{N+1} = \underline{u}_{(\ell-1)}^{N+1} + A_\ell\underline{u}^{N} , \ \ell=2,3,\ldots,q; \ \underline{u}^{N+1}=\underline{u}_{(q)}^{N+1}$$

and showed that it is equivalent to (3.21.1) with a perturbed right hand side.

Mitchell and Fairweather (1964) developed generalised Peaceman-Rachford and Douglas-Rachford schemes and then derived the corresponding formulae which led to a better accuracy than the original P.R. and D.R. methods.

Several Russian mathematicians have studied some closely related techniques (to the A.D.I.) under the name of the "method of fractional steps". In particular, *the locally one-dimensional* (L.O.D.) methods have been developed by D'Yakonov, Marchuk, Samarskii and Yanenko to solve time dependent partial differential equations in two or more space variables. In the two-dimensional case, the equation

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} \tag{3.21.2}$$

is written as the pair,

$$\tfrac{1}{2}\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial y^2} \tag{3.21.2a}$$

and

$$\tfrac{1}{2}\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} \; . \tag{3.21.2b}$$

In advancing the solution from time $t=k\Delta t$ to $t=(k+1)\Delta t$, it is assumed that equation (3.21.2a) holds from $t=k\Delta t$ to $t=(k+\frac{1}{2})\Delta t$ and equation (3.21.2b) holds from $t=(k+\frac{1}{2})\Delta t$ to $t=(k+1)\Delta t$. The derivatives in (3.21.2a/2b) are discretised accordingly. For example, the Crank-Nicolson discretisations on the square region give the following L.O.D. split schemes,

$$(1-\tfrac{1}{2}\lambda\delta_y^2)u_{i,j,k+\frac{1}{2}} = (1+\tfrac{1}{2}\lambda\delta_y^2)u_{i,j,k}$$

and

$$(1-\tfrac{1}{2}\lambda\delta_x^2)u_{i,j,k+1} = (1+\tfrac{1}{2}\lambda\delta_x^2)u_{i,j,k+\frac{1}{2}} \; .$$

Finally, we briefly mention *the hopscotch method* which attempts to reduce the implicitness of difference schemes whilst maintaining their order of accuracy and stability. The θ-weighted scheme to approximate (3.18.1) is given by,

$$u_{i,j,k+1}-u_{i,j,k} = \lambda\{\theta(\delta_x^2+\delta_y^2)u_{i,j,k+1}+(1-\theta)(\delta_x^2+\delta_y^2)u_{i,j,k}\} \; .$$

The values of $\theta=0,\frac{1}{2}$ and 1 give the classical explicit, the Crank-Nicolson and the fully implicit scheme respectively. Hopscotch generalises the role of θ by making θ a function of space and time, i.e. $\theta=\theta_{i,j,k}$. To preserve accuracy and stability, conditions must be placed on the values of $\theta_{i,j,k}$. The following formula,

$$u_{i,j,k+1}-u_{i,j,k} = \lambda\{\theta_{i,j,k+1}(\delta_x^2+\delta_y^2)u_{i,j,k+1}+\theta_{i,j,k}(\delta_x^2+\delta_y^2)u_{i,j,k}\}$$

where

$$\theta_{i,j,k} = \begin{cases} 1 \text{ if } i+j+k \text{ is an even integer} \\ 0 \text{ otherwise} \end{cases}$$

was introduced by Gourlay (1970) as *the odd-even hopscotch* method. Another choice of θ, known as *the line hopscotch*, is given by,

$$\theta_{i,j,k} = \begin{cases} 1 \text{ if } i+k \text{ is even} \\ 0 \text{ otherwise.} \end{cases}$$

In addition, we also have *the peripheral hopscotch* which can be carried out on either a circular or a rectangular grid. The reader who wishes to have further details on hopscotch methods is recommended to refer to Gourlay (1970), Gourlay and McGuire (1971), Gane (1974), Gane and Gourlay (1977), Gourlay and McKee (1977) and Greig and Morris (1976).

CHAPTER FOUR

THE GROUP EXPLICIT METHODS FOR HYPERBOLIC EQUATIONS

4.1 INTRODUCTION

In this chapter, *the group explicit (GE) methods* which were first introduced by Evans and Abdullah (1983) to solve parabolic problems will be extended to hyperbolic equations of first and second order. The development of these methods stems from the general observation that the *alternate* use of different algorithms with truncation errors of opposite signs can lead to the *cancellations* of the error terms at most points on the mesh lines. Although this alternating strategy does not necessarily amount to the upgrading of the order of the approximation, it is, however, expected to provide a better accuracy than the individual algorithms themselves as well as other schemes that are traditionally employed to solve the given differential equation.

The GE techniques involve the utilisation of *asymmetric* approximations which when coupled in groups of two adjacent points on the mesh result in *implicit* equations. These equations will then be converted to *explicit* ones which produce the numerical solutions of the differential equation thus exhibiting the simple nature of the methods.

This chapter will deal with the construction of two different schemes in which the GE procedure is used to solve the simple hyperbolic equation of first order of the form,

$$\frac{\partial U}{\partial t} = -\frac{\partial U}{\partial x}\ , \quad 0\leq x\leq 1,\ t\geq 0\ . \tag{4.1.1}$$

The application of the procedure will then be extended to the wave equation,

$$\frac{\partial^2 U}{\partial t^2} = \frac{\partial^2 U}{\partial x^2}\ , \quad 0\leq x\leq 1,\ t\geq 0\ . \tag{4.1.2}$$

An analysis of the local truncation errors will also be performed followed by an investigation of the stability requirements of the various schemes.

4.2 GE METHODS FOR THE GENERALISED WEIGHTED APPROXIMATION TO THE FIRST ORDER EQUATION

From equation (2.9.8) of Chapter 2, the generalised weighted finite-difference analogue for (4.1.1) at the point $(x_i,t_{j+\theta})=(i\Delta x,(j+\theta)\Delta t)$ is given by,

$$-\frac{1}{\Delta x}[\theta\{(1-w)u_{i+1,j+1}+(2w-1)u_{i,j+1}-wu_{i-1,j+1}\}+(1-\theta)\{(1-w)u_{i+1,j}+(2w-1)u_{ij}-wu_{i-1,j}\}] = \frac{(u_{i,j+1}-u_{ij})}{\Delta t}, \tag{4.2.1}$$

or

$$-\lambda[\theta\{(1-w)u_{i+1,j+1}+(2w-1)u_{i,j+1}-wu_{i-1,j+1}\}+(1-\theta)\{(1-w)u_{i+1,j}+(2w-1)u_{ij}-wu_{i-1,j}\}] = u_{i,j+1}-u_{ij}, \tag{4.2.2}$$

where $\lambda=\frac{\Delta t}{\Delta x}$, the mesh ratio.

With w=1, this equation reduces to

$$(1+\lambda\theta)u_{i,j+1}-\lambda\theta u_{i-1,j+1} = (1-\lambda(1-\theta))u_{i,j}+\lambda(1-\theta)u_{i-1,j} \tag{4.2.3}$$

and for w=0 equation (4.2.2) becomes,

$$(1-\lambda\theta)u_{i,j+1}+\lambda\theta u_{i+1,j+1} = (1+\lambda(1-\theta))u_{i,j}-\lambda(1-\theta)u_{i+1,j}. \tag{4.2.4}$$

The local truncation error representations can be obtained by expanding the terms $U_{i,j+1}$, $U_{i-1,j+1}$, $U_{i-1,j}$ and U_{ij} about the point $(i\Delta x,(j+\frac{1}{2})\Delta t)$ using the Taylor series. The expansion for (4.2.3) leads to,

$$T_{4.2.3} = \left(\frac{\partial U}{\partial x}+\frac{\partial U}{\partial t}\right)_{i,j+\frac{1}{2}}+\Delta x\left(-\frac{1}{2}\frac{\partial^2 U}{\partial x^2}-\frac{(\Delta t)^2}{16}\frac{\partial^4 U}{\partial x^2\partial t^2}\right)_{i,j+\frac{1}{2}}+\Delta t\left(-\frac{1}{2}(1-2\theta)\frac{\partial^2 U}{\partial x\partial t}-\frac{(\Delta x)^2}{12}(1-2\theta)\frac{\partial^4 U}{\partial x^3\partial t}\right)_{i,j+\frac{1}{2}}+(\Delta x)(\Delta t)\left(\frac{1}{4}(1-2\theta)\frac{\partial^3 U}{\partial x^2\partial t}\right)_{i,j+\frac{1}{2}}+$$
$$(\Delta x)^2\left(\frac{1}{6}\frac{\partial^3 U}{\partial x^3}-\frac{\Delta x}{24}\frac{\partial^4 U}{\partial x^4}\right)_{i,j+\frac{1}{2}}+(\Delta t)^2\left(\frac{1}{8}\frac{\partial^3 U}{\partial x\partial t^2}+\frac{1}{24}\frac{\partial^3 U}{\partial t^3}-\frac{\Delta t}{48}(1-2\theta)\frac{\partial^4 U}{\partial x\partial t^3}\right)_{i,j+\frac{1}{2}}+\left(\frac{(\Delta x)^4}{5!}\frac{\partial^5 U}{\partial x^5}+\frac{5(\Delta x)^3(\Delta t)}{5!2}(1-2\theta)\frac{\partial^5 U}{\partial x^4\partial t}+\frac{5(\Delta x)^2(\Delta t)^2}{5!2}\right.$$

$$\frac{\partial^5 U}{\partial x^3 \partial t^2} + \frac{5(\Delta x)(\Delta t)^3}{5!4}(1-2\theta)\frac{\partial^5 U}{\partial x^2 \partial t^3} + \frac{5(\Delta t)^4}{5!16}\frac{\partial^5 U}{\partial x \partial t^4} +$$

$$\frac{(\Delta t)^4}{5!16}\frac{\partial^5 U}{\partial t^5})_{i,j+\frac{1}{2}} + \dots ,$$

$$\text{i.e. } T_{4.2.3} = \Delta x[-\tfrac{1}{2}\frac{\partial^2 U}{\partial x^2} - \frac{(\Delta t)^2}{16}\frac{\partial^4 U}{\partial x^2 \partial t^2}]_{i,j+\frac{1}{2}} + \Delta t[-\tfrac{1}{2}(1-2\theta)\frac{\partial^2 U}{\partial x \partial t} -$$

$$\frac{(\Delta x)^2}{12}(1-2\theta)\frac{\partial^4 U}{\partial x^3 \partial t}]_{i,j+\frac{1}{2}} + (\Delta x)(\Delta t)[\frac{1}{4}(1-2\theta)\frac{\partial^3 U}{\partial x^2 \partial t}]_{i,j+\frac{1}{2}}$$

$$+(\Delta x)^2[\frac{1}{6}\frac{\partial^3 U}{\partial x^3} - \frac{\Delta x}{24}\frac{\partial^4 U}{\partial x^4}]_{i,j+\frac{1}{2}} + (\Delta t)^2[\frac{1}{8}\frac{\partial^3 U}{\partial x \partial t^2} + \frac{1}{24}\frac{\partial^3 U}{\partial t^3} -$$

$$\frac{\Delta t}{48}(1-2\theta)\frac{\partial^4 U}{\partial x \partial t^3}]_{i,j+\frac{1}{2}} + O((\Delta x)^{\alpha_1}(\Delta t)^{\alpha_2}) , \qquad (4.2.5)$$

with $\alpha_1+\alpha_2=4$ and $0 \leq \theta \leq 1$. A similar expansion for the terms $U_{i,j+1}$, $U_{i+1,j+1}$, U_{ij} and $U_{i+1,j}$ about the point $(i\Delta x,(j+\frac{1}{2})\Delta t)$ provides the following truncation error expression for the formula (4.2.4),

$$T_{4.2.4} = (\frac{\partial U}{\partial x} + \frac{\partial U}{\partial t})_{i,j+\frac{1}{2}} + \Delta x(\tfrac{1}{2}\frac{\partial^2 U}{\partial x^2} + \frac{(\Delta t)^2}{16}\frac{\partial^4 U}{\partial x^2 \partial t^2})_{i,j+\frac{1}{2}} + \Delta t(-\tfrac{1}{2}(1-2\theta)$$

$$\frac{\partial^2 U}{\partial x \partial t} - \frac{1}{12}(\Delta x)^2(1-2\theta)\frac{\partial^4 U}{\partial x^3 \partial t})_{i,j+\frac{1}{2}} + (\Delta x)(\Delta t)(-\frac{1}{4}(1-2\theta)\frac{\partial^3 U}{\partial x^2 \partial t})_{i,j+\frac{1}{2}}$$

$$+(\Delta x)^2(\frac{1}{6}\frac{\partial^3 U}{\partial x^3} + \frac{\Delta x}{24}\frac{\partial^4 U}{\partial x^4})_{i,j+\frac{1}{2}} + (\Delta t)^2(\frac{1}{8}\frac{\partial^3 U}{\partial x \partial t^2} + \frac{1}{24}\frac{\partial^3 U}{\partial t^3} - \frac{\Delta t}{48}(1-2\theta)$$

$$\frac{\partial^4 U}{\partial x \partial t^3})_{i,j+\frac{1}{2}} + (\frac{(\Delta x)^4}{5!}\frac{\partial^5 U}{\partial x^5} - \frac{5(\Delta x)^3 \Delta t}{5!2}(1-2\theta)\frac{\partial^5 U}{\partial x^4 \partial t} + \frac{5(\Delta x)^2(\Delta t)^2}{5!2}$$

$$\frac{\partial^5 U}{\partial x^3 \partial t^2} - \frac{5(\Delta x)(\Delta t)^3(1-2\theta)}{5!4}\frac{\partial^5 U}{\partial x^2 \partial t^3} + \frac{5(\Delta t)^4}{5!16}\frac{\partial^5 U}{\partial x \partial t^4} + \frac{(\Delta t)^4}{5!16}\frac{\partial^5 U}{\partial t^2})_{i,j+\frac{1}{2}}$$

$$+ \dots ,$$

i.e.

$$T_{4.2.4} = \Delta x(\tfrac{1}{2}\frac{\partial^2 U}{\partial x^2} + \frac{(\Delta t)^2}{16}\frac{\partial^4 U}{\partial x^2 \partial t^2})_{i,j+\frac{1}{2}} + \Delta t(-\tfrac{1}{2}(1-2\theta)\frac{\partial^2 U}{\partial x \partial t} - \frac{1}{12}(\Delta x)^2$$

$$(1-2\theta)\frac{\partial^4 U}{\partial x^3 \partial t})_{i,j+\frac{1}{2}} + (\Delta x)(\Delta t)(-\frac{1}{4}(1-2\theta)\frac{\partial^3 U}{\partial x^2 \partial t})_{i,j+\frac{1}{2}} + (\Delta x)^2$$

$$\left(\frac{1}{6}\frac{\partial^3 U}{\partial x^3}+\frac{\Delta x}{24}\frac{\partial^4 U}{\partial x^4}\right)_{i,j+\frac{1}{2}}+(\Delta t)^2\left(\frac{1}{8}\frac{\partial^3 U}{\partial x\partial t^2}+\frac{1}{24}\frac{\partial^3 U}{\partial t^3}-\frac{\Delta t}{48}(1-2\theta)\frac{\partial^4 U}{\partial x\partial t^3}\right)_{i,j+\frac{1}{2}}+O((\Delta x)^{\alpha_1}(\Delta t)^{\alpha_2})\ ; \qquad (4.2.6)$$

with $\alpha_1+\alpha_2=4$ and $0\leq\theta\leq 1$.

Now, at the point $((i-1)\Delta x,(j+\theta)\Delta t)$, equation (4.2.4) takes the form,

$$\lambda\theta u_{i,j+1}+(1-\lambda\theta)u_{i-1,j+1}=-\lambda(1-\theta)u_{i,j}+(1+\lambda(1-\theta))u_{i-1,j} \qquad (4.2.7)$$

By coupling the equations (4.2.3) and (4.2.7) the two formulae can be written simultaneously in matrix form as,

$$\begin{bmatrix}-\lambda\theta & (1+\lambda\theta)\\ (1-\lambda\theta) & \lambda\theta\end{bmatrix}\begin{bmatrix}u_{i-1,j+1}\\ u_{i,j+1}\end{bmatrix}=\begin{bmatrix}\lambda(1-\theta) & 1-\lambda(1-\theta)\\ 1+\lambda(1-\theta) & -\lambda(1-\theta)\end{bmatrix}\begin{bmatrix}u_{i-1,j}\\ u_{ij}\end{bmatrix} \qquad (4.2.8)$$

i.e.,

$$A\underline{u}_{j+1}=B\underline{u}_j\ , \qquad (4.2.9)$$

where,

$$A=\begin{bmatrix}-\lambda\theta & (1+\lambda\theta)\\ (1-\lambda\theta) & \lambda\theta\end{bmatrix} \qquad B=\begin{bmatrix}\lambda(1-\theta) & 1-\lambda(1-\theta)\\ 1+\lambda(1-\theta) & -\lambda(1-\theta)\end{bmatrix}$$

and $\underline{u}_j=(u_{i-1,j},u_{i,j})^T$.

The (2×2) matrix A can be easily inverted. Hence from (4.2.9), we have,

$$\underline{u}_{j+1}=A^{-1}B\underline{u}_j\ , \qquad (4.2.10)$$

with,

$$A^{-1}=\begin{bmatrix}-\lambda\theta & (1+\lambda\theta)\\ (1-\lambda\theta) & \lambda\theta\end{bmatrix} \text{ and } A^{-1}B=\begin{bmatrix}(1+\lambda) & -\lambda\\ \lambda & (1-\lambda)\end{bmatrix}$$

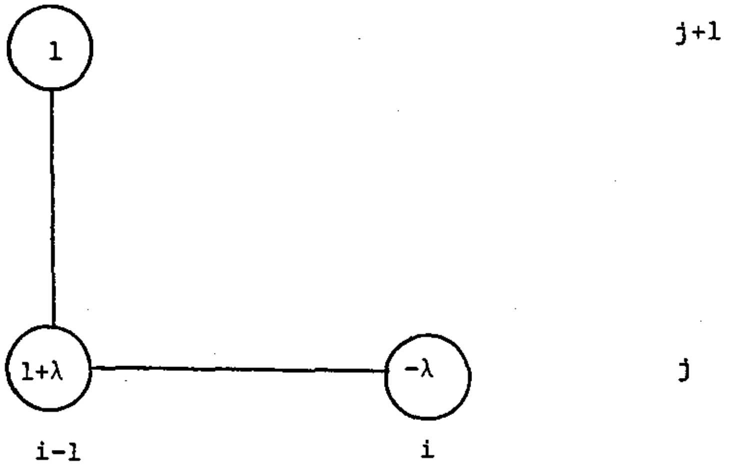
From equation (4.2.10), this gives rise to the following *set of explicit* equations,

$$u_{i-1,j+1} = (1+\lambda)u_{i-1,j}-\lambda u_{i,j} , \tag{4.2.11}$$

and $$u_{i,j+1} = \lambda u_{i-1,j}+(1-\lambda)u_{ij} . \tag{4.2.12}$$

whose computational molecules are shown in Figure (4.2.1).

Equation (4.2.11)



Equation (4.2.12)

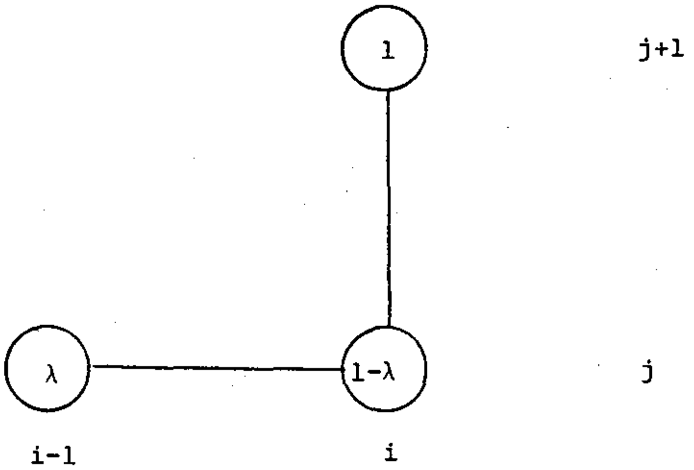


FIGURE 4.2.1

Equations (4.2.11) and (4.2.12) are for adjacent points which are grouped two at a time on the mesh line. Special formulae are needed to cope with the possibility of the existence of *ungrouped* points near the boundaries. The solution at the ungrouped point near *the right boundary* at the advanced time level can be computed from the equation (4.2.4) by putting i=m-1. This leads to,

$$u_{m-1,j+1} = [(1+\lambda(1-\theta))u_{m-1,j}-\lambda(1-\theta)u_{m,j}-\lambda\theta u_{m,j+1}]/(1-\lambda\theta), \quad (4.2.13)$$

where $\lambda\theta\neq 1$.

Equation (4.2.3) with i=1, deals with the value of u at the ungrouped point near *the left boundary*. Thus, we have,

$$u_{1,j+1} = [\lambda(1-\theta)u_{0,j}+(1-\lambda(1-\theta))u_{1,j}+\lambda\theta u_{0,j+1}]/(1+\lambda\theta). \quad (4.2.14)$$

Since the initial line $0\leq x\leq 1$ is uniformly divided with a spacing or increment Δx, the manner in which the above points are grouped very much depends on whether the number m of intervals of the line segment is *even* or *odd*. On this basis, a variety of group explicit schemes can be devised - as we will presently see.

Even Number of Intervals

When m is even, we will have an odd number (m-1) of internal points (i.e. points that do not include the left and right boundaries whose values are given by u_0 and u_m respectively at every time level). Consequently, the single ungrouped point will be located near either boundary.

(i) The GER Scheme

This refers to the Group Explicit with Right ungrouped point (GER) scheme. It results in the consecutive application for $\frac{1}{2}(m-2)$ times of the equations (4.2.11/12) for the first (m-2) points grouped two at a time. This is followed by a final use of equation (4.2.13) for the (m-1)th point at every time level as shown in Figure (4.2.2). Thus, we have the following set of equations,

$$\left.\begin{aligned}
-\lambda\theta u_{i-1,j+1}+(1+\lambda\theta)u_{i,j+1} &= \lambda(1-\theta)u_{i-1,j}+(1-\lambda(1-\theta))u_{i,j}, \\
(1-\lambda\theta)u_{i-1,j+1}+\lambda\theta u_{i,j+1} &= (1+\lambda(1-\theta))u_{i-1,j}-\lambda(1-\theta)u_{i,j},
\end{aligned}\right\} \; i=2,4,\ldots,(m-2)$$

and

$$(1-\lambda\theta)u_{m-1,j+1} = -\lambda\theta u_{m,j+1}-\lambda(1-\theta)u_{m,j}+(1+\lambda(1-\theta))u_{m-1,j}, \; \lambda\theta\neq 1.$$

which can be written in the more compact, *implicit* matrix form as,

$$\left[\begin{array}{cc:cc:c:cc:c}
-\lambda\theta & (1+\lambda\theta) & & & & & & \\
(1-\lambda\theta) & \lambda\theta & & & & & & \\ \hdashline
 & & -\lambda\theta & (1+\lambda\theta) & & & \bigcirc & \\
 & & (1-\lambda\theta) & \lambda\theta & & & & \\ \hdashline
 & & & & \ddots & & & \\ \hdashline
 & & & & & -\lambda\theta & (1+\lambda\theta) & \\
 & \bigcirc & & & & (1-\lambda\theta) & \lambda\theta & \\ \hdashline
 & & & & & & & (1-\lambda\theta)
\end{array}\right]
\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_{m-3} \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j+1}$$

$$= \left[\begin{array}{cc:cc:c:cc:c}
\lambda(1-\theta) & (1-\lambda(1-\theta)) & & & & & & \\
(1+\lambda(1-\theta)) & -\lambda(1-\theta) & & & & & & \\ \hdashline
 & & \lambda(1-\theta) & (1-\lambda(1-\theta)) & & & \bigcirc & \\
 & & (1+\lambda(1-\theta)) & -\lambda(1-\theta) & & & & \\ \hdashline
 & & & & \ddots & & & \\ \hdashline
 & & & & & \lambda(1-\theta) & (1-\lambda(1-\theta)) & \\
 & & \bigcirc & & & (1+\lambda(1-\theta)) & -\lambda(1-\theta) & \\ \hdashline
 & & & & & & & (1+\lambda(1-\theta))
\end{array}\right]$$

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_{m-3} \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_j + \underline{b}_1 \qquad (4.2.15)$$

where $\underline{b}_1=(0,0,\ldots,-\lambda(1-\theta)u_{m,j}-\lambda\theta u_{m,j+1})^T$ which consists of known boundary values. Now, if we define,

$$E_1 = \left[\begin{array}{cc|cc|c|cc|c} 0 & 1 & & & & & & \\ 1 & 0 & & & & & & \\ \hline & & 0 & 1 & & & \bigcirc & \\ & & 1 & 0 & & & & \\ \hline & & & & \ddots & & & \\ \hline & & & & & 0 & 1 & \\ & \bigcirc & & & & 1 & 0 & \\ \hline & & & & & & & 1 \end{array}\right]_{(m-1)\times(m-1)} \qquad (4.2.16)$$

and

$$G_1 = \left[\begin{array}{cc|cc|c|cc|c} -1 & 1 & & & & & & \\ -1 & 1 & & & & & & \\ \hline & & -1 & 1 & & & \bigcirc & \\ & & -1 & 1 & & & & \\ \hline & & & & \ddots & & & \\ \hline & & & & & -1 & 1 & \\ & \bigcirc & & & & -1 & 1 & \\ \hline & & & & & & & -1 \end{array}\right]_{((m-1)\times(m-1))} \qquad (4.2.17)$$

then we have,

$$(E_1+\lambda\theta G_1)\underline{u}_{j+1} = (E_1-\lambda(1-\theta)G_1)\underline{u}_j+\underline{b}_1 \ .$$

On premultiplying this equation by E_1^{-1} provides us with

$$E_1^{-1}(E_1+\lambda\theta G_1)\underline{u}_{j+1} = E_1^{-1}(E_1-\lambda(1-\theta)G_1)\underline{u}_j + E_1^{-1}\underline{b}_1, \text{ i.e.,}$$

$$(E_1^{-1}E_1+\lambda\theta E_1^{-1}G_1)\underline{u}_{j+1} = (E_1^{-1}E_1-\lambda(1-\theta)E_1^{-1}G_1)\underline{u}_j + \underline{b}_1 .$$

But $E_1^{-1}E_1=I$ and $E_1^{-1}G_1=G_1$ which implies that,

$$(I+\lambda\theta G_1)\underline{u}_{j+1} = (I-\lambda(1-\theta)G_1)\underline{u}_j + \underline{b}_1 , \tag{4.2.18}$$

where I is the identity matrix of order $((m-1)\times(m-1))$.

Hence, we obtain,

$$\underline{u}_{j+1} = (I+\lambda\theta G_1)^{-1}(I-\lambda(1-\theta)G_1)\underline{u}_j + \hat{\underline{b}}_1 , \tag{4.2.19}$$

where, $\hat{\underline{b}}_1 = (I+\lambda\theta G_1)^{-1}\underline{b}_1$.

The *explicit* equation (4.2.19) is the governing equation for the computation of the GER scheme.

(ii) *The GEL Scheme*

This is an abbreviation for the Group Explicit with Left ungrouped point scheme and it is in fact a reverse of the GER scheme. It is obtained by the use of equation (4.2.14) for the first internal point followed by the application of equations (4.2.11/12) for ½(m-2) times for the remaining points on the mesh line. The scheme is displayed diagrammatically in Figure (4.2.3) and is determined by the following set of linear equations,

$$(1+\lambda\theta)u_{1,j+1} = (1-\lambda(1-\theta)u_{1j} + \lambda(1-\theta)u_{0j} +\lambda\theta u_{0,j+1} ,$$

$$\left.\begin{aligned} -\lambda\theta u_{i-1,j+1}+(1+\lambda\theta)u_{i,j+1} &= \lambda(1-\theta)u_{i-1,j}+(1-\lambda(1-\theta))u_{ij}, \\ \text{and } (1-\lambda\theta)u_{i-1,j+1}+\lambda\theta u_{i,j+1} &= (1+\lambda(1-\theta))u_{i-1,j}-\lambda(1-\theta)u_{ij} . \end{aligned}\right\} \begin{aligned} &i=3,5,\ldots,m-1; \\ &\lambda\theta\neq 1 . \end{aligned}$$

In the *implicit* matrix form, these equations can be written as,

$$\left[\begin{array}{c|cc|cc|c|cc}
1+\lambda\theta & & & & & & & \\ \hline
 & -\lambda\theta & (1+\lambda\theta) & & & & & \\
 & (1-\lambda\theta) & \lambda\theta & & & & & \\ \hline
 & & & -\lambda\theta & (1+\lambda\theta) & & & \\
 & & & (1-\lambda\theta) & \lambda\theta & & \multicolumn{2}{c}{\bigcirc} \\ \hline
 & & & & & \ddots & & \\ \hline
 & & \multicolumn{2}{c}{\bigcirc} & & & -\lambda\theta & (1+\lambda\theta) \\
 & & & & & & (1-\lambda\theta) & \lambda\theta
\end{array}\right]
\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ \vdots \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j+1} =$$

$$\left[\begin{array}{c|cc|c|cc}
(1-\lambda(1-\theta)) & & & & & \\ \hline
 & \lambda(1-\theta) & (1-\lambda(1-\theta)) & & & \\
 & (1+\lambda(1-\theta)) & -\lambda(1-\theta) & & \multicolumn{2}{c}{\bigcirc} \\ \hline
 & & & \ddots & & \\ \hline
 & & \bigcirc & & \lambda(1-\theta) & (1-\lambda(1-\theta)) \\
 & & & & (1+\lambda(1-\theta)) & -\lambda(1-\theta)
\end{array}\right]
\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j} + \underline{b}_2 \qquad (4.2.20)$$

where $\underline{b}_2=(\lambda(1-\theta)u_{0,j}+\lambda\theta u_{0,j+1},0,\ldots,0)^T$.

If we define,

$$E_2=\left[\begin{array}{c|cc|cc|c|cc}
1 & & & & & & & \\ \hline
 & 0 & 1 & & & & & \\
 & 1 & 0 & & & & & \\ \hline
 & & & 0 & 1 & & & \\
 & & & 1 & 0 & \bigcirc & & \\ \hline
 & & & & & \ddots & & \\
 & & \bigcirc & & & & & \\ \hline
 & & & & & & 0 & 1 \\
 & & & & & & 1 & 0
\end{array}\right], \qquad (4.2.21)$$

$$\text{and}\quad G_2 = \left[\begin{array}{c|cc|cc|c|cc}
1 & & & & & & & \\
\hline
 & -1 & 1 & & & & & \\
 & -1 & 1 & & & & & \\
\hline
 & & & -1 & 1 & & & \\
 & & & -1 & 1 & & O & \\
\hline
 & & & & & \ddots & & \\
\hline
 & & & & & & -1 & 1 \\
 & & & O & & & -1 & 1
\end{array}\right] \qquad (4.2.22)$$

then the GEL scheme takes the form,

$$(E_2+\lambda\theta G_2)\underline{u}_{j+1} = (E_2-\lambda(1-\theta)G_2)\underline{u}_j + \underline{b}_2 \ .$$

If we premultiply this equation by E_2^{-1}, we get

$$(I+\lambda\theta G_2)\underline{u}_{j+1} = (I-\lambda(1-\theta)G_2)\underline{u}_j + \underline{b}_2 \ , \qquad (4.2.23)$$

and this leads to the following explicit formula for the computation of the GEL scheme,

$$\underline{u}_{j+1} = (I+\lambda\theta G_2)^{-1}(I-\lambda(1-\theta)G_2)\underline{u}_j + \hat{\underline{b}}_2 \ , \qquad (4.2.24)$$

where $\hat{\underline{b}}_2 = (I+\lambda\theta G_2)^{-1}\underline{b}_2$.

(iii) The (S)AGE Scheme

As the name suggests, this (Single) Alternating Group Explicit scheme entails the *alternate* use of the GER and the GEL formulae (i.e. equations (4.2.18) and (4.2.23)) as we march our solutions forward with respect to time as illustrated in Figure (4.2.4). Thus, *the two time-level process* of the S(AGE) scheme is given by,

$$\left.\begin{array}{ll} & (I+\lambda\theta G_1)\underline{u}_{j+1} = (I-\lambda(1-\theta)G_1)\underline{u}_j + \underline{b}_1 \\ \text{and} & (I+\lambda\theta G_2)\underline{u}_{j+2} = (I-\lambda(1-\theta)G_2)\underline{u}_{j+1} + \underline{b}_2 \end{array}\right\} j=0,2,4,\ldots \qquad (4.2.25)$$

(iv) *The (D)AGE Scheme*

The (Double) Alternating Group Explicit scheme is a *four-time level process*. A single application of the algorithm requires the utilisation of the GER and the GEL schemes for the first two time levels followed by the employment of the same formulae but now in reverse order for the second two time levels. This alternating procedure is repeated as the solutions are progressed vertically. Thus, we see that the (D)AGE scheme is basically a periodic rotation of the (S)AGE scheme with the second half cycle implemented in opposite order to complete the four time level process and is given by,

$$\left.\begin{array}{l} (I+\lambda\theta G_1)\underline{u}_{j+1} = (I-\lambda(1-\theta)G_1)\underline{u}_j + \underline{b}_1 \\ (I+\lambda\theta G_2)\underline{u}_{j+2} = (I-\lambda(1-\theta)G_2)\underline{u}_{j+1} + \underline{b}_2 \\ (I+\lambda\theta G_2)\underline{u}_{j+3} = (I-\lambda(1-\theta)G_2)\underline{u}_{j+2} + \underline{b}_2 \\ (I+\lambda\theta G_1)\underline{u}_{j+4} = (I-\lambda(1-\theta)G_1)\underline{u}_{j+3} + \underline{b}_1 \end{array}\right\} j=0,4,8,\ldots \qquad (4.2.26)$$

The scheme is represented diagrammatically by Figure (4.2.5).

Before we proceed to establish the corresponding finite-difference analogues for the case when the line segment $0 \leq x \leq 1$ has an *odd* number of intervals, we observe that if we denote,

$$G^{(i)} = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, \quad i=1,2,\ldots,\tfrac{1}{2}(m-2) \qquad (4.2.27)$$

then from (4.2.17) and (4.2.22), we have,

$$G_1 = \begin{bmatrix} G^{(1)} & & & & & \\ & G^{(2)} & & & O & \\ & & \ddots & & & \\ & & & G^{(\frac{1}{2}(m-2)-1)} & & \\ & & & & G^{(\frac{1}{2}(m-2))} & \\ & O & & & & -1 \end{bmatrix}_{((m-1)\times(m-1))} \qquad (4.2.28)$$

$$\text{and } G_2 = \begin{bmatrix} 1 & & & & & \\ & G^{(1)} & & & O & \\ & & G^{(2)} & & & \\ & O & & \ddots & & \\ & & & & G^{(\frac{1}{2}(m-2)-1)} & \\ & & & & & G^{(\frac{1}{2}(m-2))} \end{bmatrix}_{(m-1)\times(m-1))} \qquad (4.2.29)$$

Odd Number of Intervals

We will have an odd number of intervals when m is odd. Therefore, at every time level, the number of internal points is even. Accordingly, again there are two possibilities which determine the manner in which the points are grouped on the mesh line. In the first possibility, we will have $\frac{1}{2}(m-1)$ *complete* groups of two points. In the second possibility, however, we are led to $(\frac{m-3}{2})$ groups of two points and one point which is ungrouped adjacent to *each* boundary. Based on these observations, the following group explicit schemes can be constructed in an analogous fashion as in the *even* case.

(i) *The GEU Scheme*

In this scheme, there are two points which are ungrouped, one each which is adjacent to the left and right boundary. Thus, for the *left* ungrouped point (the second point), we use equation (4.2.14) whilst the solution at the *right* ungrouped point (the (m-1)th point) is determined by equation (4.2.13). For the grouped points in between, we apply equations (4.2.11) and (4.2.12) in succession for $\frac{1}{2}(m-3)$ times to give the solutions at these points. This is repeated for progressive time levels and the whole procedure is known as the Group Explicit with Ungrouped ends method. Thus, the GEU method requires the solution of,

$$(1+\lambda\theta)u_{1,j+1} = (1-\lambda(1-\theta))u_{1,j}+\lambda(1-\theta)u_{0,j}+\lambda\theta u_{0,j+1},$$

$$\left.\begin{aligned}-\lambda\theta u_{i-1,j+1}+(1+\lambda\theta)u_{i,j+1} &= \lambda(1-\theta)u_{i-1,j}+(1-\lambda(1-\theta))u_{i,j}\\ (1-\lambda\theta)u_{i-1,j+1}+\lambda\theta u_{i,j+1} &= (1+\lambda(1-\theta))u_{i-1,j}-\lambda(1-\theta)u_{i,j}\end{aligned}\right\}\begin{aligned}&i=3,5,\\&\quad\ldots,m-2;\\&\lambda\theta\neq 1\end{aligned}$$

$$(1-\lambda\theta)u_{m-1,j+1} = -\lambda\theta u_{m,j+1}-\lambda(1-\theta)u_{m,j}+(1+\lambda(1-\theta))u_{m-1,j}$$

which can be written in the implicit matrix form as,

$$\left[\begin{array}{c|cc|c|cc|c}
(1+\lambda\theta) & & & & & & \\ \hline
 & -\lambda\theta & (1+\lambda\theta) & & & & \\
 & (1-\lambda\theta) & \lambda\theta & & & & \\ \hline
 & & & \ddots & & O & \\ \hline
 & & & & -\lambda\theta & (1+\lambda\theta) & \\
 & & O & & (1-\lambda\theta) & \lambda\theta & \\ \hline
 & & & & & & 1-\lambda\theta
\end{array}\right]
\left[\begin{array}{c} u_1\\ u_2\\ u_3\\ \vdots\\ u_{m-3}\\ u_{m-2}\\ u_{m-1}\end{array}\right]_{j+1}$$

$$=\left[\begin{array}{c|cc|c|cc|c}
1-\lambda(1-\theta) & & & & & & \\ \hline
 & \lambda(1-\theta) & (1-\lambda(1-\theta)) & & & O & \\
 & (1+\lambda(1-\theta)) & -\lambda(1-\theta) & & & & \\ \hline
 & & & \ddots & & & \\ \hline
 & & & & \lambda(1-\theta) & (1-\lambda(1-\theta)) & \\
 & & O & & (1+\lambda(1-\theta)) & -\lambda(1-\theta) & \\ \hline
 & & & & & & (1+\lambda(1-\theta))
\end{array}\right]
\left[\begin{array}{c} u_1\\ u_2\\ u_3\\ \vdots\\ u_{m-3}\\ u_{m-2}\\ u_{m-1}\end{array}\right]_{j}+\underline{b}_3, \tag{4.2.30}$$

where $\underline{b}_3 = (\lambda(1-\theta)u_{0,j}+\lambda\theta u_{0,j+1},0,\ldots,0,-\lambda(1-\theta)u_{m,j}-\lambda\theta u_{m,j+1})^T$.

Now if we let

$$\hat{G}_1 = \begin{bmatrix} 1 & & & & & \\ & G^{(1)} & & & \bigcirc & \\ & & G^{(2)} & & & \\ & & & \ddots \; G^{(\frac{1}{2}(m-2)-1)} & & \\ & & \bigcirc & & G^{(\frac{1}{2}(m-3))} & \\ & & & & & -1 \end{bmatrix} \qquad (4.2.31)$$

and

$$\hat{G}_2 = \begin{bmatrix} G^{(1)} & & & \\ & G^{(2)} & & \bigcirc \\ & & \ddots \; G^{(\frac{1}{2}(m-3))} & \\ \bigcirc & & & G^{(\frac{1}{2}(m-1))} \end{bmatrix} \qquad (4.2.32)$$

where the (2×2) matrices $G^{(i)}$, $i=1,2,\ldots,\frac{1}{2}(m-1)$ are defined as in (4.2.25), then the GEU scheme is given by,

$$(I+\lambda\theta\hat{G}_1)\underline{u}_{j+1} = (I-\lambda(1-\theta)\hat{G}_1)\underline{u}_j + \underline{b}_3 \; , \qquad (4.2.33)$$

and is described by Figure (4.2.6).

(ii) *The GEC Scheme*

This scheme, known as the Group Explicit Complete method is obtained by applying successively $\frac{1}{2}(m-1)$ times equations (4.2.11) and (4.2.12) for the first to $(m-1)^{th}$ point along each progressive mesh line as displayed in Figure (4.2.7). Thus, the relevant implicit equations are,

$$\left.\begin{aligned} -\lambda\theta u_{i-1,j+1}+(1+\lambda\theta)u_{i,j+1} &= \lambda(1-\theta)u_{i-1,j}+(1-\lambda(1-\theta))u_{i,j} \\ (1-\lambda\theta)u_{i-1,j+1}+\lambda\theta u_{i,j+1} &= (1+\lambda(1-\theta))u_{i-1,j}-\lambda(1-\theta)u_{i,j} \end{aligned}\right\} \begin{array}{l} i=2,4,\ldots,(m-1); \\ \lambda\theta\neq 1 \end{array}$$

which in the matrix form are written as,

$$\left[\begin{array}{cc|c|cc} -\lambda\theta & (1+\lambda\theta) & & & \\ (1-\lambda\theta) & \lambda\theta & & & \\ \hline & & \ddots & & O \\ \hline & & O & -\lambda\theta & (1+\lambda\theta) \\ & & & (1-\lambda\theta) & \lambda\theta \end{array}\right] \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j+1}$$

$$= \left[\begin{array}{cc|c|cc} \lambda(1-\theta) & (1-\lambda(1-\theta)) & & & \\ (1+\lambda(1-\theta)) & -\lambda(1-\theta) & & & \\ \hline & & \ddots & & O \\ \hline & & O & \lambda(1-\theta) & (1-\lambda(1-\theta)) \\ & & & (1+\lambda(1-\theta)) & -\lambda(1-\theta) \end{array}\right] \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j} \qquad (4.2.34)$$

Therefore, by using (4.2.32), the GEC scheme can be expressed as,

$$(I+\lambda\theta\hat{G}_2)\underline{u}_{j+1} = (I-\lambda(1-\theta)\hat{G}_2)\underline{u}_j \;. \qquad (4.2.35)$$

(iii) *The (S)AGE and (D)AGE Scheme*

The alternating schemes corresponding to the ones that we have developed for the even case are given by,

$$\left.\begin{aligned} (I+\lambda\theta\hat{G}_1)\underline{u}_{j+1} &= (I-\lambda(1-\theta)\hat{G}_1)\underline{u}_j + \underline{b}_3 \\ (I+\lambda\theta\hat{G}_2)\underline{u}_{j+2} &= (I-\lambda(1-\theta)\hat{G}_2)\underline{u}_{j+1} \end{aligned}\right\} \qquad (4.2.36)$$

for (S)AGE (Figure 4.2.8) and,

$$\left.\begin{aligned} (I+\lambda\theta\hat{G}_1)\underline{u}_{j+1} &= (I-\lambda(1-\theta)\hat{G}_1)\underline{u}_j + \underline{b}_3 \\ (I+\lambda\theta\hat{G}_2)\underline{u}_{j+2} &= (I-\lambda(1-\theta)\hat{G}_2)\underline{u}_{j+1} \\ (I+\lambda\theta\hat{G}_2)\underline{u}_{j+3} &= (I-\lambda(1-\theta)\hat{G}_2)\underline{u}_{j+2} \\ (I+\lambda\theta\hat{G}_1)\underline{u}_{j+4} &= (I-\lambda(1-\theta)\hat{G}_1)\underline{u}_{j+3} + \underline{b}_3 \end{aligned}\right\} \qquad (4.2.37)$$

for (D)AGE (Figure 4.2.9).

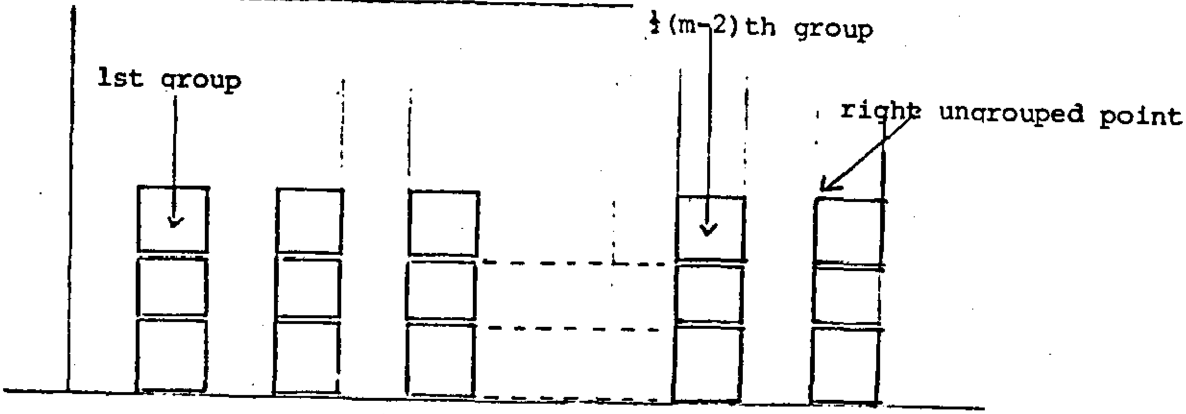
Notice that the above GE formulae are slightly different from those obtained by Evans and Abdullah (1983) for parabolic problems in the sense that the same group $\hat{G}_i$ (i=1 or 2) appears in both sides of the equations. The following conclusions may therefore be drawn:

(a) the GE schemes can be derived from the class of locally one dimensional methods (LOD),

(b) there is no overlapping of the grouping of points. Rather they are disjoint as shown in Figures 4.2.2-4.2.9,

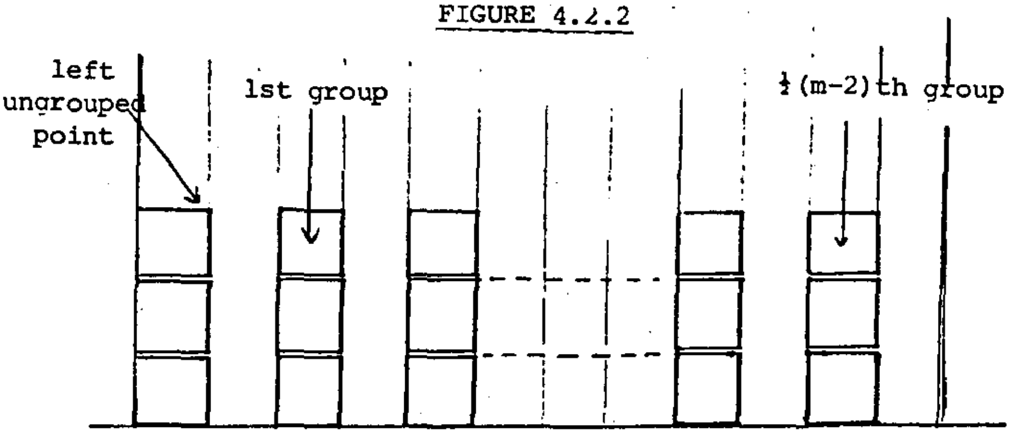
and (c) there is no longer a need for the commutativity of the matrices $\hat{G}_1$ and $\hat{G}_2$.

Even Number of Intervals



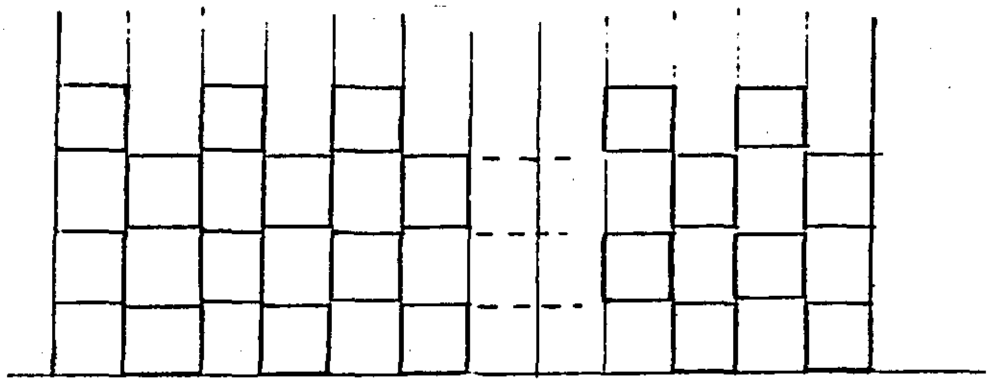
The GER Scheme

FIGURE 4.2.2



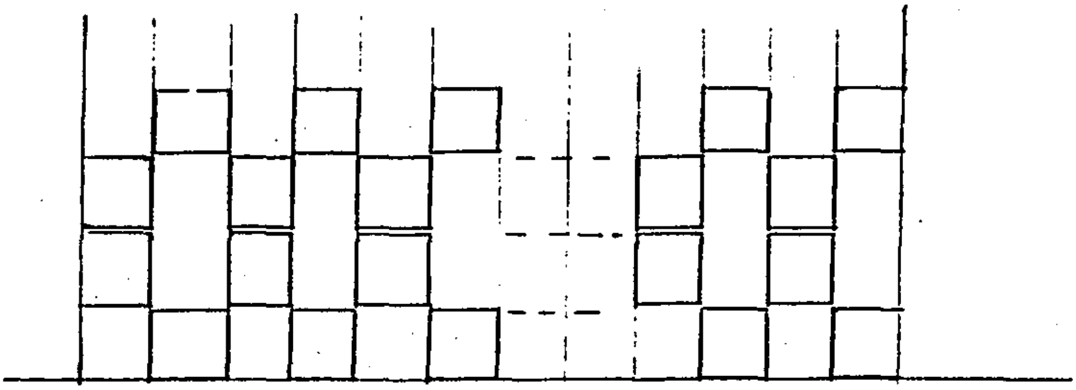
The GEL Scheme

FIGURE 4.2.3



The (S)AGE Scheme

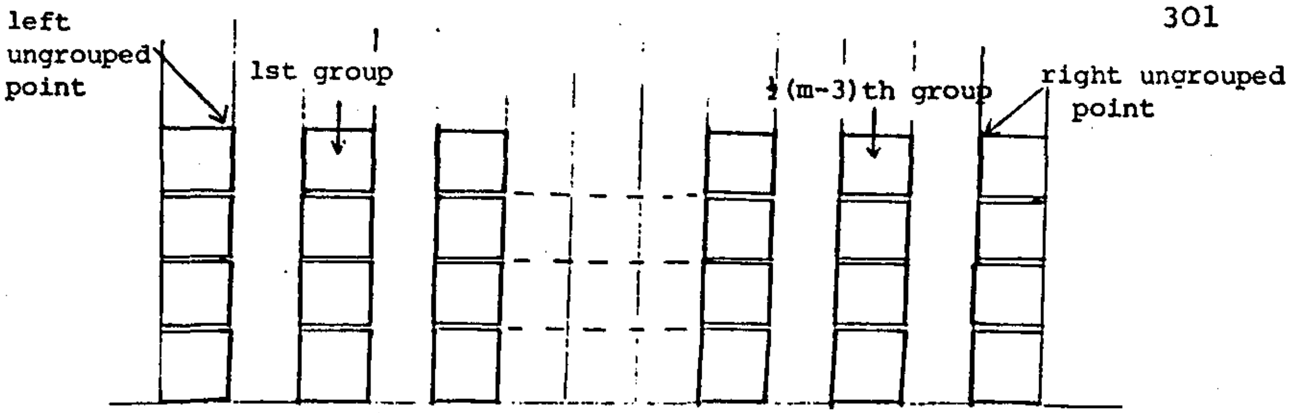
FIGURE 4.2.4



The (D)AGE Scheme

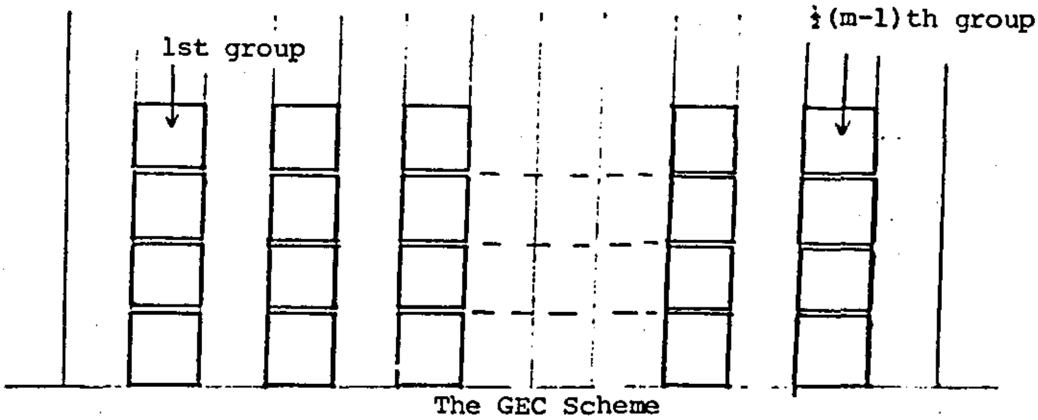
FIGURE 4.2.5

Odd Number of Intervals



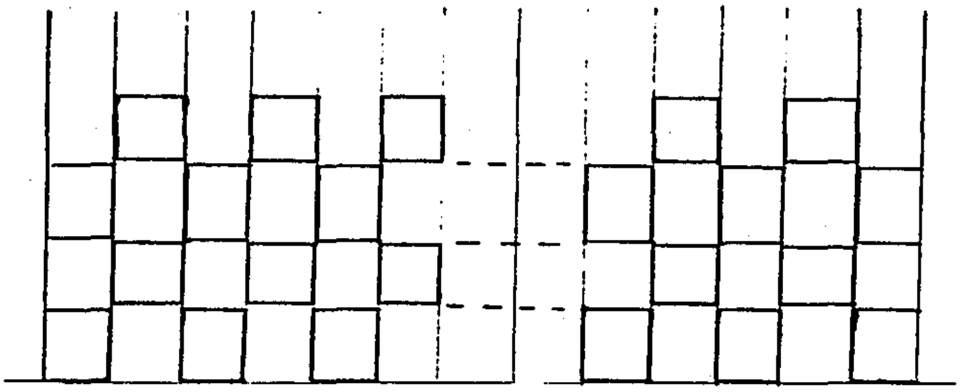
The GEU Scheme

FIGURE 4.2.6



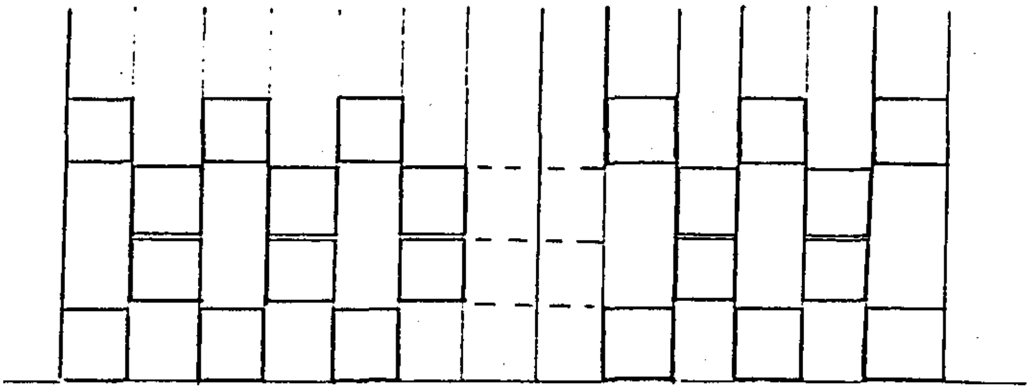
The GEC Scheme

FIGURE 4.2.7



The (S)AGE Scheme

FIGURE 4.2.8



The (D)AGE Scheme

FIGURE 4.2.9

4.3 TRUNCATION ERROR ANALYSIS FOR THE GE METHODS

(i) *Truncation Error for the GER Scheme*

The set of explicit equations obtained by coupling equations (4.2.3) and (4.2.4) are,

$$u_{i-1,j+1}+\lambda u_{ij}-(1+\lambda)u_{i-1,j} = 0 \qquad (4.3.1)$$

$$u_{i,j+1}-\lambda u_{i-1,j}-(1-\lambda)u_{ij} = 0 \,. \qquad (4.3.2)$$

The truncation errors for any two grouped points are given by the truncation errors of equations (4.3.1) and (4.3.2) for i=2,4,...,m-2. By expanding the terms $U_{i-1,j+1}, U_{ij}, U_{i-1,j}$ in equation (4.3.1) about the point $((i-1)\Delta x,(j+\frac{1}{2})\Delta t)$, we get,

$$\begin{aligned}
T_{4.3.1} = {} & (\frac{\partial U}{\partial x} + \frac{\partial U}{\partial t})_{i-1,j+\frac{1}{2}}+\Delta x(\tfrac{1}{2}\,\frac{\partial^2 U}{\partial x^2} + \frac{(\Delta t)^2}{16}\,\frac{\partial^4 U}{\partial x^2\partial t^2})_{i-1,j+\frac{1}{2}}+\Delta t(-\tfrac{1}{2}\,\frac{\partial^2 U}{\partial x\partial t} - \\
& \frac{1}{12}(\Delta x)^2\,\frac{\partial^4 U}{\partial x^3\partial t})_{i-1,j+\frac{1}{2}}+(\Delta x)(\Delta t)(-\frac{1}{4}\,\frac{\partial^3 U}{\partial x^2\partial t})_{i-1,j+\frac{1}{2}}+(\Delta x)^2(\frac{1}{6}\,\frac{\partial^3 U}{\partial x^3} + \\
& \frac{\Delta x}{24}\,\frac{\partial^4 U}{\partial x^4})_{i-1,j+\frac{1}{2}}+(\Delta t)^2(\frac{1}{8}\,\frac{\partial^3 U}{\partial x\partial t^2} + \frac{1}{24}\,\frac{\partial^3 U}{\partial t^3} - \frac{\Delta t}{48}\,\frac{\partial^4 U}{\partial x\partial t^3})_{i-1,j+\frac{1}{2}} \\
& +(\frac{(\Delta x)^4}{120}\,\frac{\partial^5 U}{\partial x^5} - \frac{(\Delta t)(\Delta x)^3}{48}\,\frac{\partial^5 U}{\partial x^4\partial t} + \frac{1}{48}\,(\Delta x)^2(\Delta t)^2\,\frac{\partial^5 U}{\partial x^3\partial t^2} - \\
& \frac{1}{96}(\Delta x)(\Delta t)^3\,\frac{\partial^5 U}{\partial x^2\partial t^3} + \frac{1}{384}\,(\Delta t)^4\,\frac{\partial^5 U}{\partial x\partial t^4} + \frac{1}{1920}\,(\Delta t)^4\,\frac{\partial^5 U}{\partial t^5})_{i-1,j+\frac{1}{2}} \\
& + \dots ,
\end{aligned}$$

i.e.,

$$\begin{aligned}
T_{4.3.1} = {} & \Delta x(\tfrac{1}{2}\,\frac{\partial^2 U}{\partial x^2} + \frac{(\Delta t)^2}{16}\,\frac{\partial^4 U}{\partial x^2\partial t^2})_{i-1,j+\frac{1}{2}}+\Delta t(-\tfrac{1}{2}\,\frac{\partial^2 U}{\partial x\partial t} - \frac{1}{12}(\Delta x)^2\,\frac{\partial^4 U}{\partial x^3\partial t})_{i-1,j+\frac{1}{2}} \\
& +(\Delta x)(\Delta t)(-\frac{1}{4}\,\frac{\partial^3 U}{\partial x^2\partial t})_{i-1,j+\frac{1}{2}}+(\Delta x)^2(\frac{1}{6}\,\frac{\partial^3 U}{\partial x^3} + \frac{\Delta x}{24}\,\frac{\partial^4 U}{\partial x^4})_{i-1,j+\frac{1}{2}} \\
& +(\Delta t)^2(\frac{1}{8}\,\frac{\partial^3 U}{\partial x\partial t^2} + \frac{1}{24}\,\frac{\partial^3 U}{\partial t^3} - \frac{\Delta t}{48}\,\frac{\partial^4 U}{\partial x\partial t^3})_{i-1,j+\frac{1}{2}}+O((\Delta x)^{\alpha_1}(\Delta t)^{\alpha_2});
\end{aligned}$$

$$\alpha_1+\alpha_2 = 4 \,. \qquad (4.3.3a)$$

Similarly by expanding the terms $U_{i,j+1}, U_{i-1,j}$ and U_{ij} about the point $(i\Delta x,(j+\frac{1}{2})\Delta t)$ leads to,

$$\begin{aligned}T_{4.3.2} = {} & \left(\frac{\partial U}{\partial x} + \frac{\partial U}{\partial t}\right)_{i,j+\frac{1}{2}} + \Delta x\left(-\tfrac{1}{2}\,\frac{\partial^2 U}{\partial x^2} - \frac{(\Delta t)^2}{16}\,\frac{\partial^4 U}{\partial x^2\partial t^2}\right)_{i,j+\frac{1}{2}} + \Delta t\left(-\tfrac{1}{2}\,\frac{\partial^2 U}{\partial x\partial t} - \right. \\ & \left.\frac{(\Delta x)^2}{12}\,\frac{\partial^4 U}{\partial x^3\partial t}\right)_{i,j+\frac{1}{2}} + (\Delta t)(\Delta x)\left(\frac{1}{4}\,\frac{\partial^3 U}{\partial x^2\partial t}\right)_{i,j+\frac{1}{2}} + (\Delta x)^2\left(\frac{1}{6}\,\frac{\partial^3 U}{\partial x^3} - \frac{\Delta x}{24}\,\frac{\partial^4 U}{\partial x^4}\right)_{i,j+\frac{1}{2}} \\ & + (\Delta t)^2\left(\frac{1}{8}\,\frac{\partial^3 U}{\partial x\partial t^2} + \frac{1}{24}\,\frac{\partial^3 U}{\partial t^3} - \frac{\Delta t}{48}\,\frac{\partial^4 U}{\partial x\partial t^3}\right)_{i,j+\frac{1}{2}} + \left(\frac{(\Delta x)^4}{120}\,\frac{\partial^5 U}{\partial x^5} + \frac{(\Delta t)(\Delta x)^3}{48}\right. \\ & \frac{\partial^5 U}{\partial x^4\partial t} + \frac{(\Delta t)^2(\Delta x)^2}{48}\,\frac{\partial^5 U}{\partial x^3\partial t^2} + \frac{(\Delta t)^3\Delta x}{96}\,\frac{\partial^5 U}{\partial x^2\partial t^3} + \frac{1}{384}(\Delta t)^4\,\frac{\partial^5 U}{\partial x\partial t^4} + \frac{(\Delta t)^4}{1920} \\ & \left.\frac{\partial^5 U}{\partial t^5}\right)_{i,j+\frac{1}{2}} + \ldots,\end{aligned}$$

i.e.,

$$\begin{aligned}T_{4.3.2} = {} & \Delta x\left(-\tfrac{1}{2}\,\frac{\partial^2 U}{\partial x^2} - \frac{(\Delta t)^2}{16}\,\frac{\partial^4 U}{\partial x^2\partial t^2}\right)_{i,j+\frac{1}{2}} + \Delta t\left(-\tfrac{1}{2}\,\frac{\partial^2 U}{\partial x\partial t} - \frac{(\Delta x)^2}{12}\,\frac{\partial^4 U}{\partial x^3\partial t}\right)_{i,j+\frac{1}{2}} \\ & + (\Delta t)(\Delta x)\left(\frac{1}{4}\,\frac{\partial^3 U}{\partial x^2\partial t}\right)_{i,j+\frac{1}{2}} + (\Delta x)^2\left(\frac{1}{6}\,\frac{\partial^3 U}{\partial x^3} - \frac{\Delta x}{24}\,\frac{\partial^4 U}{\partial x^4}\right)_{i,j+\frac{1}{2}} \\ & + (\Delta t)^2\left(\frac{1}{8}\,\frac{\partial^3 U}{\partial x\partial t^2} + \frac{1}{24}\,\frac{\partial^3 U}{\partial t^3} - \frac{\Delta t}{48}\,\frac{\partial^4 U}{\partial x\partial t^3}\right)_{i,j+\frac{1}{2}} + O((\Delta x)^{\alpha_1}(\Delta t)^{\alpha_2})\,; \\ & \alpha_1+\alpha_2=4\ . \qquad (4.3.3b)\end{aligned}$$

The truncation error for the single ungrouped point near the right end is given by the truncation error incurred for equation (4.2.4). This is obtained directly by putting i=m-1 in equation (4.2.6) which gives,

$$\begin{aligned}T_R = {} & \Delta x\left(\tfrac{1}{2}\,\frac{\partial^2 U}{\partial x^2} + \frac{(\Delta t)^2}{16}\,\frac{\partial^4 U}{\partial x^2\partial t^2}\right)_{m-1,j+\frac{1}{2}} + \Delta t\left(-\tfrac{1}{2}(1-2\theta)\frac{\partial^2 U}{\partial x\partial t} - \frac{1}{12}(\Delta x)^2(1-2\theta)\right. \\ & \left.\frac{\partial^4 U}{\partial x^3\partial t}\right)_{m-1,j+\frac{1}{2}} + (\Delta x)(\Delta t)\left(-\frac{1}{4}(1-2\theta)\frac{\partial^3 U}{\partial x^2\partial t}\right)_{m-1,j+\frac{1}{2}} + (\Delta x)^2\left(\frac{1}{6}\,\frac{\partial^3 U}{\partial x^3} + \frac{\Delta x}{24}\right. \\ & \left.\frac{\partial^4 U}{\partial x^4}\right)_{m-1,j+\frac{1}{2}} + (\Delta t)^2\left(\frac{1}{8}\,\frac{\partial^3 U}{\partial x\partial t^2} + \frac{1}{24}\,\frac{\partial^3 U}{\partial t^3} - \frac{\Delta t}{48}(1-2\theta)\,\frac{\partial^4 U}{\partial x\partial t^3}\right)_{m-1,j+\frac{1}{2}} + \\ & O((\Delta x)^{\alpha_1}(\Delta t)^{\alpha_2})\ ;\quad \alpha_1+\alpha_2=4\ . \qquad (4.3.3c)\end{aligned}$$

(ii) Truncation Error for the GEL Scheme

The truncation error for the single ungrouped point near the left boundary is given by the truncation error for equation (4.2.3). Hence with i=1, the expression (4.2.5) gives,

$$\begin{aligned} T_L = {} & \Delta x[-\tfrac{1}{2}\frac{\partial^2 U}{\partial x^2} - \frac{(\Delta t)^2}{16}\frac{\partial^4 U}{\partial x^2 \partial t^2}]_{1,j+\frac{1}{2}} + \Delta t[-\tfrac{1}{2}(1-2\theta)\frac{\partial^2 U}{\partial x \partial t} - \frac{(\Delta x)^2}{12}(1-2\theta) \\ & \frac{\partial^4 U}{\partial x^3 \partial t}]_{1,j+\frac{1}{2}} + (\Delta x)(\Delta t)[\frac{1}{4}(1-2\theta)\frac{\partial^3 U}{\partial x^2 \partial t}]_{1,j+\frac{1}{2}} + (\Delta x)^2[\frac{1}{6}\frac{\partial^3 U}{\partial x^3} - \frac{\Delta x}{24} \\ & \frac{\partial^4 U}{\partial x^4}]_{1,j+\frac{1}{2}} + (\Delta t)^2[\frac{1}{8}\frac{\partial^3 U}{\partial x \partial t^2} + \frac{1}{24}\frac{\partial^3 U}{\partial t^3} - \frac{\Delta t}{48}(1-2\theta)\frac{\partial^4 U}{\partial x \partial t^3}]_{1,j+\frac{1}{2}} + \\ & O((\Delta x)^{\alpha_1}(\Delta t)^{\alpha_2}); \quad \alpha_1+\alpha_2=4,\ 0\leqslant\theta\leqslant 1. \end{aligned} \tag{4.3.4}$$

We note that the truncation errors for any two grouped points of the GEL scheme are given by $T_{4.3.1}$ and $T_{4.3.2}$ of the equations (4.3.3a) and (4.3.3b) respectively.

(iii) *Truncation Error for the GEU Scheme*

As indicated by Figure (4.2.6) for the case when an odd number of intervals is used, the truncation error of the scheme at the left ungrouped point is given by T_L of (4.3.4) whilst the error at the single ungrouped point near the right end is T_R of (4.3.3c). For the points in between the boundaries which are grouped two at a time, the truncation errors are given by $T_{4.3.1}$ and $T_{4.3.2}$ of equations (4.3.3a) and (4.3.3b) respectively.

(iv) *Truncation Error for the GEC Scheme*

In this scheme, the grouping of two points at a time along each mesh line is complete as shown by Figure (4.2.7). Hence, the truncation error for this scheme is given by $T_{4.3.1}$ and $T_{4.3.2}$ respectively for i=1,3,...,m-4,m-2 when m is odd.

(v) Truncation Error for the S(AGE) Scheme

If we assume that the number of intervals is even, then as we know from Figure (4.2.4), this scheme entails the alternate use of the GER and the GEL schemes along the vertical direction. Accordingly, its truncation error is given by the truncation errors of the GER and the GEL schemes along the alternate time level. This produces the possible effect of the cancellation of the component error terms at most internal points. A more accurate solution with this scheme is therefore expected than any of the previous GE methods. A similar argument holds when m is odd.

(vi) Truncation Error for the D(AGE) Scheme

The GER, GEL, GEL AND GER methods, in that order, are employed at each of every four time levels. By the same reasoning as above, we will expect this four-step process to be as accurate if not better than the S(AGE) method. In fact, our numerical experiment will show that the D(AGE) procedure can be more superior than the S(AGE) scheme or any of the other GE methods.

4.4 STABILITY ANALYSIS FOR THE GE METHODS

We shall now proceed to establish the stability requirement of the GE methods. From the formulae (4.2.19), (4.2.24), (4.2.33) and (4.2.35), we present below the explicit expressions for the GER, GEL, GEU and the GEC schemes:

$$\underline{u}_{j+1} = (I+\lambda\theta G_1)^{-1}(I-\lambda(1-\theta)G_1)\underline{u}_j + \hat{\underline{b}}_1 , \tag{4.4.1}$$

$$\underline{u}_{j+1} = (I+\lambda\theta G_2)^{-1}(I-\lambda(1-\theta)G_2)\underline{u}_j + \hat{\underline{b}}_2 , \tag{4.4.2}$$

$$\underline{u}_{j+1} = (I+\lambda\theta \hat{G}_1)^{-1}(I-\lambda(1-\theta)\hat{G}_1)\underline{u}_j + \hat{\underline{b}}_3 , \tag{4.4.3}$$

and $$\underline{u}_{j+1} = (I+\lambda\theta \hat{G}_2)^{-1}(I-\lambda(1-\theta)\hat{G}_2)\underline{u}_j . \tag{4.4.4}$$

These processes may be written in a single form as,

$$\underline{u}_{j+1} = \Gamma\underline{u}_j + \widehat{\underline{b}} , \tag{4.4.5}$$

where Γ, the amplification matrix, corresponds to the scheme employed and $\hat{\underline{b}}$ is the relevant column vector of order (m-1) as indicated in the formulae above.

(i) Stability of the GER Scheme

From the equations (4.4.1) and (4.4.5), the GER amplification matrix is given by,

$$\Gamma_{GER} = \left[\begin{array}{cc|cc|c|cc|c}
(1+\lambda) & -\lambda & & & & & & \\
\lambda & (1-\lambda) & & & & & & \\ \hline
 & & (1+\lambda) & -\lambda & & & O & \\
 & & \lambda & (1-\lambda) & & & & \\ \hline
 & & & & \ddots & & & \\ \hline
 & & & & & (1+\lambda) & -\lambda & \\
 & & O & & & \lambda & (1-\lambda) & \\ \hline
 & & & & & & & 1+\dfrac{\lambda}{(1-\lambda\theta)}
\end{array}\right]_{(m-1)\times(m-1)} \tag{4.4.6}$$

with $\lambda\theta \neq 1$. It can be easily shown that Γ_{GER} possesses the eigenvalues 1, of multiplicity (m-2), and $(1+\frac{\lambda}{(1-\lambda\theta)})$. If we denote $\rho(\Gamma_{GER})$ as the spectral radius of Γ_{GER}, then for the stability of the GER scheme, we require $\rho(\Gamma_{GER}) \leq 1$. This implies that

$$\left|1+\frac{\lambda}{(1-\lambda\theta)}\right| \leq 1 \tag{4.4.7}$$

which gives,

$$-2 \leq \frac{\lambda}{(1-\lambda\theta)} \leq 0 . \tag{4.4.8}$$

Since λ is non-negative, then $(1-\lambda\theta)<0$ or

$$\lambda\theta>1 \text{ and } \lambda\leq -2(1-\lambda\theta) . \tag{4.4.9}$$

Different cases of θ are now treated to investigate the condition of stability of the GER scheme.

(a) For $\theta=0$, we have

$\left|1+\frac{\lambda}{(1-\lambda\theta)}\right| = 1+\lambda$ for all positive values of λ. Therefore,

$$\rho(\Gamma_{GER}) > 1,$$

which shows that the GER scheme is always unstable.

(b) For $0<\theta<\frac{1}{2}$, (4.4.9) gives,

$$\lambda\theta>1 \text{ and } \lambda\leq\frac{2}{(2\theta-1)} .$$

The second inequality can never be satisfied since λ is non-negative whilst $(2\theta-1)$ is always negative. Hence, for this particular case of θ, the GER method is always unstable.

(c) For $\theta=\frac{1}{2}$, we obtain,

$$\left|1+\frac{\lambda}{(1-\lambda\theta)}\right| = \left|1+\frac{\lambda}{(1-\frac{1}{2}\lambda)}\right| > 1,$$

and as in case (a), the GER scheme is absolutely unstable.

(d) For $\frac{1}{2}<\theta\leq 1$, (4.4.9) becomes,

$$\lambda\theta>1 \quad \text{and} \quad \lambda\geq\frac{2}{(2\theta-1)}$$

or $$\lambda>\frac{1}{\theta} \quad \text{and} \quad \lambda\geq\frac{1}{(\theta-\frac{1}{2})} \quad .$$

We deduce that, the scheme is conditionally stable for $\lambda\geq\frac{2}{(2\theta-1)}$.

We conclude from the cases (a),(b),(c) and (d) that the GER scheme is:

(1) always unstable for $0\leq\theta\leq\frac{1}{2}$, and

(2) it is conditionally stable for $\lambda\geq\frac{2}{(2\theta-1)}$ when $\theta\in(\frac{1}{2},1]$.

It may therefore be summarised that none of the cases above can really be considered useful either because of their unconditional instability (when $0\leq\theta\leq\frac{1}{2}$) or due to their "inverse" conditional stability (when $\frac{1}{2}<\theta\leq 1$ and which could lead to excessively large time steps).

(ii) Stability of the GEL Scheme

From the equations (4.4.2) and (4.4.5), the GEL amplification matrix is given by,

$$\Gamma_{GEL} = \left[\begin{array}{c|cc|c|cc} 1-\frac{\lambda}{(1+\lambda\theta)} & & & & & \\ \hline & (1+\lambda) & -\lambda & & & \\ & \lambda & (1-\lambda) & & O & \\ \hline & & & \ddots & & \\ \hline & & O & & (1+\lambda) & -\lambda \\ & & & & \lambda & (1-\lambda) \end{array}\right]_{(m-1)\times(m-1)} \tag{4.4.10}$$

The eigenvalues of Γ_{GEL} are 1 (of multiplicity (m-2)) and $(1-\frac{\lambda}{(1+\lambda\theta)})$ and the GEL scheme is stable if $\rho(\Gamma_{GEL})\leq 1$. This requires that

$$\left|1-\frac{\lambda}{(1+\lambda\theta)}\right| \leq 1 \ , \tag{4.4.11}$$

giving,

$$0 \leq \frac{\lambda}{(1+\lambda\theta)} \leq 2. \tag{4.4.12}$$

Since λ is non-negative then from (4.4.12), we must have $(1+\lambda\theta)>0$. Hence,

$$\lambda \leq 2(1+\lambda\theta) \ . \qquad (4.4.13)$$

(a) For $\theta=0$, we have

$$\left|1 - \frac{\lambda}{(1+\lambda\theta)}\right| = |1-\lambda| \ .$$

In order that $\rho(\Gamma_{GEL})\leq 1$, we must have $|1-\lambda|\leq 1$ which is satisfied for $\lambda\leq 2$. Therefore, for this particular case of θ, the condition of stability is $\lambda\leq 2$.

(b) If $0<\theta<\frac{1}{2}$, then from (4.4.13) we obtain $\lambda(1-2\theta)\leq 2$ which leads to the following condition of stability,

$$\lambda \leq \frac{2}{(1-2\theta)} \ .$$

(c) For $\theta=\frac{1}{2}$, we get,

$$\left|1 - \frac{\lambda}{(1+\lambda\theta)}\right| = \left|1 - \frac{\lambda}{(1+\frac{1}{2}\lambda)}\right| < 1 \text{ for every positive value of } \lambda.$$

This implies that the scheme is always stable for $\theta=\frac{1}{2}$.

(d) For $\frac{1}{2}<\theta\leq 1$, inequality (4.4.13) leads to,

$$\lambda \geq \frac{2}{(1-2\theta)} \ .$$

Now, the quantity $2/(1-2\theta)$ is always negative whilst λ is non-negative. Hence, the scheme is absolutely stable for all values of λ. From all the cases above, we conclude that the GEL scheme is:

(1) conditionally stable for $\lambda\leq 2/(1-2\theta)$ with $0\leq\theta<\frac{1}{2}$, and

(2) it is absolutely stable for all values of λ when $\frac{1}{2}\leq\theta\leq 1$.

(iii) Stability of the GEU Scheme

From the equations (4.4.3) and (4.4.5), the GEU amplification matrix is given by,

$$\Gamma_{GEU}=\left[\begin{array}{c|cc|c|cc|c} 1-\frac{\lambda}{(1+\lambda\theta)} & & & & & & \\ \hline & (1+\lambda) & -\lambda & & & & \\ & \lambda & (1-\lambda) & & & & \\ \hline & & & \ddots & & O & \\ \hline & & O & & (1+\lambda) & -\lambda & \\ & & & & \lambda & (1-\lambda) & \\ \hline & & & & & & 1+\frac{\lambda}{(1-\lambda\theta)} \end{array}\right]_{(m-1)\times(m-1)} \qquad (4.4.14)$$

whose eigenvalues are $1-\frac{\lambda}{(1+\lambda\theta)}$, 1 (of multiplicity (m-3)) and $1+\frac{\lambda}{(1-\lambda\theta)}$. Hence we can easily deduce from the conclusions drawn on the stability analysis of the GER and the GEL schemes that the GEU method is conditionally stable for $\lambda \geq \frac{2}{(2\theta-1)}$ when $\theta \in (\frac{1}{2},1]$.

(iv) *Stability of the GEC Scheme*

From the equations (4.4.4) and (4.4.5), the GEC amplification matrix is given by,

$$\Gamma_{GEC}=\left[\begin{array}{cc|cc|c|cc|cc} (1+\lambda) & -\lambda & & & & & & & \\ \lambda & (1-\lambda) & & & & & & & \\ \hline & & (1+\lambda) & -\lambda & & & O & & \\ & & \lambda & (1-\lambda) & & & & & \\ \hline & & & & \ddots & & & & \\ \hline & & & O & & (1+\lambda) & -\lambda & & \\ & & & & & \lambda & (1-\lambda) & & \\ \hline & & & & & & & (1+\lambda) & -\lambda \\ & & & & & & & \lambda & (1-\lambda) \end{array}\right]_{(m-1)\times(m-1)} \qquad (4.4.15)$$

Since Γ_{GEC} has (m-1) eigenvalues, each equals to 1, then clearly the GEC scheme is always stable with no restrictions on λ and $\theta \in [0,1]$.

(v) Stability of the (S)AGE Scheme

We shall first consider the case when m is even. By means of equations (4.2.25), we obtain,

$$\begin{aligned}\underset{\sim}{u}_{j+2} &= (I+\lambda\theta G_2)^{-1}(I-\lambda(1-\theta)G_2)\underset{\sim}{u}_{j+1}+(I+\lambda\theta G_2)^{-1}\underset{\sim}{b}_2 \ ,\\ &= (I+\lambda\theta G_2)^{-1}(I-\lambda(1-\theta)G_2)(I+\lambda\theta G_1)^{-1}(I-\lambda(1-\theta)G_1)\underset{\sim}{u}_j + \underset{\sim}{b}_2',\end{aligned} \tag{4.4.16}$$

i.e., $\underset{\sim}{u}_{j+2} = \Gamma_{SAGE}\underset{\sim}{u}_j + \underset{\sim}{b}_2'$, (4.4.17)

where,

$$\begin{aligned}\Gamma_{SAGE} &= \{(I+\lambda\theta G_2)^{-1}(I-\lambda(1-\theta)G_2)\}\{(I+\lambda\theta G_1)^{-1}(I-\lambda(1-\theta)G_1)\}\\ &= \Gamma_1\Gamma_2\end{aligned} \tag{4.4.18}$$

and $\underset{\sim}{b}_2'$ is the appropriate column vector of order (m-1). We observe that Γ_1 and Γ_2 are exactly the amplification matrices of the GEL (4.4.10) and the GER (4.4.6) schemes respectively. Hence, by multiplying these matrices, we obtain,

$$\Gamma_{SAGE} = \begin{bmatrix} a & b & & & & & & & & \\ c & d & -c & e & & & & & & \\ e & f & d & -f & & & & O & & \\ & & c & d & -c & e & & & & \\ & & e & f & d & -f & & & & \\ & & & & & \ddots & & & & \\ & & & & & c & d & -c & e & \\ & & O & & & e & f & d & -f & \\ & & & & & & & c & d & g \\ & & & & & & & e & f & h \end{bmatrix}_{(m-1)\times(m-1)} \tag{4.4.19}$$

where $a = (1+\lambda)(1-\lambda/(1+\lambda\theta))$,

$b = -\lambda(1-\lambda/(1+\lambda\theta))$,

$c = \lambda(1+\lambda)$, $d = 1-\lambda^2$, $e = \lambda^2$, (4.4.20)

$f = \lambda(1-\lambda)$, $g = -\lambda(1+\lambda/(1-\lambda\theta))$ and $h= (1-\lambda)(1+\lambda/(1-\lambda\theta))$.

Note that diag$(\Gamma_{SAGE})=(a,d,d,\ldots,d,h)$ with d occurring (m-3) times. It is difficult to evaluate directly the eigenvalues of Γ_{SAGE} in a closed form. However, we know from Chapter 1 that if the eigenvalues of Γ_{SAGE} are denoted by μ_i, $i=1,2,\ldots,m-1$, then,

$$\sum_{i=1}^{m-1}\mu_i = \text{tr.}(\Gamma_{SAGE}) \text{ where tr.}(\Gamma_{SAGE}) \text{ is the trace of } \Gamma_{SAGE}$$

which is the sum of the diagonal elements of Γ_{SAGE},

i.e., $\mu_1+\mu_2+\ldots+\mu_{m-1} = a+(m-3)d+h$.

Now if we insist $\rho(\Gamma_{SAGE})\leqslant 1$ it follows that,

$$|\mu_1+\mu_2+\ldots+\mu_{m-1}| = |a+(m-3)d+h| \leqslant |\mu_1|+|\mu_2|+\ldots+|\mu_{m-1}| \leqslant (m-1).$$

Hence we seek the values of λ such that,

$$|a+(m-3)d+h| \leqslant |a|+(m-3)|d|+|h| \leqslant (m-1) ,$$

i.e., $(1+\lambda)|(1-\lambda/(1+\lambda\theta)|+(m-3)|(1-\lambda^2)|+|(1-\lambda)|(1+\lambda/(1-\lambda\theta))| \leqslant (m-1)$.

Let $\phi(\lambda) = (1+\lambda)|(1-\lambda/(1+\lambda\theta))|+(m-3)|(1-\lambda^2)|+|(1-\lambda)||(1+\lambda/(1-\lambda\theta))|$.

$\phi(\lambda)$ is non-negative and if $\lambda\leqslant 1$ for $\theta \in [0,1]$, $\lambda\theta\neq 1$ we find that $\phi(\lambda)$ will be a continuous function of λ,

i.e. $\phi(\lambda) = (1+\lambda)(1-\lambda/(1+\lambda\theta))+(m-3)(1-\lambda^2)+(1-\lambda)(1+\lambda/(1-\lambda\theta))$.

$\phi(\lambda)$ attains its greatest value of (m-1) at $\lambda=0$ and $\phi(\lambda)<m-1$ in the range $0<\lambda\leqslant 1$.

Therefore, if, $\rho(\Gamma_{SAGE})\leqslant 1$, then $\lambda\leqslant 1$, for $\theta\in[0,1]$, $\lambda\theta\neq 1$.

Now suppose that we form the sequence $\Gamma^2_{SAGE},\Gamma^3_{SAGE},\ldots$ It is observed that the entries of the product Γ^k_{SAGE} contain combinations of powers in $\lambda,(1-\lambda)$ and $(1-\lambda/(1+\lambda\theta))$. Hence if $\lambda\leqslant 1$, $\lim_{k\to\infty}\Gamma^k_{SAGE}=0$ which implies that Γ_{SAGE} is convergent. A necessary and sufficient condition for this to be so is $\rho(\Gamma_{SAGE})<1$. We conclude that the (S)AGE scheme is stable for $\lambda\leqslant 1$.

When m is odd, the equations constituting the (S)AGE procedure are given by (4.2.36). This time, however, the amplification matrix Γ_{SAGE} is the product of the amplification matrices of the GEC and the GEU schemes and it takes the form,

$$\Gamma_{SAGE}=\begin{bmatrix}
a & -c & e & & & & & & & \\
-b & d & -f & & & & & & & \\
 & c & d & -c & e & & & & & \\
 & e & f & d & -f & & & & & \\
 & & & c & d & -c & e & & \bigcirc & \\
 & & & e & f & d & -f & & & \\
 & & & & & & \ddots & & & \\
 & \bigcirc & & & & c & d & -c & e & \\
 & & & & & e & f & d & -f & \\
 & & & & & & & c & d & g \\
 & & & & & & & e & f & h
\end{bmatrix}_{(m-1)\times(m-1)} . \qquad (4.4.21)$$

The matrix has the same diagonal elements as in (4.4.19) and clearly the (S)AGE scheme for the odd case is also conditionally stable for $\lambda\leqslant 1$.

(vi) Stability of the D(AGE) Scheme

The four-step process of (4.2.26) can also be written in the form,

$$\underline{u}_{j+4} = \Gamma_{DAGE}\, \underline{u}_j + \underline{b}_1' \quad , \qquad (4.4.22)$$

where the amplification matrix Γ_{DAGE} is given by,

$$\Gamma_{DAGE} = \{(I+\lambda\theta G_1)^{-1}(I-\lambda(1-\theta)G_1)\}\{(I+\lambda\theta G_2)^{-1}(I-\lambda(1-\theta)G_2\}\{(I+\lambda\theta G_2)^{-1}$$
$$(I-\lambda(1-\theta)G_2)(I+\lambda\theta G_1)^{-1}(I-\lambda(1-\theta)G_1)\}$$
$$= \Gamma_{GER}\Gamma_{GEL}\Gamma_{SAGE} \qquad (4.4.23)$$

and $\underline{b}_1'$ is the relevant column vector of order (m-1). By multiplying the matrices in (4.4.23), we find that,

$$|p+t+\tfrac{1}{2}(m-4)w+\tfrac{1}{2}(m-4)z+p_1| \leqslant |p|+|t|+\tfrac{1}{2}(m-4)|w|+\tfrac{1}{2}(m-4)|z|+|p_1|$$

and we seek values of λ such that

$$|p|+|t|+\tfrac{1}{2}(m-4)|w|+\tfrac{1}{2}(m-4)|z|+|p_1| \leqslant m-1.$$

Now, $|p|+|t|+\tfrac{1}{2}(m-4)|w|+\tfrac{1}{2}(m-4)|z|+|p_1|$

$$\begin{aligned}
&= |(1+\lambda)^2(1-\frac{\lambda}{(1+\lambda\theta)})^2 -\lambda^2(1+2\lambda)|+|-\lambda^2(1-\lambda(1+\lambda\theta))^2+(1-\lambda)^2(1+2\lambda)| \\
&\quad +\tfrac{1}{2}(m-4)|(1+\lambda)^2(1-2\lambda)-\lambda^2(1+2\lambda)|+\tfrac{1}{2}(m-4)|(1-\lambda)^2(1+2\lambda)-\lambda^2(1-2\lambda)| \\
&\quad +|(1-2\lambda)(1+\lambda(1-\lambda\theta))^2|, \\
&\leqslant |(1+\lambda)^2(1-\lambda/(1+\lambda\theta))^2|+|\lambda^2(1+2\lambda)|+|-\lambda^2(1-\lambda/(1+\lambda\theta))^2|+ \\
&\qquad\qquad |(1-\lambda)^2(1+2\lambda)| \\
&\quad +\tfrac{1}{2}(m-4)\{|(1+\lambda)^2(1-2\lambda)|+|\lambda^2(1+2\lambda)|\}+\tfrac{1}{2}(m-4)\{|(1-\lambda)^2(1+2\lambda)| \\
&\quad +|\lambda^2(1-2\lambda)|\}+|(1-2\lambda)(1+\lambda/(1-\lambda\theta))^2|, \\
&= \psi(\lambda).
\end{aligned}$$

$\psi(\lambda)$ is non-negative and if $\lambda \leqslant \tfrac{1}{2}$ for $\theta \in [0,1]$ we observe that $\psi(\lambda)$ is a continuous function of λ,

i.e.
$$\begin{aligned}
\psi(\lambda) &= (1+\lambda)^2(1-\lambda/(1+\lambda\theta))^2+\lambda^2(1+2\lambda)+\lambda^2(1-\lambda/(1+\lambda\theta))^2+(1-\lambda)^2(1+2\lambda) \\
&\quad +\tfrac{1}{2}(m-4)\{(1+\lambda)^2(1-2\lambda)+\lambda^2(1+2\lambda)\}+\tfrac{1}{2}(m-4)\{(1-\lambda)^2(1+2\lambda)+\lambda^2(1-2\lambda)\} \\
&\quad +(1-2\lambda)(1+\lambda/(1-\lambda\theta))^2 \\
&= (\lambda^2+(1+\lambda)^2)(1-\lambda/(1+\lambda\theta))^2+(\lambda^2+(1-\lambda)^2)(1+2\lambda)+(m-4)(1-2\lambda^2) \\
&\quad +(1-2\lambda)(1+\lambda/(1-\lambda\theta))^2.
\end{aligned}$$

Our problem is now reduced to seeking λ such that $\psi(\lambda) \leqslant (m-1)$. $\psi(\lambda)$ achieves its greatest value of $(m-1)$ at $\lambda=0$ and in the range $0<\lambda \leqslant \tfrac{1}{2}$, $\psi(\lambda)<m-1$. Therefore, if $\rho(\Gamma_{DAGE}) \leqslant 1$ then $\lambda \leqslant \tfrac{1}{2}$ for $\theta \in [0,1]$.

For convenience, let us replace $\Gamma_{GEL}, \Gamma_{GER}, \Gamma_{SAGE}$ and Γ_{DAGE} by $\Gamma_1, \Gamma_2, \Gamma_3$ and Γ_4 respectively. We now construct the sequence of matrices $\Gamma_4, \Gamma_4^2, \Gamma_4^3, \ldots, \Gamma_4^r, \ldots$. Consider,

$$\Gamma_{DAGE} = \begin{bmatrix} p & q & r & s & & & & & & & & & \\ -q & t & u & v & & & & & & & & & \\ r & -u & w & x & r & s & & & & & & & \\ -s & v & -x & z & u & v & & & & & & O & \\ & & r & -u & w & x & r & s & & & & & \\ & & -s & v & -x & z & u & v & & & & & \\ & & & & & & \ddots & & \ddots & & & & \\ & & & & & & r & -u & w & x & r & s & \\ & & & & & & -s & v & -x & z & u & v & \\ & & O & & & & & & r & -u & w & x & y \\ & & & & & & & & -s & v & -x & z & q_1 \\ & & & & & & & & & & y & -q_1 & p_1 \end{bmatrix}_{(m-1)\times(m-1)} \qquad (4.4.24)$$

where

$$p = (1+\lambda)^2(1-\frac{\lambda}{(1+\lambda\theta)})^2-\lambda^2(1+2\lambda),\ q = -(\lambda(1+\lambda)(1-\lambda/(1+\lambda\theta))^2+\lambda(1-\lambda)(1+2\lambda)),$$

$$r = 2(1+\lambda)\lambda^2,\ s = -2\lambda^3,\ t = -\lambda^2(1-\lambda/(1+\lambda\theta))^2+(1-\lambda)^2(1+2\lambda),$$

$$u = -2\lambda(1-\lambda^2),\ v = 2(1-\lambda)\lambda^2\ ,\ w = (1+\lambda)^2(1-2\lambda)-\lambda^2(1+2\lambda)\ , \qquad (4.4.25)$$

$$x = 2\lambda(2\lambda^2-1),\ y = 2\lambda^2(1+\lambda/(1-\lambda\theta)),\ z = (1-\lambda)^2(1+2\lambda)-\lambda^2(1-2\lambda),$$

$$q_1 = -2\lambda(1-\lambda)(1+\lambda/(1-\lambda\theta)),\ \ p_1 = (1-2\lambda)(1+\lambda/(1-\lambda\theta))^2.$$

Now, diag.$(\Gamma_{DAGE}) = (p,t,w,z,w,z,\ldots,w,z,p_1)$ with w,z each occurring alternately for (m-4)/2 times.

If μ_i, i=1,...,m-1 represents the eigenvalues of Γ_{DAGE} then

$$\sum_{i=1}^{m-1}\mu_i = \text{tr.}(\Gamma_{DAGE})\ ,$$

i.e., $\mu_1+\mu_2+\ldots+\mu_{m-1} = p+t+\frac{1}{2}(m-4)w+\frac{1}{2}(m-4)z+p_1$

$$\therefore\quad |\mu_1+\mu_2+\ldots+\mu_{m-1}| = |p+t+\tfrac{1}{2}(m-4)w+\tfrac{1}{2}(m-4)z+p_1|$$
$$\leq |\mu_1|+|\mu_2|+\ldots+|\mu_{m-1}|\quad .$$

If we require that $\rho(\Gamma_{DAGE})\leq 1$ then we have,

$$\begin{aligned}\Gamma_4 &= \Gamma_2\Gamma_1\Gamma_3 \text{ , (from equation (4.4.23))}\\ &= \Gamma_2\Gamma_1(\Gamma_1\Gamma_2) \text{ ,}\\ &= \Gamma_2(\Gamma_1^2)\Gamma_2 .\end{aligned}$$

Hence,

$$\begin{aligned}\Gamma_4^2 &= (\Gamma_2\Gamma_1^2\Gamma_2)(\Gamma_2\Gamma_1\Gamma_3) \text{ ,}\\ &= \Gamma_2\Gamma_1^2\Gamma_2^2\Gamma_1(\Gamma_1\Gamma_2) \text{ ,}\\ &= \Gamma_2(\Gamma_1^2\Gamma_2^2)\Gamma_1^2\Gamma_2 \text{ .}\end{aligned}$$

Similarly, we also have,

$$\begin{aligned}\Gamma_4^3 &= \Gamma_2(\Gamma_1^2\Gamma_2^2)^2\Gamma_1^2\Gamma_2 \text{ ,}\\ \Gamma_4^4 &= \Gamma_2(\Gamma_1^2\Gamma_2^2)^3\Gamma_1^2\Gamma_2 \text{ ,}\end{aligned}$$

and continuing in this manner, we find that,

$$\Gamma_4^r = \Gamma_2(\Gamma_1^2\Gamma_2^2)^{r-1}\Gamma_1^2\Gamma_2 \text{ .} \qquad (4.4.26)$$

Combinations of powers in 2λ, $(1-2\lambda)$ and $(1-\frac{\lambda}{(1+\lambda\theta)})$ appear in the entries of $(\Gamma_1^2\Gamma_2^2)^r$. Therefore, if $\lambda\leq\frac{1}{2}$ then $\lim\limits_{r\to\infty}(\Gamma_1^2\Gamma_2^2)^r=0$ and from (4.4.26), $\lim\limits_{r\to\infty}\Gamma_4^r=0$, the null matrix.

Hence the sequence $\Gamma_4,\Gamma_4^2,\Gamma_4^3,\ldots$ converges and a necessary and sufficient condition for this to be so is $\rho(\Gamma_4)<1$.

We conclude that the DAGE scheme is conditionally stable for $\lambda\leq\frac{1}{2}$.

The corresponding amplification matrix for the (D)AGE scheme when m is odd is found to be,

$$\Gamma_{DAGE} = \left[\begin{array}{cccccccccccc}
p_2 & q_2 & r_1 & & & & & & \bigcirc & & & \\
c_1 & t_1 & x & r & s & & & & & & & \\
r_2 & -x & w_1 & u & v & & & & & & & \\
 & r & -u & t_1 & x & r & s & & & & & \\
 & -s & v & -x & w_1 & u & v & & & & & \\
 & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & & \\
 & & \bigcirc & & & r & -u & t_1 & x & r & s & \\
 & & & & & -s & v & -x & w_1 & u & v & \\
 & & & & & & & r & -u & t_1 & x & y \\
 & & & & & & & -s & v & -x & w_1 & q_1 \\
 & & & & & & & & & y & -q_1 & p_1
\end{array}\right]_{(m-1)\times(m-1)} \qquad (4.4.27)$$

where $p_2 = (1+2\lambda)(1-\frac{\lambda}{(1+\lambda\theta)})^2$, $q_2 = -2\lambda(1+\lambda)(1-\frac{\lambda}{(1+\lambda\theta)})$, $r_1 = 2\lambda^2(1-\frac{\lambda}{(1+\lambda\theta)})$,

$c_1 = 2\lambda(1+\lambda)(1-\frac{\lambda}{(1+\lambda\theta)})$, $t_1 = 1-4\lambda^2(1+\lambda)$, $w_1 = 1+4\lambda^2(\lambda-1)$,

$r_2 = 2\lambda^2(1-\frac{\lambda}{(1+\lambda\theta)})$ (4.4.28)

and the other entries of the matrix take on the same values as in (4.4.25). As previously, it can be established that the scheme has conditional stability for $\lambda \leq \frac{1}{2}$.

4.5 APPLICATION OF THE GE METHODS TO A MORE GENERAL FIRST ORDER EQUATION

Our discussion on the GE methods easily carries over to the case of solving a more general equation of the form,

$$\frac{\partial U}{\partial t} + \frac{\partial U}{\partial x} = k(x,t) \ . \tag{4.5.1}$$

The basic implicit formulae defining the GE schemes now become,

$$(1+\lambda\theta)u_{i,j+1}-\lambda\theta u_{i-1,j+1} = (1-\lambda(1-\theta))u_{i,j}+\lambda(1-\theta)u_{i-1,j}+\Delta tk_{i,j+\theta} \tag{4.5.2}$$

and $$\lambda\theta u_{i,j+1}+(1-\lambda\theta)u_{i-1,j+1} = -\lambda(1-\theta)u_{ij}+(1+\lambda(1-\theta))u_{i-1,j}+\Delta tk_{i-1,j+\theta} \tag{4.5.3}$$

from which the following set of explicit equations determining the solutions at the grouped points are derived,

$$u_{i-1,j+1} = (1+\lambda)u_{i-1,j}-\lambda u_{ij}+\Delta t((1+\lambda\theta)k_{i-1,j+\theta}-\lambda\theta k_{i,j+\theta}) \tag{4.5.4}$$

and $$u_{i,j+1} = \lambda u_{i-1,j}+(1-\lambda)u_{ij}+\Delta t(\lambda\theta k_{i-1,j+\theta}+(1-\lambda\theta)k_{i,j+\theta}) . \tag{4.5.5}$$

The equations describing the u-values at the left and right ungrouped points are given respectively by,

$$u_{1,j+1} = [\lambda(1-\theta)u_{0,j}+(1-\lambda(1-\theta))u_{1j}+\lambda\theta u_{0,j+1}+\Delta tk_{1,j+\theta}]/(1+\lambda\theta) \tag{4.5.6}$$

and $$u_{m-1,j+1} = [(1+\lambda(1-\theta))u_{m-1,j}-\lambda(1-\theta)u_{mj}-\lambda\theta u_{m,j+1}+\Delta tk_{m-1,j+\theta}]/(1-\lambda\theta) , \quad \lambda\theta\neq 1 . \tag{4.5.7}$$

The one-step, two-step and four-step processes are then developed in exactly the same manner as before.

4.6 GE METHODS FOR THE SPATIALLY-CENTRED APPROXIMATION TO THE FIRST ORDER EQUATION

Let us now consider the hyperbolic equation of first order of the form,

$$\frac{\partial U}{\partial t} + \frac{\partial U}{\partial x} = 0 \ , \qquad (4.6.1)$$

If we approximate the time and spatial derivatives by the forward and central difference formulae respectively at the point (x_i,t_j) we obtain,

$$\frac{\partial U}{\partial t} = \frac{U_{i,j+1}-U_{i,j}}{\Delta t} + O(\Delta t) \ , \qquad (4.6.2)$$

and $$\frac{\partial U}{\partial x} = \frac{U_{i+1,j}-U_{i-1,j}}{2\Delta x} + O([\Delta x]^2) \ , \qquad (4.6.3)$$

Now, by using the Taylor's series about the point (x_{i+1},t_j) we have,

$$U_{i+1,j+1} = U_{i+1,j}+(\Delta t)(\frac{\partial U}{\partial x})_{i+1,j} + O([\Delta t]^2) \qquad (4.6.4)$$

or

$$U_{i+1,j} = U_{i+1,j+1}-(\Delta t)(\frac{\partial U}{\partial x})_{i+1,j}+O([\Delta t]^2) \ . \qquad (4.6.5)$$

If we substitute this expression into (4.6.3), we get

$$\frac{\partial U}{\partial x} = \frac{[U_{i+1,j+1}-(\Delta t)(\frac{\partial U}{\partial x})_{i+1,j}-U_{i-1,j}]}{2\Delta x} + O(\frac{[\Delta t]^2}{\Delta x})+O([\Delta x]^2) . \qquad (4.6.6)$$

By virtue of (4.6.1), equation (4.6.6) together with the equation (4.6.2) leads to the following finite difference analogue,

$$\frac{u_{i,j+1}-u_{i,j}}{\Delta t} = \frac{-(u_{i+1,j+1}-u_{i-1,j})}{2\Delta x} \ ,$$

or

$$u_{i,j+1}+ru_{i+1,j+1} = u_{ij}+ru_{i-1,j} \ , \qquad (4.6.7)$$

where we have assumed the consistency relation $\frac{\Delta t}{\Delta x} \to 0$ as $\Delta t,\Delta x \to 0$ and $r=\frac{1}{2}\lambda=\frac{1}{2}\frac{\Delta t}{\Delta x}$. The formula (4.6.7) has the following computational molecule,

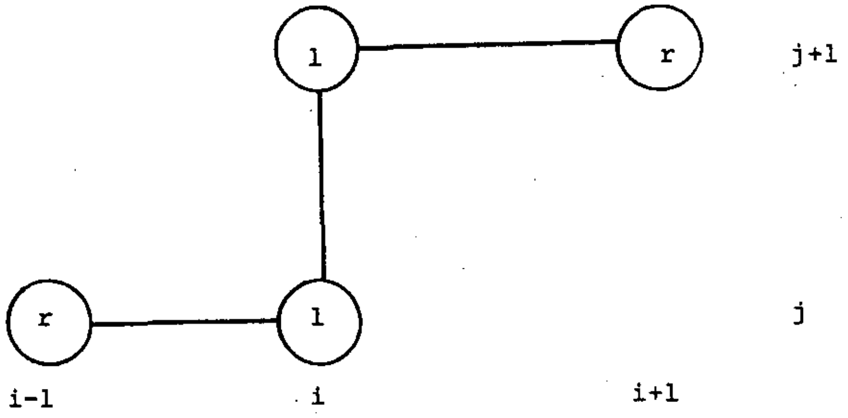


FIGURE 4.6.1

and the approximation is known as the RL (right to left) type since its computation at the mesh points proceeds from the right boundary. Similarly, if we reverse the above procedure, we obtain the following LR approximation,

$$ru_{i-1,j+1}+u_{i,j+1} = u_{ij} + ru_{i+1,j} \qquad (4.6.8)$$

and its computational molecule is given by,

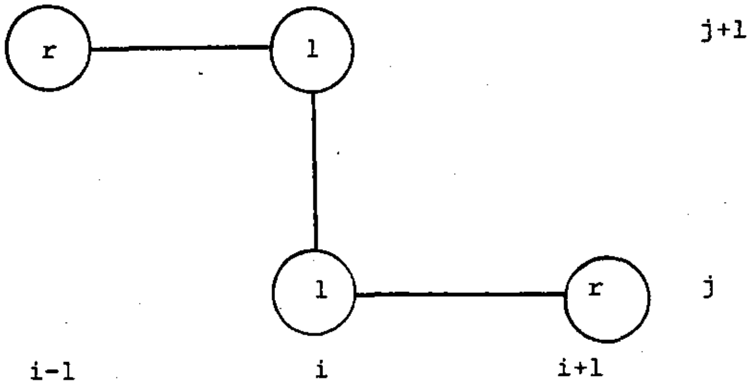


FIGURE 4.6.2

The local truncation error for equation (4.6.7) is obtained from the following Taylor series expansion about the point $(x_i,t_{j+\frac{1}{2}})$:

$$T_{4.6.7} = [(\frac{\partial U}{\partial t} + \frac{\partial U}{\partial x})+\frac{1}{2}\frac{\Delta t}{\Delta x}\frac{\partial U}{\partial t} + \frac{1}{8.3!}\frac{(\Delta t)^3}{\Delta x}\frac{\partial^3 U}{\partial t^3} + \frac{1}{3!}(\Delta x)^2\frac{\partial^3 U}{\partial x^3} + \frac{3}{2.3!}$$

$$(\Delta t)(\Delta x)\ \frac{\partial^3 U}{\partial x^2 \partial t} - \frac{3}{4.3!}(\Delta t)^2\ \frac{\partial^3 U}{\partial x \partial t^2} + \frac{1}{4.3!}(\Delta t)^2\ \frac{\partial^3 U}{\partial t^3} + \frac{1}{5!}\ (\Delta x)^4$$
$$\frac{\partial^5 U}{\partial x^5} + \frac{5}{2.5!}\ (\Delta x)^3(\Delta t)\ \frac{\partial^5 U}{\partial x^4 \partial t} + \frac{5}{2.5!}\ (\Delta x)^2(\Delta t)^2\ \frac{\partial^5 U}{\partial x^3 \partial t^2} + \frac{5}{4.5!}(\Delta x)(\Delta t)\frac{\partial^5 U}{\partial x^2 \partial t}$$
$$+\frac{5}{16.5!}(\Delta t)^4\ \frac{\partial^5 U}{\partial x \partial t^4} + \frac{1}{32.5!}\ \frac{(\Delta t)^5}{\Delta x}\ \frac{\partial^5 U}{\partial t^5} + \ldots]_{i,j+\frac{1}{2}}$$

i.e.,

$$T_{4.6.7} = \frac{\Delta t}{\Delta x}[\tfrac{1}{2}\ \frac{\partial U}{\partial t} + \frac{1}{48}(\Delta t)^2\ \frac{\partial^3 U}{\partial t^3} + \frac{1}{4}(\Delta x)^2\ \frac{\partial^3 U}{\partial x^2 \partial t}]_{i,j+\frac{1}{2}} + \frac{1}{6}(\Delta x)^2(\frac{\partial^3 U}{\partial x^3})_{i,j+\frac{1}{2}}$$
$$+(\Delta t)^2[\frac{1}{24}\ \frac{\partial^3 U}{\partial t^3} - \frac{1}{8}\ \frac{\partial^3 U}{\partial x \partial t^2}]_{i,j+\frac{1}{2}} + \frac{1}{(\Delta x)}\ O((\Delta x)^{\alpha_1}(\Delta t)^{\alpha_2}),$$
$$\alpha_1+\alpha_2=5; \qquad (4.6.9)$$

and the truncation error for (4.6.8) is given by,

$$T_{4.6.8} = [(\frac{\partial U}{\partial t} - \frac{\partial U}{\partial x})+\tfrac{1}{2}\ \frac{\Delta t}{\Delta x}\ \frac{\partial U}{\partial t} - \frac{1}{3}(\Delta t)^2\ \frac{\partial^3 U}{\partial x^3} + \frac{3}{2.3!}\ (\Delta t)(\Delta x)\ \frac{\partial^3 U}{\partial x^2 \partial t} - \frac{3}{4.3!}$$
$$(\Delta t)^2\ \frac{\partial^3 U}{\partial x \partial t^2} +(\frac{1}{4.3!}\ (\Delta t)^2 + \frac{1}{8.3!}\ \frac{(\Delta t)^3}{\Delta x})\frac{\partial^3 U}{\partial t^3} - \frac{1}{5!}(\Delta x)^4\ \frac{\partial^5 U}{\partial x^5} + \frac{5}{2.5!}$$
$$(\Delta x)^3(\Delta t)\frac{\partial^5 U}{\partial x^4 \partial t} - \frac{5}{2.5!}(\Delta x)^2(\Delta t)^2\frac{\partial^5 U}{\partial x^3 \partial t^2} + \frac{5}{4.5!}(\Delta x)(\Delta t)^3\ \frac{\partial^5 U}{\partial x^2 \partial t^3}$$
$$-\ (\frac{1}{16.5!}\ (\Delta t)^4 + \frac{1}{32.5!}\ \frac{(\Delta t)^5}{\Delta x})\ \frac{\partial^5 U}{\partial t^5} + \ \ldots]_{i,j+\frac{1}{2}}$$

i.e.,

$$T_{4.6.8} = (\frac{\partial U}{\partial t} - \frac{\partial U}{\partial x})_{i,j+\frac{1}{2}} + \frac{\Delta t}{\Delta x}(\tfrac{1}{2}\ \frac{\partial U}{\partial t} + \frac{1}{48}\ (\Delta t)^2\ \frac{\partial^3 U}{\partial t^3} + \frac{1}{4}\ (\Delta x)^2\ \frac{\partial^3 U}{\partial x^2 \partial t})_{i,j+\frac{1}{2}}$$
$$-\ \frac{1}{6}(\Delta x)^2\ (\frac{\partial^3 U}{\partial x^3})_{i,j+\frac{1}{2}} + (\Delta t)^2\ (\frac{1}{24}\ \frac{\partial^3 U}{\partial t^3} - \frac{1}{8}\ \frac{\partial^3 U}{\partial x \partial t^2})_{i,j+\frac{1}{2}} +$$
$$\frac{1}{\Delta x}\ O((\Delta x)^{\alpha_1}(\Delta t)^{\alpha_2})\ ; \quad \alpha_1+\alpha_2 = 5\ , \qquad (4.6.10)$$

As previously, the two equations in implicit form, i.e. equations (4.6.7) and (4.6.8) can be coupled to produce the following set of explicit equations,

$$\begin{bmatrix} u_{i-1,j+1} \\ \\ u_{i,j+1} \end{bmatrix} = \frac{1}{(1-r^2)} \begin{bmatrix} u_{i-1,j}-ru_{ij}+ru_{i-2,j}-r^2u_{i+1,j} \\ \\ -ru_{i-1,j}+u_{ij}-r^2u_{i-2,j}+ru_{i+1,j} \end{bmatrix}$$

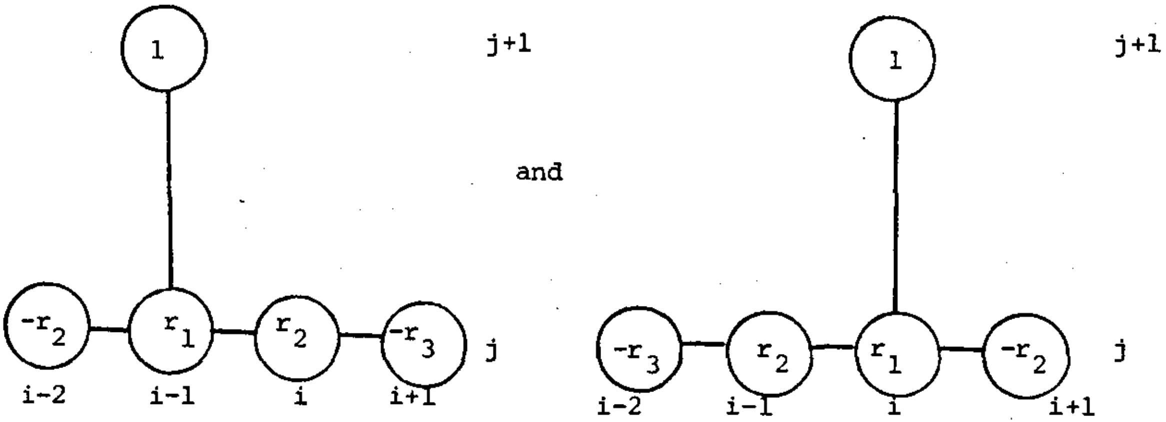
or

$$u_{i-1,j+1} = r_1u_{i-1,j}+r_2(u_{ij}-u_{i-2,j})-r_3u_{i+1,j} \tag{4.6.11}$$

and

$$u_{i,j+1} = r_1u_{ij}+r_2(u_{i-1,j}-u_{i+1,j})-r_3u_{i-2,j} \tag{4.6.12}$$

where $r_1 = \frac{1}{(1-r^2)}$, $r_2 = \frac{-r}{(1-r^2)}$ and $r_3 = \frac{r^2}{(1-r^2)}$ with $r\neq 1$. The corresponding computational molecules are



Equation (4.6.11) Equation (4.6.12)

FIGURE 4.6.3

The solution at the right ungrouped point is given by equation (4.6.7) with i=m-1,i.e.,

$$u_{m-1,j+1} = -ru_{m,j+1}+u_{m-1,j}+ru_{m-2,j} \ , \tag{4.6.13}$$

whilst the solution at the ungrouped point near the left boundary is determined by equation (4.6.8) with i=1, i.e.,

$$u_{1,j+1} = -ru_{0,j+1}+u_{1j}+ru_{2j} \ . \tag{4.6.14}$$

Without loss of generality, we shall consider only the GER, GEL, S(AGE) and D(AGE) schemes of the GE class of methods. The truncation errors and stability of these methods will be investigated in some detail.

(i) The GER Scheme

From (4.6.7), (4.6.8) and (4.6.13), the implicit equations constituting the GER scheme are given by,

$$u_{i-1,j+1}+ru_{i,j+1} = u_{i-1,j}+ru_{i-2,j}\ ,$$
$$ru_{i-1,j+1}+u_{i,j+1} = u_{i,j}+ru_{i+1,j}\ , \quad i=2,4,\ldots,(m-2)$$

and $$u_{m-1,j+1} = -ru_{m,j+1}+u_{m-1,j}+ru_{m-2,j}\ ,$$

which, in matrix form, can be written as,

$$\left[\begin{array}{cc|cc|c|cc|c} 1 & r & & & & & & \\ r & 1 & & & & & & \\ \hline & & 1 & r & & & & \\ & & r & 1 & & & O & \\ \hline & & & & \ddots & & & \\ \hline & & & O & & 1 & r & \\ & & & & & r & 1 & \\ \hline & & & & & & & 1 \end{array}\right] \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_{m-3} \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j+1} =$$

$$\left[\begin{array}{c|cc|c|cc} 1 & & & & & \\ \hline & 1 & r & & & \\ & r & 1 & & & \\ \hline & & & \ddots & O & \\ & & O & & & \\ \hline & & & & 1 & r \\ & & & & r & 1 \end{array}\right] \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j} + \underline{b}_1\ , \qquad (4.6.15)$$

where $\underline{b}_1 = (ru_{0j},0,\ldots,-ru_{m,j+1})^T$. The matrix equation (4.6.15) can be represented in the more compact form by,

$$(I+rG_1)\underline{u}_{j+1} = (I+rG_2)\underline{u}_j + \underline{b}_1 \; , \qquad (4.6.16)$$

where,

$$G_1 = \left[\begin{array}{cc|cc|c|cc|c} 0 & 1 & & & & & & \\ 1 & 0 & & & & & & \\ \hline & & 0 & 1 & & & \bigcirc & \\ & & 1 & 0 & & & & \\ \hline & & & & \ddots & & & \\ \hline & & & & & 0 & 1 & \\ & & \bigcirc & & & 1 & 0 & \\ \hline & & & & & & & 0 \end{array}\right] = \left[\begin{array}{c|c|c|c|c} G^{(1)} & & & & \\ \hline & G^{(2)} & & \bigcirc & \\ \hline & & \ddots & & \\ \hline & \bigcirc & & G^{\frac{1}{2}(m-2)} & \\ \hline & & & & 0 \end{array}\right]$$

and

$$G_2 = \left[\begin{array}{c|c|c|c|c} 0 & & & & \\ \hline & G^{(1)} & & & \\ \hline & & G^{(2)} & & \bigcirc \\ \hline & & & \ddots & \\ \hline & & \bigcirc & & G^{\frac{1}{2}(m-2)} \end{array}\right] \qquad (4.6.17)$$

and $G^{(i)} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $i=1,2,\ldots,\frac{1}{2}(m-2)$; m being even.

(ii) *The GEL Scheme*

The implicit equations for the GEL scheme are given by,

$$u_{1,j+1} = u_{1j}+ru_{2j}-ru_{0,j+1} \; ,$$

$$u_{i-1,j+1}+ru_{i,j+1} = u_{i-1,j}+ru_{i-2,j} \; ,$$

$$ru_{i-1,j+1}+u_{i,j+1} = u_{ij}+ru_{i+1,j} \; , \quad i=3,5,\ldots,(m-1)$$

or in matrix form as,

$$\left[\begin{array}{c|cc|cc|c|cc} 1 & & & & & & & \\ \hline & 1 & r & & & & & \\ & r & 1 & & & & & \\ \hline & & & 1 & r & & & \\ & & & r & 1 & & \bigcirc & \\ \hline & & & & & \ddots & & \\ \hline & & & \bigcirc & & & 1 & r \\ & & & & & & r & 1 \end{array}\right] \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ \vdots \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j+1} =$$

$$\left[\begin{array}{cc|cc|c|cc|c} 1 & r & & & & & & \\ r & 1 & & & & & & \\ \hline & & 1 & r & & & & \\ & & r & 1 & & \bigcirc & & \\ \hline & & & & \ddots & & & \\ \hline & & \bigcirc & & & 1 & r & \\ & & & & & r & 1 & \\ \hline & & & & & & & 1 \end{array}\right] \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ \vdots \\ u_{m-3} \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j} + \underset{\sim}{b}_2 \ , \qquad (4.6.18)$$

where, $\underset{\sim}{b}_2 = (-ru_{0,j+1},0,0,\ldots,ru_{mj})^T$, i.e.,

$$(I+rG_2)\underset{\sim}{u}_{j+1} = (I+rG_1)\underset{\sim}{u}_j + \underset{\sim}{b}_2 \ . \qquad (4.6.19)$$

(iii) *The (S)AGE and (D)AGE Schemes*

The two-step (S)AGE scheme is given by the following equations,

$$\left.\begin{array}{ll} (I+rG_1)\underset{\sim}{u}_{j+1} = (I+rG_2)\underset{\sim}{u}_j + \underset{\sim}{b}_1 \ , & \\ & j=0,2,4,\ldots \\ \text{and} \quad (I+rG_2)\underset{\sim}{u}_{j+2} = (I+rG_1)\underset{\sim}{u}_{j+1} + \underset{\sim}{b}_2 \ , & \end{array}\right\} \qquad (4.6.20)$$

while the four-step (D)AGE method is computed from

and

$$\left.\begin{aligned}(I+rG_1)\underline{u}_{j+1} &= (I+rG_2)\underline{u}_j+\underline{b}_1\\(I+rG_2)\underline{u}_{j+2} &= (I+rG_1)\underline{u}_{j+1}+\underline{b}_2\\(I+rG_2)\underline{u}_{j+3} &= (I+rG_1)\underline{u}_{j+2}+\underline{b}_2\\(I+rG_1)\underline{u}_{j+4} &= (I+rG_2)\underline{u}_{j+3}+\underline{b}_1\end{aligned}\right\}. \quad j=0,4,8,\ldots \qquad (4.6.21)$$

4.7 TRUNCATION ERROR ANALYSIS FOR THE GE METHODS

(i) Truncation Error for the GER Scheme

From (4.6.16), the explicit form of the GER scheme is given by,

$$\underline{u}_{j+1} = (I+rG_1)^{-1}(I+rG_2)\underline{u}_j + \hat{\underline{b}}_1 \ , \tag{4.7.1}$$

where $\hat{\underline{b}}_1=(I+rG_1)^{-1}\underline{b}_1$. The matrix representation (4.7.1) consists of the following equations,

$$u_{i-1,j+1} = r_1u_{i-1,j}+r_2(u_{ij}-u_{i-2,j})-r_3u_{i+1,j} \ , \tag{4.7.2}$$

$$u_{i,j+1} = r_1u_{ij}+r_2(u_{i-1,j}-u_{i+1,j})-r_3u_{i-2,j} \ , \text{ for } i=2,4,\ldots,(m-2), \tag{4.7.3}$$

and $$u_{m-1,j+1} = -ru_{m,j+1}+u_{m-1,j}+ru_{m-2,j} \ . \tag{4.7.4}$$

The truncation error for (4.7.2) is obtained by expanding the terms $U_{i-1,j+1}$, U_{ij}, $U_{i-2,j}$, $U_{i-1,j}$ and $U_{i+1,j}$ using the Taylor's series about the point $((i-1)\Delta x,(j+\frac{1}{2})\Delta t)$:

$$\begin{aligned}T_{4.7.2} = {} & (\Delta t)(\frac{\partial U}{\partial t})_{(i-1,j+\frac{1}{2})} + \frac{(\Delta x)(\Delta t)}{(\Delta t-2\Delta x)}[-2\frac{\partial U}{\partial x} - \frac{1}{4}(\Delta t)^2\frac{\partial^3 U}{\partial x\partial t^2} + \frac{1}{24} \\ & (\Delta t)^3\frac{\partial^4 U}{\partial x\partial t^3}]_{(i-1,j+\frac{1}{2})} + \frac{(\Delta x)(\Delta t)^2}{(2\Delta x+\Delta t)}(\frac{\partial^2 U}{\partial x\partial t})_{(i-1,j+\frac{1}{2})} + \frac{(\Delta t)^3}{24} \\ & (\frac{\partial^3 U}{\partial t^3})_{(i-1,j+\frac{1}{2})} + \frac{(\Delta t)^2(\Delta x)^2}{(4(\Delta x)^2-(\Delta t)^2)}[2\frac{\partial^2 U}{\partial x^2} - (\Delta t)\frac{\partial^3 U}{\partial x^2\partial t} + \frac{1}{4} \\ & (\Delta t)^2\frac{\partial^4 U}{\partial x^2\partial t^2} + \frac{2}{3}(\Delta x)^2\frac{\partial^4 U}{\partial x^4}]_{(i-1,j+\frac{1}{2})} + \frac{2}{3}\frac{(\Delta x)^3(\Delta t)(\Delta x+2\Delta t)}{(4(\Delta x)^2+(\Delta t)^2)} \\ & [-\frac{\partial^3 U}{\partial x^3} + \frac{(\Delta t)}{2}\frac{\partial^4 U}{\partial x^3\partial t}]_{(i-1,j+\frac{1}{2})} + \ldots\end{aligned} \tag{4.7.5}$$

In the same way, a Taylor series expansion for $U_{i,j+1}$, $U_{i-1,j}$, $U_{i+1,j}$, $U_{i,j}$ and $U_{i-2,j}$ about the point $(i\Delta x,(j+\frac{1}{2})\Delta t)$ gives the following truncation error expression for equation (4.7.3):

$$T_{4.7.3} = (\Delta t)(\frac{\partial U}{\partial t})_{(i,j+\frac{1}{2})} + \frac{(\Delta x)(\Delta t)}{(\Delta t-2\Delta x)}[2\frac{\partial U}{\partial x} + \frac{1}{4}(\Delta t)^2\frac{\partial^3 U}{\partial x\partial t^3} - \frac{1}{24}(\Delta t)^3\frac{\partial^4 U}{\partial x\partial t^3}]_{(i,j+\frac{1}{2})}+$$

$$\frac{(\Delta x)(\Delta t)^2}{(2\Delta x+\Delta t)}\left(\frac{\partial^2 U}{\partial x\partial t}\right)_{(i,j+\frac{1}{2})} + \frac{1}{24}(\Delta t)^3\left(\frac{\partial^3 U}{\partial t^3}\right)_{(i,j+\frac{1}{2})}$$

$$+ \frac{(\Delta x)^2(\Delta t)^2}{(4(\Delta x)^2-(\Delta t)^2)}\left[2\frac{\partial^2 U}{\partial x^2} - (\Delta t)\frac{\partial^3 U}{\partial x^2\partial t} + \frac{1}{4}(\Delta t)^2\frac{\partial^4 U}{\partial x^2\partial t^2} + \frac{2}{3}\right.$$

$$\left.(\Delta x)^2\frac{\partial^4 U}{\partial x^4}\right]_{(i,j+\frac{1}{2})} + \frac{2}{3}\frac{(\Delta x)^3(\Delta t)(\Delta x+2\Delta t)}{(4(\Delta x)^2+(\Delta t)^2)}\left[\frac{\partial^3 U}{\partial x^3} - \frac{1}{2}(\Delta t)\frac{\partial^4 U}{\partial x^3\partial t}\right]_{(i,j+\frac{1}{2})}$$

$$+ \ldots \qquad (4.7.6)$$

The truncation error for the single ungrouped point near the right end is given by the equation (4.6.9) with $i=m-1$, i.e.,

$$T_R = \frac{\Delta t}{\Delta x}\left[\frac{1}{2}\frac{\partial U}{\partial t} + \frac{1}{48}(\Delta t)^2\frac{\partial^3 U}{\partial t^3} + \frac{1}{4}(\Delta x)^2\frac{\partial^3 U}{\partial x^2\partial t}\right]_{(m-1,j+\frac{1}{2})} + \frac{1}{6}(\Delta x)^2\left(\frac{\partial^3 U}{\partial x^3}\right)_{(m-1,j+\frac{1}{2})}$$

$$+(\Delta t)^2\left[\frac{1}{24}\frac{\partial^3 U}{\partial t^3} - \frac{1}{8}\frac{\partial^3 U}{\partial x\partial t^2}\right]_{(m-1,j+\frac{1}{2})} + \ldots \qquad (4.7.7)$$

(ii) *Truncation Error for the GEL Scheme*

From (4.6.19), the GEL scheme takes the explicit matrix form,

$$\underline{u}_{j+1} = (I+rG_2)^{-1}(I+rG_1)\underline{u}_j + \hat{\underline{b}}_2 , \qquad (4.7.8)$$

where $\hat{\underline{b}}_2=(I+rG_2)^{-1}\underline{b}_2$. When written component-wise, (4.7.8) becomes,

$$u_{1,j+1} = -ru_{0,j+1}+u_{1j}+ru_{2j} , \qquad (4.7.9)$$

$$u_{i-1,j+1} = r_1u_{i-1,j}+r_2(u_{ij}-u_{i-2,j})-r_3u_{i+1,j} , \qquad (4.7.10)$$

$$i=3,5,\ldots m-1,$$

and
$$u_{i,j+1} = r_1u_{ij}+r_2(u_{i-1,j}-u_{i+1,j})-r_3u_{i-2,j} . \qquad (4.7.11)$$

The truncation error for the equation (4.7.9) can be obtained directly from (4.6.10) by putting $i=1$, to give,

$$T_L = \left(\frac{\partial U}{\partial t} - \frac{\partial U}{\partial x}\right)_{1,j+\frac{1}{2}} + \frac{\Delta t}{\Delta x}\left(\frac{1}{2}\frac{\partial U}{\partial t} + \frac{1}{48}(\Delta t)^2\frac{\partial^3 U}{\partial t^3} + \frac{1}{4}(\Delta x)^2\frac{\partial^3 U}{\partial x^2\partial t}\right)_{1,j+\frac{1}{2}}$$

$$- \frac{1}{6}(\Delta x)^2\left(\frac{\partial^3 U}{\partial x^3}\right)_{1,j+\frac{1}{2}}+(\Delta t)^2\left(\frac{1}{24}\frac{\partial^3 U}{\partial t^3} - \frac{1}{8}\frac{\partial^3 U}{\partial x\partial t^2}\right)_{1,j+\frac{1}{2}}+ \ldots \qquad (4.7.12)$$

The truncation errors for any two grouped points (equations (4.7.10) and (4.7.11)) are given by $T_{4.7.2}$ and $T_{4.7.3}$ of the expressions (4.7.5) and (4.7.6) respectively for $i=3,5,\ldots,m-1$.

(iii) Truncation Errors for the (S)AGE and (D)AGE Schemes

Based on the truncation errors of the GER and the GEL schemes, the truncation errors of the (S)AGE and (D)AGE methods are analysed in exactly the same manner as for the alternating schemes of the generalised weighted approximation of Section 4.3.

4.8 STABILITY ANALYSIS FOR THE GE METHODS

Before we proceed to investigate the stability condition of the GE schemes, we shall first of all establish the following results which will be used in our stability analysis.

From (4.6.17), we find that the matrices G_1 and G_2 have the same set of eigenvalues which consists of 0,1 (of multiplicity $\frac{1}{2}(m-2)$) and -1 (of multiplicity $\frac{1}{2}(m-2)$). Hence, we observe that although these matrices are symmetric, they are not positive definite. Furthermore, each of the matrices $(I+rG_1)$ and $(I+rG_2)$ has the eigenvalues $0,1+r$ (of multiplicity $\frac{1}{2}(m-2)$) and $1-r$ (of multiplicity $\frac{1}{2}(m-2)$). Hence, the spectral radius of $(I+rG_k)$ $(k=1,2)$ is given by,

$$\begin{aligned}\rho(I+rG_k) &= ||(I+rG_k)||_2 \\ &= 1+r \text{ for all positive values of } r.\end{aligned} \tag{4.8.1}$$

The inverses of $(I+rG_k)$ take the form,

$$(I+rG_1)^{-1} = \frac{1}{(1-r^2)}\left[\begin{array}{cc|cc|c|cc|c} 1 & -r & & & & & & \\ -r & 1 & & & & & & \\ \hline & & 1 & -r & & & O & \\ & & -r & 1 & & & & \\ \hline & & & & \ddots & & & \\ \hline & & & & & 1 & -r & \\ & O & & & & -r & 1 & \\ \hline & & & & & & & (1-r^2) \end{array}\right]_{(m-1)\times(m-1)} \tag{4.8.2}$$

and

$$(I+rG_2)^{-1} = \frac{1}{(1-r^2)}\left[\begin{array}{c|cc|ccc|cc} (1-r^2) & & & & & & & \\ \hline & 1 & -r & & & & & \\ & -r & 1 & & & & & \\ \hline & & & \ddots & & O & & \\ & & & & \ddots & & & \\ & & & O & & \ddots & & \\ \hline & & & & & & 1 & -r \\ & & & & & & -r & 1 \end{array}\right]_{(m-1)\times(m-1)} \tag{4.8.3}$$

A direct evaluation of the eigenvalues of $(I+rG_k)^{-1}$ shows that these eigenvalues are 1, $1/(1-r)$ of multiplicity $\frac{1}{2}(m-2)$ and $1/(1+r)$ of multiplicity $\frac{1}{2}(m-2)$.

(i) *Stability of the GER Scheme*

From the equation (4.7.1), the amplification matrix of the GER scheme is known to be

$$\Gamma_{GER} = (I+rG_1)^{-1}(I+rG_2) \tag{4.8.4}$$

Hence,

$$\begin{aligned} \rho(\Gamma_{GER}) &= ||\Gamma_{GER}||_2 \;, \\ &= ||(I+rG_1)^{-1}(I+rG_2)||_2 \;, \\ &\leq ||(I+rG_1)^{-1}||_2||(I+rG_2)||_2 \;, && (4.8.5) \\ &= \alpha_1 \;. && (4.8.6) \end{aligned}$$

To enable us to find the spectral radius of Γ_{GER}, we shall now consider (4.8.5) for different cases of r.

(a) For $0<r<1$, we have

$$\begin{aligned} \rho(rG_1) &= ||rG_1||_2 \\ &= r. \end{aligned}$$

Therefore, if $||rG_1||_2<1$, then it follows from Corollary 1.25 that,

$$\frac{1}{1+||rG_1||_2} \leq ||(I+rG_1)^{-1}||_2 \leq \frac{1}{1-||rG_1||_2} \;. \tag{4.8.7}$$

Hence, using (4.8.5) and (4.8.7) we find that

$$\rho(\Gamma_{GER}) \leqslant \frac{||(I+rG_2)||_2}{1-||rG_1||_2}$$

$$= \alpha_1 \ .$$

It is clear that $\rho(\Gamma_{GER}) \leqslant \alpha_1$ with $\alpha_1 = \frac{1+r}{1-r} > 1$.

(b) For $r>1$, we have

$$(I+rG_1)^{-1}(I+rG_1) = I$$

and $||I||_2 \leqslant ||(I+rG_1)^{-1}||_2 ||(I+rG_1||_2$, which implies,

$$||(I+rG_1)^{-1}||_2 \geqslant \frac{1}{||(I+rG_1)||_2} \quad . \tag{4.8.8}$$

Hence, we get from (4.8.5) and (4.8.8) that $\rho(\Gamma_{GER}) \leqslant \alpha_1$ with

$$\alpha_1 \geqslant \frac{||(I+rG_2)||_2}{||(I+rG_1)||_2} = \frac{1+r}{1+r} = 1.$$

We deduce from the cases (a) and (b) that for all values of r, the GER scheme is always unstable.

Alternatively, this condition of stability can also be established by first considering the eigenvalues of $(I+rG_k)^{-1}$ which are 1, $1/(1-r)$ of multiplicity $\frac{1}{2}(m-2)$ and $1/(1+r)$ of multiplicity $\frac{1}{2}(m-2)$ for $k=1,2$. It is seen that,

$$\rho((I+rG_k)^{-1}) = \begin{cases} \frac{1}{(1-r)} \ , \text{ if } 0<r<1 \\ \frac{1}{|(1-r)|} \ , \text{ if } 1<r\leqslant 2, \\ 1, \text{ if } r>2. \end{cases} \tag{4.8.9}$$

Now, using (4.8.5), (4.8.6) and (4.8.9) we find that:

(a) for $0<r<1$,

$$\alpha_1 = \frac{(1+r)}{(1-r)} > 1 \text{ and } ||\Gamma_{GER}||_2 \leqslant \alpha_1 \text{ with } \alpha_1 > 1;$$

(b) for $1<r\leq 2$,

$$\alpha_1 = \frac{(1+r)}{|(1-r)|} > 1 \text{ and } ||\Gamma_{GER}||_2 \leq \alpha_1 \text{ with } \alpha_1>1;$$

and

(c) for $r>2$,

$$\alpha_1 = 1+r > 1 \text{ and } ||\Gamma_{GER}||_2 \leq \alpha_1 \text{ with } \alpha_1>1.$$

From (a),(b) and (c) we deduce that the GER scheme is always unstable.

(ii) *Stability of the GEL Scheme*

From the equation (4.7.8), the amplification matrix of the GEL scheme is given by,

$$\Gamma_{GEL} = (I+rG_2)^{-1}(I+rG_1) \text{ ,} \tag{4.8.10}$$

and

$$\begin{aligned}\rho(\Gamma_{GEL}) &= ||\Gamma_{GEL}||_2 \text{ ,}\\ &= ||(I+rG_2)^{-1}(I+rG_1)||_2\\ &\leq ||(I+rG_2)^{-1}||_2||(I+rG_1)||_2 \text{ .}\end{aligned} \tag{4.8.11}$$

Since $||(I+rG_2)^{-1}||_2=||(I+rG_1)^{-1}||_2$ and $||(I+rG_1)||_2=||(I+rG_2)||_2$, the analysis of the stability of the GEL method will be the same as that of the GER scheme and we therefore conclude that the GEL scheme is also absolutely unstable.

(iii) *Stability of the (S)AGE and (D)AGE Schemes*

The second equation of (4.6.20) gives us,

$$\underline{u}_{j+2} = (I+rG_2)^{-1}(I+rG_1)\underline{u}_{j+1}+(I+rG_2)^{-1}\underline{b}_2 \text{ .} \tag{4.8.12}$$

By inserting $\underline{u}_{j+1}$ obtained from the first equation leads to,

$$\begin{aligned}\underline{u}_{j+2} &= (I+rG_2)^{-1}(I+rG_1)\{(I+rG_1)^{-1}(I+rG_2)\}\underline{u}_j + \underline{b}_2' \text{ ,}\\ &= (I+rG_2)^{-1}I(I+rG_2)\underline{u}_j + \underline{b}_2' \text{ ,}\\ &= I\underline{u}_j + \underline{b}_2' \text{ .}\end{aligned} \tag{4.8.13}$$

Hence, the amplification matrix of the (S)AGE scheme is $\Gamma_{SAGE}=I$

with eigenvalues equal to 1 of multiciplicity (m-1). (S)AGE is therefore stable (weakly) for whatever choice of r or λ.

Similarly, from the last two equations of (4.6.21) we obtain

$$\begin{aligned}\underline{u}_{j+4} &= (I+rG_1)^{-1}(I+rG_2)\{(I+rG_2)^{-1}(I+rG_1)\underline{u}_{j+2}+(I+rG_2)^{-1}\underline{b}_2\}+(I+rG_1)^{-1}\underline{b}_1 ,\\ &= (I+rG_1)^{-1}\{(I+rG_2)(I+rG_2)^{-1}\}(I+rG_1)\underline{u}_{j+2}+\underline{b}_2''' ,\\ &= (I+rG_1)^{-1}I(I+rG_1)\underline{u}_{j+2}+\underline{b}_2''' ,\\ &= (I+rG_1)^{-1}(I+rG_1)\underline{u}_{j+2}+\underline{b}_2''' ,\\ &= I\underline{u}_{j+2}+\underline{b}_2''' .\end{aligned} \tag{4.8.14}$$

The vector $\underline{u}_{j+2}$ of equation (4.8.13) is then inserted into (4.8.14) to give,

$$\underline{u}_{j+4} = I\underline{u}_j + \underline{b}_2'' . \tag{4.8.15}$$

Again, the amplification matrix Γ_{DAGE} is the identity matrix with (m-1) eigenvalues, each equal to 1 implying that the (D)AGE scheme is also weakly stable.

4.9 GE METHODS FOR THE ROBERTS AND WEISS APPROXIMATION TO FIRST ORDER EQUATION

A similar semi-implicit method for solving the first order convection equation has been developed by Roberts and Weiss (1966). It is based on the following discretisation of the equation at the point $(x_i,t_{j+\frac12}) = (i\Delta x,(j+\frac12)\Delta t)$ using only values of u_{ij}, $u_{i+1,j}$, $u_{i\ \ ,j+1}$ and $u_{i-1,j+1}$. Let us consider the first order equation at the point $(x_i,t_{j+\frac12})$, i.e.,

$$(\frac{\partial U}{\partial x})_{i,j+\frac12} + (\frac{\partial U}{\partial t})_{i,j+\frac12} = 0\ . \tag{4.9.1}$$

If we replace the space derivatives as follows,

$$(\frac{\partial U}{\partial t})_{i,j+\frac12} + \frac12((\frac{\partial U}{\partial x})_{i,j+1} + (\frac{\partial U}{\partial x})_{i,j}) + O([\Delta t]^2) = 0,$$

followed by the discretisation

$$\frac{(U_{i,j+1}-U_{ij})}{\Delta t} + O([\Delta t]^2) + \frac12\{\frac{(U_{i,j+1}-U_{i-1,j+1})}{\Delta x} + \frac{(U_{i+1,j}-U_{ij})}{\Delta x}$$
$$+ O([\Delta t]^2+\Delta x)\} = 0\ , \tag{4.9.2}$$

in which the backward difference form replaces $(\frac{\partial U}{\partial x})_{i,j+1}$ and the forward difference form replaces $(\frac{\partial U}{\partial x})_{i,j}$, we get the analogue

$$u_{i,j+1}-ru_{i-1,j+1} = u_{ij}-ru_{i+1,j}\ , \tag{4.9.3}$$

where $r=\lambda/(2+\lambda)$ and $\lambda=\Delta t/\Delta x$.

The computational molecule of equation (4.9.3) is given by,

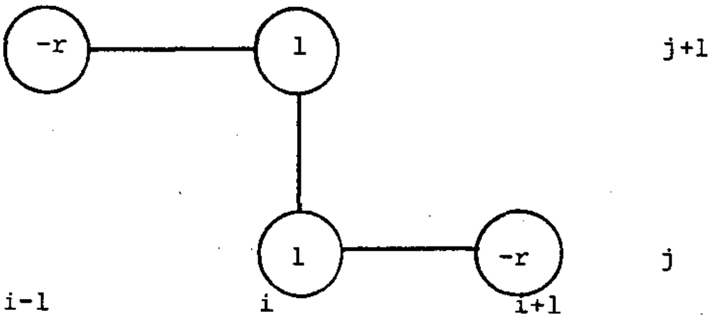


FIGURE 4.9.1

If we reverse the above procedure, we obtain the following approximation,

$$-ru_{i,j+1}+u_{i-1,j+1} = u_{i-1,j}-ru_{i-2,j} \ , \tag{4.9.4}$$

with its computational molecule given by,

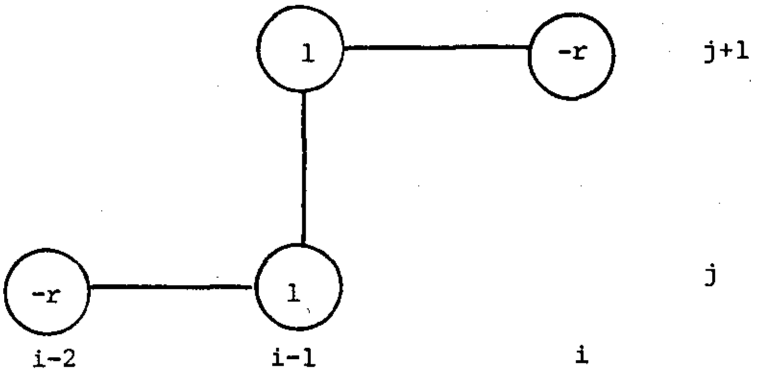


FIGURE 4.9.2

By using the familiar Taylor series expansion about the point $(x_i,t_{j+\frac{1}{2}})$, we obtain the following truncation error for the $L\mathcal{R}$ approximation (4.9.3),

$$T_{4.9.3} = \tfrac{1}{3}(\Delta x)^2\left(\frac{\partial^3U}{\partial x^3}\right)_{i,j+\frac{1}{2}}-\tfrac{1}{2}(\Delta x)(\Delta t)\left(\frac{\partial^3U}{\partial x^2\partial t}\right)_{i,j+\frac{1}{2}} + \tfrac{1}{4}(\Delta t)^2\left(\frac{\partial^3U}{\partial x\partial t^2}\right)_{i,j+\frac{1}{2}}$$
$$+\ \tfrac{1}{12}(\Delta t)^2\left(\frac{\partial^3U}{\partial t^3}\right)_{i,j+\frac{1}{2}} + \ldots \tag{4.9.5}$$

Similarly, the truncation error for the $R L$ approximation (4.9.4) is given by,

$$T_{4.9.4} = 2\left(\frac{\partial U}{\partial t} - \frac{\partial U}{\partial x}\right)_{i,j+\frac{1}{2}} - \tfrac{8}{3}(\Delta x)^2\left(\frac{\partial^3U}{\partial x^3}\right)_{i,j+\frac{1}{2}}+[(\Delta x)^2-\tfrac{1}{2}(\Delta x)(\Delta t)]\left(\frac{\partial^3U}{\partial x^2\partial t}\right)_{i,j+\frac{1}{2}}$$
$$-\ \tfrac{1}{4}(\Delta t)^2\left(\frac{\partial^3U}{\partial x\partial t^2}\right)_{i,j+\frac{1}{2}} + \tfrac{1}{12}(\Delta t)^2\left(\frac{\partial^3U}{\partial t^3}\right)_{i,j+\frac{1}{2}}+[\tfrac{1}{2}(\Delta x)^2(\Delta t)-\tfrac{1}{3}(\Delta x)^3]$$
$$\left(\frac{\partial^4U}{\partial x^3\partial t}\right)_{i,j+\frac{1}{2}} - \tfrac{1}{12}(\Delta x)(\Delta t)^2\left(\frac{\partial^4U}{\partial x\partial t^3}\right)_{i,j+\frac{1}{2}} + \tfrac{1}{4}(\Delta x)(\Delta t)^2\left(\frac{\partial^4U}{\partial x^2\partial t^2}\right)_{i,j+\frac{1}{2}}$$
$$+\ \tfrac{2}{3}(\Delta x)^3\left(\frac{\partial^4U}{\partial x^4}\right)_{i,j+\frac{1}{2}} + \ldots \tag{4.9.6}$$

When the RL and LR formulae are coupled and the mesh points are grouped two at a time as before, we obtain the following set of explicit equations,

$$u_{i-1,j+1} = r_1u_{i-1,j}+r_2(u_{ij}-u_{i-2,j})-r_3u_{i+1,j} , \tag{4.9.7}$$

and

$$u_{i,j+1} = r_2(u_{i-1,j}-u_{i+1,j})+r_1u_{ij}-r_3u_{i-2,j} , \tag{4.9.8}$$

where $r_1 = 1/(1-r^2)$, $r_2 = r/(1-r^2)$ and $r_3 = r^2/(1-r^2)$.

The solutions at the right and left ungrouped points are given respectively by,

$$u_{m-1,j+1} = u_{m-1,j}+r(u_{m,j+1}-u_{m-2,j}) \tag{4.9.9}$$

and

$$u_{1,j+1} = r(u_{0,j+1}-u_{2j}) + u_{1j} . \tag{4.9.10}$$

Further discussions on the GE schemes of the Roberts and Weiss approximation are abandoned because as will be explained in a later section, these schemes produce low accuracies in their solutions.

4.10 GE METHODS FOR THE SECOND-ORDER WAVE EQUATION

Let us now consider solving the following second-order wave equation,

$$\frac{\partial^2 U}{\partial x^2} = \frac{\partial^2 U}{\partial t^2} \quad , \tag{4.10.1}$$

subject to the initial conditions,

$$\begin{aligned} U(x,0) &= f_1(x) \quad , \\ \frac{\partial U}{\partial t}(x,0) &= f_2(x) \quad , \end{aligned} \tag{4.10.2}$$

and the boundary conditions,

$$U(0,t) = g_1(t)$$

and $\quad U(1,t) = g_2(t) \quad .$ (4.10.3)

The wave equation (4.10.1) can be reduced to a system of simultaneous differential equations of *first order* by the following substitutions,

$$U^{(1)} = \frac{\partial U}{\partial t}$$

and $\quad U^{(2)} = \frac{\partial U}{\partial x} \quad .$ (4.10.4)

In the more general case, first-order systems of equations can be written in matrix form as,

$$\frac{\partial \underline{\hat{U}}}{\partial t} + A\,\frac{\partial \underline{\hat{U}}}{\partial x} = \underline{0} \quad , \tag{4.10.5}$$

where A is an n×n real matrix (not necessarily symmetric) and $\underline{\hat{U}}$ is an n-component column vector $\underline{\hat{U}}=(U^{(1)},U^{(2)},\ldots,U^{(n)})^T$.

A non-singular matrix P exists through *the similarity transformation*,

$$PAP^{-1} = D \quad , \tag{4.10.6}$$

where D is a diagonal matrix having the real eigenvalues of A as its elements (i.e. D=diag(μ_i), the μ_i being the eigenvalues of A). On premultiplying (4.10.5) by P, we get

$$\frac{\partial}{\partial t}(P\hat{\underline{U}})+PAP^{-1}\,\frac{\partial}{\partial x}(P\hat{\underline{U}}) = \underline{0}\ ,$$

i.e.,

$$\frac{\partial \underline{V}}{\partial t} + D\,\frac{\partial \underline{V}}{\partial x} = 0\ , \qquad (4.10.7)$$

where $\underline{V}=P\hat{\underline{U}}$. Hence, *the decoupled scalar form* of (4.10.7) is given by

$$\frac{\partial V^{(i)}}{\partial t} + \mu_i\,\frac{\partial V^{(i)}}{\partial x} = 0,\ i=1,2,\ldots,n. \qquad (4.10.8)$$

For our particular problem, if

$$P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\ ,$$

then equation (4.10.5) takes the form (4.10.7) where

$$\underline{V} = (V^{(1)},V^{(2)})^T = (U^{(1)}-U^{(2)},U^{(1)}+U^{(2)})^T,\ \text{i.e.,}$$

$$\begin{aligned} V^{(1)} &= U^{(1)} - U^{(2)} & & U^{(1)} = \tfrac{1}{2}(V^{(1)} + V^{(2)}) \\ V^{(2)} &= U^{(1)} + U^{(2)} & \text{and}\quad & U^{(2)} = \tfrac{1}{2}(V^{(2)} - V^{(1)})\ , \end{aligned} \qquad (4.10.9)$$

and

$$D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\ .$$

Hence, the decoupled scalar equations for $V^{(1)}$ and $V^{(2)}$ are

$$\frac{\partial V^{(1)}}{\partial t} + \frac{\partial V^{(1)}}{\partial x} = 0 \qquad (4.10.10a)$$

and

$$\frac{\partial V^{(2)}}{\partial t} - \frac{\partial V^{(2)}}{\partial x} = 0 \qquad (4.10.10b)$$

respectively. The system (4.10.10) can be rewritten as

$$\frac{\partial V^{(p)}}{\partial t} = a\,\frac{\partial V^{(p)}}{\partial x}\ , \qquad (4.10.11a)$$

where,

$$a = \begin{cases} -1 & \text{when } p=1 \\ 1 & \text{when } p=2\ . \end{cases} \qquad (4.10.11b)$$

These *first order* differential equations in $\underline{V}=(V^{(1)},V^{(2)})^T$ will be solved by using the *weighted difference analogues*,

$$\lambda[\theta\{(1-w)v^{(p)}_{i+1,j+1}+(2w-1)v^{(p)}_{i,j+1}-wv^{(p)}_{i-1,j+1}\}+(1-\theta)\{(1-w)v^{(p)}_{i+1,j}+(2w-1)v^{(p)}_{ij}$$
$$-wv^{(p)}_{i-1,j}\}] = a(v^{(p)}_{i,j+1}-v^{(p)}_{ij});$$
$$0\leq\theta,w\leq 1. \qquad (4.10.12)$$

These equations reduce to

$$\lambda\theta v^{(p)}_{i-1,j+1}+(a-\lambda\theta)v^{(p)}_{i,j+1} = -\lambda(1-\theta)v^{(p)}_{i-1,j}+(a+\lambda(1-\theta))v^{(p)}_{ij} \qquad (4.10.13)$$

and

$$(a+\lambda\theta)v^{(p)}_{i,j+1}-\lambda\theta v^{(p)}_{i+1,j+1} = (a-\lambda(1-\theta))v^{(p)}_{ij}+\lambda(1-\theta)v^{(p)}_{i+1,j} \qquad (4.10.14)$$

when w takes the values 1 and 0 respectively. The local truncation errors of (4.10.13) and (4.10.14) at the point $(x_i,t_{j+\frac{1}{2}})$ are given respectively by,

$$T_{4.10.13} = \Delta x[-\tfrac{1}{2}\frac{\partial^2 v^{(p)}}{\partial x^2} - \frac{(\Delta t)^2}{16}\frac{\partial^4 v^{(p)}}{\partial x^2\partial t^2}]_{i,j+\frac{1}{2}}+\Delta t[\tfrac{1}{2}(1-2\theta)\frac{\partial^2 v^{(p)}}{\partial x\partial t}+\frac{(\Delta x)^2}{12}(1-2\theta)$$
$$\frac{\partial^4 v^{(p)}}{\partial x^3\partial t}]_{i,j+\frac{1}{2}}+(\Delta x)(\Delta t)[-\frac{1}{4}(1-2\theta)\frac{\partial^3 v^{(p)}}{\partial x^2\partial t}]_{i,j+\frac{1}{2}}+(\Delta x)^2[-\frac{1}{6}$$
$$\frac{\partial^3 v^{(p)}}{\partial x^3}+\frac{\Delta x}{24}\frac{\partial^4 v^{(p)}}{\partial x^4}]_{i,j+\frac{1}{2}}+(\Delta t)^2[-\frac{1}{8}(1-2\theta)\frac{\partial^3 v^{(p)}}{\partial x\partial t^2}+\frac{a}{24}$$
$$\frac{\partial^3 v^{(p)}}{\partial t^3}+\frac{\Delta t}{48}(1-2\theta)\frac{\partial^4 v^{(p)}}{\partial x\partial t^3}]_{i,j+\frac{1}{2}}+O((\Delta x)^{\alpha_1}(\Delta t)^{\alpha_2}) \qquad (4.10.15)$$

and

$$T_{4.10.14} = \Delta x[\tfrac{1}{2}\frac{\partial^2 v^{(p)}}{\partial x^2}+\frac{(\Delta t)^2}{16}\frac{\partial^4 v^{(p)}}{\partial x^2\partial t^2}]_{i,j+\frac{1}{2}}+\Delta t[\tfrac{1}{2}(1-2\theta)\frac{\partial^2 v^{(p)}}{\partial x\partial t}+$$
$$\frac{(\Delta x)^2}{12}(1-2\theta)\frac{\partial^4 v^{(p)}}{\partial x^3\partial t}]_{i,j+\frac{1}{2}}+(\Delta x)(\Delta t)[\frac{1}{4}(1-2\theta)\frac{\partial^3 v^{(p)}}{\partial x^2\partial t}]_{i,j+\frac{1}{2}}+$$
$$+(\Delta x)^2[-\frac{1}{6}\frac{\partial^3 v^{(p)}}{\partial x^3}-\frac{\Delta x}{24}\frac{\partial^4 v^{(p)}}{\partial x^4}]_{i,j+\frac{1}{2}}+(\Delta t)^2(-\frac{1}{8}(1-2\theta)\frac{\partial^3 v^{(p)}}{\partial x\partial t^2}$$
$$+\frac{a}{24}\frac{\partial^3 v^{(p)}}{\partial t^3}+\frac{\Delta t}{48}(1-2\theta)\frac{\partial^4 v^{(p)}}{\partial x\partial t^3})_{i,j+\frac{1}{2}}+O((\Delta x)^{\alpha_1}(\Delta t)^{\alpha_2}).$$
$$(4.10.16)$$

with $\alpha_1+\alpha_2=4$ and $0\leq\theta\leq1$. If we apply the formula (4.10.14), at the point $(x_{i-1},t_{j+\theta})$, we obtain,

$$(a+\lambda\theta)v_{i-1,j+1}^{(p)}-\lambda\theta v_{i,j+1}^{(p)} = (a-\lambda(1-\theta))v_{i-1,j}^{(p)}+\lambda(1-\theta)v_{ij}^{(p)} \quad (4.10.17)$$

By coupling equations (4.10.3) and (4.10.17), we arrive at the following set of explicit equations (we have omitted the details to avoid repetition),

$$v_{i-1,j+1}^{(p)} = \frac{(a-\lambda)}{a}v_{i-1,j}^{(p)}+\frac{\lambda}{a}v_{ij}^{(p)} , \quad (4.10.18a)$$

and

$$v_{i,j+1}^{(p)} = -\frac{\lambda}{a}v_{i-1,j}^{(p)}+\frac{(a+\lambda)}{a}v_{ij}^{(p)} . \quad (4.10.18b)$$

These equations must be solved simultaneously to give the values of $v^{(1)}$ and $v^{(2)}$ at the grid points along each j-line. From (4.10.14), the equation determining the values of $v^{(p)}$ at the ungrouped point adjacent to the right boundary is given by,

$$v_{m-1,j+1}^{(p)} = [(a-\lambda(1-\theta))v_{m-1,j}^{(p)}+\lambda(1-\theta)v_{mj}^{(p)}+\lambda\theta v_{m,j+1}^{(p)}]/(a+\lambda\theta) \quad (4.10.19)$$

whilst from (4.10.13) we obtain the following formula for the ungrouped point at the left end,

$$v_{1,j+1}^{(p)} = [-\lambda(1-\theta)v_{0j}^{(p)}-\lambda\theta v_{0,j+1}^{(p)}+(a+\lambda(1-\theta))v_{1j}^{(p)}]/(a-\lambda\theta), \quad (4.10.20)$$

$$a\neq\lambda\theta.$$

The GE schemes are then constructed along a similar line as before - and without loss of generality we assume that we will be using an even number of intervals of the line segment $0\leq x\leq1$.

(i) *The GER Scheme*

By means of equations (4.10.13), (4.10.17) and (4.10.19), the GER scheme is represented by the formula,

$$(aI+\lambda\theta G_1)\underline{v}_{j+1}^{(p)} = (aI-\lambda(1-\theta)G_1)\underline{v}_j^{(p)}+\underline{b}_1 , \quad (4.10.21)$$

where

$$G_1 = \left[\begin{array}{cc|cc|c|cc|c}
1 & -1 & & & & & & \\
1 & -1 & & & & & & \\ \hline
 & & 1 & -1 & & & & \\
 & & 1 & -1 & & & & \\ \hline
 & & & & \ddots & & O & \\ \hline
 & & & O & & 1 & -1 & \\
 & & & & & 1 & -1 & \\ \hline
 & & & & & & & 1
\end{array}\right]_{((m-1)\times(m-1))} \tag{4.10.22}$$

and $\underline{b}_1 = (0,0,\ldots,\lambda(1-\theta)v_{mj}^{(p)}+\lambda\theta v_{m,j+1}^{(p)})^T$.

(ii) *The GEL Scheme*

The GEL scheme is determined by the equations (4.10.20), (4.10.13) and (4.10.17) which can be expressed in a more compact form as,

$$(aI+\lambda\theta G_2)\underline{v}_{j+1}^{(p)} = (aI-\lambda(1-\theta)G_2)\underline{v}_j^{(p)} + \underline{b}_2 , \tag{4.10.23}$$

where,

$$G_2 = \left[\begin{array}{c|cc|c|cc}
-1 & & & & & \\ \hline
 & 1 & -1 & & & \\
 & 1 & -1 & & O & \\ \hline
 & & & \ddots & & \\ \hline
 & & O & & 1 & -1 \\
 & & & & 1 & -1
\end{array}\right]_{(m-1)\times(m-1)} \tag{4.10.24}$$

and $\underline{b}_2 = (-\lambda(1-\theta)v_{0j}^{(p)}-\lambda\theta v_{0,j+1}^{(p)},0,0,\ldots,0)^T$.

(iii) *The (S)AGE and (D)AGE Schemes*

The alternative use of the GER and the GEL methods leads to the following (S)AGE formulae,

$$\left.\begin{aligned}(aI+\lambda\theta G_1)\underline{v}_{j+1}^{(p)} &= (aI-\lambda(1-\theta)G_1)\underline{v}_j^{(p)} + \underline{b}_1\\(aI+\lambda\theta G_2)\underline{v}_{j+2}^{(p)} &= (aI-\lambda(1-\theta)G_2)\underline{v}_{j+1}^{(p)} + \underline{b}_2\end{aligned}\right\}$$

$$j=0,2,4,\ldots \tag{4.10.25}$$

and the (D)AGE four-step process,

$$\left.\begin{aligned}(aI+\lambda\theta G_1)\underline{v}_{j+1}^{(p)} &= (aI-\lambda(1-\theta)G_1)\underline{v}_j^{(p)} + \underline{b}_1\\(aI+\lambda\theta G_2)\underline{v}_{j+2}^{(p)} &= (aI-\lambda(1-\theta)G_2)\underline{v}_{j+1}^{(p)} + \underline{b}_2\\(aI+\lambda\theta G_2)\underline{v}_{j+3}^{(p)} &= (aI-\lambda(1-\theta)G_2)\underline{v}_{j+2}^{(p)} + \underline{b}_2\\(aI+\lambda\theta G_1)\underline{v}_{j+4}^{(p)} &= (aI-\lambda(1-\theta)G_1)\underline{v}_{j+3}^{(p)} + \underline{b}_1 .\end{aligned}\right\} \tag{4.10.26}$$

$$j=0,4,\ldots$$

All of the GE schemes employed above provide us with the values of $V^{(1)}$ and $V^{(2)}$ at the mesh points. The solution u of the wave equation (4.10.1) can then be computed using the relations in (4.10.9) i.e.,

$$U^{(1)} \simeq \tfrac{1}{2}(V^{(1)}+V^{(2)}) \tag{4.10.27}$$

and

$$U^{(2)} \simeq \tfrac{1}{2}(V^{(2)}-V^{(1)}) \ . \tag{4.10.28}$$

From (4.10.27), for example, we have, at the point (x_i,t_j),

$$\left(\frac{\partial U}{\partial t}\right)_{i,j} \simeq \tfrac{1}{2}(v_{ij}^{(1)} + v_{ij}^{(2)}) \tag{4.10.29}$$

and a first-order explicit approximation is obtained from the equation,

$$\frac{(u_{i,j+1}-u_{ij})}{\Delta t} = \tfrac{1}{2}(v_{ij}^{(1)} + v_{ij}^{(2)}) \ ,$$

or

$$u_{i,j+1} = u_{ij} + \tfrac{1}{2}\Delta t(v_{ij}^{(1)}+v_{ij}^{(2)}) \ . \tag{4.10.30}$$

On the other hand, if we add equations (4.10.27) and (4.10.28), we find that,

$$U^{(1)} + U^{(2)} \simeq V^{(2)} \tag{4.10.31}$$

and at the point (x_i,t_j) we have

$$(\frac{\partial U}{\partial t})_{i,j} + (\frac{\partial U}{\partial x})_{ij} \simeq v_{ij}^{(2)} . \qquad (4.10.32)$$

This can be solved by the second-order Lax-Wendroff explicit analogue given by,

$$u_{i,j+1} = \tfrac{1}{2}\lambda(1+\lambda)u_{i-1,j}+(1-\lambda^2)u_{ij}-\tfrac{1}{2}\lambda(1-\lambda)u_{i+1,j}+\Delta tv_{ij}^{(2)} \quad . \qquad (4.10.33)$$

At the point $(x_i,t_{j+\frac{1}{2}})$ however, (4.10.31) becomes

$$(\frac{\partial U}{\partial t})_{i,j+\frac{1}{2}} + (\frac{\partial U}{\partial x})_{i,j+\frac{1}{2}} \simeq v_{i,j+\frac{1}{2}}^{(2)} , \qquad (4.10.34)$$

and the following second-order accurate Crank-Nicolson type implicit approximation can be used,

$$-\frac{1}{4}\lambda u_{i-1,j+1}+u_{i,j+1}+\frac{1}{4}\lambda u_{i+1,j+1} = \frac{1}{4}\lambda u_{i-1,j}+u_{ij} - \frac{1}{4}\lambda u_{i+1,j}+\tfrac{1}{2}\Delta t(v_{ij}^{(2)}+ v_{i,j+1}^{(2)}) \quad . \qquad (4.10.35)$$

In employing the method of solution to (4.10.31), we must of course bear in mind its stability requirements as well as its order of accuracy.

4.11 TRUNCATION ERROR ANALYSIS OF THE GE SCHEMES

(i) Truncation Error for the GER Scheme

The set of explicit equations obtained by coupling equations (4.10.13) and (4.10.17) are

$$v_{i-1,j+1}^{(p)} - \frac{(a-\lambda)}{a} v_{i-1,j}^{(p)} - \frac{\lambda}{a} v_{ij}^{(p)} = 0 \tag{4.11.1}$$

and
$$v_{i,j+1}^{(p)} + \frac{\lambda}{a} v_{i-1,j}^{(p)} - \frac{(a+\lambda)}{a} v_{ij}^{(p)} = 0\ . \tag{4.11.2}$$

The truncation errors for any two grouped points are given by the truncation errors of equations (4.11.1) and (4.11.2) for $i=2,4,\ldots,m-2$. Thus we have,

$$\begin{aligned}T_{4.11.1} = {} & \Delta x[\tfrac{1}{2} \frac{\partial^2 v^{(p)}}{\partial x^2} + \frac{(\Delta t)^2}{16} \frac{\partial^4 v^{(p)}}{\partial x^2 \partial t^2}]_{i-1,j+\frac{1}{2}} + \Delta t[\tfrac{1}{2} \frac{\partial^2 v^{(p)}}{\partial x \partial t} + \frac{(\Delta x)^2}{12} \\ & \frac{\partial^4 v^{(p)}}{\partial x^3 \partial t}]_{i-1,j+\frac{1}{2}} + (\Delta x)(\Delta t)[\frac{1}{4} \frac{\partial^3 v^{(p)}}{\partial x^2 \partial t}]_{i-1,j+\frac{1}{2}} + (\Delta x)^2[-\frac{1}{6} \frac{\partial^3 v^{(p)}}{\partial x^3} \\ & - \frac{\Delta x}{24} \frac{\partial^4 v^{(p)}}{\partial x^4}]_{i-1,j+\frac{1}{2}} + (\Delta t)^2(-\frac{1}{8} \frac{\partial^3 v^{(p)}}{\partial x \partial t^2} + \frac{a}{24} \frac{\partial^3 v^{(p)}}{\partial t^3} + \frac{\Delta t}{48} \frac{\partial^4 v^{(p)}}{\partial x \partial t^3})_{i-1,j+\frac{1}{2}} \\ & + O(\Delta x)^{\alpha_1} (\Delta t)^{\alpha_2}) \ , \quad \alpha_1 + \alpha_2 = 4 \ ,\end{aligned} \tag{4.11.3}$$

and

$$\begin{aligned}T_{4.11.2} = {} & \Delta x[-\tfrac{1}{2} \frac{\partial^2 v^{(p)}}{\partial x^2} - \frac{(\Delta t)^2}{16} \frac{\partial^4 v^{(p)}}{\partial x^2 \partial t^2}]_{i,j+\frac{1}{2}} + \Delta t[\tfrac{1}{2} \frac{\partial^2 v^{(p)}}{\partial x \partial t} + \frac{(\Delta x)^2}{12} \\ & \frac{\partial^4 v^{(p)}}{\partial x^3 \partial t}]_{i,j+\frac{1}{2}} + (\Delta x)(\Delta t)[-\frac{1}{4} \frac{\partial^3 v^{(p)}}{\partial x^2 \partial t}]_{i,j+\frac{1}{2}} + (\Delta x)^2[-\frac{1}{6} \frac{\partial^3 v^{(p)}}{\partial x^3} \\ & + \frac{\Delta x}{24} \frac{\partial^4 v^{(p)}}{\partial x^4}]_{i,j+\frac{1}{2}} + (\Delta t)^2[-\frac{1}{8} \frac{\partial^3 v^{(p)}}{\partial x \partial t^2} + \frac{a}{24} \frac{\partial^3 v^{(p)}}{\partial t^3} + \frac{\Delta t}{48} \\ & \frac{\partial^4 v^{(p)}}{\partial x \partial t^3}]_{i,j+\frac{1}{2}} + O((\Delta x)^{\alpha_1} (\Delta t)^{\alpha_2}) \ , \text{ with } \alpha_1 + \alpha_2 = 4\end{aligned} \tag{4.11.4}$$

The truncation error for the single ungrouped point near the right end is obtained from (4.10.16) by putting $i=m-1$. This gives,

$$T_R = \Delta x[\tfrac{1}{2}\ \frac{\partial^2 v^{(p)}}{\partial x^2} + \frac{(\Delta t)^2}{16}\ \frac{\partial^4 v^{(p)}}{\partial x^2 \partial t^2}]_{m-1,j+\frac{1}{2}} + \Delta t[\tfrac{1}{2}(1-2\theta)\frac{\partial^2 v^{(p)}}{\partial x \partial t} +$$

$$\frac{(\Delta x)^2}{12}\ \frac{\partial^4 v^{(p)}}{\partial x^3 \partial t}]_{m-1,j+\frac{1}{2}} + (\Delta x)(\Delta t)[\frac{1}{4}(1-2\theta)\ \frac{\partial^3 v^{(p)}}{\partial x^2 \partial t}]_{m-1,j+\frac{1}{2}} +$$

$$(\Delta x)^2[-\frac{1}{6}\ \frac{\partial^3 v^{(p)}}{\partial x^3} - \frac{\Delta x}{24}\ \frac{\partial^4 v^{(p)}}{\partial x^4}]_{m-1,j+\frac{1}{2}} + (\Delta t)^2(-\frac{1}{8}(1-2\theta)\frac{\partial^3 v^{(p)}}{\partial x \partial t^2}$$

$$+ \frac{a}{24}\ \frac{\partial^3 v^{(p)}}{\partial t^3} + \frac{\Delta t}{48}\ (1-2\theta)\ \frac{\partial^4 v^{(p)}}{\partial x \partial t^3})_{m-1,j+\frac{1}{2}} + O((\Delta x)^{\alpha_1}(\Delta t)^{\alpha_2}),$$

$$\alpha_1+\alpha_2=4\ . \qquad (4.11.5)$$

(ii) *Truncation Error for the GEL Scheme*

The truncation error for the single ungrouped point near the left boundary is obtained from (4.10.15) with $i=1$ and this gives the expression,

$$T_L = \Delta x[-\tfrac{1}{2}\ \frac{\partial^2 v^{(p)}}{\partial x^2} - \frac{(\Delta t)^2}{16}\ \frac{\partial^4 v^{(p)}}{\partial x^2 \partial t^2}]_{1,j+\frac{1}{2}} + \Delta t[\tfrac{1}{2}(1-2\theta)\frac{\partial^2 v^{(p)}}{\partial x \partial t} + \frac{(\Delta x)^2}{12}\ (1-2\theta)$$

$$\frac{\partial^4 v^{(p)}}{\partial x^3 \partial t}]_{1,j+\frac{1}{2}} + (\Delta x)(\Delta t)[-\frac{1}{4}(1-2\theta)\ \frac{\partial^3 v^{(p)}}{\partial x^2 \partial t}]_{1,j+\frac{1}{2}} + (\Delta x)^2[-\frac{1}{6}\ \frac{\partial^3 v^{(p)}}{\partial x^3}$$

$$+ \frac{\Delta x}{24}\ \frac{\partial^4 v^{(p)}}{\partial x^4}\ ]_{1,j+\frac{1}{2}} + (\Delta t)^2[-\frac{1}{8}(1-2\theta)\ \frac{\partial^3 v^{(p)}}{\partial x \partial t^2} + \frac{a}{24}\ \frac{\partial^3 v^{(p)}}{\partial t^3} + \frac{\Delta t}{48}$$

$$(1-2\theta)\ \frac{\partial^4 v^{(p)}}{\partial x \partial t^3}]_{1,j+\frac{1}{2}} + O((\Delta x)^{\alpha_1}(\Delta t)^{\alpha_2}\ ,\ \text{with}\ \alpha_1+\alpha_2=4\ . \qquad (4.11.6)$$

The truncation errors for any two grouped points are given by $T_{4.11.1}$ and $T_{4.11.2}$ respectively.

(iii) *Truncation Error for the (S)AGE and (D)AGE Scheme*

As we have already seen, the truncation errors of the GER and GEL schemes (in their appropriate order of alternation) constitute the overall truncation errors of the two and four-step processes. Thus, there will be cancellations of errors at most points leading to some improvement in the solutions of the methods when compared with the constituent GER and GEL schemes.

4.12 STABILITY ANALYSIS OF THE GE SCHEMES

It is clear from (4.10.1) and (4.10.9) that to reach an *overall stability*, the GE schemes applied to the decoupled equations in (4.10.10a/b) must be stable simultaneously.

(i) Stability of the GER Scheme.

From equation (4.10.21), we have

$$\underset{\sim}{v}_{j+1}^{(p)} = \Gamma_{GER}\underset{\sim}{v}_{j}^{(p)} + \hat{\underset{\sim}{b}}_1 , \tag{4.12.1}$$

where Γ_{GER} is the amplification matrix given by

$$\Gamma_{GER} = (aI+\lambda\theta G_1)^{-1}(aI-\lambda(1-\theta)G_1) \tag{4.12.2}$$

and

$$\hat{\underset{\sim}{b}}_1 = (aI+\lambda\theta G_1)^{-1}\underset{\sim}{b}_1 . \tag{4.12.3}$$

For the case p=1 (and a=-1) we have already established in Section 4.4 that the GER scheme for (4.10.10a) is absolutely unstable in the range $0\leq\theta\leq\frac{1}{2}$ and is conditionally stable for $\lambda\geq\frac{2}{(2\theta-1)}$ when $\frac{1}{2}<\theta\leq1$. For the case p=2 (and a=1) we have,

$$\Gamma_{GER}=\left[\begin{array}{cc|cc|c|cc|c}
(1-\lambda) & \lambda & & & & & & \\
-\lambda & (1+\lambda) & & & & & & \\ \hline
 & & (1-\lambda) & \lambda & & & & \\
 & & -\lambda & (1+\lambda) & & & O & \\ \hline
 & & & & \ddots & & & \\ \hline
 & & & O & & (1-\lambda) & \lambda & \\
 & & & & & -\lambda & (1+\lambda) & \\ \hline
 & & & & & & & 1-\frac{\lambda}{(1+\lambda\theta)}
\end{array}\right]_{(m-1)\times(m-1)} \tag{4.12.4}$$

whose eigenvalues are 1 (of multiplicity (m-2)) and $1-\frac{\lambda}{(1+\lambda\theta)}$.

For stability, we require that

$$\left|1-\frac{\lambda}{(1+\lambda\theta)}\right| \leq 1$$

giving $0 \leq \frac{\lambda}{(1+\lambda\theta)} \leq 2$, which is just the inequality (4.4.12). Hence we deduce from Section 4.4 that the scheme is conditionally stable for $\lambda \leq \frac{2}{(1-2\theta)}$ with $0 \leq \theta < \frac{1}{2}$ and it is absolutely stable for all values of λ when $\frac{1}{2} \leq \theta \leq 1$. From the two stability requirements above, we therefore conclude that for *overall stability*, the GER scheme is stable only for $\lambda \geq \frac{2}{(2\theta-1)}$ when $\frac{1}{2} < \theta \leq 1$.

(ii) Stability of the GEL Scheme

From equation (4.10.23), the GEL scheme can be explicitly expressed as,

$$\underline{v}_{j+1}^{(p)} = \Gamma_{GEL}\underline{v}_{j}^{(p)} + \hat{\underline{b}}_2 , \qquad (4.12.5)$$

where Γ_{GEL} is the amplification matrix given by $\Gamma_{GEL} = (aI+\lambda\theta G_2)^{-1}(aI-\lambda(1-\theta)G_2)$ and $\hat{\underline{b}}_2 = (aI+\lambda\theta G_2)^{-1}\underline{b}_2$. We have already seen in Section 4.4 that the GEL scheme when applied to the differential equation (4.10.10a) (when p=1 and a=-1) is conditionally stable for $\lambda \leq \frac{2}{(1-2\theta)}$ and is always stable when $\frac{1}{2} \leq \theta \leq 1$. The amplification matrix of the GEL scheme for (4.10.10b) (when p=2 and a=1) is

$$\Gamma_{GEL} = \left[\begin{array}{c|cc|c|cc}
1+\frac{\lambda}{(1-\lambda\theta)} & & & & & \\ \hline
 & (1-\lambda) & \lambda & & & \\
 & -\lambda & (1+\lambda) & & \multicolumn{2}{c}{O} \\ \hline
 & & & \ddots & & \\ \hline
 & & & & (1-\lambda) & \lambda \\
 & \multicolumn{2}{c|}{O} & & -\lambda & (1+\lambda)
\end{array}\right]_{(m-1)\times(m-1)} \qquad (4.12.6)$$

whose eigenvalues are 1 (of multiplicity (m-2)) and $1 + \frac{\lambda}{(1-\lambda\theta)}$. For stability, we require,

$$\left|1 + \frac{\lambda}{(1+\lambda\theta)}\right| \leq 1$$

or

$$-2 \leq \frac{\lambda}{(1-\lambda\theta)} \leq 0$$

which is the same inequality given by (4.4.8) of Section 4.4. From the argument that followed, we deduce that the GEL scheme for (4.10.10b) is always unstable when $0\leq\theta\leq\frac{1}{2}$ and is stable only for $\lambda\geq\frac{2}{(2\theta-1)}$ when $\frac{1}{2}<\theta\leq 1$. Again for *overall stability*, we are led to the same stability conditions as that which was concluded for the GER scheme above.

(iii) Stability of the (S)AGE and (D)AGE Schemes

From (4.10.25) and (4.10.26) we find that the amplification matrices of the S(AGE) and D(AGE) schemes are given respectively by

$$\Gamma_{SAGE} = \Gamma_{GEL}\Gamma_{GER}$$

and

$$\Gamma_{DAGE} = \Gamma_{GER}\Gamma_{GEL}\Gamma_{SAGE}\ .$$

It has already been proved in Section 4.4 that the (S)AGE scheme for (4.10.10a) (when p=1 and a=-1) is conditionally stable for $\lambda\leq 1$. Similarly, the (D)AGE scheme is found to have conditional stability only for $\lambda\leq\frac{1}{2}$. For the case p=2 (and a=1), we have,

$$\Gamma_{SAGE} = \begin{bmatrix} a' & b' & & & & & & & \\ c' & d' & -c' & e' & & & & & \\ e' & f' & d' & -f' & & & & & \\ & & c' & d' & -c' & e' & & O & \\ & & e' & f' & d' & -f' & & & \\ & & & & \ddots & & & & \\ & O & & & c' & d' & -c' & e' & \\ & & & & e' & f' & d' & -f' & \\ & & & & & & c' & d' & g' \\ & & & & & & e' & f' & h' \end{bmatrix}_{(m-1)\times(m-1)} \tag{4.12.7}$$

where

$$a' = (1-\lambda)(1+\frac{\lambda}{(1-\lambda\theta)})$$

$$b' = \lambda(1+\frac{\lambda}{(1-\lambda\theta)})$$

$$c' = -\lambda(1-\lambda),\ d' = 1-\lambda^2,\ \ e' = \lambda^2$$
$$f' = -\lambda(1+\lambda),\ g' = \lambda(1-\frac{\lambda}{(1+\lambda\theta)})\text{ and } h' = (1+\lambda)(1-\frac{\lambda}{(1+\lambda\theta)}) \qquad (4.12.8)$$

and

$$|a'|+(m-3)|d'|+|h'|\leqslant\psi(\lambda)$$

where $\psi(\lambda) = (1-\lambda)(1+\frac{\lambda}{(1-\lambda\theta)})+(m-3)(1-\lambda^2)+(1+\lambda)(1-\frac{\lambda}{(1+\lambda\theta)})$

for $\lambda\leqslant 1$, $\theta \in [0,1]$ and $\lambda\theta\neq 1$. It can be shown as in Section 4.4 that the (S)AGE method is stable for $\lambda\leqslant 1$.

The (D)AGE amplification matrix (for p=2 and a=1), however, takes the form,

$$\Gamma_{DAGE}=\left[\begin{array}{cccccccccccc}
p' & q' & r' & s' & & & & & & & & \\
-q' & t' & u' & v' & & & & & & & & \\
r' & -u' & w' & x' & r' & s' & & & & & & \\
-s' & v' & -x' & z' & u' & v' & & & & & & \\
 & & r' & -u' & w' & x' & r' & s' & & & \bigcirc & \\
 & & -s' & v' & -x' & z' & u' & v' & & & & \\
 & & & & & \ddots & & & & & & \\
 & \bigcirc & & & & r' & -u' & w' & x' & r' & s' & \\
 & & & & & -s' & v' & -x' & z' & u' & v' & \\
 & & & & & & & r' & -u' & w' & x' & y' \\
 & & & & & & & -s' & v' & -x' & z' & q_1' \\
 & & & & & & & & & y' & -q_1' & p_1'
\end{array}\right]_{(m-1)\times(m-1)} \qquad (4.12.9)$$

where $p' = (1-\lambda)^2(1+\frac{\lambda}{(1-\lambda\theta)})^2-\lambda^2(1-2\lambda)$, $q' = -[-\lambda(1-\lambda)(1+\frac{\lambda}{(1-\lambda\theta)})^2$

$-\lambda(1+\lambda)(1-2\lambda)]$,

$r' = 2(1-\lambda)\lambda^2$, $s' = 2\lambda^3$,

$t' = -\lambda^2(1+\frac{\lambda}{(1-\lambda\theta)})^2+(1+\lambda)^2(1-2\lambda)$, $u' = 2\lambda(1-\lambda^2)$, $v' = 2(1+\lambda)\lambda^2$

$w' = (1-\lambda)^2(1+2\lambda)-\lambda^2(1-2\lambda)$, $x' = -2\lambda(2\lambda^2-1)$, $y' = 2\lambda^2(1-\frac{\lambda}{(1+\lambda\theta)})$

$z' = (1+\lambda)^2(1-2\lambda)-\lambda^2(1+2\lambda)$, $q_1' = 2\lambda(1+\lambda)(1-\frac{\lambda}{(1+\theta\lambda)})$,

$p_1' = (1+2\lambda)(1-\frac{\lambda}{(1+\theta\lambda)})^2$ (4.12.10)

and

$$|p'|+|t'|+\tfrac{1}{2}(m-4)|w'|+\tfrac{1}{2}(m-4)|z'|+|p_1'|\leq\psi(\lambda)$$

where,

$$\begin{aligned}\psi(\lambda) &= (1-\lambda)^2\left(1+\frac{\lambda}{(1-\lambda\theta)}\right)^2-\lambda^2(1-2\lambda)+\lambda^2\left(1+\frac{\lambda}{(1-\lambda\theta)}\right)^2+(1+\lambda)^2(1-2\lambda)\\ &\quad+\tfrac{1}{2}(m-4)\{(1-\lambda)^2(1+2\lambda)+\lambda^2(1-2\lambda)\}+\tfrac{1}{2}(m-4)\{(1+\lambda)^2(1-2\lambda)+\lambda^2(1+2\lambda)\}\\ &\quad+(1+2\lambda)\left(1-\frac{\lambda}{(1+\lambda\theta)}\right)^2\\ &= (\lambda^2+(1-\lambda)^2)\left(1+\frac{\lambda}{(1-\lambda\theta)}\right)^2+(\lambda^2+(1+\lambda)^2)(1-2\lambda)+(m-4)(1-2\lambda^2)\\ &\quad+(1+2\lambda)\left(1-\frac{\lambda}{(1+\lambda\theta)}\right)^2\end{aligned}$$

for $\lambda\leq\frac{1}{2}$, $\theta\in[0,1]$ and $\lambda\theta\neq1$. It can be shown in a similar manner as before that the (D)AGE method is stable for $\lambda\leq\frac{1}{2}$. For an overall stability, we conclude that the (S)AGE and (D)AGE processes are conditionally stable for $\lambda\leq1$ and $\lambda\leq\frac{1}{2}$ respectively. Therefore, it is recommended that for practical purposes, only (S)AGE is used.

4.13 NUMERICAL EXAMPLES AND COMPARATIVE RESULTS

To demonstrate the application of the GE schemes on hyperbolic problems, four numerical experiments were conducted.

Experiment 1

The weighted GE algorithms of Section 4.2 were implemented on the following two first-order hyperbolic problems:

(a) Problem 1

$$\frac{\partial U}{\partial t} + \frac{\partial U}{\partial x} = 0 \ ,$$

subject to,

$$U(x,0) = \cos x, \qquad (4.13.1)$$

$$U(0,t) = \cos t,$$

and

$$U(1,t) = \cos(1-t) \ .$$

The analytical solution is given by,

$$U(x,t) = \cos(x-t) \ , \qquad (4.13.2)$$

and

(b) Problem 2

$$\frac{\partial U}{\partial t} + \frac{\partial U}{\partial x} = k(x,t) \ ,$$

$(k(x,t) = -2\sin(x-t)e^{-2t})$,

subject to,

$$U(x,0) = \sin x,$$

$$U(0,t) = -\sin t\, e^{-2t} \ , \qquad (4.13.3)$$

and

$$U(1,t) = \sin(1-t)e^{-2t} \ .$$

The analytical solution is given by,

$$U(x,t) = \sin(x-t)e^{-2t} \ . \qquad (4.13.4)$$

The GE solutions to Problems 1 and 2 are compared with the solutions obtained from some of the standard methods, such as the classical explicit scheme (EXP) and the schemes of Lax-Wendroff (L-W), Roberts-Weiss (R-W) and Crank-Nicolson (C-N) (or the Centred-In-Distance, Centred-In-Time (CD-CT) scheme).

A comparison of their accuracies is obtained by computing the absolute error (A.E.),

$$\text{A.E.} = |e_{i,j}| = |u_{ij}-U_{ij}|, \tag{4.13.5}$$

or the percentage error (P.E.),

$$\text{P.E.} = \frac{|e_{ij}|}{|U_{i,j}|} \times 100 \tag{4.13.6}$$

at each point along the mesh line where u and U are the numerical and the analytical (exact) solutions respectively. Tables 4.13.1 and 4.13.2 provide the absolute errors of the numerical solutions to Problem 1 at t=0.4 and t=1.0 for λ=0.5 and θ=0.5. Similarly, the absolute errors for the numerical solutions to Problem 2 are shown by Tables 4.13.3 and 4.13.4. The average of all the absolute errors along the time levels t=0.4 and t=1.0 for each of the schemes involved is also entered in the tables.

Experiment 2

Several runs were made on the implementation of the (S)AGE and (D)AGE schemes for a range of values of θ in [0,1] and for t=0.2(0.2)1.0. For each particular value of θ, the entries in Tables 4.13.5 (Problem 1) and 4.13.6 (Problem 2) give the average of the absolute errors along each of the chosen time levels.

Experiment 3

The (S)AGE and (D)AGE schemes of the spatially-centred approximations of Section 4.6 were applied on Problem 1 and the absolute and percentage errors calculated. Table 4.13.7 displays these errors at each mesh point on the time level t=1.0 for λ=0.1.

Experiment 4

In this experiment, we proceeded with the application of the GE techniques on the second-order wave equation,

$$\frac{\partial^2 U}{\partial x^2} = \frac{\partial^2 U}{\partial t^2} ,$$

subject to,

$$U(x,0) = \frac{1}{8}\sin(\pi x) , \qquad (4.13.7)$$

$$\frac{\partial U}{\partial t}(x,0) = 0 ,$$

$$U(0,t) = 0 ,$$

and

$$U(1,t) = 0.$$

The analytical solution is given by,

$$U(x,t) = \frac{1}{8}\sin(\pi x)\cos(\pi t) . \qquad (4.13.8)$$

Again, we display the absolute errors of the numerical solutions along the mesh line t=1.0 for $\lambda=0.5$ and $\theta=0.5$ in Table 4.13.8.

4.14 DISCUSSION OF NUMERICAL RESULTS

It is clear from Tables 4.13.1-4.13.4 that the (S)AGE and (D)AGE schemes are more accurate than the GEL method in solving Problems 1 and 2. This result is expected because of the cancellation of error terms at most points of the grid system when the GER and the GEL schemes are applied in their appropriate order of alternation for the (S)AGE and (D)AGE processes. We also find that at some of the mesh points (along t=0.4 and t=1.0), the (S)AGE and (D)AGE schemes can have about the same magnitude of absolute errors as that of the high-order Lax-Wendroff, Roberts-Weiss and the Crank-Nicolson methods. In fact, an examination of the average of absolute errors for Problem 2 (Tables 4.13.3 and 4.13.4) clearly shows that the (S)AGE and (D)AGE schemes are more superior than the other methods that we have considered. Furthermore, the computational complexity incurred in solving the first-order hyperbolic equation (4.1.1) is also considerably less than that of, say, the Crank-Nicolson method. The following Table 4.14.1 gives us a comparison of the amount of arithmetic involved at m internal mesh points *along each time row* where the solutions of the various difference schemes are determined. It is seen that the (S)AGE and (D)AGE schemes even compare well with the explicit, second-order accurate Lax-Wendroff formula.

Method	Number of Multiplications	Number of Divisions	Number of Additions (Subtractions)
EXP	m	-	2m
L-W	2m	m	5m
GER/GEL/ (S)AGE/(D)AGE	m+1	1	2m+1
C-N(CD-CT)	8m-1	3m-2	7m-3

TABLE 4.14.1

We observe from the entries in Tables 4.13.5 and 4.13.6 that the (S)AGE and (D)AGE schemes are most accurate along the time rows t= 0.2(0.2)1.0 for $\lambda=0.5$ when the time weighting θ takes the value of about 0.5. A possible explanation of this result is that, the terms involving the coefficients $(1-2\theta)$ in the truncation errors in (4.3.3c) and (4.3.4) vanish when θ is exactly 0.5. This leads to a considerable increase in the accuracy of the solutions at the ungrouped points and the overall effect of the cancellation of errors due to the alternate use of the GER and GEL algorithms is the improvement in the solutions as they progress forward in time.

Table 4.13.7 obviously shows that the stability advantage of the (S)AGE and (D)AGE schemes is clearly overriden by their very poor accuracy when applied to the spatially-centred approximations of Section 4.6. This stems from the consistency difficulty of the two asymmetric formulae (equations (4.6.7) and (4.6.8)) which when coupled together determine the basic equations of the GE schemes. From (4.6.9) we see that in order for $T_{4.6.7} \longrightarrow 0$ as $\Delta x,\Delta t\to 0$, it is essential $\Delta t\to 0$ *faster* than $\Delta x\to 0$. Even if we assume that this consistency requirement is accomplished, we still find from (4.6.10) that the difference equation (4.6.8) would be consistent with the differential equation

$$\frac{\partial U}{\partial t}+\frac{\partial U}{\partial x}+\frac{\partial U}{\partial t}-\frac{\partial U}{\partial x}=0\ ,\ \text{i.e.,}$$

$$\frac{\partial U}{\partial t}=0\ ,$$

rather than with the hyperbolic equation (4.1.1). The truncation error expressions for the GER and GEL further confirm the above consistency problem.

The GE methods for the Roberts and Weiss approximations also suffer with the same consistency inadequacy as above. While the truncation error of the RL approximation (4.9.3) tends to O as $\Delta x, \Delta t \to 0$, the truncation error of the LR approximation (4.9.4), however, appears to tend to $2(\frac{\partial U}{\partial t} - \frac{\partial U}{\partial x})$ at the mesh point. This consistency difficulty only serves to produce very low accuracies in the (S)AGE and (D)AGE solutions when the RL and LR formulae are coupled in the same way as before. Any further theoretical treatment on these methods is therefore not pursued. However, it seems only sensible to suggest that due to the wave like nature of the solution then methods based on the Saulev semi-implicit strategy are more favourable than the GE type methods.

To arrive at the solution of the second-order wave equation (4.13.7), Experiment 4 necessitates us to solve two different sets of first-order differential equations. The first set involves $V^{(1)}$ and $V^{(2)}$ whose approximations at the mesh points are obtained by applying the GE techniques on (4.10.10a/10b). The solutions u are then computed by means of the explicit ((4.10.30)), Lax-Wendroff ((4.10.33)) and the Crank-Nicolson type ((4.10.35)) formulae. These solutions are compared in Table 4.13.8. No attempt is made to compute the GER and the GEL solutions as these schemes have a rather rigid stability requirement. It becomes apparent from the table that the (S)AGE-LW methods provide the most accurate solution. The stability restrictions of the (S)AGE-CN and (D)AGE-CN methods are $\lambda \leq 1$ and $\lambda \leq \frac{1}{2}$ respectively and besides incurring a comparatively heavier computational load, these methods also happen to produce a less accurate solution for our particular problem. Hence, for its simplicity and accuracy, the (S)AGE-LW (also stable for $\lambda \leq 1$) combination is favoured.

t=0.4, λ=0.5, Δt=0.05, Δx=0.1, θ=0.5

x / Method	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of All Absolu Errors
GEL	4.75×10^{-3}	9.89×10^{-2}	5.93×10^{-2}	9.5×10^{-2}	5.65×10^{-2}	8.72×10^{-2}	5.15×10^{-2}	7.6×10^{-2}	4.44×10^{-2}	6.37×10
(S)AGE	6.38×10^{-4}	9.92×10^{-4}	2.95×10^{-4}	7.87×10^{-4}	2.42×10^{-4}	9.86×10^{-4}	5.98×10^{-4}	1.25×10^{-3}	6.99×10^{-4}	$7.21\times10^{-}$
(D)AGE	1.02×10^{-3}	1.02×10^{-3}	1.0×10^{-3}	2.11×10^{-4}	2.04×10^{-4}	2.04×10^{-4}	5.43×10^{-4}	5.43×10^{-4}	$1.04\ 10^{-3}$	$6.43\times10^{-}$
EXP	2.41×10^{-3}	4.83×10^{-3}	7.0×10^{-3}	8.6×10^{-3}	9.45×10^{-3}	9.66×10^{-3}	9.5×10^{-3}	9.17×10^{-3}	8.74×10^{-3}	$7.71\times10^{-}$
L-W	3.29×10^{-5}	3.87×10^{-5}	1.76×10^{-5}	2.18×10^{-5}	6.91×10^{-5}	1.17×10^{-4}	1.69×10^{-4}	1.93×10^{-4}	3.39×10^{-4}	$1.11\times10^{-}$
R-W	9.49×10^{-5}	1.11×10^{-4}	6.28×10^{-4}	2.58×10^{-5}	1.36×10^{-4}	2.42×10^{-4}	4.28×10^{-4}	2.75×10^{-4}	1.09×10^{-3}	$2.74\times10^{-}$
C-N(CD-CT)	5.77×10^{-5}	7.2×10^{-5}	4.18×10^{-5}	5.06×10^{-6}	9.73×10^{-5}	9.82×10^{-5}	3.44×10^{-4}	5.78×10^{-5}	7.15×10^{-4}	$1.65\times10^{-}$
EXACT SOLUTION	0.9553365	0.9800666	0.9950042	1.0	0.9950042	0.9800666	0.9553365	0.9210610	0.8775826	-

TABLE 4.13.1: Absolute Errors of the Numerical Solutions to Problem 1

$t=1.0,\ \lambda=0.5,\ \Delta t=0.05,\ \ \Delta x=0.1,\ \theta=0.5$

x Method	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of All Absolute Errors
GEL	3.33×10^{-3}	5.31×10^{-1}	4.38×10^{-1}	5.31×10^{-1}	4.35×10^{-1}	5.09×10^{-1}	4.14×10^{-1}	4.86×10^{-1}	3.78×10^{-1}	4.12×10^{-1}
(S)AGE	2.18×10^{-4}	1.58×10^{-3}	2.46×10^{-4}	2.07×10^{-3}	4.82×10^{-4}	2.27×10^{-3}	5.0×10^{-4}	1.89×10^{-3}	1.18×10^{-4}	1.04×10^{-3}
(D)AGE	3.27×10^{-4}	1.12×10^{-3}	1.87×10^{-4}	1.36×10^{-3}	1.66×10^{-4}	1.50×10^{-3}	1.55×10^{-4}	1.88×10^{-3}	2.39×10^{-4}	7.71×10^{-4}
EXP	1.65×10^{-3}	3.64×10^{-3}	5.94×10^{-3}	8.47×10^{-3}	1.11×10^{-2}	1.39×10^{-2}	1.66×10^{-2}	1.90×10^{-2}	2.11×10^{-2}	1.13×10^{-2}
L-W	9.52×10^{-5}	1.73×10^{-4}	2.31×10^{-4}	2.68×10^{-4}	2.8×10^{-4}	2.72×10^{-4}	2.15×10^{-4}	1.80×10^{-4}	4.89×10^{-6}	1.91×10^{-4}
R-W	2.38×10^{-4}	4.7×10^{-4}	5.29×10^{-4}	9.27×10^{-4}	3.53×10^{-4}	1.39×10^{-3}	1.44×10^{-4}	1.12×10^{-3}	1.93×10^{-4}	5.97×10^{-4}
C-N(CD-CT)	1.14×10^{-6}	4.58×10^{-4}	4.97×10^{-5}	8.29×10^{-4}	7.91×10^{-6}	9.04×10^{-4}	5.1×10^{-6}	5.75×10^{-4}	3.55×10^{-5}	3.18×10^{-4}
EXACT SOLUTION	0.6216100	0.6967067	0.7648422	0.8253356	0.8775826	0.9210610	0.9553365	0.9800666	0.9950042	-

TABLE 4.13.2: Absolute Errors of the Numerical Solutions to Problem 1

$t=0.4,\ \lambda=0.5,\ \Delta t=0.05,\ \Delta x=0.1,\ \theta=0.5$

x Method	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of All Absolute Errors
GEL	5.94×10^{-4}	9.06×10^{-3}	6.99×10^{-3}	2.39×10^{-2}	1.64×10^{-2}	3.78×10^{-2}	2.52×10^{-2}	5.01×10^{-2}	3.3×10^{-2}	1.85×10^{-2}
(S)AGE	9.13×10^{-4}	7.71×10^{-4}	8.33×10^{-4}	9.11×10^{-4}	8.14×10^{-4}	6.27×10^{-4}	7.08×10^{-4}	1.37×10^{-4}	6.84×10^{-4}	5.82×10^{-4}
(D)AGE	5.04×10^{-4}	2.74×10^{-4}	4.41×10^{-4}	1.65×10^{-5}	3.57×10^{-5}	5.19×10^{-5}	5.08×10^{-4}	4.69×10^{-4}	3.68×10^{-4}	2.43×10^{-4}
EXP	3.76×10^{-3}	9.7×10^{-3}	1.71×10^{-2}	2.46×10^{-2}	3.05×10^{-2}	3.47×10^{-2}	3.76×10^{-2}	3.99×10^{-2}	4.17×10^{-2}	2.18×10^{-2}
L-W	3.02×10^{-4}	7.62×10^{-4}	1.26×10^{-3}	1.61×10^{-3}	1.75×10^{-3}	1.77×10^{-3}	1.77×10^{-3}	1.62×10^{-3}	2.04×10^{-3}	1.17×10^{-3}
R-W	1.11×10^{-4}	3.35×10^{-4}	5.72×10^{-4}	7.7×10^{-4}	9.25×10^{-4}	1.02×10^{-3}	1.24×10^{-3}	8.2×10^{-4}	2.18×10^{-3}	7.25×10^{-4}
C-N(CD-CT)	7.0×10^{-5}	2.08×10^{-4}	3.58×10^{-4}	4.56×10^{-4}	5.61×10^{-4}	5.05×10^{-4}	7.85×10^{-4}	2.9×10^{-4}	1.17×10^{-3}	4.89×10^{-4}
EXACT SOLUTION	-0.1327158	-0.0890597	-0.044858	0	0.0448580	0.0892679	0.1327858	0.1749769	0.2154198	-

TABLE 4.13.3: Absolute Errors of the Numerical Solution to Problem 2

$t=1.0,\ \lambda=0.5,\ \Delta t=0.05,\ \Delta x=0.1,\ \theta=0.5$

x Method	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all Absolute Errors
GEL	5.93×10^{-4}	1.12×10^{-3}	4.96×10^{-3}	6.06×10^{-2}	5.61×10^{-2}	1.18×10^{-1}	1.05×10^{-1}	1.70×10^{-1}	1.5×10^{-1}	6.06×10^{-2}
(S)AGE	1.38×10^{-3}	2.48×10^{-4}	1.23×10^{-3}	2.53×10^{-4}	1.16×10^{-3}	4.85×10^{-4}	1.01×10^{-3}	4.15×10^{-4}	8.9×10^{-4}	6.43×10^{-4}
(D)AGE	7.06×10^{-4}	2.99×10^{-4}	6.61×10^{-4}	2.83×10^{-5}	5.24×10^{-4}	1.82×10^{-4}	4.28×10^{-4}	3.34×10^{-4}	3.48×10^{-4}	3.19×10^{-4}
EXP	3.64×10^{-4}	1.99×10^{-4}	6.98×10^{-4}	2.57×10^{-3}	5.68×10^{-3}	1.02×10^{-2}	1.63×10^{-2}	2.35×10^{-2}	3.15×10^{-2}	8.28×10^{-3}
L-W	5.24×10^{-5}	1.38×10^{-4}	2.54×10^{-4}	3.9×10^{-4}	5.93×10^{-4}	8.71×10^{-4}	1.43×10^{-3}	1.57×10^{-3}	3.16×10^{-3}	7.69×10^{-4}
R-W	6.86×10^{-6}	6.5×10^{-5}	9.85×10^{-5}	4.22×10^{-4}	8.38×10^{-4}	1.01×10^{-3}	2.06×10^{-3}	8.27×10^{-4}	2.48×10^{-3}	7.1×10^{-4}
C-N(CD-CT)	2.21×10^{-4}	2.63×10^{-4}	4.99×10^{-4}	5.96×10^{-4}	8.46×10^{-4}	6.46×10^{-4}	$1.17\ 10^{-3}$	$3.34\ 10^{-4}$	1.44×10^{-3}	6.68×10^{-4}
EXACT SOLUTION	-0.1060118	-0.0970836	-0.0871854	-0.0764160	-0.0648832	-0.0527020	-0.03999431	-0.0268870	-0.0135110	-

TABLE 4.13.4: Absolute Errors of the Numerical Solutions to Problem 2

$\lambda=0.5,\ \Delta t=0.05$

Method / t \ θ	(S)AGE					(D)AGE				
	0	0.25	0.5	0.75	1	0	0.25	0.5	0.75	1
0.2	3.80×10^{-4}	3.5×10^{-4}	3.48×10^{-4}	3.71×10^{-4}	4.39×10^{-4}	2.43×10^{-4}	2.09×10^{-4}	2.88×10^{-4}	3.48×10^{-4}	3.88×10^{-4}
0.4	7.66×10^{-4}	7.23×10^{-4}	7.21×10^{-4}	7.88×10^{-4}	1.02×10^{-3}	5.72×10^{-4}	4.63×10^{-4}	6.43×10^{-4}	7.62×10^{-4}	1.01×10^{-3}
0.6	1.01×10^{-3}	9.43×10^{-4}	9.43×10^{-4}	1.12×10^{-3}	1.62×10^{-3}	8.55×10^{-4}	6.43×10^{-4}	8.19×10^{-4}	1.03×10^{-3}	1.38×10^{-3}
0.8	1.18×10^{-3}	1.07×10^{-3}	1.02×10^{-3}	1.27×10^{-3}	2.13×10^{-3}	1.01×10^{-3}	7.90×10^{-4}	8.55×10^{-4}	1.29×10^{-3}	1.67×10^{-3}
1.0	1.34×10^{-3}	1.17×10^{-3}	1.04×10^{-3}	1.3×10^{-3}	2.45×10^{-3}	1.21×10^{-3}	9.49×10^{-4}	7.71×10^{-4}	1.38×10^{-3}	2.09×10^{-3}

TABLE 4.13.5: Average of Absolute Errors for Problem 1

$\lambda=0.5$, $\Delta t=0.05$

Method	(S)AGE					(D)AGE				
t \ θ	0	0.25	0.5	0.75	1	0	0.25	0.5	0.75	1
0.2	1.60×10^{-2}	7.81×10^{-3}	3.82×10^{-4}	7.56×10^{-3}	1.47×10^{-2}	1.67×10^{-2}	8.32×10^{-3}	1.41×10^{-4}	8.4×10^{-3}	1.69×10^{-2}
0.4	2.05×10^{-2}	9.8×10^{-3}	5.82×10^{-4}	9.88×10^{-3}	2.16×10^{-2}	2.29×10^{-2}	1.13×10^{-2}	2.43×10^{-4}	1.15×10^{-2}	2.34×10^{-2}
0.6	1.66×10^{-2}	8.19×10^{-3}	6.68×10^{-4}	1.07×10^{-2}	2.72×10^{-2}	2.18×10^{-2}	1.07×10^{-2}	3.19×10^{-4}	1.13×10^{-2}	2.64×10^{-2}
0.8	1.58×10^{-2}	9.07×10^{-3}	6.77×10^{-4}	1.57×10^{-2}	4.31×10^{-2}	1.82×10^{-2}	9.62×10^{-3}	2.87×10^{-4}	1.38×10^{-2}	3.56×10^{-2}
1.0	2.09×10^{-2}	1.2×10^{-2}	6.43×10^{-4}	2.05×10^{-2}	5.73×10^{-2}	2.11×10^{-2}	1.20×10^{-2}	3.19×10^{-4}	1.8×10^{-2}	4.65×10^{-2}

TABLE 4.13.6: Average of Absolute Errors for Problem 2

t=1.0, λ=0.1, Δt=0.01, Δx=0.1

Method \ x		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all Errors
(S)AGE	A.E.	3.96×10^{-1}	2.83×10^{-1}	1.9×10^{-1}	9.57×10^{-2}	3.16×10^{-12}	9.57×10^{-2}	1.9×10^{-1}	2.83×10^{-1}	3.73×10^{-1}	1.73×10^{-1}
	P.E.(%)	63.77	40.67	24.91	11.6	3.6×10^{-10}	10.39	19.94	28.91	37.53	21.61
(D)AGE	A.E.	3.85×10^{-1}	2.83×10^{-1}	1.9×10^{-1}	9.57×10^{-2}	2.89×10^{-12}	9.57×10^{-2}	1.9×10^{-1}	2.83×10^{-1}	3.85×10^{-1}	1.74×10^{-1}
	P.E.(%)	61.88	40.67	24.91	11.6	3.3×10^{-10}	10.39	19.94	28.91	38.66	21.54
EXACT SOLUTION		0.6216100	0.6967067	0.7648422	0.8253356	0.8775826	0.9210610	0.9553365	0.9800666	0.9950042	-

TABLE 4.13.7: Absolute and Percentage Errors of the Numerical Solutions to Problem 1

$t=1.0$, $\lambda=0.5$, $\Delta t=0.05$, $\Delta x=0.1$, $\theta=0.5$

x Method	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all Absolute Errors
(S)AGE-EXP	4.02×10^{-3}	2.6×10^{-3}	1.61×10^{-3}	3.74×10^{-3}	6.18×10^{-4}	3.38×10^{-3}	1.63×10^{-3}	1.86×10^{-3}	4.06×10^{-3}	2.61×10^{-3}
(D)AGE-EXP	7.66×10^{-4}	6.25×10^{-4}	1.27×10^{-3}	1.27×10^{-3}	1.22×10^{-3}	1.55×10^{-3}	7.35×10^{-4}	1.46×10^{-3}	1.14×10^{-4}	1.0×10^{-3}
(S)AGE-LW	8.35×10^{-4}	1.11×10^{-3}	4.34×10^{-4}	1.17×10^{-3}	3.61×10^{-4}	2.52×10^{-4}	7.79×10^{-4}	1.5×10^{-3}	2.09×10^{-3}	9.47×10^{-4}
(D)AGE-LW	1.13×10^{-3}	8.85×10^{-4}	1.83×10^{-3}	1.66×10^{-3}	1.24×10^{-3}	1.15×10^{-3}	3.88×10^{-4}	1.76×10^{-4}	1.98×10^{-3}	1.16×10^{-3}
(S)AGE-CN	1.46×10^{-3}	1.78×10^{-3}	4.69×10^{-3}	2.44×10^{-3}	1.09×10^{-2}	1.67×10^{-3}	1.2×10^{-2}	6.54×10^{-4}	1.02×10^{-2}	5.09×10^{-3}
(D)AGE-CN	1.95×10^{-3}	4.35×10^{-4}	5.96×10^{-3}	9.9×10^{-4}	9.17×10^{-3}	1.49×10^{-3}	9.3×10^{-3}	1.42×10^{-3}	7.86×10^{-3}	4.29×10^{-3}
EXACT SOLUTION	-0.0386272	-0.0734732	-0.1011271	-0.1188821	-0.125000	-0.1188821	-0.1011271	-0.0734732	-0.0386272	-

TABLE 4.13.8: Absolute Errors of the Numerical Solutions to the Wave Equation

CHAPTER FIVE

THE GROUP EXPLICIT METHODS FOR PARABOLIC PROBLEMS WITH SPECIAL GEOMETRIES

5.1 INTRODUCTION

In this chapter, we consider applying the GE strategy to solve parabolic problems with special geometries involving one-space dimension given by,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial r^2} + \frac{\alpha}{r}\frac{\partial U}{\partial r} \qquad (5.1.1)$$

together with the following initial-boundary conditions,

$$U(r,0) = f(r)\ ,\ 0\leq r\leq 1$$

and $$\frac{\partial U}{\partial r}(0,t) = 0,\ U(1,t) = 0 \text{ for } 0\leq t\leq T\ . \qquad (5.1.1a)$$

As we have seen earlier, equation (5.1.1) reduces to the simple diffusion equation of (3.1.2) when $\alpha=0$. Evans and Abdullah (1983) have successfully implemented the GE algorithms for this equation and showed them to be more superior than most of the difference schemes that are currently in use. We shall now extend the GE application to parabolic problems that possess *cylindrical and spherical symmetry* by putting $\alpha=1$ and 2 respectively in (5.1.1). The GER, GEL, (S)AGE and (D)AGE schemes will be developed and the stability requirements established. We will also investigate the truncation errors of the methods and perform some numerical experiments to compare their accuracy with that of existing schemes.

5.2 DERIVATION OF SOME OF THE COMMONLY USED SCHEMES

We shall follow the approach adopted by Evans and Abdullah (1983) by utilising the following generalised formulae to approximate the derivatives in (5.1.1) at the point $(r_i,t_{j+\frac{1}{2}})=(i\Delta r,(j+\frac{1}{2})\Delta t)$, i.e.,

$$\left(\frac{\partial^2 U}{\partial r^2}\right)_{i,j+\frac{1}{2}} \simeq (\theta_1\delta_r u_{i+\frac{1}{2},j+1}-\theta_2\delta_r u_{i-\frac{1}{2},j+1}+\theta_3\delta_r u_{i+\frac{1}{2},j}-\theta_4\delta_r u_{i-\frac{1}{2},j})/(\Delta r)^2, \tag{5.2.1}$$

$$\left(\frac{\partial U}{\partial r}\right)_{i,j+\frac{1}{2}} \simeq (\alpha_1\Delta_r u_{i,j+1}+\alpha_2\nabla_r u_{ij}+\alpha_3\nabla_r u_{i,j+1}+\alpha_4\Delta_r u_{ij})/2\Delta r \tag{5.2.2}$$

and

$$\left(\frac{\partial U}{\partial t}\right)_{i,j+\frac{1}{2}} \simeq (u_{i,j+1}-u_{ij})/\Delta t\ , \tag{5.2.3}$$

where Δr and Δt are the increments with respect to the r- and t-axes and,

$$\begin{aligned} \Delta_r u_{ij} &= u_{i+1,j}-u_{ij} && \text{(forward difference)}\\ \nabla_r u_{ij} &= u_{ij}-u_{i-1,j} && \text{(backward difference)}\\ \text{and}\quad \delta_r u_{ij} &= u_{i+\frac{1}{2},j}-u_{i-\frac{1}{2},j} && \text{(central difference).}\end{aligned} \tag{5.2.4}$$

The finite-difference analogue of (5.1.1) is therefore given by,

$$\begin{aligned} u_{i,j+1} &= u_{ij}+\lambda(\theta_1\delta_r u_{i+\frac{1}{2},j+1}-\theta_2\delta_r u_{i-\frac{1}{2},j+1}+\theta_3\delta_r u_{i+\frac{1}{2},j}-\theta_4\delta_r u_{i-\frac{1}{2},j})\\ &\quad + \frac{\alpha}{2i}\lambda(\alpha_1\Delta_r u_{i,j+1}+\alpha_2\nabla_r u_{ij}+\alpha_3\nabla_r u_{i,j+1}+\alpha_4\Delta_r u_{ij}) \end{aligned} \tag{5.2.5}$$

where $\lambda=\frac{\Delta t}{(\Delta r)^2}$ the mesh ratio. Some of the well-known methods to approximate (5.1.1) may be obtained by an appropriate choice of the weighting parameters. For points not on the axis ($r\neq 0$), we have the following examples:

(a) The Classical Explicit Scheme

By taking $\theta_1=\theta_2=0$; $\theta_3=\theta_4=1$; $\alpha_1=\alpha_3=0$ and $\alpha_2=\alpha_4=1$, equation (5.2.5) becomes,

$$u_{i,j+1} = (1-2\lambda)u_{ij}+(1-\frac{\alpha}{2i})\lambda u_{i-1,j}+(1+\frac{\alpha}{2i})\lambda u_{i+1,j} \tag{5.2.6}$$

or
$$u_{i,j+1} = (1-2\lambda)u_{ij}+p_i u_{i-1,j}+q_i u_{i+1,j} \text{ ,} \tag{5.2.7}$$

where,

$$\begin{aligned} p_i &= (1 - \frac{\alpha}{2i})\lambda \\ q_i &= (1 + \frac{\alpha}{2i})\lambda \text{ .} \end{aligned} \tag{5.2.8}$$

We note that the formula (5.2.6/7) is just the explicit scheme of (3.13.25).

(b) The Fully Implicit Scheme

With $\theta_1=\theta_2=1$; $\theta_3=\theta_4=0$; $\alpha_1=\alpha_3=1$ and $\alpha_2=\alpha_4=0$, equation (5.2.5) reduces to

$$-(1+ \frac{\alpha}{2i})\lambda u_{i+1,j+1}+(1+2\lambda)u_{i,j+1}-(1- \frac{\alpha}{2i})\lambda u_{i-1,j+1} = u_{ij} \text{ ,} \quad i=1,2,\ldots,m-1 \text{ ,} \tag{5.2.9}$$

or
$$-q_i u_{i+1,j+1}+(1+2\lambda)u_{i,j+1}-p_i u_{i-1,j+1} = u_{ij} \text{ ,} \quad i=1,2,\ldots,m-1, \tag{5.2.10}$$

which coincides with equation (3.14.1a).

(c) The Crank-Nicolson Scheme

The Crank-Nicolson scheme is obtained by putting $\theta_1=\theta_2=\theta_3=\theta_4=\frac{1}{2}$ and $\alpha_1=\alpha_2=\alpha_3=\alpha_4=\frac{1}{2}$, which leads to the formula

$$\tfrac{1}{2}q_i u_{i+1,j+1}-(1+\lambda)u_{i,j+1}+\tfrac{1}{2}p_i u_{i-1,j+1} = -\tfrac{1}{2}q_i u_{i+1,j}-(1-\lambda)u_{ij}-\tfrac{1}{2}p_i u_{i-1,j}\text{,} \quad \text{for } i=1,2,\ldots,m-1, \tag{5.2.11}$$

and this bears the same form as in (3.14.10).

(d) The Asymmetric RL Approximation

If we choose $\theta_1=\theta_4=1$; $\theta_2=\theta_3=0$; $\alpha_1=\alpha_2=1$ and $\alpha_3=\alpha_4=0$, equation (5.2.5) gives us the following RL approximation,

$$(1+q_i)u_{i,j+1}-q_iu_{i+1,j+1} = (1-p_i)u_{ij}+p_iu_{i-1,j}, \quad i=1,2,\ldots,m-1 \tag{5.2.12}$$

whose asymmetric computational molecule is given by,

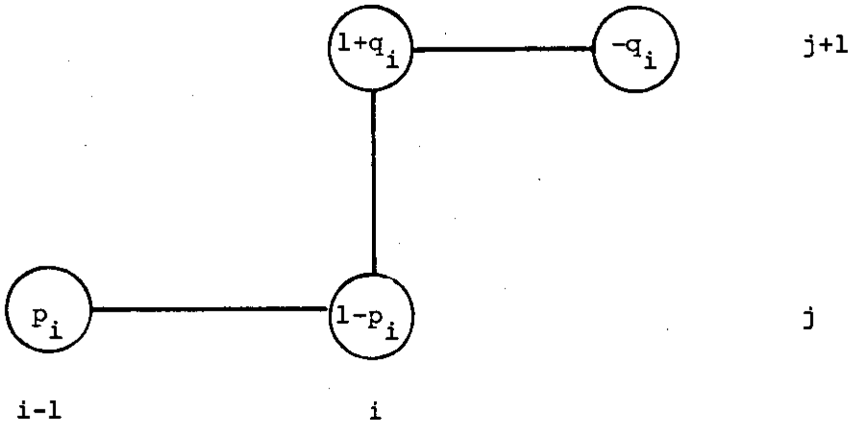


FIGURE 5.2.1

(e) The Asymmetric LR Approximation

By inserting $\theta_1=\theta_4=0$; $\theta_2=\theta_3=1$, $\alpha_1=\alpha_2=0$ and $\alpha_3=\alpha_4=1$, we get the following LR approximation,

$$(1+p_i)u_{i,j+1}-p_iu_{i-1,j+1} = (1-q_i)u_{ij}+q_iu_{i+1,j}, \quad i=1,2,\ldots,m-1 \tag{5.2.13}$$

whose asymmetric computational molecule is given by,

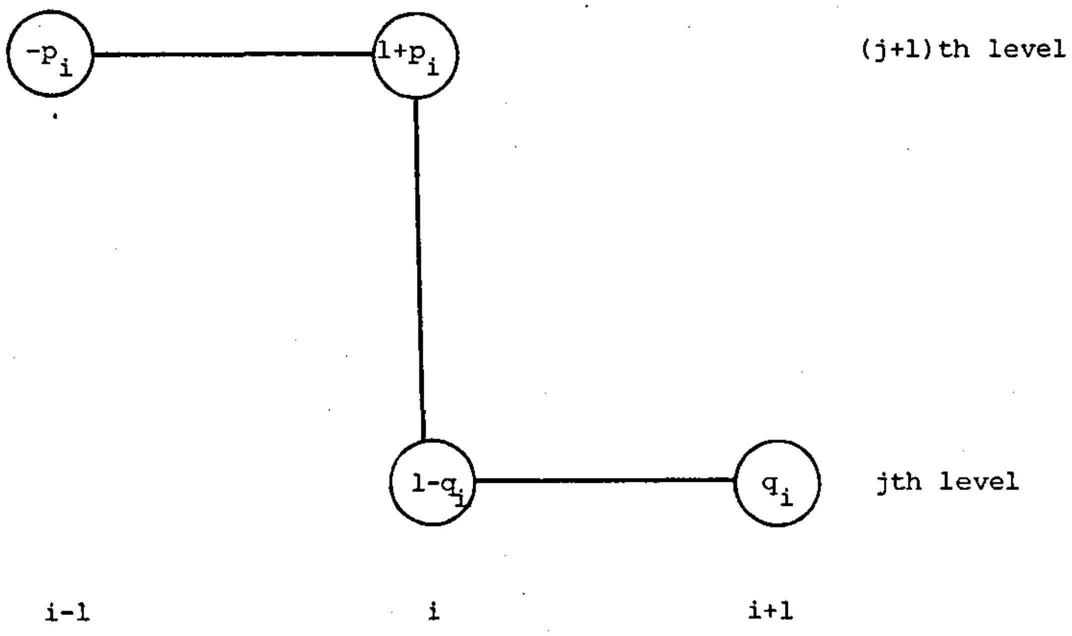


FIGURE 5.2.2

5.3 DERIVATION OF THE GE SCHEMES

If we apply equation (5.2.12) at the point $(r_{i-1},t_{j+\frac{1}{2}})$ we find that,

$$-q_{i-1}u_{i,j+1}+(1+q_{i-1})u_{i-1,j+1} = (1-p_{i-1})u_{i-1,j}+p_{i-1}u_{i-2,j} \qquad (5.3.1)$$

and together with equation (5.2.13) form the following stencil,

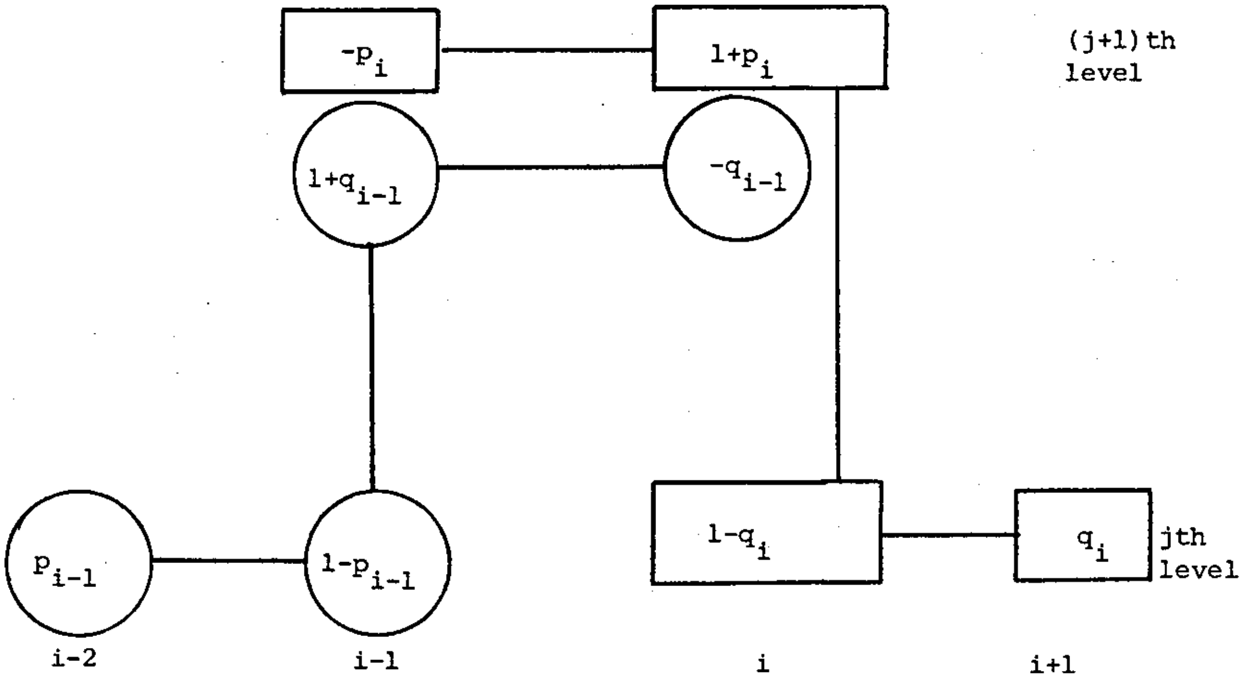


FIGURE 5.3.1

as well as the system,

$$\begin{bmatrix} (1+q_{i-1}) & -q_{i-1} \\ -p_i & (1+p_i) \end{bmatrix} \begin{bmatrix} u_{i-1,j+1} \\ u_{i,j+1} \end{bmatrix} = \begin{bmatrix} (1-p_{i-1}) & 0 \\ 0 & (1-q_i) \end{bmatrix} \begin{bmatrix} u_{i-1,j} \\ u_{ij} \end{bmatrix} + \begin{bmatrix} p_{i-1}u_{i-2,j} \\ q_iu_{i+1,j} \end{bmatrix}$$

(5.3.2)

i.e., $$A\underline{u}_{j+1} = B\underline{u}_j + \hat{\underline{u}}_j\ , \qquad (5.3.3)$$

or $$\underline{u}_{j+1} = A^{-1}B\underline{u}_j + A^{-1}\hat{\underline{u}}_j\ . \qquad (5.3.4)$$

Since

$$A^{-1} = \frac{1}{(1+p_i+q_{i-1})}\begin{bmatrix}(1+p_i) & q_{i-1}\\ & \\ p_i & (1+q_{i-1})\end{bmatrix} \qquad (5.3.5)$$

equation (5.3.4) leads to the following explicit equations for *general points* not on the axis,

$$u_{i-1,j+1} = \frac{1}{(1+p_i+q_{i-1})}[(1+p_i)p_{i-1}u_{i-2,j}+(1+p_i)(1-p_{i-1})u_{i-1,j}+q_{i-1}$$

$$(1-q_i)u_{ij}+q_iq_{i-1}u_{i+1,j}] \qquad (5.3.6)$$

and

$$u_{i,j+1} = \frac{1}{(1+p_i+q_{i-1})}[p_ip_{i-1}u_{i-2,j}+p_i(1-p_{i-1})u_{i-1,j}+(1+q_{i-1})(1-q_i)u_{ij}$$

$$+(1+q_{i-1})q_iu_{i+1,j}] \qquad (5.3.7)$$

whose computational molecules are given respectively by

Equation (5.3.6)

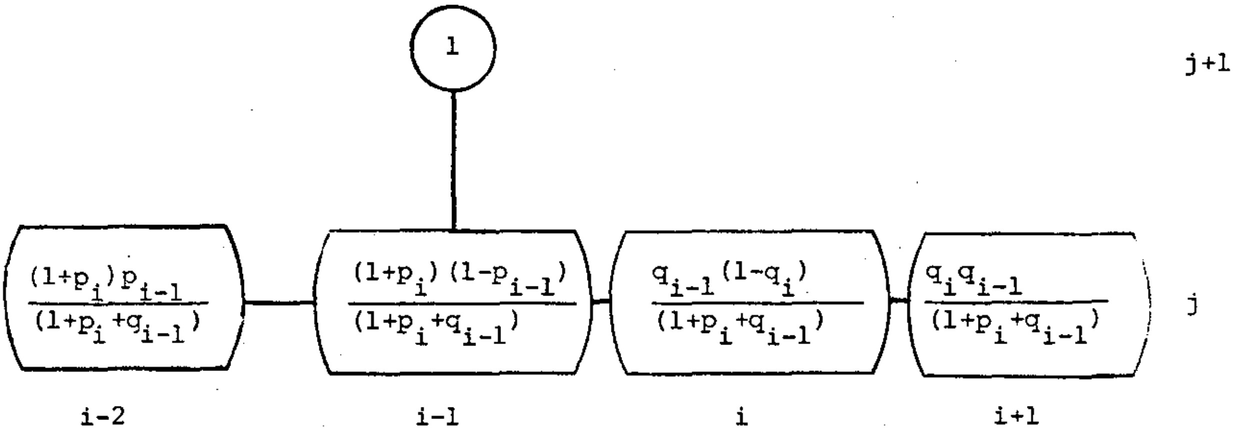


FIGURE 5.3.2

and

Equation (5.3.7)

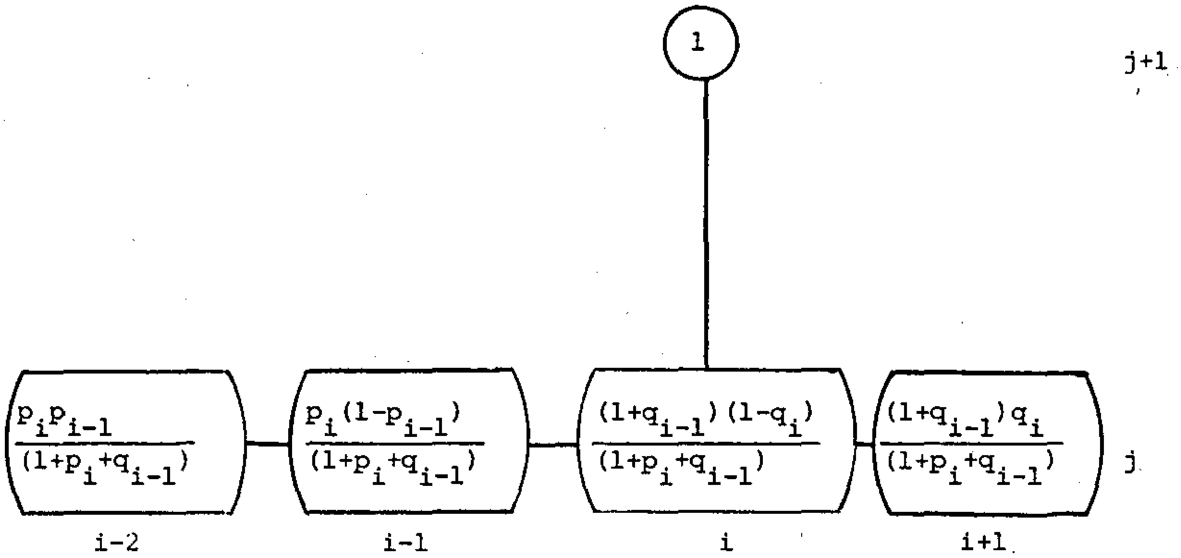


FIGURE 5.3.3

The diagrammatic representation of the GER scheme is given by Figure 5.3.4 below:

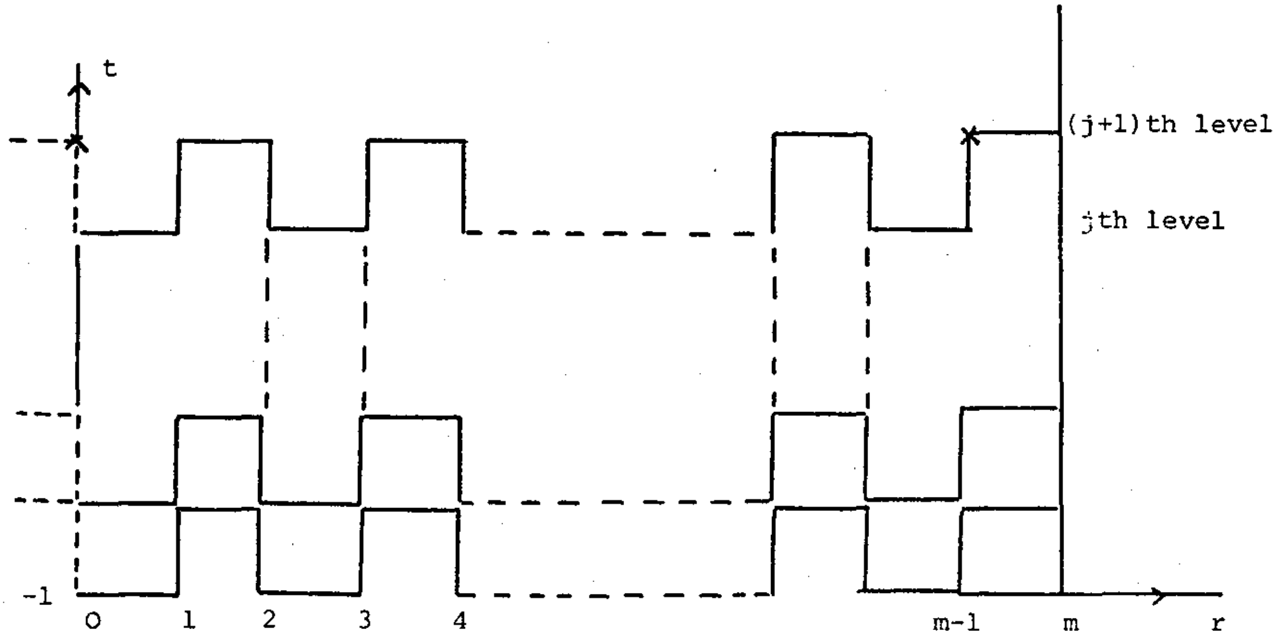


FIGURE 5.3.4: The GER scheme

By choosing $\theta_1=\theta_4=0$, $\theta_2=\theta_3=1$, $\alpha_3=\alpha_4=1$ and $\alpha_1=\alpha_2=0$ in (5.2.1/2), we obtain the following approximations of the derivatives at the points $(0,j+\frac{1}{2})$ on the axis:

$$\left(\frac{\partial^2 U}{\partial r^2}\right)_{(0,j+\frac{1}{2})} \simeq (u_{1,j}-u_{0,j}-u_{0,j+1}+u_{-1,j+1})/(\Delta r)^2 \ , \tag{5.3.8}$$

$$\left(\frac{\partial U}{\partial t}\right)_{(0,j+\frac{1}{2})} \simeq (u_{0,j+1}-u_{0,j})/\Delta t \ , \tag{5.3.9}$$

and

$$\left(\frac{\partial U}{\partial r}\right)_{(0,j+\frac{1}{2})} \simeq (u_{0,j+1}-u_{-1,j+1}+u_{1,j}-u_{0j})/2\Delta r \tag{5.3.10}$$

From (3.13.28), the following relationship holds on the axis,

$$\left(\frac{\partial U}{\partial t}\right)_{(0,j+\frac{1}{2})} = (1+\alpha)\left(\frac{\partial^2 U}{\partial r^2}\right)_{(0,j+\frac{1}{2})} \tag{5.3.11}$$

and the substitution of (5.3.8) and (5.3.9) into (5.3.11) leads to the approximations

$$u_{0,j+1}-u_{0,j} = (1+\alpha)\lambda(u_{1j}-u_{0j}-u_{0,j+1}+u_{-1,j+1}) \ . \tag{5.3.12}$$

By utilising (5.3.10) and the boundary condition (5.1.1a) in which $\frac{\partial U}{\partial r}(0,t) = 0$, we arrive at the following formula for the fictitious values $u_{-1,j+1}$,

$$u_{-1,j+1} = u_{0,j+1}+u_{1j}-u_{0j} \ . \tag{5.3.13}$$

Now by inserting $u_{-1,j+1}$ into (5.3.12), we therefore obtain the approximations to *the left boundary values* as,

$$u_{0,j+1} = (1-2\hat{\alpha})u_{0j} + 2\hat{\alpha}u_{1j} \ , \tag{5.3.14}$$

where $\hat{\alpha} = (1+\alpha)\lambda$ and $\lambda = \frac{\Delta t}{(\Delta r)^2}$, the mesh ratio. The solutions at the single ungrouped point near the right boundary can be obtained from (5.2.12) by taking i=m-1 to give,

$$u_{m-1,j+1} = (p_{m-1}u_{m-2,j}+(1-p_{m-1})u_{m-1,j}+q_{m-1}u_{m,j+1})/(1+q_{m-1}) \qquad (5.3.15)$$

Hence the GER scheme is expressed by the following implicit equations,

$$u_{0,j+1} = (1-2\hat{\alpha})u_{0j} + 2\hat{\alpha}u_{1j} \text{ ,}$$

$$\left.\begin{array}{l} (1+q_{i-1})u_{i-1,j+1}-q_{i-1}u_{i,j+1} = p_{i-1}u_{i-2,j}+(1-p_{i-1})u_{i-1,j} \text{ ,} \\ \\ -p_iu_{i-1,j+1}+(1+p_i)u_{i,j+1} = (1-q_i)u_{ij}+q_iu_{i+1,j} \text{ ,} \end{array}\right\} \begin{array}{l} i=2,4,\ldots, \\ \quad (m-4),(m-2), \\ m \text{ even and } m>4, \end{array}$$

and

$$u_{m-1,j+1} = \frac{q_{m-1}}{(1+q_{m-1})}u_{m,j+1} + \frac{(1-p_{m-1})}{(1+q_{m-1})}u_{m-1,j} + \frac{p_{m-1}}{(1+q_{m-1})}u_{m-2,j} \text{ ,}$$

or in matrix form as,

$$\left[\begin{array}{c|cc|cc|c|cc|c}
1 & & & & & & & & \\ \hline
 & (1+q_1) & -q_1 & & & & & & \\
 & -p_2 & (1+p_2) & & & & & & \\ \hline
 & & & (1+q_3) & -q_3 & & & O & \\
 & & & -p_4 & (1+p_4) & & & & \\ \hline
 & & & & & \ddots & & & \\ \hline
 & & & & & & (1+q_{m-3}) & -q_{m-3} & \\
 & & O & & & & -p_{m-2} & (1+p_{m-2}) & \\ \hline
 & & & & & & & & 1
\end{array}\right]
\left[\begin{array}{c} u_0 \\ u_1 \\ u_2 \\ u_3 \\ u_4 \\ \vdots \\ u_{m-3} \\ u_{m-2} \\ u_{m-1} \end{array}\right]_{j+1}$$

$$= \left[\begin{array}{cc|cc|cc|c|cc}
(1-2\hat{\alpha}) & 2\hat{\alpha} & & & & & & & \\
p_1 & (1-p_1) & & & & & & & \\ \hline
 & & (1-q_2) & q_2 & & & & & \\
 & & p_3 & (1-p_3) & & & & & \\ \hline
 & & & & (1-q_4) & q_4 & & & \\
 & & & & p_5 & (1-p_5) & & O & \\ \hline
 & & & & & & \ddots & & \\ \hline
 & & & & & & & (1-q_{m-2}) & q_{m-2} \\
 & & & O & & & & \frac{p_{m-1}}{(1+q_{m-1})} & \frac{(1-p_{m-1})}{(1+q_{m-1})}
\end{array}\right]
\left[\begin{array}{c} u_0 \\ u_1 \\ u_2 \\ u_3 \\ u_4 \\ \vdots \\ u_{m-2} \\ u_{m-1} \end{array}\right]_j$$

$$+ \underline{b}_1 \qquad (5.3.16)$$

where $\underline{b}_1 = (0,0,\ldots,\frac{q_{m-1}}{(1+q_{m-1})}u_{m,j+1})^T$.

Now, if we let,

$$G_1 = \left[\begin{array}{c|c|c|c|c|c}
0 & & & & & \\ \hline
 & G_1^{(1)} & & & & \\ \hline
 & & G_1^{(2)} & & O & \\ \hline
 & & & \ddots & & \\ \hline
 & & O & & G_1^{(\frac{1}{2}(m-2))} & \\ \hline
 & & & & & 0
\end{array}\right]_{(m\times m)} \qquad (5.3.17)$$

and

$$G_2=\left[\begin{array}{cc|c|c|c|c|cc} -2\hat{\alpha} & 2\hat{\alpha} & & & & & & \\ p_1 & -p_1 & & & & & & \\ \hline & & G_2^{(1)} & & & & & \\ \hline & & & G_2^{(2)} & & & \multicolumn{2}{c}{O} \\ \hline & & & & \ddots & & & \\ \hline & & & & & G_2^{(\frac{1}{2}(m-4))} & & \\ \hline & & & O & & & -q_{m-2} & q_{m-2} \\ & & & & & & \frac{p_{m-1}}{(1+q_{m-1})} & \frac{-(p_{m-1}+q_{m-1})}{(1+q_{m-1})} \end{array}\right]_{m\times m}$$

(5.3.18)

with

$$G_1^{(i)} = \begin{bmatrix} q_{2i-1} & -q_{2i-1} \\ -p_{2i} & p_{2i} \end{bmatrix}, \quad i=1,2,\ldots,\tfrac{1}{2}(m-2)$$

and

$$G_2^{(i)} = \begin{bmatrix} -q_{2i} & q_{2i} \\ p_{2i+1} & -p_{2i+1} \end{bmatrix}, \quad i=1,2,\ldots,\tfrac{1}{2}(m-4)$$

then, the GER scheme can be written as,

$$(I+G_1)\underline{u}_{j+1} = (I+G_2)\underline{u}_j + \underline{b}_1 , \qquad (5.3.19)$$

or $$\underline{u}_{j+1} = (I+G_1)^{-1}(I+G_2)\underline{u}_j + \hat{\underline{b}}_1 , \qquad (5.3.20)$$

where $\underline{u}_j = (u_{0,j},u_{1,j},\ldots,u_{m-1,j})^T$ and $\hat{\underline{b}}_1 = (I+G_1)^{-1}\underline{b}_1$.

(ii) The GEL Scheme

The GEL scheme is represented diagrammatically by Figure 5.3.5 below,

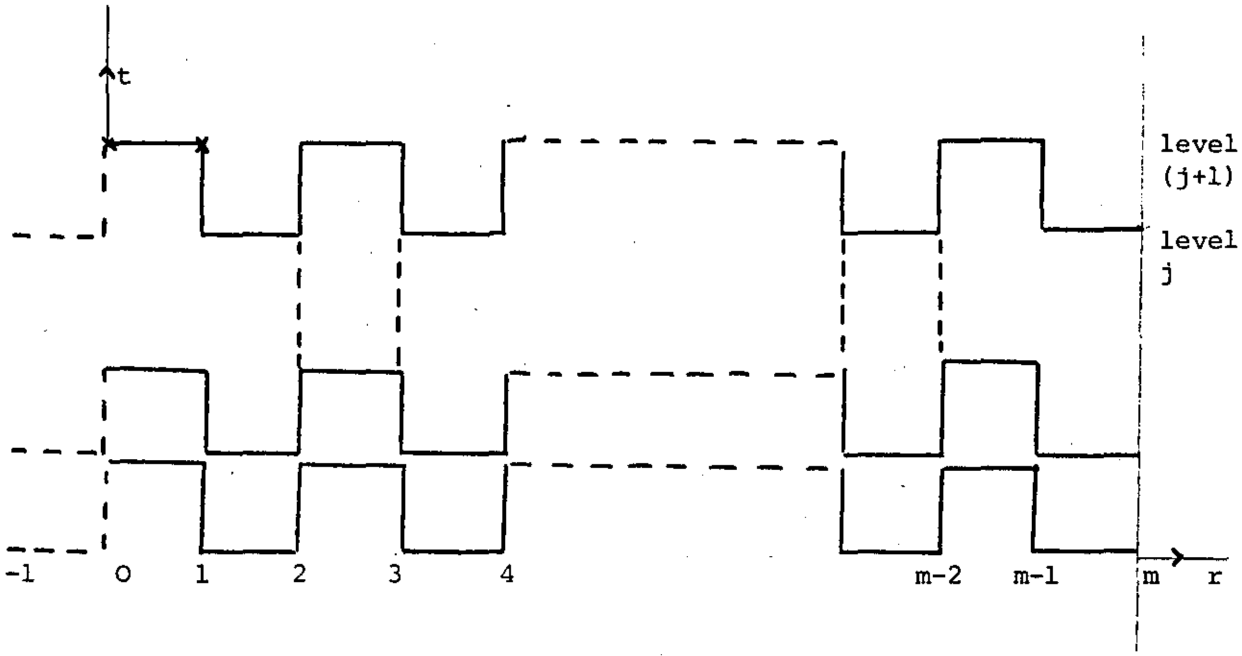


FIGURE 5.3.5: The GEL Scheme

By choosing $\theta_1=\theta_4=1$ and $\theta_2=\theta_3=0$ in (5.2.1), we have as approximations to (5.3.11) at r=0, the formula,

$$(1+\hat{\alpha})u_{0,j+1}-\hat{\alpha}u_{1,j+1} = (1-\hat{\alpha})u_{0j}+\hat{\alpha}u_{-1,j} \ . \qquad (5.3.21)$$

From (5.2.13), we also obtain, with i=1, the relation,

$$-p_1u_{0,j+1}+(1+p_1)u_{1,j+1} = (1-q_1)u_{1j}+q_1u_{2j} \ . \qquad (5.3.22)$$

Equations (5.3.21) and (5.3.22) form the coupled system:

$$\begin{bmatrix}(1+\hat{\alpha}) & -\hat{\alpha}\\ -p_1 & (1+p_1)\end{bmatrix}\begin{bmatrix}u_{0,j+1}\\ u_{1,j+1}\end{bmatrix} = \begin{bmatrix}(1-\hat{\alpha}) & 0\\ 0 & (1-q_1)\end{bmatrix}\begin{bmatrix}u_{0j}\\ u_{1j}\end{bmatrix} + \begin{bmatrix}\hat{\alpha}u_{-1,j}\\ q_1u_{2,j}\end{bmatrix} \quad (5.3.23)$$

i.e., $$A\underline{u}_{j+1} = B\underline{u}_j + \hat{\underline{u}}_j \ ,$$

or $$\underline{u}_{j+1} = A^{-1}B\underline{u}_j + A^{-1}\hat{\underline{u}}_j \ , \qquad (5.3.24)$$

and with

$$A^{-1} = \frac{1}{(1+p_1+\hat{\alpha})}\begin{bmatrix}(1+p_1) & \hat{\alpha}\\ p_1 & (1+\hat{\alpha})\end{bmatrix} \ ,$$

this leads to,

$$u_{0,j+1} = \frac{1}{(1+p_1+\hat{\alpha})} \{(1+p_1)(1-\hat{\alpha})u_{0j}+(1-q_1)\hat{\alpha}u_{1j}+(1+p_1)\hat{\alpha}u_{-1,j}+\hat{\alpha}q_1u_{2j}\} \tag{5.3.25}$$

and

$$u_{1,j+1} = \frac{1}{(1+p_1+\hat{\alpha})} \{p_1(1-\hat{\alpha})u_{0j}+(1+\hat{\alpha})(1-q_1)u_{1j}+p_1\hat{\alpha}u_{-1,j}+(1+\hat{\alpha})q_1u_{2j}\}. \tag{5.3.26}$$

From the boundary condition (5.1.1a), we have $(\frac{\partial U}{\partial r})_{(0,j+\frac{1}{2})}=0$ and if we choose $\alpha_1=\alpha_3=0$ and $\alpha_2=\alpha_4=1$ in (5.2.2), we obtain the following approximation on the axis,

$$(\frac{\partial U}{\partial r})_{(0,j+\frac{1}{2})} \approx (u_{1,j}-u_{-1,j})/2\Delta r = 0$$

and hence $u_{-1,j}=u_{1,j}$. With these values, the required equations for the left boundary as well as the single ungrouped point adjacent to it are given respectively (from (5.3.25) and (5.3.26)) by,

$$u_{0,j+1} = \frac{1}{(1+p_1+\hat{\alpha})} \{(1+p_1)(1-\hat{\alpha})u_{0j}+(2+p_1-q_1)\hat{\alpha}u_{1j}+\hat{\alpha}q_1u_{2j}\} \tag{5.3.27}$$

and

$$u_{1,j+1} = \frac{1}{(1+p_1+\hat{\alpha})} \{p_1(1-\hat{\alpha})u_{0j}+((1+\hat{\alpha})(1-q_1)+p_1\hat{\alpha})u_{1j}+(1+\hat{\alpha})q_1u_{2j}\}. \tag{5.3.28}$$

These equations together with the system (5.3.2/3) for i=3,5,...,(m-3), (m-1) describe the GEL scheme which can be written in matrix form as,

$$\left[\begin{array}{c|c|cc|cc|c|cc}
1 & 0 & & & & & & & \\
\hline
0 & 1 & & & & & & & \\
\hline
 & & (1+q_2) & -q_2 & & & & & \\
 & & -p_3 & (1+p_3) & & & & & \\
\hline
 & & & & (1+q_4) & -q_4 & & & \\
 & & & & -p_5 & (1+p_5) & & \bigcirc & \\
\hline
 & & & & & & \ddots & & \\
\hline
 & & & & & & & (1+q_{m-2}) & -q_{m-2} \\
 & & & & \bigcirc & & & -p_{m-1} & (1+p_{m-1})
\end{array}\right]
\begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ \vdots \\ u_{m-2} \\ u_{m-1} \end{bmatrix}_{j+1}$$

$$
= \left[\begin{array}{ccc|cc|c|cc|c}
1-\frac{(2+p_1)\hat{\alpha}}{(1+p_1+\hat{\alpha})} & \frac{(2+p_1-q_1)\hat{\alpha}}{(1+p_1+\hat{\alpha})} & \frac{\hat{\alpha}q_1}{(1+p_1+\hat{\alpha})} & & & & & & \\
\frac{p_1(1-\hat{\alpha})}{(1+p_1+\hat{\alpha})} & 1+\frac{p_1(\hat{\alpha}-1)-q_1(1+\hat{\alpha})}{(1+p_1+\hat{\alpha})} & \frac{(1+\hat{\alpha})q_1}{(1+p_1+\hat{\alpha})} & & & & & & \\
0 & p_2 & (1-p_2) & & & & & & \\
\hline
 & & & (1-q_3) & q_3 & & & & \\
 & & & p_4 & (1-p_4) & & & O & \\
\hline
 & & & & & \ddots & & & \\
\hline
 & & & O & & & (1-q_{m-3}) & q_{m-3} & \\
 & & & & & & p_{m-2} & (1-p_{m-2}) & \\
\hline
 & & & & & & & & (1-q_{m-1})
\end{array}\right]
\left[\begin{array}{c}
u_0 \\ u_1 \\ u_2 \\ u_3 \\ u_4 \\ \vdots \\ u_{m-3} \\ u_{m-2} \\ u_{m-1}
\end{array}\right]_j
$$

$$
+ \underline{b}_2 , \qquad (5.3.29)
$$

where $\underline{b}_2 = (0,0,0,\ldots,q_{m-1}u_{mj})^T$.

If we define,

$$\hat{G}_1 = \left[\begin{array}{ccc|c|c|c|c|c}
\frac{-(2+p_1)\hat{\alpha}}{(1+p_1+\hat{\alpha})} & \frac{(2+p_1-q_1)\hat{\alpha}}{(1+p_1+\hat{\alpha})} & \frac{\hat{\alpha}q_1}{(1+p_1+\hat{\alpha})} & & & & & \\
\frac{p_1(1-\hat{\alpha})}{(1+p_1+\hat{\alpha})} & \frac{p_1(\hat{\alpha}-1)-q_1(1+\hat{\alpha})}{(1+p_1+\hat{\alpha})} & \frac{(1+\hat{\alpha})q_1}{(1+p_1+\hat{\alpha})} & & & & & \\
0 & p_2 & -p_2 & & & & & \\ \hline
 & & & -G_1^{(2)} & & & & \\ \hline
 & & & & -G_1^{(3)} & & O & \\ \hline
 & & & & & \ddots & & \\ \hline
 & & & O & & & -G_1^{(\frac{1}{2}(m-2))} & \\ \hline
 & & & & & & & -q_{m-1}
\end{array}\right] \qquad (5.3.30) \qquad (m\times m$$

and

$$\hat{G}_2 = \left[\begin{array}{cc|c|c|c|c}
0 & 0 & & & & \\
0 & 0 & & & & \\ \hline
 & & -G_2^{(1)} & & & \\ \hline
 & & & -G_2^{(2)} & & O \\ \hline
 & & & & \ddots & \\ \hline
 & & & O & & -G_2^{(\frac{1}{2}(m-2))}
\end{array}\right]_{(m\times m)} \qquad (5.3.31)$$

then the GEL scheme takes the form,

$$(I+\hat{G}_2)\underline{u}_{j+1} = (I+\hat{G}_1)\underline{u}_j + \underline{b}_2 , \qquad (5.3.32)$$

or

$$\underline{u}_{j+1} = (I+\hat{G}_2)^{-1}(I+\hat{G}_1)\underline{u}_j + \hat{\underline{b}}_2 , \qquad (5.3.33)$$

where $\underline{u}_j = (u_{0j}, u_{1j}, \ldots, u_{m-1,j})^T$ and $\hat{\underline{b}}_2 = (I+\hat{G}_2)^{-1}\underline{b}_2$.

(iii) The (S)AGE and (D)AGE Schemes

The alternating group explicit schemes are formed by the application of the GER and the GEL processes in their appropriate sequences. Thus, the following formulae constitute the (S)AGE two-step scheme:

$$\left.\begin{aligned} &(I+G_1)\underline{u}_{j+1} = (I+G_2)\underline{u}_j + \underline{b}_1 \\ \text{and} \quad &(I+\hat{G}_2)\underline{u}_{j+2} = (I+\hat{G}_1)\underline{u}_{j+1} + \underline{b}_2 \end{aligned}\right\} j=0,2,4,\ldots \qquad (5.3.34)$$

and the (D)AGE four-step scheme is given by,

$$\left.\begin{aligned} &(I+G_1)\underline{u}_{j+1} = (I+G_2)\underline{u}_j + \underline{b}_1 , \\ &(I+\hat{G}_2)\underline{u}_{j+2} = (I+\hat{G}_1)\underline{u}_{j+1} + \underline{b}_2 , \\ &(I+\hat{G}_2)\underline{u}_{j+3} = (I+\hat{G}_1)\underline{u}_{j+2} + \underline{b}_2 , \\ \text{and} \quad &(I+G_1)\underline{u}_{j+4} = (I+G_2)\underline{u}_{j+3} + \underline{b}_1 . \end{aligned}\right\} j=0,4,8, \qquad (5.3.35)$$

5.4 TRUNCATION ERROR ANALYSIS OF THE GE METHODS

(i) Truncation Error for the GER Scheme

The truncation errors of the approximations of the *left boundary* values as given by equation (5.3.14) can be estimated by expanding $U_{0,j+1}, U_{0,j}$ and $U_{1,j}$ about the point $(r_0, t_{j+\frac{1}{2}})$. This leads to the following expression of the truncation error:

$$
\begin{aligned}
T_{LB} &= T_{5.3.14} \\
&= -2\hat{\alpha}(\Delta r)(\frac{\partial U}{\partial r})_{0,j+\frac{1}{2}}+(\Delta t)(\frac{\partial U}{\partial t})_{0,j+\frac{1}{2}}-\hat{\alpha}(\Delta r)^2(\frac{\partial^2 U}{\partial r^2})_{0,j+\frac{1}{2}}+\hat{\alpha}(\Delta r)(\Delta t)(\frac{\partial^2 U}{\partial r\partial t})_{0,j+\frac{1}{2}} \\
&\quad -\frac{1}{3}\hat{\alpha}(\Delta r)^3(\frac{\partial^3 U}{\partial r^3})_{0,j+\frac{1}{2}}+\frac{1}{2}\hat{\alpha}(\Delta r)^2(\Delta t)(\frac{\partial^3 U}{\partial r^2\partial t})_{0,j+\frac{1}{2}}-\frac{1}{4}\hat{\alpha}(\Delta r)(\Delta t)^2(\frac{\partial^3 U}{\partial r\partial t^2})_{0,j+\frac{1}{2}} \\
&\quad +\frac{1}{24}(\Delta t)^3(\frac{\partial^3 U}{\partial t^3})_{0,j+\frac{1}{2}}-\frac{\hat{\alpha}}{12}(\Delta r)^4(\frac{\partial^4 U}{\partial r^4})_{0,j+\frac{1}{2}}+\frac{\hat{\alpha}}{6}(\Delta r)^3(\Delta t)(\frac{\partial^4 U}{\partial r^3\partial t})_{0,j+\frac{1}{2}} \\
&\quad -\frac{\hat{\alpha}}{8}(\Delta r)^2(\Delta t)^2(\frac{\partial^4 U}{\partial r^2\partial t^2})_{0,j+\frac{1}{2}}+\frac{\hat{\alpha}}{24}(\Delta r)(\Delta t)^3(\frac{\partial^4 U}{\partial r\partial t^3})_{0,j+\frac{1}{2}}+\frac{\hat{\alpha}}{192}(\Delta t)^4 \\
&\quad (\frac{\partial^4 U}{\partial t^4})_{0,j+\frac{1}{2}}+O((\Delta r)^{\alpha_1}(\Delta t)^{\alpha_2})\ ,\ \alpha_1+\alpha_2=5\ . \qquad (5.4.1)
\end{aligned}
$$

The truncation errors for the GER scheme at *the general grouped* points can be derived by resorting to Taylor's series expansions of equations (5.3.6) and (5.3.7) about the points $(r_{i-1}, t_{j+\frac{1}{2}})$ and $(r_i, t_{j+\frac{1}{2}})$ respectively. Thus, we obtain

$$
\begin{aligned}
T_{5.3.6} = &\{(1+p_i)p_{i-1}+(1+q_i)q_{i-1}\}(\Delta r)(\frac{\partial U}{\partial r})_{i-1,j+\frac{1}{2}}+(1+p_i+q_{i-1})(\Delta t)(\frac{\partial U}{\partial t})_{i-1,j+\frac{1}{2}} \\
&-\frac{1}{2}\{(1+p_i)p_{i-1}+(1+3q_i)q_{i-1}\}(\Delta r)^2(\frac{\partial^2 U}{\partial r^2})_{i-1,j+\frac{1}{2}} \\
&-\frac{1}{2}\{(1+p_i)p_{i-1}-(1+q_i)q_{i-1}\}(\Delta r)(\Delta t)(\frac{\partial^2 U}{\partial r\partial t})_{i-1,j+\frac{1}{2}}+\frac{1}{6}\{(1+p_i)p_{i-1}- \\
&(1+7q_i)q_{i-1}\}(\Delta r)^3(\frac{\partial^3 U}{\partial r^3})_{i-1,j+\frac{1}{2}}+\frac{1}{4}\{(1+p_i)p_{i-1}+(1+3q_i)q_{i-1}\}(\Delta r)^2(\Delta t) \\
&(\frac{\partial^3 U}{\partial r^2\partial t})_{i-1,j+\frac{1}{2}}+\frac{1}{8}\{(1+p_i)p_{i-1}-(1+q_i)q_{i-1}\}(\Delta r)(\Delta t)^2(\frac{\partial^3 U}{\partial r\partial t^2})_{i-1,j+\frac{1}{2}} \\
&+\frac{1}{24}(1+p_i+q_{i-1})(\Delta t)^3(\frac{\partial^3 U}{\partial t^3})_{i-1,j+\frac{1}{2}}-\frac{1}{24}\{(1+p_i)p_{i-1}+(1+15q_i)q_{i-1}\}
\end{aligned}
$$

$$(\Delta r)^4\left(\frac{\partial^4 U}{\partial r^4}\right)_{i-1,j+\frac{1}{2}} - \frac{1}{12}\{(1+p_i)p_{i-1}-(1+7q_i)q_{i-1}\}(\Delta r)^3(\Delta t)\left(\frac{\partial^4 U}{\partial r^3\partial t}\right)_{i-1,j+\frac{1}{2}}$$

$$-\frac{1}{16}\{(1+p_i)p_{i-1}+(1+3q_i)q_{i-1}\}(\Delta r)^2(\Delta t)^2\left(\frac{\partial^4 U}{\partial r^2\partial t^2}\right)_{i-1,j+\frac{1}{2}} - \frac{1}{48}\{(1+p_i)p_{i-1}$$

$$-(1+q_i)q_{i-1}\}(\Delta r)(\Delta t)^3\left(\frac{\partial^4 U}{\partial r\partial t^3}\right)_{i-1,j+\frac{1}{2}}+O((\Delta r)^{\alpha_1}(\Delta t)),$$

where $\alpha_1+\alpha_2=5$; (5.4.2)

and

$$T_{5.3.7} = \{(1+p_i)p_{i-1}-(1+q_i)_{i-1}q_i\}(\Delta r)\left(\frac{\partial U}{\partial r}\right)_{i,j+\frac{1}{2}}+(1+p_i+q_{i-1})(\Delta t)\left(\frac{\partial U}{\partial t}\right)_{i,j+\frac{1}{2}}$$

$$-\tfrac{1}{2}\{p_i(1+3p_{i-1})+(1+q_{i-1})q_i\}(\Delta r)^2\left(\frac{\partial^2 U}{\partial r^2}\right)_{i,j+\frac{1}{2}}-\tfrac{1}{2}\{p_i(1+p_{i-1})-(1+q_{i-1})q_i\}$$

$$(\Delta r)(\Delta t)\left(\frac{\partial^2 U}{\partial r\partial t}\right)_{i,j+\frac{1}{2}} + \frac{1}{6}\{p_i(1+7p_{i-1})-(1+q_{i-1})q_i\}(\Delta r)^3\left(\frac{\partial^3 U}{\partial r^3}\right)_{i,j+\frac{1}{2}}$$

$$+\frac{1}{4}\{p_i(1+3p_{i-1})+(1+q_{i-1})q_i\}(\Delta r)^2(\Delta t)\left(\frac{\partial^3 U}{\partial r^2\partial t}\right)_{i,j+\frac{1}{2}} + \frac{1}{8}\{p_i(1+p_{i-1})-$$

$$(1+q_{i-1})q_i\}(\Delta r)(\Delta t)^2\left(\frac{\partial^3 U}{\partial r\partial t^2}\right)_{i,j+\frac{1}{2}}+ \frac{1}{24}(1+p_i+q_{i-1})(\Delta t)^3\left(\frac{\partial^3 U}{\partial t^3}\right)_{i,j+\frac{1}{2}}$$

$$-\frac{1}{24}\{p_i(1+15p_{i-1})+(1+q_{i-1})q_i\}(\Delta r)^4\left(\frac{\partial^4 U}{\partial r^4}\right)_{i,j+\frac{1}{2}} - \frac{1}{12}\{p_i(1+7p_{i-1})-$$

$$(1+q_{i-1})q_i\}(\Delta r)^3(\Delta t)\left(\frac{\partial^4 U}{\partial r^3\partial t}\right)_{i,j+\frac{1}{2}} - \frac{1}{16}\{p_i(1+3p_{i-1})+(1+q_{i-1})q_i\}$$

$$(\Delta r)^2(\Delta t)^2\left(\frac{\partial^4 U}{\partial r^2\partial t^2}\right)_{i,j+\frac{1}{2}} - \frac{1}{48}\{p_i(1+p_{i-1})-(1+q_{i-1})q_i\}(\Delta r)(\Delta t)^3$$

$$\left(\frac{\partial^4 U}{\partial r\partial t^3}\right)_{i,j+\frac{1}{2}} + O((\Delta r)^{\alpha_1}(\Delta t)^{\alpha_2})\ ,\ \alpha_1+\alpha_2=5. \qquad (5.4.3)$$

Finally, from equation (5.3.15), the truncation error of the approximation near to the right boundary is given by,

$$T_R = T_{5.3.15}$$

$$= (p_i-q_i)(\Delta r)\left(\frac{\partial U}{\partial r}\right)_{m-1,j+\frac{1}{2}}+(\Delta t)\left(\frac{\partial U}{\partial t}\right)_{m-1,j+\frac{1}{2}}-\tfrac{1}{2}(p_i+q_i)(\Delta r)^2\left(\frac{\partial^2 U}{\partial r^2}\right)_{m-1,j+\frac{1}{2}}$$

$$-\tfrac{1}{2}(p_i+q_i)(\Delta r)(\Delta t)\left(\frac{\partial^2 U}{\partial r\partial t}\right)_{m-1,j+\frac{1}{2}} + \frac{1}{6}(p_i-q_i)(\Delta r)^3\left(\frac{\partial^3 U}{\partial r^3}\right)_{m-1,j+\frac{1}{2}}$$

$$+\frac{1}{4}(p_i-q_i)(\Delta r)^2(\Delta t)\left(\frac{\partial^3 U}{\partial r^2\partial t}\right)_{m-1,j+\frac{1}{2}}+\frac{1}{8}(p_i-q_i)(\Delta r)(\Delta t)^2\left(\frac{\partial^3 U}{\partial r\partial t^2}\right)_{m-1,j+\frac{1}{2}}$$

$$+\frac{1}{24}(\Delta t)^3\left(\frac{\partial^3 U}{\partial t^3}\right)_{m-1,j+\frac{1}{2}}-\frac{1}{24}(p_i+q_i)(\Delta r)^4\left(\frac{\partial^4 U}{\partial r^4}\right)_{m-1,j+\frac{1}{2}}$$
$$-\frac{1}{12}(p_i+q_i)(\Delta r)^3(\Delta t)\left(\frac{\partial^4 U}{\partial r^3\partial t}\right)_{m-1,j+\frac{1}{2}}-\frac{1}{8}(p_i+q_i)(\Delta r)^2(\Delta t)^2\left(\frac{\partial^4 U}{\partial r^2\partial t^2}\right)_{m-1,j+\frac{1}{2}}$$
$$-\frac{1}{48}(p_i+q_i)(\Delta r)(\Delta t)^3\left(\frac{\partial^4 U}{\partial r\partial t^3}\right)+O((\Delta r)^{\alpha_1}(\Delta t)^{\alpha_2}),\ \text{where } \alpha_1+\alpha_2=5. \tag{5.4.4}$$

(ii) Truncation Error for the GEL Scheme

The truncation errors at the left boundary and at the ungrouped point near to it can be derived by means of equations (5.3.27) and (5.3.28). By expanding these equations about the points $(r_0,t_{j+\frac{1}{2}})$ and $(r_1,t_{j+\frac{1}{2}})$, we obtain,

$$T_{LB}=T_{5.3.27}$$
$$=(1+p_1+\hat{\alpha})(\Delta t)\left(\frac{\partial U}{\partial t}\right)_{0,j+\frac{1}{2}}-(2+p_1+q_1)\hat{\alpha}(\Delta r)\left(\frac{\partial U}{\partial r}\right)_{0,j+\frac{1}{2}}-\tfrac{1}{2}(2+p_1+3q_1)\hat{\alpha}(\Delta r)^2$$
$$\left(\frac{\partial^2 U}{\partial r^2}\right)_{0,j+\frac{1}{2}}+\tfrac{1}{2}(2+p_1+q_1)\hat{\alpha}(\Delta r)(\Delta t)\left(\frac{\partial^2 U}{\partial r\partial t}\right)_{0,j+\frac{1}{2}}-\frac{1}{8}\{(1+p_1)(1-\hat{\alpha})+\hat{\alpha}q_1\}(\Delta t)^2$$
$$\left(\frac{\partial^2 U}{\partial t^2}\right)_{0,j+\frac{1}{2}}-\frac{1}{6}(2+p_1+7q_1)\hat{\alpha}(\Delta r)^3\left(\frac{\partial^3 U}{\partial r^3}\right)_{0,j+\frac{1}{2}}+\frac{1}{4}(2+p_1+3q_1)\hat{\alpha}(\Delta r)^2(\Delta t)$$
$$\left(\frac{\partial^3 U}{\partial r^2\partial t}\right)_{0,j+\frac{1}{2}}-\frac{1}{8}(2+p_1+q_1)\hat{\alpha}(\Delta r)(\Delta t)^2\left(\frac{\partial^3 U}{\partial r\partial t^2}\right)_{0,j+\frac{1}{2}}+\frac{1}{24}(1+p_1+\hat{\alpha})\ (\Delta t)^3$$
$$\left(\frac{\partial^3 U}{\partial t^3}\right)_{0,j+\frac{1}{2}}-\frac{1}{24}(2+p_1+q_1)\hat{\alpha}(\Delta r)^4\left(\frac{\partial^4 U}{\partial r^4}\right)_{0,j+\frac{1}{2}}+\frac{1}{12}(2+p_1+7q_1)\hat{\alpha}(\Delta r)^3(\Delta t)$$
$$\left(\frac{\partial^4 U}{\partial r^3\partial t}\right)_{0,j+\frac{1}{2}}-\frac{1}{16}(2+p_1+3q_1)\hat{\alpha}(\Delta r)^2(\Delta t)^2\left(\frac{\partial^4 U}{\partial r^2\partial t^2}\right)_{0,j+\frac{1}{2}}+\frac{1}{48}(2+p_1+q_1)$$
$$\hat{\alpha}(\Delta r)(\Delta t)^3\left(\frac{\partial^4 U}{\partial r\partial t^3}\right)_{0,j+\frac{1}{2}}-\frac{1}{384}(2+p_1)\hat{\alpha}(\Delta t)^4\ \left(\frac{\partial^4 U}{\partial t^4}\right)_{0,j+\frac{1}{2}}+$$
$$O((\Delta r)^{\alpha_1}(\Delta t)^{\alpha_2})\ \text{and } \alpha_1+\alpha_2=5; \tag{5.4.5}$$

and

$$T_L=T_{5.3.28}$$
$$=\{p_1(1-\hat{\alpha})-q_1(1+\hat{\alpha})\}(\Delta r)\left(\frac{\partial U}{\partial r}\right)_{1,j+\frac{1}{2}}+(1+p_1+\hat{\alpha})(\Delta t)\left(\frac{\partial U}{\partial t}\right)_{1,j+\frac{1}{2}}$$

$$
\begin{aligned}
&-\tfrac{1}{2}\{p_1(1-\hat{\alpha})+(1+\hat{\alpha})q_1\}(\Delta r)^2\left(\frac{\partial^2 U}{\partial r^2}\right)_{1,j+\frac{1}{2}} - \tfrac{1}{2}\{p_1(1-\hat{\alpha})-(1+\hat{\alpha})q_1\}(\Delta r)(\Delta t)\\
&\left(\frac{\partial^2 U}{\partial r\partial t}\right)_{1,j+\frac{1}{2}} + \frac{1}{6}\{p_1(1-\hat{\alpha})-(1+\hat{\alpha})q_1\}(\Delta r)^3\left(\frac{\partial^3 U}{\partial r^3}\right)_{1,j+\frac{1}{2}} + \frac{1}{4}\{p_1(1-\hat{\alpha})+(1+\hat{\alpha})q_1\}\\
&(\Delta r)^2(\Delta t)\left(\frac{\partial^3 U}{\partial r^2\partial t}\right)_{1,j+\frac{1}{2}} + \frac{1}{8}\{p_1(1-\hat{\alpha})-(1+\hat{\alpha})q_1\}(\Delta r)(\Delta t)^2\left(\frac{\partial^3 U}{\partial r\partial t^2}\right)_{1,j+\frac{1}{2}}\\
&+\frac{1}{24}(1+p_1+\hat{\alpha})(\Delta t)^3\left(\frac{\partial^3 U}{\partial t^3}\right)_{1,j+\frac{1}{2}} - \frac{1}{24}\{p_1(1-\hat{\alpha})+(1+\hat{\alpha})q_1\}(\Delta r)^4)\left(\frac{\partial^4 U}{\partial r^4}\right)_{1,j+\frac{1}{2}}\\
&-\frac{1}{12}\{p_1(1-\hat{\alpha})-(1+\hat{\alpha})q_1\}(\Delta r)^3(\Delta t)\left(\frac{\partial^4 U}{\partial r^3\partial t}\right)_{1,j+\frac{1}{2}} - \frac{1}{16}\{p_1(1-\hat{\alpha})+(1+\hat{\alpha})q_1\}\\
&(\Delta r)^2(\Delta t)^2\left(\frac{\partial^4 U}{\partial r^2\partial t^2}\right)_{1,j+\frac{1}{2}} - \frac{1}{48}\{p_1(1-\hat{\alpha})-(1+\hat{\alpha})q_1\}(\Delta r)(\Delta t)^3\left(\frac{\partial^4 U}{\partial r\partial t^3}\right)_{1,j+\frac{1}{2}}\\
&+O((\Delta r)^{\alpha_1}(\Delta t)^{\alpha_2}) \text{ and } \alpha_1+\alpha_2=5\ . \qquad (5.4.6)
\end{aligned}
$$

The truncation errors of the scheme at the remaining points *grouped* two at a time are given by $T_{5.3.6}$ (equation (5.4.2) and $T_{5.3.7}$ (equation 5.4.3) respectively for $i=3,5,\ldots,(m-3)$ and $(m-1)$.

(iii) Truncation Errors for the (S)AGE and (D)AGE Schemes

The truncation errors of these alternating group explicit methods are given by the truncation errors of the GER and the GEL schemes when they are applied in their correct sequence.

5.5 STABILITY ANALYSIS FOR THE GE METHODS

(i) Stability of the GER Scheme

From equation (5.3.20), the explicit form of the GER scheme is

$$\underline{u}_{j+1} = \Gamma_{GER}\underline{u}_j + \hat{\underline{b}}_1 , \qquad (5.5.1)$$

where $\Gamma_{GER} = (I+G_1)^{-1}(I+G_2)$, the amplification matrix which is given by,

$$
\Gamma_{GER} = \left[\begin{array}{cccccccccc}
(1-2\hat{\alpha}) & 2\hat{\alpha} & & & & & & & & \\
\frac{(1+p_2)p_1}{(1+p_2+q_1)} & \frac{(1+p_2)(1-p_1)}{(1+p_2+q_1)} & \frac{q_1(1-q_2)}{(1+p_2+q_1)} & \frac{q_2q_1}{(1+p_2+q_1)} & & & & & & \\
\frac{p_2p_1}{(1+p_2+q_1)} & \frac{p_2(1-p_1)}{(1+p_2+q_1)} & \frac{(1+q_1)(1-q_2)}{(1+p_2+q_1)} & \frac{(1+q_1)q_2}{(1+p_2+q_1)} & & & & & 0 & \\
0 & 0 & \frac{(1+p_4)p_3}{(1+p_4+q_3)} & \frac{(1+p_4)(1-p_3)}{(1+p_4+q_3)} & \frac{q_3(1-q_4)}{(1+p_4+q_3)} & \frac{q_4q_3}{(1+p_4+q_3)} & & & & \\
0 & 0 & \frac{p_4p_3}{(1+p_4+q_3)} & \frac{p_4(1-p_3)}{(1+p_4+q_3)} & \frac{(1+q_3)(1-q_4)}{(1+p_4+q_3)} & \frac{(1+q_3)q_4}{(1+p_4+q_3)} & & & & \\
& & & \ddots & \ddots & \ddots & \ddots & & & \\
& & & & & & \frac{(1+p_{m-2})p_{m-3}}{(1+p_{m-2}+q_{m-3})} & \frac{(1+p_{m-2})(1-p_{m-3})}{(1+p_{m-2}+q_{m-3})} & \frac{q_{m-3}(1-q_{m-2})}{(1+p_{m-2}+q_{m-3})} & \frac{q_{m-2}q_{m-3}}{(1+p_{m-2}+q_{m-3})} \\
& & & 0 & & & \frac{p_{m-2}p_{m-3}}{(1+p_{m-2}+q_{m-3})} & \frac{p_{m-2}(1-p_{m-3})}{(1+p_{m-2}+q_{m-3})} & \frac{(1+q_{m-3})(1-q_{m-2})}{(1+p_{m-2}+q_{m-3})} & \frac{(1+q_{m-3})q_{m-2}}{(1+p_{m-2}+q_{m-3})} \\
& & & & & & 0 & 0 & \frac{p_{m-1}}{(1+q_{m-1})} & \frac{(1-p_{m-1})}{(1+q_{m-1})}
\end{array}\right]_{(m\times m}
$$

m even, m>4 (5.5.2)

The stability of the GER scheme will now be investigated for two different cases of α.

(a) For the case $\alpha=2$ (spherical symmetry)

We have $\hat{\alpha}=3\lambda$, $p_1=0$, $p_2=\frac{\lambda}{2}$, $q_1=2\lambda$ and $q_2=\frac{3\lambda}{2}$. The characteristic equation of the matrix Γ_{GER} is $|D|=\det(\Gamma_{GER}-\mu I)=0$ where μ are the eigenvalues of Γ_{GER}. Now, if we expand the determinant by the first column, we get $\mu=(1-6\lambda)$ as an eigenvalue of Γ_{GER}. For stability, we require $|\mu|\leq 1$ which leads to the condition

$$\lambda\leq\frac{1}{3}\ . \tag{5.5.3}$$

For the second and subsequent even rows of Γ_{GER} we have as the sum of the moduli of the elements of the row,

$$S_{i+1} = \frac{|(1+p_{i+1})||p_i|}{|1+p_{i+1}+q_i|} + \frac{|(1+p_{i+1})||(1-p_i)|}{|1+p_{i+1}+q_i|} + \frac{|q_i||1-q_{i+1}|}{|1+p_{i+1}+q_i|} + \frac{|q_{i+1}||q_i|}{|1+p_{i+1}+q_i|}$$

$$i=1,3,\ldots,(m-3),\ m>4$$

$$= \frac{(1+p_{i+1})p_i+(1+p_{i+1})|1-p_i|+q_i|1-q_{i+1}|+q_{i+1}q_i}{(1+p_{i+1}+q_i)}$$

since p_i and q_i are non-negative.

For i=1 we obtain,

$$S_2 = \frac{1+p_2+q_1|1-q_2|+q_2q_1}{(1+p_2+q_1)}$$

If $q_2\leq 1$ then $\frac{3}{2}\lambda\leq 1$

$$\text{or } \lambda\leq\frac{2}{3}\ . \tag{5.5.4}$$

For these values of λ,

$$S_1 = \frac{1+p_2+q_1-q_1q_2+q_2q_1}{(1+p_2+q_1)}$$

$$= 1\ .$$

For i=3,...,(m-3), we find that,

if $p_i \leq 1$ then $\lambda \leq \frac{i}{(i-1)}$;

and if $q_{i+1} \leq 1$ then $\lambda \leq \frac{i+1}{(i+2)}$.

Now for $\lambda \leq \min.(\frac{i}{(i-1)}, \frac{i+1}{i+2})$,

$$= \frac{i+1}{(i+2)} , \qquad (5.5.5)$$

$$s_{i+1} = \frac{(1+p_{i+1})p_i+(1+p_{i+1})(1-p_i)+q_i(1-q_{i+1})+q_{i+1}q_i}{(1+p_{i+1}+q_i)}$$

$$= \frac{p_i+p_ip_{i+1}+1-p_i+p_{i+1}-p_ip_{i+1}+q_i-q_{i+1}q_i+q_{i+1}q_i}{(1+p_{i+1}+q_i)}$$

$$= (1+p_{i+1}+q_i)/(1+p_{i+1}+q_i) ,$$

$$= 1.$$

For the third row, the sum of the moduli of the elements of the row is

$$s_3 = \frac{p_2+(1+q_1)|1-q_2|+(1+q_1)q_2}{(1+p_2+q_1)} .$$

Again, if $\lambda \leq \frac{2}{3}$ as in (5.5.4) we have,

$$s_3 = (p_2+(1+q_1)(1-q_2)+(1+q_1)q_2)/(1+p_2+q_1) ,$$

$$= (p_2+1+q_1-q_2-q_1q_2+q_2+q_1q_2)/(1+p_2+q_1) ,$$

$$= (1+p_2+q_1)/(1+p_2+q_1) ,$$

$$= 1 .$$

For the subsequent odd rows of Γ_{GER} we get

$$s_{i+2} = \frac{p_{i+1}p_i+p_{i+1}|1-p_i|+(1+q_i)|1-q_{i+1}|+(1+q_i)q_{i+1}}{(1+p_{i+1}+q_i)} , \quad i=3,\ldots,(m-3)$$

with $m>4$.

Again using (5.5.5), if $\lambda \leq \frac{(i+1)}{(i+2)}$ then

$$s_{i+2} = \frac{p_{i+1}p_i+p_{i+1}(1-p_i)+(1+q_i)(1-q_{i+1})+(1+q_i)q_{i+1}}{(1+p_{i+1}+q_i)} ,$$

$$= \frac{p_{i+1}p_i+p_{i+1}-p_{i+1}p_i+1-q_{i+1}+q_i-q_iq_{i+1}+q_{i+1}+q_iq_{i+1}}{(1+p_{i+1}+q_i)} ,$$

$$= (1+p_{i+1}+q_i)/(1+p_{i+1}+q_i) \; ,$$
$$= 1.$$

An application of Brauer's theorem to the last row leads to

$$\left|\mu - \frac{(1-p_{m-1})}{(1+q_{m-1})}\right| \leq \frac{p_{m-1}}{(1+q_{m-1})} \quad ,$$

i.e.,

$$\frac{-p_{m-1}}{(1+q_{m-1})} \leq \mu - \frac{(1-p_{m-1})}{(1+q_{m-1})} \leq \frac{p_{m-1}}{(1+q_{m-1})} \quad ,$$

or

$$\frac{1-2p_{m-1}}{(1+q_{m-1})} \leq \mu \leq \frac{1}{(1+q_{m-1})} \; .$$

Let

$$\mu_1 = \frac{1-2p_{m-1}}{(1+q_{m-1})} \quad \text{and} \quad \mu_2 = \frac{1}{(1+q_{m-1})} \; .$$

The requirement $|\mu_1| \leq 1$ implies $-1 \leq \frac{1-2p_{m-1}}{1+q_{m-1}} \leq 1,$

$$\text{or } -1-q_{m-1} \leq 1-2p_{m-1} \leq 1+q_{m-1} \; .$$

The left-hand side inequality gives

$$2\frac{(m-2)}{(m-1)}\lambda - \frac{m}{(m-1)}\lambda \leq 2$$

or

$$\lambda \leq \frac{(2m-2)}{(m-4)} \; . \tag{5.5.6}$$

The second requirement

$|\mu_2| \leq 1$ implies $\frac{1}{(1+q_{m-1})} \leq 1$ which is true for all $\lambda > 0$.

Therefore, for overall stability we have using (5.5.3), (5.5.4), (5.5.5) and (5.5.6),

$$\lambda \leq \min.\{\frac{1}{3}, \frac{2}{3}, \min.(\frac{i+1}{i+2} \; , \; i=3,\ldots,(m-3)) \, , \; \frac{2m-2}{m-4}\} \; ; \; m>4$$
$$= \frac{1}{3} \; .$$

Hence the GER scheme for the spherically symmetric parabolic problem is conditionally stable for $\lambda \leq \frac{1}{3}$.

(b) For the case $\alpha=1$ (cylindrical symmetry)

The characteristic equation is $|D|=0$, where,

$$|D| = \begin{vmatrix}
(1-4\lambda)-\mu & 4\lambda & & & & & \\
\frac{(4+3\lambda)\lambda}{2(4+9\lambda)} & \frac{(4+3\lambda)(2-\lambda)}{2(4+9\lambda)}-\mu & \frac{3(4-5\lambda)\lambda}{2(4+9\lambda)} & \frac{15\lambda^2}{2(4+9\lambda)} & & & \\
\frac{3\lambda^2}{2(4+9\lambda)} & \frac{3(2-\lambda)\lambda}{2(4+9\lambda)} & \frac{(2+3\lambda)(4-5\lambda)}{2(4+9\lambda)}-\mu & \frac{5(2+3\lambda)\lambda}{2(4+9\lambda)} & & & \\
0 & 0 & \frac{(1+p_4)p_3}{(1+p_4+q_3)} & \frac{(1+p_4)(1-p_3)}{(1+p_4+q_3)}-\mu & \frac{q_3(1-q_4)}{(1+p_4+q_3)} & \frac{q_4q_3}{(1+p_4+q_3)} & \\
0 & 0 & \frac{p_4p_3}{(1+p_4+q_3)} & \frac{p_4(1-p_3)}{(1+p_4+q_3)} & \frac{(1+q_3)(1-q_4)}{(1+p_4+q_3)}-\mu & \frac{(1+q_3)q_4}{(1+p_4+q_3)} & \bigcirc \\
 & & & \ddots & \ddots & \ddots & \\
 & & \bigcirc & \frac{(1+p_{m-2})p_{m-3}}{(1+p_{m-2}+q_{m-3})} & \frac{(1+p_{m-2})(1-p_{m-3})}{(1+p_{m-2}+q_{m-3})}-\mu & \frac{q_{m-3}(1-q_{m-2})}{(1+p_{m-2}+q_{m-3})} & \frac{q_{m-2}q_{m-3}}{(1+p_{m-2}+q_{m-3})} \\
 & & & \frac{p_{m-2}p_{m-3}}{(1+p_{m-2}+q_{m-3})} & \frac{p_{m-2}(1-p_{m-3})}{(1+p_{m-2}+q_{m-3})} & \frac{(1+q_{m-3})(1-q_{m-2})}{(1+p_{m-2}+q_{m-3})}-\mu & \frac{(1+q_{m-3})q_{m-2}}{(1+p_{m-2}+q_{m-3})} \\
 & & & & & \frac{p_{m-1}}{(1+q_{m-1})} & \frac{(1-p_{m-1})}{(1+q_{m-1})}-\mu
\end{vmatrix}$$

(5.5.7)

By carrying out the following set of row operations R_i on the determinant $|D|$ we obtain:

(i) $R_1 \rightarrow \frac{(4+3\lambda)\lambda}{2(4+9\lambda)} \frac{1}{(1-4\lambda-\mu)} R_1$,

$$\text{i.e.} \quad R_1 = (\frac{(4+3\lambda)\lambda}{2(4+9\lambda)}, \frac{2(4+3\lambda)\lambda^2}{(4+9\lambda)(1-4\lambda-\mu)}, 0,\ldots,0),$$

$$\text{and} \quad |D| \rightarrow \frac{(4+3\lambda)\lambda}{2(4+9\lambda)(1-4\lambda-\mu)}|D|;$$

(ii) $R_2 \rightarrow R_2-R_1$,

$$\text{i.e.,} \quad R_2 = (0,\beta_1, \frac{3(4-5\lambda)\lambda}{2(4+9\lambda)}, \frac{15\lambda^2}{2(4+9\lambda)}, 0,\ldots,0);$$

(iii) $R_1 \rightarrow \frac{3\lambda^2}{(4+3\lambda)\lambda}R_1$,

$$\text{i.e.} \quad R_1 = (3\lambda^2/2(4+9\lambda), 6\lambda^3/(4+9\lambda)(1-4\lambda-\mu), 0,\ldots,0),$$

$$\text{and} \quad |D| \rightarrow \frac{3\lambda^2}{2(4+9\lambda)(1-4\lambda-\mu)}|D|; \text{ and}$$

(iv) $R_3 \rightarrow R_3-R_1$

$$\text{i.e.} \quad R_3 = (0,\beta_2, \frac{(2+3\lambda)(4-5\lambda)}{2(4+9\lambda)} -\mu , \frac{5(2+3\lambda)\lambda}{2(4+9\lambda)}, 0,\ldots,0).$$

Thus,

$$|D_1| = \frac{3\lambda^2}{2(4+9\lambda)(1-4\lambda-\mu)}|D|$$

or

$$|D| = \frac{2(4+9\lambda)(1-4\lambda-\mu)}{3\lambda^2} |D_1|, \tag{5.5.8}$$

where,

$$|D_1| = \begin{vmatrix} \frac{3\lambda^2}{2(4+9\lambda)} & \frac{6\lambda^3}{(4+9\lambda)(1-4\lambda-\mu)} & & & \\ 0 & \beta_1 & \frac{3(4-5\lambda)\lambda}{2(4+9\lambda)} & \frac{15\lambda^2}{2(4+9\lambda)} & \bigcirc \\ & \beta_2 & \frac{(2+3\lambda)(4-5\lambda)}{2(4+9\lambda)} - \mu & \frac{5(2+3\lambda)\lambda}{2(4+9\lambda)} & 0 \\ & \bigcirc & \ddots & \ddots & \ddots \quad 0 \\ & & \frac{p_{m-1}}{(1+q_{m-1})} & \frac{(1-p_{m-1})}{(1+q_{m-1})} - \mu & \end{vmatrix} \qquad (5.5.9)$$

with

$$\beta_1 = \frac{(4+3\lambda)(2-9\lambda) - \{(16-12\lambda-75\lambda^2) - 2(4+9\lambda)\mu\}\mu}{2(4+9\lambda)(1-4\lambda-\mu)}$$

and

$$\beta_2 = \frac{3(2-\lambda)\lambda(1-4\lambda-\mu) - 12\lambda^3}{2(4+9\lambda)(1-4\lambda-\mu)} \quad .$$

The determinant $|D|$ now takes the form

$$|D| = \begin{vmatrix} (1-4\lambda-\mu) & 4\lambda & 0 & & \\ 0 & \beta_1 & \frac{3(4-5\lambda)\lambda}{2(4+9\lambda)} & \frac{15\lambda^2}{2(4+9\lambda)} & \bigcirc \\ & \beta_2 & \frac{(2+3\lambda)(4-5\lambda)}{2(4+9\lambda)} - \mu & \frac{5(2+3\lambda)\lambda}{2(4+9\lambda)} & 0 \\ & \bigcirc & \ddots & \ddots & \ddots \quad 0 \\ & & \frac{p_{m-1}}{(1+q_{m-1})} & \frac{(1-p_{m-1})}{(1+q_{m-1})} - \mu & \end{vmatrix} \qquad (5.5.10)$$

and on expanding the determinant $|D|=0$ about the first column we find that $(1-4\lambda-\mu)|D_2|=0$ where $|D_2|$ is now the minor obtained from $|D|$ by deleting the first row and column. This gives us the result,

$$1-4\lambda-\mu = 0 \text{ and } |D_2|=0 \ .$$

From the first equation, the stability requirement that $|\mu| \leqslant 1$ leads to

$$\lambda \leqslant \tfrac{1}{2} . \qquad (5.5.11)$$

By repeating the same procedure on $|D_2|$ we also have

$$\beta_1 = 0 ,$$

or $$(4+3\lambda)(2-9\lambda)-\{(16-12\lambda-75\lambda^2\}-2(4+9\lambda)\mu\}\mu = 0 ,$$

i.e., $$2(4+9\lambda)\mu^2-(16-12\lambda-75\lambda^2)\mu+(4+3\lambda)(2-9\lambda) = 0 ,$$

whose solutions are given by,

$$\mu = \frac{(16-12\lambda-75\lambda^2)\pm\sqrt{(16-12\lambda-75\lambda^2)^2-8(4+9\lambda)(4+3\lambda)(2-9\lambda)}}{4(4+9\lambda)} .$$

If the solutions are μ_1 and μ_2, we find by inspection that for $\lambda=0.454$,

$|\mu_1|=0.685006039$

and $|\mu_2|= 0.997073088$

giving $\max\{|\mu_1|,|\mu_2|\} = |\mu_2|< 1$;

and for $\lambda = 0.455$, $|\mu_1|= 0.693311298$

and $|\mu_2|= 1.001333226$

which imply $\max.\{|\mu_1|,|\mu_2|\} = |\mu_2| >1$.

Thus for stability, we require

$$\lambda \leqslant 0.454 . \qquad (5.5.12)$$

Now the sum of the moduli of the elements of the third row to row $(m-1)$ is each equal to 1 if $p_i \leqslant 1$ and $q_{i+1} \leqslant 1$ for $i=1,3,\dots,(m-3)$. Since $p_i=(1-\frac{1}{2i})\lambda$ and $q_{i+1} = (1+\frac{1}{2(i+1)})\lambda$, the conditions on p_i and q_{i+1} require,

$$\lambda \leqslant \min.\left(\frac{2i}{(2i-1)} , \frac{(2i+2)}{(2i+3)}\right)$$

$$= \frac{(2i+2)}{(2i+3)} . \qquad (5.5.13)$$

By applying Brauer's theorem to the last row, we have

$$2p_{m-1}-q_{m-1} \leqslant 2,$$

i.e., $$\frac{2(2m-3)}{2(m-1)}\lambda - \frac{(2m-1)\lambda}{2(m-1)} \leqslant 2$$

giving $\lambda \leqslant (4m-4)/(2m-5)$. (5.5.14)

Therefore, for the stability of the entire set of equations, we must have using (5.5.11), (5.5.12), (5.5.13) and (5.5.14),

$$\lambda \leqslant \min.\{\tfrac{1}{2}, 0.45, \min.(\frac{2i+2}{2i+3},\ i=1,3,\ldots,(m-3)),\ \frac{4m-4}{2m-5}\},\ m>4$$
$$= 0.454.$$

(ii) Stability of the GEL Scheme

From equation (5.3.33), the explicit form of the GEL scheme is

$$\underline{u}_{j+1} = \Gamma_{GEL}\underline{u}_j + \underline{\hat{b}}_2 \ , \tag{5.5.15}$$

where $\Gamma_{GEL} = (I+\hat{G}_2)^{-1}(I+\hat{G}_1)$, the amplification matrix which is given by,

$$\Gamma_{GEL} = \begin{bmatrix} \frac{(1+p_1)(1-\hat{\alpha})}{(1+p_1+\hat{\alpha})} & \frac{(2+p_1-q_1)\hat{\alpha}}{(1+p_1+\hat{\alpha})} & \frac{\hat{\alpha}q_1}{(1+p_1+\hat{\alpha})} & 0 & & & & & & \\ \frac{p_1(1-\hat{\alpha})}{(1+p_1+\hat{\alpha})} & \frac{(1+\hat{\alpha})(1-q_1)+p_1\hat{\alpha}}{(1+p_1+\hat{\alpha})} & \frac{(1+\hat{\alpha})q_1}{(1+p_1+\hat{\alpha})} & 0 & & & & & & \\ 0 & \frac{(1+p_3)p_2}{(1+p_3+q_2)} & \frac{(1+p_3)(1-p_2)}{(1+p_3+q_2)} & \frac{q_2(1-q_3)}{(1+p_3+q_3)} & \frac{q_3q_2}{(1+p_3+q_3)} & & & & & \\ 0 & \frac{p_3p_2}{(1+p_3+q_2)} & \frac{p_3(1-p_2)}{(1+p_3+q_2)} & \frac{(1+q_2)(1-q_3)}{(1+p_3+q_3)} & \frac{(1+q_2)q_3}{(1+p_3+q_2)} & & & & \bigcirc & \\ 0 & 0 & 0 & \frac{(1+p_5)p_4}{(1+p_5+q_4)} & \frac{(1+p_5)(1-p_4)}{(1+p_5+q_4)} & \frac{q_4(1-q_5)}{(1+p_5+q_4)} & \frac{q_5q_4}{(1+p_5+q_4)} & & & \\ 0 & 0 & 0 & \frac{p_5p_4}{(1+p_5+q_4)} & \frac{p_5(1-p_4)}{(1+p_5+q_4)} & \frac{(1+q_4)(1-q_5)}{(1+p_5+q_4)} & \frac{(1+q_4)q_5}{(1+p_5+q_4)} & & & \\ & & & & & \ddots & \ddots & \ddots & & \\ & \bigcirc & & & & & & \frac{(1+p_{m-1})p_{m-2}}{(1+p_{m-1}+q_{m-2})} & \frac{(1+p_{m-1})(1-p_{m-2})}{(1+p_{m-1}+q_{m-2})} & \frac{q_{m-2}(1-q_{m-1})}{(1+p_{m-1}+q_{m-2})} \\ & & & & & & & \frac{p_{m-1}p_{m-2}}{(1+p_{m-1}+q_{m-2})} & \frac{p_{m-1}(1-p_{m-2})}{(1+p_{m-1}+q_{m-2})} & \frac{(1+q_{m-2})(1-q_{m-1})}{(1+p_{m-1}+q_{m-2})} \end{bmatrix}_{(m\times}$$

m even, m>4, (5.5.16)

(a) For the case α=2 (spherical symmetry)

The characteristic equation is $|D|=0$ where,

$$|D| = \begin{vmatrix}
\frac{(1-3\lambda)}{(1+3\lambda)}-\mu & \frac{(2-2\lambda)3\lambda}{(1+3\lambda)} & \frac{6\lambda^2}{(1+3\lambda)} & & & & & & \\
0 & (1-2\lambda)-\mu & 2\lambda & & & & & & \\
0 & \frac{(3+2\lambda)\lambda}{(6+13\lambda)} & \frac{(3+2\lambda)(2-\lambda)}{(6+13\lambda)}-\mu & \frac{3(3-4\lambda)\lambda}{(6+13\lambda)} & \frac{12\lambda^2}{(6+13\lambda)} & & & & \\
0 & \frac{2\lambda^2}{(6+13\lambda)} & \frac{2(2-\lambda)\lambda}{(6+13\lambda)} & \frac{(2+3\lambda)(3-4\lambda)}{(6+13\lambda)}-\mu & \frac{4(2+3\lambda)\lambda}{(6+13\lambda)} & & & & \\
0 & 0 & 0 & \frac{(5+4\lambda)3\lambda}{(20+41\lambda)} & \frac{(5+4\lambda)(4-3\lambda)}{(20+41\lambda)}-\mu & \frac{5\lambda(5-6\lambda)}{(20+41\lambda)} & \frac{30\lambda^2}{(20+41\lambda)} & & \\
0 & 0 & 0 & \frac{12\lambda^2}{(20+41\lambda)} & \frac{4\lambda(4-3\lambda)}{(20+41\lambda)} & \frac{(4+5\lambda)(5-6\lambda)}{(20+41\lambda)}-\mu & \frac{(4+5\lambda)6\lambda}{(20+41\lambda)} & & \bigcirc \\
 & & & & \ddots & \ddots & \ddots & & \\
 & & \bigcirc & & & & \frac{(1+p_{m-1})p_{m-2}}{(1+p_{m-1}+q_{m-2})} & \frac{(1+p_{m-1})(1-p_{m-2})}{(1+p_{m-1}+q_{m-2})}-\mu & \frac{q_{m-2}(1-q_{m-1})}{(1+p_{m-1}+q_{m-2})} \\
 & & & & & & \frac{p_{m-1}p_{m-2}}{(1+p_{m-1}+q_{m-2})} & \frac{p_{m-1}(1-p_{m-2})}{(1+p_{m-1}+q_{m-2})} & \frac{(1+q_{m-2})(1-q_{m-1})}{(1+p_{m-1}+q_{m-2})}-\mu
\end{vmatrix}$$

(5.5.17)

If we expand the determinant about the first column, we obtain $\{\frac{(1-3\lambda)}{(1+3\lambda)}-\mu\}|D_1| = 0$ where $|D_1|$ is the minor obtained from $|D|$ by deleting the first row and the first column. This gives us,

$$\frac{(1-3\lambda)}{(1+3\lambda)} -\mu = 0 \text{ and } |D_1| = 0.$$

From the first equation, $\mu = \frac{(1-3\lambda)}{(1+3\lambda)}$ and $|\mu|<1$ for every $\lambda>0$.

Now we carry out the following row operations on $|D_1|$:

(i) $R_1 \to \frac{(3+2\lambda)\lambda}{(6+13\lambda)(1-2\lambda-\mu)} R_1$,

i.e. $R_1 \to (\frac{(3+2\lambda)\lambda}{(6+13\lambda)}, \frac{2(3+2\lambda)\lambda^2}{(6+13\lambda)(1-2\lambda-\mu)}, 0,\ldots,0)$,

and $|D_1| \to \frac{(3+2\lambda)\lambda}{(6+13\lambda)(1-2\lambda-\mu)} |D_1|$;

(ii) $R_2 \to R_2-R_1$,

i.e. $R_2 \to (0,\beta_1, \frac{3(3-4\lambda)\lambda}{(6+13\lambda)}, \frac{12\lambda^2}{(6+13\lambda)}, 0,0,\ldots,0)$;

(iii) $R_1 \to \frac{2\lambda}{(3+2\lambda)} R_1$,

i.e. $R_1 = (\frac{2\lambda^2}{(6+13\lambda)}, \frac{4\lambda^3}{(6+13\lambda)(1-2\lambda-\mu)},0,\ldots,0)$,

and $|D_1| \to \frac{2\lambda^2}{(6+13\lambda)(1-2\lambda-\mu)}|D_1|$; and

(iv) $R_3 \to R_3-R_1$

i.e. $R_3 = (0,\beta_2,\frac{(2+3\lambda)(3-4\lambda)}{(6+13\lambda)}-\mu, \frac{4(2+3\lambda)\lambda}{(6+13\lambda)},0,\ldots,0)$.

Thus, we have

$$|D_2| = \frac{2\lambda^2}{(6+13\lambda)(1-2\lambda-\mu)}|D_1|$$

or

$$|D_1| = \frac{(6+13\lambda)(1-2\lambda-\mu)}{2\lambda^2}|D_2|$$

which takes the form,

$$|D_1| = \left| \begin{array}{cccccccc}
(1-2\lambda-\mu) & 2\lambda & & & & & & \\
0 & \beta_1 & 3(3-4\lambda)\lambda/(6+13\lambda) & \frac{12\lambda^2}{(6+13\lambda)} & & & & \\
0 & \beta_2 & \frac{(2+3\lambda)(3-4\lambda)}{(6+13\lambda)}-\mu & \frac{4(2+3\lambda)\lambda}{(6+13\lambda)} & & & & \\
 & & \frac{(5+4\lambda)3\lambda}{(20+41\lambda)} & \frac{(5+4\lambda)(4-3\lambda)}{(20+41\lambda)}-\mu & \frac{5\lambda(5-6\lambda)}{(20+41\lambda)} & \frac{30\lambda^2}{(20+41\lambda)} & & \\
 & & \frac{12\lambda^2}{(20+41\lambda)} & \frac{4\lambda(4-3\lambda)}{(20+41\lambda)} & \frac{(4+5\lambda)(5-6\lambda)}{(20+41\lambda)}-\mu & \frac{(4+5\lambda)6\lambda}{(20+41\lambda)} & & \bigcirc \\
 & & & \ddots & \ddots & \ddots & & \\
 & & \bigcirc & & & & & \\
 & & & \frac{(1+p_{m-1})p_{m-2}}{(1+p_{m-1}+q_{m-2})} & \frac{(1+p_{m-1})(1-p_{m-2})}{(1+p_{m-1}+q_{m-2})}-\mu & \frac{q_{m-2}(1-q_{m-1})}{(1+p_{m-1}+q_{m-2})} & & \\
 & & & \frac{p_{m-1}p_{m-2}}{(1+p_{m-1}+q_{m-2})} & \frac{p_{m-1}(1-p_{m-2})}{(1+p_{m-1}+q_{m-2})} & \frac{(1+q_{m-2})(1-q_{m-1})}{(1+p_{m-1}+q_{m-2})} & &
\end{array} \right| \qquad (5.5.18)$$

where

$$\beta_1 = \{(6+13\lambda)\mu^2-[(3+2\lambda)(2-\lambda)+(6+13\lambda)(1-2\lambda)]\mu+(3+2\lambda)(2-\lambda)(1-2\lambda)-2(3+2\lambda)\lambda^2\}/(6+13\lambda)(1-2\lambda-\mu)$$

and

$$\beta_2 = \{2\lambda(2-5\lambda)-2\lambda(2-\lambda)\mu\}/(6+13\lambda)(1-2\lambda-\mu) .$$

By expanding the determinant about the first column, we get

$$(1-2\lambda-\mu)|D_2| = 0$$

where $|D_2|$ is the minor obtained from $|D_1|$ by the deletion of the first row and the first column. This gives us,

$$(1-2\lambda-\mu) = 0$$

and $|D_2|=0$. Hence, $\mu=1-2\lambda$ and the stability requirement $|\mu|\leq 1$ leads to the condition,

$$\lambda \leq 1 . \qquad (5.5.19)$$

The same procedure is repeated on the minors $|D_2|$ and $|D_3|$ from which we obtain the conditions of stability

$$\lambda \leq 0.849 \quad \text{and} \quad \lambda \leq 0.848 \tag{5.5.20}$$

respectively. Finally, by operating on $|D_4|$ we obtain

$$|D_4| = \left[\begin{array}{cccccc}
\beta_5 & \frac{5\lambda(5-6\lambda)}{(20+41\lambda)} & \frac{30\lambda^2}{(20+41\lambda)} & & & \\
0 & \beta_7 & \beta_8 & & & \\
0 & \frac{5\lambda(7+6\lambda)}{(42+85\lambda)} & \frac{(7+6\lambda)(6-5\lambda)}{(42+85\lambda)} & \frac{7\lambda(7-8\lambda)}{(42+85\lambda)} & \frac{56\lambda^2}{(42+85\lambda)} & \\
 & & \ddots & \ddots & \ddots & \bigcirc \\
 & \bigcirc & & \frac{(1+p_{m-1})p_{m-2}}{(1+p_{m-1}+q_{m-2})} & \frac{(1+p_{m-1})(1-p_{m-2})}{(1+p_{m-1}+q_{m-2})}-\mu & \frac{q_{m-2}(1-q_{m-1})}{(1+p_{m-1}+q_{m-2})} \\
 & & & \frac{p_{m-1}p_{m-2}}{(1+p_{m-1}+q_{m-2})} & \frac{p_{m-1}(1-p_{m-2})}{(1+p_{m-1}+q_{m-2})} & \frac{(1+q_{m-2})(1-q_{m-1})}{(1+p_{m-1}+q_{m-2})}-\mu
\end{array}\right] \tag{5.5.21}$$

where,

$$\beta_3 = \{\{(2+3\lambda)(3-4\lambda)-(6+13\lambda)\mu\}\beta_1-3\beta_2(3-4\lambda)\lambda\}/(6+13\lambda)\beta_1,$$

$$\beta_4 = \{4\beta_1(2+3\lambda)\lambda-12\beta_2\lambda^2\}/(6+13\lambda)\beta_1 \ ,$$

$$\beta_5 = \{\{(5+4\lambda)(4-3\lambda)-(20+41\lambda)\mu\}\beta_3-(5+4\lambda)3\lambda\beta_4\}/(20+41\lambda)\beta_3 \ ,$$

$$\beta_6 = \{4\lambda(4-3\lambda)\beta_3-12\lambda^2\beta_4\}/(20+41\lambda)\beta_3 \ ,$$

$$\beta_7 = \{\{(4+5\lambda)(5-6\lambda)-(20+41\lambda)\mu\}\beta_5-5\lambda(5-6\lambda)\beta_6\}/(20+41\lambda)\beta_5$$

and $\beta_8 = \{(4+5\lambda)6\lambda\beta_5-30\lambda^2\beta_6\}/(20+41\lambda)\beta_5$.

If we expand the determinant about the first column, we get,

$$|D_4| = \beta_5|D_5|=0 \text{ which implies } \beta_5=0 \text{ and } |D_5|=0 \ .$$

From $\beta_5=0$, we obtain the following quartic equation:

$$a\mu^4+b\mu^3+c\mu^2+d\mu+e = 0 \ ,$$

where $\quad a = (20+41\lambda)(6+13\lambda)^2$

$$b = -(6+13\lambda)\{[(5+4\lambda)(4-3\lambda)(6+13\lambda)+(20+41\lambda)(2+3\lambda)(3-4\lambda)]$$
$$+(20+41\lambda)[(3+2\lambda)(2-\lambda)+(6+13\lambda)(1-2\lambda)]\}$$

$$c = (5+4\lambda)(6+13\lambda)(12-25\lambda)(2+3\lambda)$$
$$+[(3+2\lambda)(2-\lambda)+(6+13\lambda)(1-2\lambda)][(5+4\lambda)(4-3\lambda)(6+13\lambda)+(20+41\lambda)$$
$$(2+3\lambda)(3-4\lambda)]+(20+41\lambda)(6+13\lambda)(3+2\lambda)(2-5\lambda)-6\lambda^2(3-4\lambda)(2-\lambda)(20+41\lambda)$$

$$d = -(5+4\lambda)[(3+2\lambda)(2-\lambda)+(6+13\lambda)(1-2\lambda)](2+3\lambda)(12-25\lambda)$$
$$+6\lambda^2[(3-4\lambda)(2-\lambda)(5+4\lambda)(4-3\lambda)+(3-4\lambda)(2-5\lambda)(20+41\lambda)-12\lambda^2(5+4\lambda)(2-\lambda)]$$
$$-(3+2\lambda)(2-5\lambda)[(5+4\lambda)(4-3\lambda)(6+13\lambda)+(20+41\lambda)(2+3\lambda)(3-4\lambda)]$$

$$e = (5+4\lambda)(2+3\lambda)(12-25\lambda)(3+2\lambda)(2-5\lambda)$$
$$-6\lambda^2(5+4\lambda)(2-5\lambda)(12-25\lambda).$$

On solving the quartic equation algebraically by the method described in Chapter 1, we find that,

for $\lambda=0.8393$,

$\mu_1 = 0.7971763129$, $\mu_2 = 0.4630300548$, $\mu_3 = -0.9999914241$, and

$\mu_4 = -0.4854824809$,

with $\max\{|\mu_1|,|\mu_2|,|\mu_3|,|\mu_4|\}=|\mu_4|<1$;

and for $\lambda=0.8394$

$\mu_1 = 0.79716185$, $\mu_2 = 0.46300037$, $\mu_3 = -1.00022034$, and

$\mu_4 = -0.48560624$,

with $\max\{|\mu_1|,|\mu_2|,|\mu_3|,|\mu_4|\}= |\mu_4|>1$.

Hence for stability we require,

$$\lambda \leq 0.8393. \qquad (5.5.22)$$

Now for rows 6(1)(m-2), the sum of the moduli of the elements in each of these rows is equal to 1 if $p_i\leq 1$ and $q_{i+1}\leq 1$ for $i=4,6,\ldots,(m-4)$. These conditions give us,

$$\lambda \leq \frac{(i+1)}{(i+2)} \quad . \qquad (5.5.23)$$

If $p_{m-2}\leq 1$ (i.e. $\lambda\leq\frac{m-2}{m-3}$) and if $q_{m-1}\leq 1$ (i.e. $\lambda\leq\frac{m-1}{m}$) then for $\lambda\leq\min(\frac{m-2}{m-3}, \frac{m-1}{m}) = \frac{m-1}{m}$, the sum of the moduli of the elements in row (m-1) is

$$\frac{(1+p_{m-1})p_{m-2}+(1+p_{m-1})|(1-p_{m-2})|+q_{m-2}|1-q_{m-1}|}{(1+p_{m-1}+q_{m-2})}$$

$$= \{(1+p_{m-1}+q_{m-2})-q_{m-2}q_{m-1}\}/(1+p_{m-1}+q_{m-2}) = 1 - \frac{q_{m-2}q_{m-1}}{(1+p_{m-1}+q_{m-2})}$$
$$< 1 .$$

Similarly, for the sum in row m we have,

$$\{p_{m-1}p_{m-2}+p_{m-1}|1-p_{m-2}|+(1+q_{m-2})|1-q_{m-1}|\}/(1+p_{m-1}+q_{m-2})$$

$$= \{(1+p_{m-1}+q_{m-2})-q_{m-1}(1+q_{m-2})\}/(1+p_{m-1}+q_{m-2})$$

$$= 1-q_{m-1}(1+q_{m-2})/(1+p_{m-1}+q_{m-2})$$
$$< 1 \text{ for } \lambda\leq\frac{m-1}{m} . \tag{5.5.24}$$

Therefore we deduce from (5.5.19)-(5.5.24) that to achieve *overall stability*, we require,

$$\lambda \leq \min\{1,0.849,0.848,0.8393,\min(\frac{(i+1)}{(i+2)}, i=4,6,\dots,(m-4),\frac{m-1}{m}\};$$
$$= \frac{5}{6} . \qquad m>4.$$

(iii) Stability of the (S)AGE Scheme

From (5.3.34), the explicit expression for the (S)AGE scheme is given by

$$\underline{u}_{j+2} = \Gamma_{SAGE}\underline{u}_j + \underline{b}_2' , \; j=0,2,4,\dots$$

where
$$\Gamma_{SAGE} = \Gamma_{GEL}\Gamma_{GER} . \tag{5.5.25}$$

By multiplying the two matrices in (5.5.25), the general form of the amplification matrix of the (S)AGE scheme is

$$\Gamma_{SAGE} = \left[\begin{array}{cccccccccccccccc}
a_1 & a_2 & a_3 & a_4 & & & & & & & & & & & & \\
b_1 & b_2 & b_3 & b_4 & & & & & & & & & & & & \\
c_1^{(1)} & c_2^{(2)} & c_3^{(3)} & c_4^{(4)} & c_5^{(5)} & c_6^{(6)} & & & & & & & & & & \\
d_1^{(1)} & d_2^{(2)} & d_3^{(3)} & d_4^{(4)} & d_5^{(5)} & d_6^{(6)} & & & & & & & & & & \\
0 & 0 & c_3^{(1)} & c_4^{(2)} & c_5^{(3)} & c_6^{(4)} & c_7^{(5)} & c_8^{(6)} & & & & & & & & \\
0 & 0 & d_3^{(1)} & d_4^{(2)} & d_5^{(3)} & d_6^{(4)} & d_7^{(5)} & d_8^{(6)} & & & & & & & & \\
0 & 0 & 0 & 0 & c_5^{(1)} & c_6^{(2)} & c_7^{(3)} & c_8^{(4)} & c_9^{(5)} & c_{10}^{(6)} & & & & & & \\
0 & 0 & 0 & 0 & d_5^{(1)} & d_6^{(2)} & d_7^{(3)} & d_8^{(4)} & d_9^{(5)} & d_{10}^{(6)} & & & & & & \\
 & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & & & \bigcirc & & \\
 & & & \bigcirc & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & & & \\
 & & & & & & & & & & c_{m-5}^{(1)} & c_{m-4}^{(2)} & c_{m-3}^{(3)} & c_{m-2}^{(4)} & c_{m-1}^{(5)} & c_m^{(6)} \\
 & & & & & & & & & & d_{m-5}^{(1)} & d_{m-4}^{(2)} & d_{m-3}^{(3)} & d_{m-2}^{(4)} & d_{m-1}^{(5)} & d_m^{(6)} \\
 & & & & & & & & & & & & c_{m-3}^{(1)} & c_{m-2}^{(2)} & e_1 & e_2 \\
 & & & & & & & & & & & & d_{m-3}^{(1)} & d_{m-2}^{(2)} & f_1 & f_2
\end{array}\right]_{(m\times m)} \qquad (5.5.25a)$$

m even

m>4

where

$$a_1 = \frac{(1+p_1)(1-\hat{\alpha})}{(1+p_1+\hat{\alpha})}(1-2\hat{\alpha}) + \frac{\hat{\alpha}p_1\{(1+p_2)(2+p_1)-q_1\}}{(1+p_1+\hat{\alpha})(1+p_2+q_1)} \quad ;$$

$$a_2 = \frac{2\hat{\alpha}(1+p_1)(1-\hat{\alpha})}{(1+p_1+\hat{\alpha})} + \frac{\hat{\alpha}(1-p_1)\{(1+p_2)(2+p_1)-q_1\}}{(1+p_1+\hat{\alpha})(1+p_2+q_1)} \; ;$$

$$a_3 = \frac{\hat{\alpha}q_1(1-q_2)(3+p_1)}{(1+p_1+\hat{\alpha})(1+p_2+q_1)} \; ;$$

$$a_4 = \frac{\hat{\alpha}q_1q_2(3+p_1)}{(1+p_1+\hat{\alpha})(1+p_2+q_1)} \; ;$$

$$b_1 = \frac{(1-\hat{\alpha})(1-2\hat{\alpha})p_1}{(1+p_1+\hat{\alpha})} + \frac{p_1\{(1+\hat{\alpha})(1+p_2-q_1)+p_1(1+p_2)\hat{\alpha}\}}{(1+p_1+\hat{\alpha})(1+p_2+q_1)} \quad ;$$

$$b_2 = \frac{2\hat{\alpha}p_1(1-\hat{\alpha})}{(1+p_1+\hat{\alpha})} + \frac{(1-p_1)\{(1+\hat{\alpha})(1+p_2-q_1)+p_1(1+p_2)\hat{\alpha}\}}{(1+p_1+\hat{\alpha})(1+p_2+q_1)};$$

$$b_3 = \frac{q_1(1-q_2)(2(1+\alpha)+p_1\hat{\alpha})}{(1+p_2+q_1)(1+p_1+\hat{\alpha})} \quad ;$$

$$b_4 = \frac{q_1q_2(2(1+\hat{\alpha})+p_1\hat{\alpha})}{(1+p_1+\hat{\alpha})(1+p_2+q_1)} \; ;$$

$$e_1 = \frac{(1+p_{m-1})(1-q_{m-2})(1-p_{m-2}+q_{m-3})}{(1+p_{m-2}+q_{m-3})(1+p_{m-1}+q_{m-2})} + \frac{p_{m-1}q_{m-2}(1-q_{m-1})}{(1+p_{m-1}+q_{m-2})(1+q_{m-1})} \; ;$$

$$e_2 = \frac{(1+p_{m-1})q_{m-2}(1-p_{m-2}+q_{m-3})}{(1+p_{m-2}+q_{m-3})(1+p_{m-1}+q_{m-2})} + \frac{(1-p_{m-1})(1-q_{m-1})q_{m-2}}{(1+p_{m-1}+q_{m-2})(1+q_{m-1})} \; ;$$

$$f_1 = \frac{p_{m-1}(1-q_{m-2})(1-p_{m-2}+q_{m-3})}{(1+p_{m-2}+q_{m-3})(1+p_{m-1}+q_{m-2})} + \frac{p_{m-1}(1-q_{m-1})(1+q_{m-2})}{(1+p_{m-1}+q_{m-2})(1+q_{m-1})} \; ;$$

$$f_2 = \frac{p_{m-1}q_{m-2}(1-p_{m-2}+q_{m-3})}{(1+p_{m-2}+q_{m-3})(1+p_{m-1}+q_{m-2})} + \frac{(1-p_{m-1})(1+q_{m-2})(1-q_{m-1})}{(1+p_{m-1}+q_{m-2})(1+q_{m-1})} \; ;$$

$$c_i^{(1)} = \frac{2p_ip_{i+1}(1+p_{i+2})}{(1+p_{i+1}+q_i)(1+p_{i+2}+q_{i+1})} \quad , \; i=1,3,\ldots,m-5,m-3;$$

$$c_i^{(2)} = \frac{2(1-p_{i-1})p_i(1+p_{i+1})}{(1+p_i+q_{i-1})(1+p_{i+1}+q_i)} \quad , \ i=2,4,\ldots,m-4,m-2 \ ;$$

$$c_i^{(3)} = \frac{(1+p_i)(1-q_{i-1})(1-p_{i-1}+q_{i-2})}{(1+p_{i-1}+q_{i-2})(1+p_i+q_{i-1})} + \frac{p_iq_{i-1}(1+p_{i+1}-q_i)}{(1+p_i+q_{i-1})(1+p_{i+1}+q_i)}$$

$$i=3,5,\ldots,m-3 \ ;$$

$$c_i^{(4)} = \frac{(1+p_{i-1})q_{i-2}(1-p_{i-2}+q_{i-3})}{(1+p_{i-2}+q_{i-3})(1+p_{i-1}+q_{i-2})} + \frac{(1-p_{i-1})q_{i-2}(1+p_i-q_{i-1})}{(1+p_{i-1}+q_{i-2})(1+p_i+q_{i-1})} \ ,$$

$$i=4,6,\ldots,m-2 \ ;$$

$$c_i^{(5)} = \frac{2q_{i-3}q_{i-2}(1-q_{i-1})}{(1+p_{i-2}+q_{i-3})(1+p_{i-1}+q_{i-2})} \quad , \ i=5,7,\ldots,m-1 \ ;$$

$$c_i^{(6)} = \frac{2q_{i-4}q_{i-3}q_{i-2}}{(1+p_{i-3}+q_{i-4})(1+p_{i-2}+q_{i-3})} \quad , \ i=6,8,\ldots,m \ ;$$

$$d_i^{(1)} = \frac{2p_ip_{i+1}p_{i+2}}{(1+p_{i+1}+q_i)(1+p_{i+2}+q_{i+1})} \quad , \ i=1,3,\ldots,m-3 \ ;$$

$$d_i^{(2)} = \frac{2(1-p_{i-1})p_ip_{i+1}}{(1+p_i+q_{i-1})(1+p_{i+1}+q_i)} \quad , \ i=2,4,\ldots,m-2 \ ;$$

$$d_i^{(3)} = \frac{p_i(1-q_{i-1})(1-p_{i-1}+q_{i-2})}{(1+p_{i-1}+q_{i-2})(1+p_i+q_{i-1})} + \frac{p_i(1+q_{i-1})(1+p_{i+1}-q_i)}{(1+p_i+q_{i-1})(1+p_{i+1}+q_i)} \ ,$$

$$i=3,5,\ldots,m-3;$$

$$d_i^{(4)} = \frac{p_{i-1}q_{i-2}(1-p_{i-2}+q_{i-3})}{(1+p_{i-2}+q_{i-3})(1+p_{i-1}+q_{i-2})} + \frac{(1-p_{i-1})(1+q_{i-2})(1+p_i-q_{i-1})}{(1+p_{i-1}+q_{i-2})(1+p_i+q_{i-1})},$$

$$i=4,6,\ldots,m-2 \ ;$$

$$d_i^{(5)} = \frac{2q_{i-2}(1+q_{i-3})(1-q_{i-1})}{(1+p_{i-2}+q_{i-3})(1+p_{i-1}+q_{i-2})} \ , \ i=5,7,\ldots,m-1 \ ; \text{ and}$$

$$d_i^{(6)} = \frac{2q_{i-3}(1+q_{i-4})q_{i-2}}{(1+p_{i-3}+q_{i-4})(1+p_{i-2}+q_{i-3})} \ , \ i=6,8,\ldots,m \ .$$

(a) For the case α=2 (spherical symmetry)

When $\alpha=2$, the characteristic polynomial $|D|=|\Gamma_{SAGE}-\mu I|=0$ is given by,

$$|D| = \left|\begin{array}{ccccccccc}
\frac{(1-3\lambda)(1-6\lambda)}{(1+3\lambda)}-\mu & \frac{6\lambda\{2-\lambda+(1-3\lambda)(2+5\lambda)\}}{(1+3\lambda)(2+5\lambda)} & \frac{18\lambda^2(2-3\lambda)}{(1+3\lambda)(2+5\lambda)} & \frac{54\lambda^3}{(1+3\lambda)(2+5\lambda)} & & & & & \\
0 & \frac{(2-3\lambda)}{(2+5\lambda)}-\mu & \frac{4\lambda(2-3\lambda)}{(2+5\lambda)} & \frac{12\lambda^2}{(2+5\lambda)} & & & & & \\
0 & c_2^{(2)} & c_3^{(3)}-\mu & c_4^{(4)} & c_5^{(5)} & c_6^{(6)} & & & \\
0 & d_2^{(2)} & d_3^{(3)} & d_4^{(4)}-\mu & d_5^{(5)} & d_6^{(6)} & & & \\
0 & 0 & c_3^{(1)} & c_4^{(2)} & c_5^{(3)}-\mu & c_6^{(4)} & c_7^{(5)} & c_8^{(6)} & \\
0 & 0 & d_3^{(1)} & d_4^{(2)} & d_5^{(3)} & d_6^{(4)}-\mu & d_7^{(5)} & d_8^{(6)} & \bigcirc \\
 & & & & \vdots & \vdots & \vdots & & \\
 & & \bigcirc & & & & & & \\
 & & & c_{m-5}^{(1)} & c_{m-4}^{(2)} & c_{m-3}^{(3)}-\mu & c_{m-2}^{(4)} & c_{m-1}^{(5)} & c_m^{(6)} \\
 & & & d_{m-5}^{(1)} & d_{m-4}^{(2)} & d_{m-3}^{(3)} & d_{m-2}^{(4)}-\mu & d_{m-1}^{(5)} & d_m^{(6)} \\
 & & & & & c_{m-3}^{(1)} & c_{m-2}^{(2)} & e_1-\mu & e_2 \\
 & & & & & d_{m-3}^{(1)} & d_{m-2}^{(2)} & f_1 & f_2-\mu
\end{array}\right| = 0 \qquad (5.5.26)$$

By expanding the determinant about the first column, we obtain $\{\frac{(1-3\lambda)(1-6\lambda)}{(1+3\lambda)}-\mu\}|D_1|=0$ where $|D_1|$ is the minor derived from $|D|$ by deleting the first row and column. Hence, an eigenvalue of Γ_{SAGE} is

$$\mu = \frac{(1-3\lambda)(1-6\lambda)}{(1+3\lambda)}$$

and since we require $|\mu|\leq 1$ for stability, a real bound on the mesh ratio is

$$\lambda \leq \frac{2}{3} . \qquad (5.5.27)$$

Now, for these values of λ, the sum S of the moduli of the elements in each of the rows 2 to 4 is $S\leq 1$. For rows 5(2)(m-3), the sum of the moduli of the elements in each of the rows is,

$$S_{i+2} = |c_i^{(1)}|+|c_{i+1}^{(2)}|+|c_{i+2}^{(3)}|+|c_{i+3}^{(4)}|+|c_{i+4}^{(5)}|+|c_{i+5}^{(6)}| ;$$
$$i=3,5,\ldots,(m-5) . \qquad (5.5.28)$$

Similarly, the sum of the moduli of the elements in each of the rows 6(2)(m-2) is

$$S_{i+3} = |d_i^{(1)}|+|d_{i+1}^{(2)}|+|d_{i+2}^{(3)}|+|d_{i+3}^{(4)}|+|d_{i+4}^{(5)}|+|d_{i+5}^{(6)}| \quad ;$$
$$i=3,5,\ldots,(m-5) . \qquad (5.5.29)$$

It can be shown that for

$$\lambda \leq \frac{i+1}{i+2} \quad , \; i=3,5,\ldots,(m-5) \qquad (5.5.30)$$

$$S_{i+2} \leq 1 \text{ and } S_{i+3} \leq 1.$$

The absolute sums in rows (m-1) and m are given respectively by

$$S_{m-1} = |c_{m-3}^{(1)}|+|c_{m-2}^{(2)}|+|e_1|+|e_2| \; , \qquad (5.5.31)$$

and

$$S_m = |d_{m-3}^{(1)}|+|d_{m-2}^{(2)}|+|f_1|+|f_2| \; . \qquad (5.5.32)$$

Again, it can be shown that for

$$\lambda \leq \frac{m-2}{m-1} \quad , \qquad (5.5.33)$$

$S_{m-1}<1$ and $S_m<1$. For overall stability of the (S)AGE scheme, we deduce from (5.5.27), (5.5.30) and (5.5.33) the condition

$$\lambda \lesssim \min.\{\frac{2}{3},\min.(\frac{i+1}{i+2}\ ,\ i=3,5,\ldots,(m-5))\ ,\ \frac{m-2}{m-1}\}\ ,\ m \text{ even and } m>4;$$

i.e., $\lambda \lesssim \frac{2}{3}$.

(b) For the case $\alpha=1$ (cylindrical symmetry)

The characteristic polynomial of Γ_{SAGE} when $\alpha=1$ is given by,

$$|D| = \begin{vmatrix}
a_1-\mu & a_2 & a_3 & a_4 & & & & & & & & & & & \\
b_1 & b_2-\mu & b_3 & b_4 & & & & & & & & & & & \\
c_1^{(1)} & c_2^{(2)} & c_3^{(3)}-\mu & c_4^{(4)} & c_5^{(5)} & c_6^{(6)} & & & & & & & & & \\
d_1^{(1)} & d_2^{(2)} & d_3^{(3)} & d_4^{(4)}-\mu & d_5^{(5)} & d_6^{(6)} & & & & & & & & & \\
0 & 0 & c_3^{(1)} & c_4^{(2)} & c_5^{(3)}-\mu & c_6^{(4)} & c_7^{(5)} & c_8^{(6)} & & & & & \bigcirc & & \\
0 & 0 & d_3^{(1)} & d_4^{(2)} & d_5^{(3)} & d_6^{(4)}-\mu & d_7^{(5)} & d_8^{(6)} & & & & & & & \\
 & & & & \ddots & \ddots & \ddots & \ddots & \ddots & & & & & & \\
 & & & \bigcirc & & & & & & c_{m-5}^{(1)} & c_{m-4}^{(2)} & c_{m-3}^{(3)}-\mu & c_{m-2}^{(4)} & c_{m-1}^{(5)} & c_m^{(6)} \\
 & & & & & & & & & d_{m-5}^{(1)} & d_{m-4}^{(2)} & d_{m-3}^{(3)} & d_{m-2}^{(4)}-\mu & d_{m-1}^{(5)} & d_m^{(6)} \\
 & & & & & & & & & & & c_{m-3}^{(1)} & c_{m-2}^{(2)} & e_1-\mu & e_2 \\
 & & & & & & & & & & & d_{m-3}^{(1)} & d_{m-2}^{(2)} & f_1 & f_2-\mu
\end{vmatrix} = 0 \quad (5.5.34)$$

where,

$$a_1 = \{(2+\lambda)(1-2\lambda)(1-4\lambda)(4+9\lambda)+\lambda^2[(4+3\lambda)(4+\lambda)-12\lambda]\}/\{(2+5\lambda)(4+9\lambda)\};$$

$$a_2 = \{4\lambda(2+\lambda)(1-2\lambda)(4+9\lambda)+\lambda(2-\lambda)[(4+3\lambda)(4+\lambda)-12\lambda]\}/\{(2+5\lambda)(4+9\lambda)\};$$

$$a_3 = \{3\lambda^2(4-5\lambda)(6+\lambda)\}/\{(2+5\lambda)(4+9\lambda)\};$$

$$a_4 = \{15\lambda^3(6+\lambda)\}/\{(2+5\lambda)(4+9\lambda)\};$$

$$b_1 = \lambda\{(1-2\lambda)(1-4\lambda)(4+9\lambda)+[(1+2\lambda)(4-3\lambda)+(4+3\lambda)\lambda^2]\}/\{(2+5\lambda)(4+9\lambda)\};$$

$$b_2 = \{4\lambda^2(1-2\lambda)(4+9\lambda)+(2-\lambda)[(1+2\lambda)(4-3\lambda)+(4+3\lambda)\lambda^2]\}/\{(2+5\lambda)(4+9\lambda)\};$$

$$b_3 = \{3\lambda(4-5\lambda)[2(1+2\lambda)+\lambda^2]\}/\{(2+5\lambda)(4+9\lambda)\};$$

$$b_4 = \{15\lambda^2[2(1+2\lambda)+\lambda^2]\}/\{(2+5\lambda)(4+9\lambda)\};$$

$$c_1^{(1)} = \{6\lambda^2(6+5\lambda)\}/\{(4+9\lambda)(12+25\lambda)\};$$

$$c_2^{(2)} = \{6\lambda(2-\lambda)(6+5\lambda)\}/\{(4+9\lambda)(12+25\lambda)\};$$

$$c_3^{(3)} = \frac{\{(6+5\lambda)(4-5\lambda)(4+3\lambda)(24+49\lambda)+25\lambda^2(24-7\lambda)(4+9\lambda)\}}{2(4+9\lambda)(12+25\lambda)(24+49\lambda)} \quad ;$$

$$c_4^{(4)} = \frac{5\{(6+5\lambda)(4+3\lambda)(24+49\lambda)+(6-5\lambda)(4+9\lambda)(24-7\lambda)\}}{2(4+9\lambda)(24+49\lambda)(12+25\lambda)} \quad ;$$

$$c_5^{(5)} = \{105\lambda^2(8-9\lambda)\}/\{(12+25\lambda)(24+49\lambda)\} \quad ;$$

$$c_6^{(6)} = 945\lambda^3/\{(12+25\lambda)(24+49\lambda)\};$$

$$d_1^{(1)} = 30\lambda^3/\{(4+9\lambda)(12+25\lambda)\};$$

$$d_2^{(2)} = \{30\lambda^2(2-\lambda)\}/\{(4+9\lambda)(12+25\lambda)\};$$

$$d_3^{(3)} = 5\lambda(4-5\lambda)(4+3\lambda)/2(4+9\lambda)(12+25\lambda) \quad ;$$

$$d_4^{(4)} = \frac{\{25\lambda^2(4+3\lambda)(24+49\lambda)+(6-5\lambda)(4+5\lambda)(4+9\lambda)(24-7\lambda)\}}{2(4+9\lambda)(12+25\lambda)(24+49\lambda)} \quad ;$$

$$d_5^{(5)} = \{21\lambda(4+5\lambda)(8-9\lambda)\}/\{(12+25\lambda)(24+49\lambda)\} \quad ; \text{ and}$$

$$d_6^{(6)} = 189\lambda^2(4+5\lambda)/(12+25\lambda)(24+49\lambda) \quad ,$$

and values for all other remaining c's, d's, e's and f's take the same form as before but with $\alpha=1$.

By making the appropriate row operations on the determinant $|D|$ we obtain:

$$|D| = \left| \begin{array}{cccccccccccc}
a_1-\mu & a_2 & a_3 & a_4 & & & & & & & & \\
0 & \beta_1 & \beta_2 & \beta_3 & & & & & & & & \\
0 & \beta_4 & \beta_5 & \beta_6 & c_5^{(5)} & c_6^{(6)} & & & & & & \\
0 & \beta_7 & \beta_8 & \beta_9 & d_5^{(5)} & d_6^{(6)} & & & & & & \\
0 & 0 & c_3^{(1)} & c_4^{(2)} & c_5^{(3)}\mu & c_6^{(4)} & c_7^{(5)} & c_8^{(6)} & & & & \\
0 & 0 & d_3^{(1)} & d_4^{(2)} & d_5^{(3)} & d_6^{(4)}-\mu & d_7^{(5)} & d_8^{(6)} & & & & \\
 & & & & \ddots & \ddots & \ddots & & & 0 & & \\
 & & 0 & & & & c_{m-5}^{(1)} & c_{m-4}^{(2)} & c_{m-3}^{(3)}-\mu & c_{m-2}^{(4)} & c_{m-1}^{(5)} & c_m^{(6)} \\
 & & & & & & d_{m-5}^{(1)} & d_{m-4}^{(2)} & d_{m-3}^{(3)} & d_{m-2}^{(4)}-\mu & d_{m-1}^{(5)} & d_m^{(6)} \\
 & & & & & & & & c_{m-3}^{(1)} & c_{m-2}^{(2)} & e_1-\mu & e_2 \\
 & & & & & & & & d_{m-3}^{(1)} & d_{m-2}^{(2)} & f_1 & f_2-\mu
\end{array} \right| = 0 , \quad (5.5.35)$$

where,

$$\beta_1 = b_2-\mu-\frac{a_2b_1}{(a_1-\mu)}\ ; \quad \beta_4 = c_2^{(2)} - \frac{c_1^{(1)}a_2}{(a_1-\mu)}\ ;$$

$$\beta_2 = b_3-\frac{a_3b_1}{(a_1-\mu)}\ ; \quad \beta_5 = c_3^{(3)}-\mu-\frac{c_1^{(1)}a_3}{(a_1-\mu)}\ ;$$

$$\beta_3 = b_4-\frac{a_4b_1}{(a_1-\mu)}\ ; \quad \beta_6 = c_4^{(4)} - \frac{c_1^{(1)}a_4}{(a_1-\mu)}\ ;$$

$$\beta_7 = d_2^{(2)}-\frac{d_1^{(1)}a_2}{(a_1-\mu)}\ ; \quad \beta_8 = d_3^{(3)} - \frac{d_1^{(1)}a_3}{(a_1-\mu)}\ ,$$

and

$$\beta_9 = d_4^{(4)}-\mu-\frac{d_1^{(1)}a_4}{(a_1-\mu)}\ .$$

By working on the first column of $|D|$ and the other minors, we obtain the following restrictions on λ:

working on $|D|$, $\lambda\leq 0.85587$; (5.5.36)

working on $|D_1|$, $\lambda\leq 1$; (5.5.37)

working on $|D_2|$, $\lambda\leq 0.85587$; (5.5.38)

and working on $|D_3|$, $\lambda\leq 0.85587$. (5.5.39)

Now, if we let S_{i+2} and S_{i+3} $(i=3,5,\ldots,m-5)$ be the sum of the moduli of the elements in each of the rows 5(2)(m-3) and 6(2)(m-2) respectively, then following the same line of argument as in the spherically symmetric case, we note that,

$$S_{i+2} \leq 1$$

and

$$S_{i+3} \leq 1$$

for

$$\lambda \leq \frac{(2i+2)}{(2i+3)}\ ,\ i=3,5,\ldots,m-5\ . \tag{5.5.40}$$

We also find that the absolute sums in rows (m-1) and m satisfy,

$$S_{m-1} < 1\ ,$$

and

$$S_m < 1\ ,$$

when

$$\lambda \leq \frac{(2m-4)}{(2m-3)}\ . \tag{5.5.41}$$

We deduce from (5.5.36)-(5.5.39) and (5.5.40)-(5.5.41) that the (S)AGE scheme for the cylindrically symmetric case is conditionally stable for

$$\lambda \leq \min\{0.85587,\ 1.0, \min(\frac{2i+2}{2i+3},\ i=3,5,\ldots,(m-5)),\ \frac{(2m-4)}{(2m-3)}\}$$
$$= 0.85587.$$

(iv) Stability of the (D)AGE Scheme

From (5.3.35), the explicit form of the (D)AGE scheme is expressed by

$$\underline{u}_{j+4} = \Gamma_{DAGE}\underline{u}_j + \underline{b}_2'' \quad , \ j=0,4,8,\ldots$$

where $$\Gamma_{DAGE} = \Gamma_{GER}\Gamma_{GEL}\Gamma_{SAGE} \ , \qquad (5.5.42)$$

which is the amplification matrix of the (D)AGE scheme and whose characteristic polynomial is given by,

$$
|D| = \begin{vmatrix}
(\tilde{a}_1-\mu) & \tilde{a}_2 & \tilde{a}_3 & \tilde{a}_4 & \tilde{a}_5 & \tilde{a}_6 &&&&&&&&&&&&&& \\
\tilde{b}_1 & (\tilde{b}_2-\mu) & \tilde{b}_3 & \tilde{b}_4 & \tilde{b}_5 & \tilde{b}_6 & \tilde{b}_7 & \tilde{b}_8 &&&&&&&&&&&& \\
\tilde{c}_1 & \tilde{c}_2 & (\tilde{c}_3-\mu) & \tilde{c}_4 & \tilde{c}_5 & \tilde{c}_6 & \tilde{c}_7 & \tilde{c}_8 &&&&&&&&&&&& \\
\tilde{d}_1 & \tilde{d}_2 & \tilde{d}_3 & (\tilde{d}_4-\mu) & \tilde{d}_5 & \tilde{d}_6 & \tilde{d}_7 & \tilde{d}_8 & \tilde{d}_9 & \tilde{d}_{10} &&&&&&&&&& \\
\tilde{e}_1 & \tilde{e}_2 & \tilde{e}_3 & \tilde{e}_4 & (\tilde{e}_5-\mu) & \tilde{e}_6 & \tilde{e}_7 & \tilde{e}_8 & \tilde{e}_9 & \tilde{e}_{10} &&&&&&&&&& \\
\tilde{c}_1^{(1)} & \tilde{c}_2^{(2)} & \tilde{c}_3^{(3)} & \tilde{c}_4^{(4)} & \tilde{c}_5^{(5)} & (\tilde{c}_6^{(6)}-\mu) & \tilde{c}_7^{(7)} & \tilde{c}_8^{(8)} & \tilde{c}_9^{(9)} & \tilde{c}_{10}^{(10)} & \tilde{c}_{11}^{(11)} & \tilde{c}_{12}^{(12)} &&&&&&&& \\
\tilde{d}_1^{(1)} & \tilde{d}_2^{(2)} & \tilde{d}_3^{(3)} & \tilde{d}_4^{(4)} & \tilde{d}_5^{(5)} & \tilde{d}_6^{(6)} & (\tilde{d}_7^{(7)}-\mu) & \tilde{d}_8^{(8)} & \tilde{d}_9^{(9)} & \tilde{d}_{10}^{(10)} & \tilde{d}_{11}^{(11)} & \tilde{d}_{12}^{(12)} &&&&&&&& \\
0 & 0 & \tilde{c}_3^{(1)} & \tilde{c}_4^{(2)} & \tilde{c}_5^{(3)} & \tilde{c}_6^{(4)} & \tilde{c}_7^{(5)} & (\tilde{c}_8^{(6)}-\mu) & \tilde{c}_9^{(7)} & \tilde{c}_{10}^{(8)} & \tilde{c}_{11}^{(9)} & \tilde{c}_{12}^{(10)} & \tilde{c}_{13}^{(11)} & \tilde{c}_{14}^{(12)} &&&&&& \\
0 & 0 & \tilde{d}_3^{(1)} & \tilde{d}_4^{(2)} & \tilde{d}_5^{(3)} & \tilde{d}_6^{(4)} & \tilde{d}_7^{(5)} & \tilde{d}_8^{(6)} & (\tilde{d}_9^{(7)}-\mu) & \tilde{d}_{10}^{(8)} & \tilde{d}_{11}^{(9)} & \tilde{d}_{12}^{(10)} & \tilde{d}_{13}^{(11)} & \tilde{d}_{14}^{(12)} &&&&&& \\
&&&&&&&& \ddots & \ddots & \ddots &&&&&&& \bigcirc && \\
&&&& \bigcirc &&&&&&&&&&&&&&& \\
&&&&&&&& \tilde{c}_{m-11}^{(1)} & \tilde{c}_{m-10}^{(2)} & \tilde{c}_{m-9}^{(3)} & \tilde{c}_{m-8}^{(4)} & \tilde{c}_{m-7}^{(5)} & (\tilde{c}_{m-6}^{(6)}-\mu) & \tilde{c}_{m-5}^{(7)} & \tilde{c}_{m-4}^{(8)} & \tilde{c}_{m-3}^{(9)} & \tilde{c}_{m-2}^{(10)} & \tilde{c}_{m-1}^{(11)} & \tilde{c}_{m}^{(12)} \\
&&&&&&&& \tilde{d}_{m-11}^{(1)} & \tilde{d}_{m-10}^{(2)} & \tilde{d}_{m-9}^{(3)} & \tilde{d}_{m-8}^{(4)} & \tilde{d}_{m-7}^{(5)} & \tilde{d}_{m-6}^{(6)} & (\tilde{d}_{m-5}^{(7)}-\mu) & \tilde{d}_{m-4}^{(8)} & \tilde{d}_{m-3}^{(9)} & \tilde{d}_{m-2}^{(10)} & \tilde{d}_{m-1}^{(11)} & \tilde{d}_{m}^{(12)} \\
&&&&&&&& 0 & 0 & \tilde{f}_1 & \tilde{f}_2 & \tilde{f}_3 & \tilde{f}_4 & \tilde{f}_5 & (\tilde{f}_6-\mu) & \tilde{f}_7 & \tilde{f}_8 & \tilde{f}_9 & \tilde{f}_{10} \\
&&&&&&&& 0 & 0 & \tilde{g}_1 & \tilde{g}_2 & \tilde{g}_3 & \tilde{g}_4 & \tilde{g}_5 & \tilde{g}_6 & (\tilde{g}_7-\mu) & \tilde{g}_8 & \tilde{g}_9 & \tilde{g}_{10} \\
&&&&&&&&&& 0 & 0 & \tilde{r}_1 & \tilde{r}_2 & \tilde{r}_3 & \tilde{r}_4 & \tilde{r}_5 & (\tilde{r}_6-\mu) & \tilde{r}_7 & \tilde{r}_8 \\
&&&&&&&&&& 0 & 0 & \tilde{s}_1 & \tilde{s}_2 & \tilde{s}_3 & \tilde{s}_4 & \tilde{s}_5 & \tilde{s}_6 & (\tilde{s}_7-\mu) & \tilde{s}_8 \\
&&&&&&&&&&&& 0 & 0 & \tilde{t}_1 & \tilde{t}_2 & \tilde{t}_3 & \tilde{t}_4 & \tilde{t}_5 & (\tilde{t}_6-\mu)
\end{vmatrix} = 0 . \qquad (5.5.43)
$$

where

$$\tilde{a}_i = \hat{a}_1 a_i + \hat{a}_2 b_i + \hat{a}_3 c_i^{(i)} \text{ , for } i=1,2,3,4,$$

$$\tilde{a}_j = \hat{a}_3 c_j^{(j)} \text{ for } j=5,6 \text{ ;}$$

$$\tilde{b}_i = \hat{b}_1 a_i + \hat{b}_2 b_i + \hat{b}_3 c_i^{(i)} + \hat{b}_4 d_i^{(i)} \text{ for } i=1,2 \text{ ,}$$

$$\tilde{b}_j = \hat{b}_1 a_j + \hat{b}_2 b_j + \hat{b}_3 c_j^{(j)} + \hat{b}_4 d_j^{(j)} + \hat{b}_5 c_j^{(j-2)} \text{ for } j=3,4$$

$$\tilde{b}_k = \hat{b}_3 c_k^{(k)} + \hat{b}_4 d_k^{(k)} + \hat{b}_5 c_k^{(k-2)} \text{ for } k=5,6,$$

$$\tilde{b}_\ell = \hat{b}_5 c_\ell^{(\ell-2)} \text{ for } \ell=7,8;$$

$$\tilde{c}_i = \hat{c}_1 a_i + \hat{c}_2 b_i + \hat{c}_3 c_i^{(i)} + \hat{c}_4 d_i^{(i)} \text{ for } i=1,2 \text{ ,}$$

$$\tilde{c}_j = \hat{c}_1 a_j + \hat{c}_2 b_j + \hat{c}_3 c_j^{(j)} + \hat{c}_4 d_j^{(j)} + \hat{c}_5 c_j^{(j-2)} \text{ for } j=3,4,$$

$$\tilde{c}_k = \hat{c}_3 c_k^{(k)} + \hat{c}_4 d_k^{(k)} + \hat{c}_5 c_k^{(k-2)} \text{ for } k=5,6$$

$$\tilde{c}_\ell = \hat{c}_5 c_\ell^{(\ell-2)} \text{ for } \ell=7,8;$$

$$\tilde{d}_i = \hat{d}_1^{(1)} b_i + \hat{d}_2^{(2)} c_i^{(i)} + \hat{d}_3^{(3)} d_i^{(i)} \text{ for } i=1,2,$$

$$\tilde{d}_j = \hat{d}_1^{(1)} b_j + \hat{d}_2^{(2)} c_j^{(j)} + \hat{d}_3^{(3)} d_j^{(j)} + \hat{d}_4^{(4)} c_j^{(j-2)} + \hat{d}_5^{(5)} d_j^{(j-2)} \text{ for } j=3,4$$

$$\tilde{d}_k = \hat{d}_2^{(2)} c_k^{(k)} + \hat{d}_3^{(3)} d_k^{(k)} + \hat{d}_4^{(4)} c_k^{(k-2)} + \hat{d}_5^{(5)} d_k^{(k-2)} + \hat{d}_6^{(6)} c_k^{(k-4)} \text{ for } k=5,6,$$

$$\tilde{d}_\ell = \hat{d}_4^{(4)} c_\ell^{(\ell-2)} + \hat{d}_5^{(5)} d_\ell^{(\ell-2)} + \hat{d}_6^{(6)} c_\ell^{(\ell-4)} \text{ for } \ell=7,8,$$

$$\tilde{d}_n = \hat{d}_6^{(6)} c_n^{(n-4)} \text{ for } n=9,10;$$

$$\tilde{e}_i = \hat{e}_1^{(1)} b_i + \hat{e}_2^{(2)} c_i^{(i)} + \hat{e}_3^{(3)} d_i^{(i)} \text{ for } i=1,2$$

$$\tilde{e}_j = \hat{e}_1^{(1)} b_j + \hat{e}_2^{(2)} c_j^{(j)} + \hat{e}_3^{(3)} d_j^{(j)} + \hat{e}_4^{(4)} c_j^{(j-2)} + \hat{e}_5^{(5)} d_j^{(j-2)} \text{ for } j=3,4,$$

$$\tilde{e}_k = \hat{e}_2^{(2)} c_k^{(k)} + \hat{e}_3^{(3)} d_k^{(k)} + \hat{e}_4^{(4)} c_k^{(k-2)} + \hat{e}_5^{(5)} d_k^{(k-2)} + \hat{e}_6^{(6)} c_k^{(k-4)} \text{ for } k=5,6,$$

$$\tilde{e}_\ell = \hat{e}_4^{(4)} c_\ell^{(\ell-2)} + \hat{e}_5^{(5)} d_\ell^{(\ell-2)} + \hat{e}_6^{(6)} c_\ell^{(\ell-4)} \text{ for } \ell=7,8,$$

$$\tilde{e}_n = \hat{e}_6^{(6)} c_n^{(n-4)} \text{ for } n=9,10 \text{ ;}$$

$$\tilde{c}_i^{(1)} = \hat{d}_{i+2}^{(1)}\, d_i^{(1)} \quad \text{for } i=1(2)(m-11),$$

$$\tilde{c}_i^{(2)} = \hat{d}_{i+1}^{(1)} d_i^{(2)} \quad \text{for } i=2(2)(m-10),$$

$$\tilde{c}_i^{(3)} = \hat{d}_i^{(1)} d_i^{(3)} + \hat{d}_{i+1}^{(2)} c_i^{(1)} + \hat{d}_{i+2}^{(3)}\, d_i^{(1)} \quad \text{for } i=3(2)(m-9)\ ,$$

$$\tilde{c}_i^{(4)} = \hat{d}_{i-1}^{(1)} d_i^{(4)} + \hat{d}_i^{(2)} c_i^{(2)} + \hat{d}_{i+1}^{(3)} d_i^{(2)} \quad \text{for } i=4(2)(m-8)\ ,$$

$$\tilde{c}_i^{(5)} = \hat{d}_{i-2}^{(1)} d_i^{(5)} + \hat{d}_{i-1}^{(2)} c_i^{(3)} + \hat{d}_i^{(3)} d_i^{(3)} + \hat{d}_{i+1}^{(4)} c_i^{(1)} + \hat{d}_{i+2}^{(5)} d_i^{(1)} \quad \text{for } i=5(2)(m-7),$$

$$\tilde{c}_i^{(6)} = \hat{d}_{i-3}^{(1)} d_i^{(6)} + \hat{d}_{i-2}^{(2)}\, c_i^{(4)} + \hat{d}_{i-1}^{(3)} d_i^{(4)} + \hat{d}_i^{(4)} c_i^{(2)} + \hat{d}_{i+1}^{(5)} d_i^{(2)} \quad \text{for } i=6(2)(m-6),$$

$$\tilde{c}_i^{(7)} = \hat{d}_{i-3}^{(2)} c_i^{(5)} + \hat{d}_{i-2}^{(3)} d_i^{(5)} + \hat{d}_{i-1}^{(4)} c_i^{(3)} + \hat{d}_i^{(5)} d_i^{(3)} + \hat{d}_{i+1}^{(6)} c_i^{(1)} \quad \text{for } i=7(2)(m-5),$$

$$\tilde{c}_i^{(8)} = \hat{d}_{i-4}^{(2)} c_i^{(6)} + \hat{d}_{i-3}^{(3)} d_i^{(6)} + \hat{d}_{i-2}^{(4)} c_i^{(4)} + \hat{d}_{i-1}^{(5)} d_i^{(4)} + \hat{d}_i^{(6)} c_i^{(2)} \quad \text{for } i=8(2)(m-4)\ ,$$

$$\tilde{c}_i^{(9)} = \hat{d}_{i-3}^{(4)} c_i^{(5)} + \hat{d}_{i-2}^{(5)} d_i^{(5)} + \hat{d}_{i-1}^{(6)} c_i^{(3)} \quad \text{for } i=9(2)(m-3)\ ,$$

$$\tilde{c}_i^{(10)} = \hat{d}_{i-4}^{(4)} c_i^{(6)} + \hat{d}_{i-3}^{(5)} d_i^{(6)} + \hat{d}_{i-2}^{(6)} c_i^{(4)} \quad \text{for } i=10(2)(m-2)\ ,$$

$$\tilde{c}_i^{(11)} = \hat{d}_{i-3}^{(6)} c_i^{(5)} \quad \text{for } i=11(2)(m-1)\ ,$$

$$\tilde{c}_i^{(12)} = \hat{d}_{i-4} c_i^{(6)} \quad \text{for } i=12(2)(m)\ ;$$

$$\tilde{d}_i^{(1)} = \hat{e}_{i+2}^{(1)}\, d_i^{(1)} \quad \text{for } i=1(2)(m-11),$$

$$\tilde{d}_i^{(2)} = \hat{e}_{i+1}^{(1)}\, d_i^{(2)} \quad \text{for } i=2(2)(m-10),$$

$$\tilde{d}_i^{(3)} = \hat{e}_i^{(1)}\, d_i^{(3)} + \hat{e}_{i+1}^{(2)} c_i^{(1)} + \hat{e}_{i+2}^{(3)} d_i^{(1)} \quad \text{for } i=3(2)(m-9),$$

$$\tilde{d}_i^{(4)} = \hat{e}_{i-1}^{(1)}\, d_i^{(4)} + \hat{e}_i^{(2)}\, c_i^{(2)} + \hat{e}_{i+1}^{(3)} d_i^{(2)} \quad \text{for } i=4(2)(m-8),$$

$$\hat{d}_i^{(5)} = \hat{e}_{i-2}^{(1)}\, d_i^{(5)} + \hat{e}_{i-1}^{(2)} c_i^{(3)} + \hat{e}_i^{(3)} d_i^{(3)} + \hat{e}_{i+1}^{(4)} c_i^{(1)} + \hat{e}_{i+2}^{(5)} d_i^{(1)} \quad \text{for } i=5(2)(m-7)\ ,$$

$$\tilde{d}_i^{(6)} = \hat{e}_{i-3}^{(1)}\, d_i^{(6)} + \hat{e}_{i-2}^{(2)}\, c_i^{(4)} + \hat{e}_{i-1}^{(3)} d_i^{(4)} + \hat{e}_i^{(4)} c_i^{(2)} + \hat{e}_{i+1}^{(5)} d_i^{(2)} \quad \text{for } i=6(2)(m-6),$$

$$\tilde{d}_i^{(7)} = \hat{e}_{i-3}^{(2)} c_i^{(5)} + \hat{e}_{i-2}^{(3)} d_i^{(5)} + \hat{e}_{i-1}^{(4)} c_i^{(3)} + \hat{e}_i^{(5)} d_i^{(3)} + \hat{e}_{i+1}^{(6)} c_i^{(1)}$$

$$\text{for } i=7(2)(m-5),$$

$$\tilde{d}_i^{(8)} = \hat{e}_{i-4}^{(2)} c_i^{(6)} + \hat{e}_{i-3}^{(3)} d_i^{(6)} + \hat{e}_{i-2}^{(4)} c_i^{(4)} + \hat{e}_{i-1}^{(5)} d_i^{(4)} + \hat{e}_i^{(6)} c_i^{(2)}$$

$$\text{for } i=8(2)(m-4),$$

$$\tilde{d}_i^{(9)} = \hat{e}_{i-3}^{(4)} c_i^{(5)} + \hat{e}_{i-2}^{(5)} d_i^{(5)} + \hat{e}_{i-1}^{(6)} c_i^{(3)} \quad \text{for } i=9(2)(m-3),$$

$$\tilde{d}_i^{(10)} = \hat{e}_{i-4}^{(4)} c_i^{(6)} + \hat{e}_{i-3}^{(5)} d_i^{(6)} + \hat{e}_{i-2}^{(6)} c_i^{(4)} \quad \text{for } i=10(2)(m-2),$$

$$\tilde{d}_i^{(11)} = \hat{e}_{i-3}^{(6)} c_i^{(5)} \quad \text{for } i=11(2)(m-1)\ ,$$

$$\tilde{d}_i^{(12)} = \hat{e}_{i-4} c_i^{(6)} \quad \text{for } i=12(2)(m);$$

$$\tilde{f}_i = \hat{d}_{m-7}^{(1)} d_{m+i-10}^{(i)} \quad \text{for } i=1,2,$$

$$\tilde{f}_j = \hat{d}_{m-7}^{(1)} d_{m+j-10}^{(j)} + \hat{d}_{m-6}^{(2)} c_{m+j-10}^{(j-2)} + \hat{d}_{m-5}^{(3)} d_{m+j-10}^{(j-2)} \quad \text{for } j=3,4,$$

$$\tilde{f}_k = \hat{d}_{m-7}^{(1)} d_{m+k-10}^{(k)} + \hat{d}_{m-6}^{(2)} c_{m+k-10}^{(k-2)} + \hat{d}_{m-5}^{(3)} d_{m+k-10}^{(k-2)} + \hat{d}_{m-4}^{(4)} c_{m+k-10}^{(k-4)} +$$

$$\hat{d}_{m-3}^{(5)} d_{m+k-10}^{(k-4)} \quad \text{for } k=5,6,$$

$$\tilde{f}_\ell = \hat{d}_{m-6}^{(2)} c_{m+\ell-10}^{(\ell-2)} + \hat{d}_{m-5}^{(3)} d_{m+\ell-10}^{(\ell-2)} + \hat{d}_{m-4}^{(4)} c_{m+\ell-10}^{(\ell-4)} + \hat{d}_{m-3}^{(5)} d_{m+\ell-10}^{(\ell-4)} +$$

$$\hat{d}_{m-2}^{(6)} c_{m+\ell-10}^{(\ell-6)} \quad \text{for } =7,8,$$

$$\tilde{f}_9 = \hat{d}_{m-4}^{(4)} c_{m-1}^{(5)} + \hat{d}_{m-3}^{(5)} d_{m-1}^{(5)} + \hat{d}_{m-2}^{(6)} e_1\ ,$$

$$\tilde{f}_{10} = \hat{d}_{m-4}^{(4)} c_m^{(6)} + \hat{d}_{m-3}^{(5)} d_m^{(6)} + \hat{d}_{m-2}^{(6)} e_2\ ;$$

$$\tilde{g}_i = \hat{e}_{m-7}^{(1)} d_{m+i-10}^{(i)} \quad \text{for } i=1,2,$$

$$\tilde{g}_j = \hat{e}_{m-7}^{(1)} d_{m+j-10}^{(j)} + \hat{e}_{m-6}^{(2)} c_{m+j-10}^{(j-2)} + \hat{e}_{m-5}^{(3)} d_{m+j-10}^{(j-2)} \quad \text{for } j=3,4,$$

$$\tilde{g}_k = \hat{e}_{m-7}^{(1)} d_{m+k-10}^{(k)} + \hat{e}_{m-6}^{(2)} c_{m+k-10}^{(k-2)} + \hat{e}_{m-5}^{(3)} d_{m+k-10}^{(k-2)} + \hat{e}_{m-4}^{(4)} c_{m+k-10}^{(k-4)} +$$

$$\hat{e}_{m-3}^{(5)} d_{m+k-10}^{(k-4)} \quad \text{for } k=5,6,$$

$$\tilde{g}_\ell = \hat{e}_{m-6}^{(2)} c_{m+\ell-10}^{(\ell-2)} + \hat{e}_{m-5}^{(3)} d_{m+\ell-10}^{(\ell-2)} + \hat{e}_{m-4}^{(4)} c_{m+\ell-10}^{(\ell-4)} + \hat{e}_{m-3}^{(5)} d_{m+\ell-10}^{(\ell-4)} +$$

$$\hat{e}_{m-2}^{(6)} c_{m+\ell-10}^{(\ell-6)}, \quad \text{for } \ell=7,8,$$

$$\tilde{g}_9 = \hat{e}^{(4)}_{m-4}c^{(5)}_{m-1} + \hat{e}^{(5)}_{m-3}d^{(5)}_{m-1} + \hat{e}^{(6)}_{m-2}e_1 ,$$

$$\tilde{g}_{10} = \hat{e}^{(4)}_{m-4}c^{(6)}_{m} + \hat{e}^{(5)}_{m-3}d^{(6)}_{m} + \hat{e}^{(6)}_{m-2}e_2 ,$$

$$\tilde{r}_i = \hat{d}^{(1)}_{m-5}d^{(i)}_{m+i-8} \quad \text{for } i=1,2,$$

$$\tilde{r}_j = \hat{d}^{(1)}_{m-5}d^{(j)}_{m+j-8} + \hat{d}^{(2)}_{m-4}c^{(j-2)}_{m+j-8} + \hat{d}^{(3)}_{m-3}d^{(j-2)}_{m+j-8} \quad \text{for } j=3,4,$$

$$\tilde{r}_k = \hat{d}^{(1)}_{m-5}d^{(k)}_{m+k-8} + \hat{d}^{(2)}_{m-4}c^{(k-2)}_{m+k-8} + \hat{d}^{(3)}_{m-3}d^{(k-2)}_{m+k-8} + \hat{d}^{(4)}_{m-2}c^{(k-4)}_{m+k-8} + \hat{d}^{(5)}_{m-1}d^{(k-4)}_{m+k-8}$$

$$\text{for } k=5,6,$$

$$\tilde{r}_7 = \hat{d}^{(2)}_{m-4}c^{(5)}_{m-1} + \hat{d}^{(3)}_{m-3}d^{(5)}_{m-1} + \hat{d}^{(4)}_{m-2}e_1 + \hat{d}^{(5)}_{m-1}f_1 ,$$

$$\tilde{r}_8 = \hat{d}^{(2)}_{m-4}c^{(6)}_{m} + \hat{d}^{(3)}_{m-3}d^{(6)}_{m} + \hat{d}^{(4)}_{m-2}e_2 + \hat{d}^{(5)}_{m-1}f_2 ;$$

$$\tilde{s}_i = \hat{e}^{(1)}_{m-5}d^{(i)}_{m+i-8} \quad \text{for } i=1,2,$$

$$\tilde{s}_j = \hat{e}^{(1)}_{m-5}d^{(j)}_{m+j-8} + \hat{e}^{(2)}_{m-4}c^{(j-2)}_{m+j-8} + \hat{e}^{(3)}_{m-3}d^{(j-2)}_{m+j-8} \quad \text{for } j=3,4,$$

$$\tilde{s}_k = \hat{e}^{(1)}_{m-5}d^{(k)}_{m+k-8} + \hat{e}^{(2)}_{m-4}c^{(k-2)}_{m+k-8} + \hat{e}^{(3)}_{m-3}d^{(k-2)}_{m+k-8} + \hat{e}^{(4)}_{m-2}c^{(k-4)}_{m+k-8} + \hat{e}^{(5)}_{m-1}d^{(k-4)}_{m+k-8}$$

$$\text{for } k=5,6,$$

$$\tilde{s}_7 = \hat{e}^{(2)}_{m-4}c^{(5)}_{m-1} + \hat{e}^{(3)}_{m-3}d^{(5)}_{m-1} + \hat{e}^{(4)}_{m-2}e_1 + \hat{e}^{(5)}_{m-1}f_1 ,$$

$$\tilde{s}_8 = \hat{e}^{(2)}_{m-4}c^{(6)}_{m} + \hat{e}^{(3)}_{m-3}d^{(6)}_{m} + \hat{e}^{(4)}_{m-2}e_2 + \hat{e}^{(5)}_{m-1}f_2 ;$$

$$\tilde{t}_i = \hat{f}_1 d^{(i)}_{m+i-6} \quad \text{for } i=1,2,$$

$$\tilde{t}_j = \hat{f}_1 d^{(j)}_{m+j-6} + \hat{f}_2 c^{(j-2)}_{m+j-6} + \hat{f}_3 d^{(j-2)}_{m+j-6} \quad \text{for } j=3,4,$$

$$\tilde{t}_5 = \hat{f}_1 d^{(5)}_{m-1} + \hat{f}_2 e_1 + \hat{f}_3 f_1 ; \quad \tilde{t}_6 = \hat{f}_1 d^{(6)}_{m} + \hat{f}_2 e_2 + \hat{f}_3 f_2 ;$$

with,

$$\hat{a}_1 = (1-\hat{\alpha})(1+p_1-2\hat{\alpha})/(1+p_1+\hat{\alpha}) ;$$

$$\hat{a}_2 = \hat{\alpha}[(1-2\hat{\alpha})(2+p_1-q_1)+2((1+\hat{\alpha})(1-q_1)+p_1\hat{\alpha})]/(1+p_1+\hat{\alpha}) ;$$

$$\hat{a}_3 = 3\hat{\alpha}q_1/(1+p_1+\hat{\alpha}) ;$$

$$\hat{b}_1 = 2(1+p_2)p_1(1-\hat{\alpha})/(1+p_1+\hat{\alpha})(1+p_2+q_1) ;$$

$$\hat{b}_2 = \frac{(1+p_2)[(2+p_1-q_1)\hat{\alpha}p_1+((1+\hat{\alpha})(1-q_1)+p_1\hat{\alpha})(1-p_1)]}{(1+p_1+\hat{\alpha})(1+p_2+q_1)} + \frac{p_2q_1(1+p_3-q_2)}{(1+p_2+q_1)(1+p_3+q_3)};$$

$$\hat{b}_3 = \frac{(1+p_2)q_1(1-p_1+\hat{\alpha})}{(1+p_1+\hat{\alpha})(1+p_2+q_1)} + \frac{(1-p_2)q_1(1+p_3-q_2)}{(1+p_2+q_1)(1+p_3+q_2)} ;$$

$$\hat{b}_4 = 2q_1q_2(1-q_3)/(1+p_2+q_1)(1+p_3+q_2) ;$$

$$\hat{b}_5 = 2q_1q_2q_3/(1+p_2+q_1)(1+p_3+q_2) ;$$

$$\hat{c}_1 = 2p_1p_2(1-\hat{\alpha})/(1+p_1+\hat{\alpha})(1+p_2+q_1) ;$$

$$\hat{c}_2 = \frac{p_2[(2+p_1-q_1)\hat{\alpha}p_1+(1-p_1)((1+\hat{\alpha})(1-q_1)+p_1\hat{\alpha})]}{(1+p_1+\hat{\alpha})(1+p_2+q_1)} + \frac{(1+q_1)p_2(1+p_3-q_2)}{(1+p_2+q_1)(1+p_3+q_2)};$$

$$\hat{c}_3 = \frac{p_2q_1(1+\hat{\alpha}-p_1)}{(1+p_1+\hat{\alpha})(1+p_2+q_1)} + \frac{(1+q_1)(1-p_2)(1+p_3-q_2)}{(1+p_2+q_1)(1+p_3+q_2)} ;$$

$$\hat{c}_4 = \frac{2q_2(1-q_3)(1+q_1)}{(1+p_2+q_1)(1+p_3+q_2)} ;$$

$$\hat{c}_5 = 2q_2q_3(1+q_1)/(1+p_2+q_1)(1+p_3+q_2) ;$$

$$\hat{d}_i^{(1)} = \frac{2p_{i+1}p_{i+2}(1+p_{i+3})}{(1+p_{i+2}+q_{i+1})(1+p_{i+3}+q_{i+2})} \quad i=1,3,\ldots,m-5 ;$$

$$\hat{d}_i^{(2)} = \frac{2(1-p_i)p_{i+1}(1+p_{i+2})}{(1+p_{i+1}+q_i)(1+p_{i+2}+q_{i+1})} \quad i=2,4,\ldots,m-4 ;$$

$$\hat{d}_i^{(3)} = \frac{(1-p_i+q_{i-1})(1+p_{i+1})(1-q_i)}{(1+p_i+q_{i-1})(1+p_{i+1}+q_i)} + \frac{p_{i+1}(1+p_{i+2}-q_{i+1})q_i}{(1+p_{i+1}+q_i)(1+p_{i+2}+q_{i+1})}$$

$$i=3,5,\ldots,m-3;$$

$$\hat{d}_i^{(4)} = \frac{(1-p_{i-1}+q_{i-2})(1+p_i)q_{i-1}}{(1+p_{i-1}+q_{i-2})(1+p_i+q_{i-1})} + \frac{(1-p_i)(1+p_{i+1}-q_i)q_{i-1}}{(1+p_i+q_{i-1})(1+p_{i+1}+q_i)} ,$$

$$i=4,6,\ldots,m-2;$$

$$\hat{d}_i^{(5)} = \frac{2q_{i-2}q_{i-1}(1-q_i)}{(1+p_{i-1}+q_{i-2})(1+p_i+q_{i-1})} \quad i=5,7,\ldots,(m-1) ;$$

$$\hat{d}_i^{(6)} = \frac{2q_{i-3}q_{i-2}q_{i-1}}{(1+p_{i-2}+q_{i-3})(1+p_{i-1}+q_{i-2})} \qquad i=6,8,\ldots,(m-2) \;;$$

$$\hat{e}_i^{(1)} = \frac{2p_{i+1}p_{i+2}p_{i+3}}{(1+p_{i+2}+q_{i+1})(1+p_{i+3}+q_{i+2})} \qquad i=1,3,\ldots,m-5 \;;$$

$$\hat{e}_i^{(2)} = \frac{2(1-p_i)p_{i+1}p_{i+2}}{(1+p_{i+1}+q_i)(1+p_{i+2}+q_{i+1})} \qquad i=2,4,\ldots,m-4 \;;$$

$$\hat{e}_i^{(3)} = \frac{p_{i+1}(1-p_i+q_{i-1})(1-q_i)}{(1+p_i+q_{i-1})(1+p_{i+1}+q_i)} + \frac{p_{i+1}(1+p_{i+2}-q_{i+1})(1+q_i)}{(1+p_{i+1}+q_i)(1+p_{i+2}+q_{i+1})}$$
$$i=3,5,\ldots,(m-3) \;;$$

$$\hat{e}_i^{(4)} = \frac{p_i(1-p_{i-1}+q_{i-2})q_{i-1}}{(1+p_{i-1}+q_{i-2})(1+p_i+q_{i-1})} + \frac{(1-p_i)(1+p_{i+1}-q_i)(1+q_{i-1})}{(1+p_i+q_{i-1})(1+p_{i+1}+q_i)}$$
$$i=4,6,\ldots,(m-2) \;;$$

$$\hat{e}_i^{(5)} = \frac{2(1+q_{i-2})q_{i-1}(1-q_i)}{(1+p_{i-1}+q_{i-2})(1+p_i+q_{i-1})} \qquad i=5,7,\ldots,(m-1) \;;$$

$$\hat{e}_i^{(6)} = \frac{2(1+q_{i-3})q_{i-2}q_{i-1}}{(1+p_{i-2}+q_{i-3})(1+p_{i-1}+q_{i-2})} \qquad i=6,8,\ldots,(m-2) \;;$$

$$\hat{f}_1 = \frac{2p_{m-2}p_{m-1}}{(1+p_{m-1}+q_{m-2})(1+q_{m-1})} \;;$$

$$\hat{f}_2 = \frac{2(1-p_{m-2})p_{m-1}}{(1+p_{m-1}+q_{m-2})(1+q_{m-1})} \;; \text{ and}$$

$$\hat{f}_3 = \frac{(1-p_{m-1}+q_{m-2})(1-q_{m-1})}{(1+p_{m-1}+q_{m-2})(1+q_{m-1})}$$

and all the a_i's, b_i's, $c_i^{(j)}$'s, $d_i^{(j)}$'s, e_i's and f_i's have the same meaning as before.

(a) *For the case $\alpha=2$ (spherical symmetry)*

We note that in this case $\tilde{b}_1 = \tilde{c}_1 = \tilde{d}_1 = \tilde{e}_1 = \tilde{c}_1^{(1)} = \tilde{d}_1^{(1)} = 0$ and therefore, by expanding $|D|$ about the first column we get $(\tilde{a}_1-\mu)|D_1|=0$ where $|D_1|$ is the minor which resulted from the deletion of the first row and column of $|D|$. This leads to

$$\mu = \{\frac{(1-3\lambda)(1-6\lambda)}{(1+3\lambda)}\}^2$$

which for stability gives the following real bound on λ,

$$\lambda \leqslant \frac{2}{3} \quad . \tag{5.5.44}$$

By making the appropriate row operations on the minors $|D_1|$ and $|D_2|$ and then solving the resulting algebraic equations, we arrive at the following stability restrictions,

$$\lambda \lesssim 4.1652 \tag{5.5.45}$$

and

$$\lambda \lesssim 1.41927 \ . \tag{5.5.46}$$

It can be established through very extensive manipulation (the mathematics involved is too long to be reproduced here) that the sum of the moduli of the entries in each of the remaining rows except the last is less than or equal to 1 when $\lambda \leqslant \frac{2}{3}$. For the last row, the absolute sum is strictly less than 1. Hence for overall stability, we deduce from (5.5.44)-(5.5.46) that the (D)AGE scheme is conditionally stable for $\lambda \leqslant \frac{2}{3}$ when $\alpha=2$.

(b) *For the case $\alpha=1$ (cylindrical symmetry)*

To establish stability, we first perform the following set of row operations R_i on the determinant $|D|$:

(i) $R_1 \rightarrow \frac{\tilde{b}_1}{(\tilde{a}_1-\mu)} R_1$,

i.e. $R_1 = (\tilde{b}_1, \frac{\tilde{a}_2\tilde{b}_1}{(\tilde{a}_1-\mu)}, \frac{\tilde{a}_3\tilde{b}_1}{(\tilde{a}_1-\mu)}, \frac{\tilde{a}_4\tilde{b}_1}{(\tilde{a}_1-\mu)}, \frac{\tilde{a}_5\tilde{b}_1}{(\tilde{a}_1-\mu)}, \frac{\tilde{a}_6\tilde{b}_1}{(\tilde{a}_1-\mu)}, 0,\ldots,0)$,

and $|D| \rightarrow \frac{\tilde{b}_1}{(\tilde{a}_1-\mu)} |D|$;

(ii) $R_2 \rightarrow R_2 - R_1$,

i.e., $R_2 = (0,\beta_1,\beta_2,\beta_3,\beta_4,\beta_5,\tilde{b}_7,\tilde{b}_8,0,\ldots,0)$, where $\beta_1 = \tilde{b}_2-\mu - \frac{\tilde{a}_2\tilde{b}_1}{(\tilde{a}_1-\mu)}$,

$\beta_i = \tilde{b}_{i+1} - \frac{\tilde{a}_{i+1}\tilde{b}_1}{(\tilde{a}_1-\mu)}$, for i=2,3,...,5 ;

(iii) $R_1 \to \frac{\tilde{c}_1}{\tilde{b}_1} R_1$,

i.e. $R_1 = (\tilde{c}_1, \frac{\tilde{c}_1\tilde{a}_2}{(\tilde{a}_1-\mu)}, \frac{\tilde{c}_1\tilde{a}_3}{(\tilde{a}_1-\mu)}, \frac{\tilde{c}_1\tilde{a}_4}{(\tilde{a}_1-\mu)}, \frac{\tilde{c}_1\tilde{a}_5}{(\tilde{a}_1-\mu)}, \frac{\tilde{c}_1\tilde{a}_6}{(\tilde{a}_1-\mu)},0,\ldots 0)$,

and $|D| \to \frac{\tilde{c}_1}{(\tilde{a}_1-\mu)} |D|$;

(iv) $R_3 \to R_3-R_1$,

i.e. $R_3 = (0,\beta_6,\beta_7,\beta_8,\beta_9,\beta_{10},\tilde{c}_7,\tilde{c}_8,0,\ldots,0)$,

where $\beta_{i+4} = \tilde{c}_i - \frac{\tilde{c}_1\tilde{a}_i}{(\tilde{a}_1-\mu)}$, for i=2(1)6, i≠3 ,

for i=3, $\beta_7 = \tilde{c}_3-\mu- \frac{\tilde{c}_1\tilde{a}_3}{(\tilde{a}_1-\mu)}$;

(v) $R_1 \to \frac{\tilde{d}_1}{\tilde{c}_1}R_1$,

i.e. $R_1 = (\tilde{d}_1, \frac{\tilde{d}_1\tilde{a}_2}{(\tilde{a}_1-\mu)}, \frac{\tilde{d}_1\tilde{a}_3}{(\tilde{a}_1-\mu)}, \frac{\tilde{d}_1\tilde{a}_4}{(\tilde{a}_1-\mu)}, \frac{\tilde{d}_1\tilde{a}_5}{(\tilde{a}_1-\mu)}, \frac{\tilde{d}_1\tilde{a}_6}{(\tilde{a}_1-\mu)},0,\ldots,0)$,

and $|D| \to \frac{\tilde{d}_1}{(\tilde{a}_1-\mu)} |D|$;

(vi) $R_4 \to R_4-R_1$,

i.e., $R_4 = (0,\beta_{11},\beta_{12},\beta_{13},\beta_{14},\beta_{15},\tilde{d}_7,\tilde{d}_8,\tilde{d}_9,\tilde{d}_{10},0,\ldots,0)$,

where $\beta_{i+9} = \tilde{d}_i - \frac{\tilde{d}_1\tilde{a}_i}{(\tilde{a}_1-\mu)}$ for i=2(1)6 i≠4 ,

for i=4, $\beta_{13} = \tilde{d}_4-\mu - \frac{\tilde{d}_1\tilde{a}_4}{(\tilde{a}_1-\mu)}$;

(vii) $R_1 \to \frac{\tilde{e}_1}{\tilde{d}_1}R_1$,

i.e. $R_1 = (\tilde{e}_1, \frac{\tilde{e}_1\tilde{a}_2}{(\tilde{a}_1-\mu)}, \frac{\tilde{e}_1\tilde{a}_3}{(\tilde{a}_1-\mu)}, \frac{\tilde{e}_1\tilde{a}_4}{(\tilde{a}_1-\mu)}, \frac{\tilde{e}_1\tilde{a}_5}{(\tilde{a}_1-\mu)}, \frac{\tilde{e}_1\tilde{a}_6}{(\tilde{a}_1-\mu)},0,\ldots,0)$,

and $|D| \to \frac{\tilde{e}_1}{(\tilde{a}_1-\mu)}|D|$;

(viii) $R_5 \to R_5 - R_1$,

i.e. $R_5 = (0,\beta_{16},\beta_{17},\beta_{18},\beta_{19},\beta_{20},\tilde{e}_7,\tilde{e}_8,\tilde{e}_9,\tilde{e}_{10},0,\ldots,0)$,

where $\beta_{i+14} = \tilde{e}_i - \frac{\tilde{e}_1\tilde{a}_i}{(\tilde{a}_1-\mu)}$ for i=2(1)6 i≠5 ,

for i=5, $\beta_{19} = \tilde{e}_5-\mu - \frac{\tilde{e}_1\tilde{a}_5}{(a_1-\mu)}$;

(ix) $R_1 \to \frac{\tilde{c}_1^{(1)}}{\tilde{e}_1} R_1$,

i.e., $R_1 = (\tilde{c}_1^{(1)}, \frac{\tilde{c}_1^{(1)}\tilde{a}_2}{(\tilde{a}_1-\mu)}, \frac{\tilde{c}_1^{(1)}\tilde{a}_3}{(\tilde{a}_1-\mu)}, \frac{\tilde{c}_1^{(1)}\tilde{a}_4}{(\tilde{a}_1-\mu)}, \frac{\tilde{c}_1^{(1)}\tilde{a}_5}{(\tilde{a}_1-\mu)}, \frac{\tilde{c}_1^{(1)}\tilde{a}_6}{(\tilde{a}_1-\mu)},0,\ldots,0)$,

and $|D| \to \frac{\tilde{c}_1^{(1)}}{(\tilde{a}_1-\mu)}|D|$;

(x) $R_6 \to R_6 - R_1$,

i.e. $R_6 = (0,\beta_{21},\beta_{22},\beta_{23},\beta_{24},\beta_{25},\tilde{c}_7^{(7)},\tilde{c}_8^{(8)},\tilde{c}_9^{(9)},\tilde{c}_{10}^{(10)},\tilde{c}_{11}^{(11)},\tilde{c}_{12}^{(12)},$
$0,\ldots,0)$,

where $\beta_{i+19} = \tilde{c}_i^{(i)} - \frac{\tilde{c}_1^{(1)}\tilde{a}_i}{(\tilde{a}_1-\mu)}$ for i=2(1)5 ,

and $\beta_{25} = \tilde{c}_6^{(6)} - \mu - \frac{\tilde{c}_1^{(1)}\tilde{a}_6}{(\tilde{a}_1-\mu)}$;

(xi) $R_1 \to \frac{\tilde{d}_1^{(1)}}{\tilde{c}_1^{(1)}} R_1$,

i.e. $R_1 = (\tilde{d}_1^{(1)}, \frac{\tilde{d}_1^{(1)}\tilde{a}_2}{(\tilde{a}_1-\mu)}, \frac{\tilde{d}_1^{(1)}\tilde{a}_3}{(\tilde{a}_1-\mu)}, \frac{\tilde{d}_1^{(1)}\tilde{a}_4}{(\tilde{a}_1-\mu)}, \frac{\tilde{d}_1^{(1)}\tilde{a}_5}{(\tilde{a}_1-\mu)}, \frac{\tilde{d}_1^{(1)}\tilde{a}_6}{(\tilde{a}_1-\mu)},0,\ldots,0)$,

and $|D| \to \frac{\tilde{d}_1^{(1)}}{(\tilde{a}_1-\mu)}|D|$;

(xii) $R_7 \to R_7 - R_1$,

i.e. $R_7 = (0,\beta_{26},\beta_{27},\beta_{28},\beta_{29},\beta_{30},\tilde{d}^{(7)}-\mu,\tilde{d}_8^{(8)},\tilde{d}_9^{(9)},\tilde{d}_{10}^{(10)},\tilde{d}_{11}^{(11)},$
$\tilde{d}_{12}^{(12)},0,\ldots,0)$,

where $\beta_{i+24} = \tilde{d}_i^{(i)} - \frac{\tilde{d}_1^{(1)}\tilde{a}_i}{(\tilde{a}_1-\mu)}$ for i=2(1)(6).

Hence we have,

$$|D_1| = \frac{\tilde{d}_1^{(1)}}{(\tilde{a}_1-\mu)}|D|,$$

or $$|D| = \frac{(\tilde{a}_1-\mu)}{\tilde{d}_1^{(1)}}|D_1|.$$

On expanding the determinant $|D|=0$ about the first column, we find that $(\tilde{a}_1-\mu)|D_2|=0$ where $|D_2|$ is the minor derived from $|D|$ by deleting the first row and column. Thus, $\mu=\tilde{a}_1$ and $|\mu|$ attains its maximum value of 0.181722393 when $\lambda \simeq 1.0006$ which implies that for the stability of the first row of the determinantal equation, we require,

$$\lambda \leq 1.0006 \ . \tag{5.5.47}$$

By similarly manipulating the other minors, we are led to the following restrictions on the mesh ratio:

$$\lambda \leq 1.54745 \tag{5.5.48}$$

and $$\lambda \leq 1.40682 \ , \tag{5.5.49}$$

and repeating the same procedure for the remaining rows and minors results in algebraic equations with $(\tilde{a}_1-\mu)$ as their multiple factors. Again, for the stability of these rows of equations, we must have $\lambda \leq 1.0006$. To achieve overall stability, we therefore conclude from (5.5.47)-(5.5.49) that $\lambda \leq 1.0006$.

5.6 NUMERICAL EXAMPLES AND COMPARATIVE RESULTS

Two numerical experiments were carried out to demonstrate the implementation of the GE schemes on parabolic problems with special geometries.

Experiment 1

This experiment dealt with the solution of the following parabolic problem with cylindrical symmetry (Mitchell and Pearce (1963)),

$$\frac{\partial U}{\partial t} = \frac{1}{r}\frac{\partial U}{\partial r} + \frac{\partial^2 U}{\partial r^2} \quad , \ (0 \leq r \leq 1), \tag{5.6.1}$$

given the auxiliary conditions,

$$U(r,0) = J_0(\beta r),\ 0 \leq r \leq 1,$$

$$\frac{\partial U}{\partial r}(0,t) = 0\ ,\ t>0\ ,$$

and $$U(1,t) = 0,\quad t>0\ , \tag{5.6.1a}$$

where $J_0(\beta r)$ is *the Bessel function* of the first kind of order 0 and β is the first root of $J_0(\beta)=0$. The exact solution is

$$U(r,t) = J_0(\beta r)e^{-\beta^2 t} \tag{5.6.2}$$

and the values of the Bessel function at the grid points are generated using the NAG library subroutine (the first four roots of $J_0(\beta)=0$ are $\beta_1=2.405$, $\beta_2=5.520$, $\beta_3=8.654$, $\beta_4=11.79$). In Tables 5.6.1-5.6.3 are displayed a comparison of the numerical solutions of the GE schemes with the exact solutions at the appropriate grid points in terms of their absolute errors for various values of the mesh ratio λ. The absolute errors of the solutions of the explicit (EXP) scheme of (3.13.25) and (3.13.33) and the Crank-Nicolson (CN) method of (3.14.10) and (3.14.12) are also included.

Experiment 2

The following parabolic problem with spherical symmetry (Saulev (1964)) is considered,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial r^2} + \frac{2}{r}\frac{\partial U}{\partial r} + k(r,t) \ , \tag{5.6.3}$$

$$(k(r,t) = e^{-t}\{[6+(1-r^2)\pi^2 t^2-(1-r^2)]\cos(\pi rt)-[(1-r^2)r+4rt-2t(1-r^2)/r]$$
$$\pi\sin(\pi rt)\})$$

subject to the initial-boundary conditions,

$$U(r,0) = 1-r^2 \ ,$$

$$\frac{\partial U}{\partial r}(0,t) = 0 \ , \tag{5.6.3a}$$

and $\quad U(1,t) = 0,$

with the exact solution,

$$U(r,t) = (1-r^2)e^{-t}\cos(\pi rt) \ . \tag{5.6.4}$$

Since our parabolic equation incorporates a source term $k(r,t)$, some modifications on the basic equations governing the GE schemes are therefore required. For general grouped points not on the axis, instead of equations (5.3.6) and (5.3.7) we have,

$$u_{i-1,j+1} = \{(1+p_i)p_{i-1}u_{i-2,j}+(1+p_i)(1-p_{i-1})u_{i-1,j}+q_{i-1}(1-q_i)u_{ij}+q_iq_{i-1}$$
$$u_{i+1,j}+\Delta t[(1+p_i)k_{i-1,j+\frac{1}{2}}+q_{i-1}k_{i,j+\frac{1}{2}}]\}/(1+p_i+q_{i-1}) \tag{5.6.5}$$

and

$$u_{i,j+1} = \{p_ip_{i-1}u_{i-2,j}+p_i(1-p_{i-1})u_{i-1,j}+(1+q_{i-1})(1-q_i \quad)u_{ij}+(1+q_{i-1})$$
$$q_iu_{i+1,j}+\Delta t[p_ik_{i-1,j+\frac{1}{2}}+(1+q_{i-1})k_{i,j+\frac{1}{2}}]\}/(1+p_i+q_{i-1}) . \tag{5.6.6}$$

For the GER method, the equations (5.3.14) and (5.3.15) are now replaced by,

$$u_{0,j+1} = (1-2\hat{\alpha})u_{0j} + 2\hat{\alpha}u_{1j} + \Delta tk_{0,j+\frac{1}{2}} \ , \tag{5.6.7}$$

and $$u_{m-1,j+1} = (p_{m-1}u_{m-2,j}+(1-p_{m-1})u_{m-1,j}+q_{m-1}u_{m,j+1}+\Delta tk_{m-1,j+\frac{1}{2}})/(1+q_{m-1})$$
$$(5.6.8)$$

respectively. Similarly, the equations defining the left boundary as well as the point immediately next to it for the GEL scheme become (replacing equations (5.3.27) and (5.3.28)),

$$u_{0,j+1} = \{(1+p_1)(1-\hat{\alpha})u_{0j}+(2+p_1-q_1)\hat{\alpha}u_{1j}+\hat{\alpha}q_1u_{2j}+\Delta t[(1+p_1)k_{0,j+\frac{1}{2}}+\hat{\alpha}k_{1,j+\frac{1}{2}}]\}$$

$$/(1+p_1+\hat{\alpha}) \qquad (5.6.9)$$

and

$$u_{1,j+1} = \{p_1(1-\hat{\alpha})u_{0j}+((1+\hat{\alpha})(1-q_1)+p_1\hat{\alpha})u_{1j}+(1+\hat{\alpha})q_1u_{2j}+\Delta t[p_1k_{0,j+\frac{1}{2}}+(1+\hat{\alpha})$$

$$k_{1,j+\frac{1}{2}}]\}/(1+p_1+\hat{\alpha}). \qquad (5.6.10)$$

The numerical solutions of the above spherical problem using the GE schemes are obtained for various values of λ and to indicate their accuracy, Tables (5.6.4)-(5.6.6) provide a comparison with the exact solution in terms of their absolute errors.

It is observed that, presumably due to the term $\frac{\alpha}{r}\frac{\partial U}{\partial r}$ in (5.6.1) and (5.6.3) (with α=1 and 2 respectively), the solutions of the GE schemes are slightly less accurate in the vicinity of the axis (r=0) than in the remainder of the field. As we have already seen, special equations have to be formulated to cope with this difficulty at the point of singularity. In fact, an examination of the truncation errors of the GE schemes at r=0 indicates the presence of the term $\frac{\Delta t}{\Delta r}$ (equations (5.4.1) and (5.4.5)) and it is therefore essential that to attain consistency, Δt approaches 0 faster than does Δr. It is also found for our cylindrical problem that the GE class of methods are more accurate than the other schemes under investigation. From Table 5.6.2, the (S)AGE and (D)AGE methods in that order are more superior whilst in Table 5.6.3, (D)AGE has the edge on other difference formulae. This is to be expected since the truncation error expressions of the constituent GER and GEL formulae possess terms of different signs and hence the correct alternate applications of these formulae to constitute the (S)AGE and (D)AGE schemes can lead to cancellations of error terms. The same observation also applies for our spherical problem although comparative results from other schemes are not available. However, it suffices to say that among the GE class of methods, the (S)AGE method gives a more satisfactory result.

We conclude that despite the limited stability of the GE schemes, their stability ratios are not so restricted as to be impractical for implementation for special geometries when compared with other schemes. Being explicit, they are simple and incur low computational load and above all exhibit better accuracy. The use of the alternating schemes in particular, i.e. (S)AGE and (D)AGE is highly recommended.

$t=0.175$, $\lambda=0.175$, $\Delta t=0.00175$, $\Delta r=0.1$

r Method	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors
GER	2.15×10^{-3}	2.16×10^{-3}	2.0×10^{-3}	2.01×10^{-3}	1.66×10^{-3}	1.65×10^{-3}	1.17×10^{-3}	1.15×10^{-3}	6.04×10^{-4}	5.67×10^{-4}	1.51×10^{-3}
GEL	1.9×10^{-3}	1.88×10^{-3}	1.89×10^{-3}	1.63×10^{-3}	1.62×10^{-3}	1.19×10^{-3}	1.17×10^{-3}	6.34×10^{-4}	6.07×10^{-4}	6.32×10^{-5}	1.26×10^{-3}
(S)AGE	1.43×10^{-3}	1.42×10^{-3}	1.4×10^{-3}	1.27×10^{-3}	1.21×10^{-3}	9.96×10^{-4}	9.07×10^{-4}	6.34×10^{-4}	5.24×10^{-4}	2.32×10^{-4}	1.0×10^{-3}
(D)AGE	1.47×10^{-3}	1.49×10^{-3}	1.42×10^{-3}	1.36×10^{-3}	1.2×10^{-3}	1.09×10^{-3}	8.66×10^{-4}	7.23×10^{-4}	4.6×10^{-4}	2.99×10^{-4}	1.04×10^{-3}
EXP	3.19×10^{-4}	2.73×10^{-4}	2.54×10^{-4}	2.37×10^{-4}	2.19×10^{-4}	1.99×10^{-4}	1.75×10^{-4}	1.47×10^{-4}	1.15×10^{-4}	7.64×10^{-5}	2.01×10^{-4}
C-N	2.06×10^{-3}	2.04×10^{-3}	1.96×10^{-3}	1.83×10^{-3}	1.64×10^{-3}	1.42×10^{-3}	1.17×10^{-3}	8.95×10^{-4}	6.07×10^{-4}	3.16×10^{-4}	1.39×10^{-3}
EXACT SOLUTION	0.3634169	0.3581809	0.3426988	0.3176383	0.2840764	0.2434480	0.1974773	0.1480959	0.0973516	0.0473117	-

TABLE 5.6.1: Absolute errors of the numerical solutions to the cylindrical problem

t=0.6, λ=0.3, Δt=0.003, Δr=0.1

r Method	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors
GER	7.07×10^{-4}	7.02×10^{-4}	6.62×10^{-4}	6.35×10^{-4}	5.47×10^{-4}	5.03×10^{-4}	3.82×10^{-4}	3.26×10^{-4}	1.9×10^{-4}	1.32×10^{-4}	4.79×10^{-4}
GEL	5.29×10^{-4}	5.22×10^{-4}	5.12×10^{-4}	4.56×10^{-4}	4.31×10^{-4}	3.38×10^{-4}	3.04×10^{-4}	1.91×10^{-4}	1.53×10^{-4}	3.84×10^{-5}	3.47×10^{-4}
(S)AGE	1.07×10^{-4}	1.04×10^{-4}	1.09×10^{-4}	9.18×10^{-5}	9.93×10^{-5}	7.15×10^{-5}	8.06×10^{-5}	4.57×10^{-5}	5.42×10^{-5}	1.72×10^{-5}	7.8×10^{-5}
(D)AGE	2.2×10^{-4}	2.27×10^{-4}	2.12×10^{-4}	2.08×10^{-4}	1.76×10^{-4}	1.68×10^{-4}	1.23×10^{-4}	1.12×10^{-4}	6.06×10^{-5}	4.88×10^{-5}	1.56×10^{-4}
EXP	2.89×10^{-4}	2.91×10^{-4}	2.8×10^{-4}	2.6×10^{-4}	2.32×10^{-4}	1.98×10^{-4}	1.6×10^{-4}	1.19×10^{-4}	7.65×10^{-5}	3.54×10^{-5}	1.94×10^{-4}
C-N	6.2×10^{-4}	6.12×10^{-4}	5.86×10^{-4}	5.44×10^{-4}	4.88×10^{-4}	4.19×10^{-4}	3.42×10^{-4}	2.58×10^{-4}	1.71×10^{-4}	8.49×10^{-5}	4.12×10^{-4}
EXACT SOLUTION	0.0311041	0.030656	0.0293309	0.0271860	0.0243135	0.0208362	0.0169017	0.0126752	0.0083321	0.0040493	-

TABLE 5.6.2: Absolute errors of the numerical solutions to the cylindrical problem

$t=0.6,\ \lambda=0.6,\ \Delta t=0.006,\ \Delta r=0.1$

r Method	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors
(S)AGE	1.39×10^{-3}	1.39×10^{-3}	1.3×10^{-3}	1.25×10^{-3}	1.05×10^{-3}	9.58×10^{-4}	6.93×10^{-4}	5.82×10^{-4}	2.93×10^{-4}	1.84×10^{-4}	9.09×10^{-4}
(D)AGE	2.88×10^{-4}	2.72×10^{-4}	2.86×10^{-4}	2.18×10^{-4}	2.5×10^{-4}	1.43×10^{-4}	1.87×10^{-4}	6.33×10^{-5}	1.2×10^{-4}	2.35×10^{-5}	1.85×10^{-4}
C-N	9.75×10^{-4}	9.63×10^{-4}	9.23×10^{-4}	8.57×10^{-4}	7.69×10^{-4}	6.62×10^{-4}	5.4×10^{-4}	4.08×10^{-4}	2.71×10^{-4}	1.35×10^{-4}	6.50×10^{-4}
EXACT SOLUTION	0.0311041	0.030656	0.0293309	0.027186	0.0243135	0.0208362	0.0169017	0.0126752	0.0083321	0.0040493	-

TABLE 5.6.3: Absolute errors of the numerical solutions to the cylindrical problem

t=0.175, λ=0.175, Δt=0.00175, Δr=0.1

r Method	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors
GER	7.86×10^{-4}	7.88×10^{-4}	7.91×10^{-4}	7.64×10^{-4}	7.5×10^{-4}	7.23×10^{-4}	6.41×10^{-4}	6.45×10^{-4}	4.15×10^{-4}	4.9×10^{-4}	6.79×10^{-4}
GEL	6.23×10^{-4}	6.27×10^{-4}	6.01×10^{-4}	5.89×10^{-4}	5.43×10^{-4}	4.84×10^{-4}	4.45×10^{-4}	2.75×10^{-4}	2.82×10^{-4}	9.27×10^{-5}	4.56×10^{-4}
(S)AGE	5.53×10^{-4}	5.57×10^{-4}	5.37×10^{-4}	5.15×10^{-4}	4.78×10^{-4}	4.3×10^{-4}	3.91×10^{-4}	3.0×10^{-4}	2.77×10^{-4}	1.16×10^{-4}	4.16×10^{-4}
(D)AGE	5.73×10^{-4}	5.69×10^{-4}	5.56×10^{-4}	5.29×10^{-4}	4.94×10^{-4}	4.52×10^{-4}	3.88×10^{-4}	3.42×10^{-4}	2.31×10^{-4}	1.85×10^{-4}	4.32×10^{-4}
EXACT SOLUTION	0.839457	0.8298068	0.8010120	0.7535391	0.6881618	0.6059549	0.5082867	0.3968087	0.2734428	0.1403673	-

TABLE 5.6.4: Absolute errors of the numerical solutions to the spherical problem

$t=0.6,\ \lambda=0.3,\ \Delta t=0.003,\ \Delta r=0.1$

r Method	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors
GER	1.04×10^{-2}	1.03×10^{-2}	9.88×10^{-3}	8.97×10^{-3}	7.99×10^{-3}	6.66×10^{-3}	5.24×10^{-3}	3.91×10^{-3}	2.28×10^{-3}	1.39×10^{-3}	6.71×10^{-3}
GEL	9.85×10^{-3}	9.77×10^{-3}	9.19×10^{-3}	8.48×10^{-3}	7.32×10^{-3}	6.11×10^{-3}	4.74×10^{-3}	3.19×10^{-3}	2.06×10^{-3}	5.12×10^{-4}	6.12×10^{-3}
(S)AGE	9.71×10^{-3}	9.64×10^{-3}	9.04×10^{-3}	8.23×10^{-3}	7.06×10^{-3}	5.79×10^{-3}	4.44×10^{-3}	3.0×10^{-3}	1.92×10^{-3}	7.11×10^{-4}	5.96×10^{-3}
(D)AGE	9.87×10^{-3}	9.7×10^{-3}	9.2×10^{-3}	8.3×10^{-3}	7.23×10^{-3}	5.92×10^{-3}	4.51×10^{-3}	3.2×10^{-3}	1.8×10^{-3}	9.28×10^{-4}	6.07×10^{-3}
EXACT SOLUTION	0.5488116	0.5336998	0.4898613	0.4216731	0.3360558	0.2419376	0.1495505	0.0696068	0.0124057	-0.013069	-

TABLE 5.6.5: Absolute errors of the numerical solutions to the spherical problem

$t=0.6$, $\lambda=0.6$, $\Delta t=0.006$, $\Delta r=0.1$

r Method	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average o all absolute errors
(S)AGE	7.91×10^{-3}	8.45×10^{-3}	7.55×10^{-3}	6.74×10^{-3}	5.28×10^{-3}	4.01×10^{-3}	2.79×10^{-3}	1.37×10^{-3}	1.17×10^{-3}	1.35×10^{-5}	4.53×10^{-3}
(D)AGE	9.01×10^{-3}	9.16×10^{-3}	8.8×10^{-3}	7.64×10^{-3}	6.69×10^{-3}	5.24×10^{-3}	3.82×10^{-3}	2.73×10^{-3}	9.33×10^{-4}	1.12×10^{-3}	5.52×10^{-3}
EXACT SOLUTION	0.5488116	0.5336998	0.4898613	0.4216731	0.3360558	0.2419376	0.1495505	0.0696068	0.0124057	-0.013068	-

TABLE 5.6.6: Absolute errors of the numerical solutions to the spherical problem

CHAPTER SIX

THE ALTERNATING GROUP EXPLICIT ITERATIVE METHOD TO SOLVE PARABOLIC AND HYPERBOLIC PARTIAL DIFFERENTIAL EQUATIONS

6.1 INTRODUCTION

As we have already seen in Chapter 3, the ADI method was developed to deal with two-dimensional parabolic (and elliptic) problems and the solutions were obtained *implicitly* in the horizontal and vertical directions. The method could then be extended for applications to higher-dimensional problems. Thus we find that the method has no analogue for the one-dimensional case.

We will, however, show that it is possible to derive another method, the analysis of which is analogous to the ADI scheme. Initially we present the method for one-dimensional problems and then extend its implementation to higher dimensional ones. This *iterative* method employs the fractional splitting strategy which is applied alternately at each half (intermediate) iteration time step on tridiagonal systems of difference schemes and which has proved to be stable. Its rate of convergence is governed by the acceleration parameter r. The accuracy of this method is, in general, comparable if not better than that of the GE class of problems as well as other existing schemes.

6.2 THE ALTERNATING GROUP EXPLICIT METHOD TO SOLVE SECOND ORDER PARABOLIC EQUATIONS WITH DIRICHLET BOUNDARY CONDITIONS

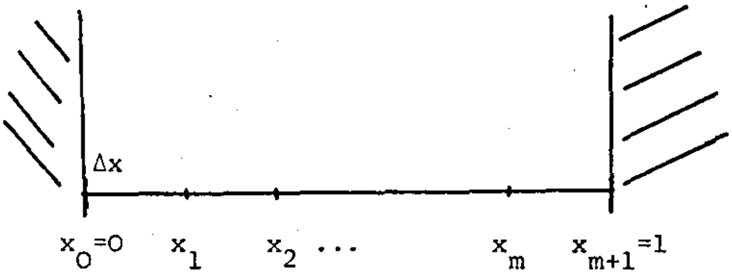
Consider the following second order parabolic equation,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2}\ ,\ 0\leq x\leq 1,\ 0<t\leq T \tag{6.2.1}$$

subject to the initial-boundary conditions

$$\begin{aligned} U(x,0) &= f(x),\ 0<x<1\ , \\ U(0,t) &= g(t),\ 0<t\leq T \qquad\qquad (6.2.1a) \\ \text{and}\qquad U(1,t) &= h(t)\ . \end{aligned}$$

A uniformly-spaced network whose mesh points are $x_i=i\Delta x$, $t_j=j\Delta t$ for $i=0,1,\ldots,m,m+1$ and $j=0,1,\ldots,n,n+1$ is used with $\Delta x=\frac{1}{(m+1)}$, $\Delta t=\frac{T}{(n+1)}$ and $\lambda = \frac{\Delta t}{(\Delta x)^2}$, the mesh ratio. The real line $0\leq x\leq 1$ is thus divided as illustrated,



From equation (3.4.4), a weighted approximation to the differential equation (6.2.1) at the point $(x_i,t_{j+\frac{1}{2}})$ is given by

$$-\lambda\theta u_{i-1,j+1}+(1+2\lambda\theta)u_{i,j+1}-\lambda\theta u_{i+1,j+1} = \lambda(1-\theta)u_{i-1,j}+(1-2\lambda(1-\theta))u_{ij}+ \lambda(1-\theta)u_{i+1,j}\ ,\ i=1,2,\ldots,m. \tag{6.2.2}$$

This approximation can be displayed in a more compact matrix form as

$$\begin{bmatrix} a & b & & & & \\ c & a & b & & \bigcirc & \\ & c & a & b & & \\ & & \ddots & \ddots & \ddots & \\ & & & c & a & b \\ & \bigcirc & & & c & a \end{bmatrix}_{(m\times m)} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_{m-1} \\ u_m \end{bmatrix}_{j+1} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ \vdots \\ f_{m-1} \\ f_m \end{bmatrix} \tag{6.2.3}$$

i.e., $$A\underline{u} = \underline{f} \, , \tag{6.2.4}$$

where,

$$\left.\begin{aligned}
c &= -\lambda\theta \; , \; a = (1+2\lambda\theta) \, , \; b = -\lambda\theta ; \\
f_1 &= \lambda(1-\theta)(u_{0j}+u_{2j})+\lambda\theta u_{0,j+1}+(1-2\lambda(1-\theta))u_{1,j} \; ; \\
f_i &= \lambda(1-\theta)(u_{i-1,j}+u_{i+1,j})+(1-2\lambda(1-\theta))u_{ij} \; , \; i=2,3,\ldots,m-2,m-1; \\
f_m &= \lambda(1-\theta)(u_{m-1,j}+u_{m+1,j})+(1-2\lambda(1-\theta))u_{mj}+\lambda(1-\theta)u_{m+1,j}+ \\
&\qquad\qquad\qquad \lambda\theta u_{m+1,j+1} \text{ and} \\
\underline{u} &= (u_{1,j+1},u_{2,j+1},\ldots,u_{m,j+1})^T \text{ and } \underline{f} = (f_1,f_2,\ldots,f_m)^T .
\end{aligned}\right\} \tag{6.2.4a}$$

We note that $\underline{f}$ is a column vector of order m consisting of the boundary values as well as known u values at time level j while $\underline{u}$ are the values at time level (j+1) which we seek. We also recall that (6.2.4) corresponds to the fully implicit, the Crank-Nicolson, the Douglas and the classical explicit methods when θ takes the values $1, \frac{1}{2}, \frac{1}{2} - \frac{1}{12\lambda}$ and 0 with accuracies of the order $O([\Delta x]^2+\Delta t)$, $O([\Delta x]^2+[\Delta t]^2)$, $O([\Delta x]^4+[\Delta t]^2)$ and $O([\Delta x]^2+\Delta t)$ respectively.

Let us first assume that we have an *even number of intervals* (corresponding to an odd number of internal points, i.e. m odd) on the real line $0 \leq x \leq 1$. We can then perform the following splitting of the coefficient matrix A:

$$A = G_1 + G_2 \quad , \tag{6.2.5}$$

where,

$$G_1 = \left[\begin{array}{c|cc|cc|c|cc}
a/2 & & & & & & & \\ \hline
 & a/2 & b & & & & & \\
 & c & a/2 & & & & & \\ \hline
 & & & a/2 & b & & & \\
 & & & c & a/2 & & O & \\ \hline
 & & & & & \ddots & & \\ \hline
 & & & O & & & a/2 & b \\
 & & & & & & c & a/2
\end{array}\right]_{(m\times m)} \tag{6.2.6}$$

and

$$G_2=\left[\begin{array}{cc|cc|c|cc|c}
a/2 & b & & & & & & \\
c & a/2 & & & & & & \\ \hline
 & & a/2 & b & & & & \\
 & & c & a/2 & & & & \\ \hline
 & & & & \ddots & & O & \\ \hline
 & & & O & & a/2 & b & \\
 & & & & & c & a/2 & \\ \hline
 & & & & & & & a/2
\end{array}\right]_{(m\times m)} . \tag{6.2.7}$$

It is assumed that the following conditions are satisfied:

(i) G_1+rI and G_2+rI are non-singular for any $r>0$,

(ii) for any vectors $\underline{f}_1$ and $\underline{f}_2$ and for any $r>0$, the systems

$$\begin{aligned} &(G_1+rI)\underline{u}_1 = \underline{f}_1 \\ \text{and}\qquad &(G_2+rI)\underline{u}_2 = \underline{f}_2 \end{aligned} \tag{6.2.8}$$

are more easily solved in explicit form since they consist of only (2×2) subsystems.

Thus, with these conditions, system (6.2.4) becomes

$$(G_1+G_2)\underline{u} = \underline{f} \ . \tag{6.2.9}$$

The Alternating Group Explicit (AGE) iteration consists of writing (6.2.9) as a pair of equations,

$$\begin{aligned} &(G_1+rI)\underline{u} = (rI-G_2)\underline{u}+\underline{f} \\ \text{and}\qquad &(G_2+rI)\underline{u} = (rI-G_1)\underline{u}+\underline{f} \ . \end{aligned} \tag{6.2.10}$$

The AGE method using *the Peaceman and Rachford variant* (cf. (3.19.12)) for the *stationary case* (r=constant) is given by,

$$\begin{aligned} &(G_1+rI)\underline{u}^{(p+\frac{1}{2})} = (rI-G_2)\underline{u}^{(p)}+\underline{f} \\ \text{and}\qquad &(G_2+rI)\underline{u}^{(p+1)} = (rI-G_1)\underline{u}^{(p+\frac{1}{2})}+\underline{f} \ , \ p\geq 0 \ , \end{aligned} \tag{6.2.11}$$

where $\underline{u}^{(0)}$ is a starting approximation and r are positive constants called acceleration parameters whose values are chosen to maximise the rate of convergence. We now seek to analyse the convergence properties of the AGE method. From (6.2.11) we can write,

$$\underline{u}^{(p+1)} = M(r)\underline{u}^{(p)} + \underline{q}(r)\ ,\ p\geq 0\ , \qquad (6.2.12)$$

where *the AGE iteration matrix* is given by,

$$M(r) = (G_2+rI)^{-1}(rI-G_1)(G_1+rI)^{-1}(rI-G_2)\ . \qquad (6.2.13)$$

If $\underline{e}$ denotes the error vector and $\underline{U}$ the exact solution of (6.2.1) then $\underline{e}^{(p)} = \underline{u}^{(p)}-\underline{U}$ and $\underline{e}^{(p+1)} = M(r)\underline{e}^{(p)}$. Hence, we have,

$$\underline{e}^{(p)} = M^p(r)\underline{e}^{(0)},\ p\geq 1\ . \qquad (6.2.14)$$

We now prove,

Theorem 6.1

If G_1 and G_2 are real positive definite matrices and if $r>0$ then

$$\rho(M(r)) < 1\ . \qquad (6.2.15)$$

Proof:

If we let $\tilde{M}(r) = (G_2+rI)M(r)(G_2+rI)^{-1}$, then by a similarity transformation, $M(r)$ and $\tilde{M}(r)$ have the same eigenvalues. Hence, from (6.2.13) we find that,

$$\begin{aligned}\rho(M(r)) &= \rho(\tilde{M}(r))\\ &\leq ||\tilde{M}(r)||\\ &\leq ||(rI-G_1)(G_1+rI)^{-1}||\ ||(rI-G_2)(G_2+rI)^{-1}|| \qquad (6.2.16)\end{aligned}$$

where $\rho(M(r))$ is the spectral radius of $M(r)$. But since G_1 and G_2 are symmetric and $(rI-G_1)$ commutes with $(G_1+rI)^{-1}$, then in the L_2 norm we have

$$\begin{aligned}||(rI-G_1)(G_1+rI)^{-1}||_2 &= \rho((rI-G_1)(G_1+rI)^{-1})\\ &= \max_{1\leq i\leq m}\left|\frac{r-\mu_i}{r+\mu_i}\right|\ , \qquad (6.2.17)\end{aligned}$$

where μ_i are the eigenvalues of G_1. But since G_1 is positive definite, its eigenvalues are positive. Therefore,

$$||(rI-G_1)(G_1+rI)^{-1}||_2 < 1 \ . \tag{6.2.18}$$

Similarly,

$$||(rI-G_2)(G_2+rI)^{-1}||_2 < 1$$

and we have,

$$\rho(M(r)) = \rho(\tilde{M}(r)) \leqslant ||M(r)||_2 < 1 \ , \tag{6.2.19}$$

and convergence is assured. We note that to establish the condition (6.2.15) for unsymmetric matrices G_1 and G_2 may require us to evaluate directly the eigenvalues of M(r) which can be difficult.

It is possible to determine the optimum parameter $\hat{r}$ such that the bound for $\rho(M(r))$ is minimised. To investigate this we assume that G_1 and G_2 are real positive definite matrices and that bounds for their eigenvalues μ_i and η_i are available, i.e.,

$$0 < \alpha \leqslant \ \mu_i,\eta_i \leqslant \beta \ . \tag{6.2.20}$$

In the L_2 norm, (6.2.16) implies

$$\begin{aligned} \rho(M(r) &\leqslant \rho((rI-G_1)(G_1+rI)^{-1})\rho((rI-G_2)(G_2+rI)^{-1}) \\ &= \left\{\max_{1\leqslant i\leqslant m}\left|\frac{r-\mu_i}{r+\mu_i}\right|\right\}\left\{\max_{1\leqslant i\leqslant m}\left|\frac{r-\eta_i}{r+\eta_i}\right|\right\} \\ &\leqslant \left\{\max_{\alpha\leqslant z\leqslant\beta}\left|\frac{r-z}{r+z}\right|\right\}^2 = \phi(\alpha,\beta;r) \ . \end{aligned} \tag{6.2.21}$$

But (r-z)/(r+z) is an increasing function of z. Therefore, we find that,

$$\max_{\alpha\leqslant z\leqslant\beta}\left|\frac{r-z}{r+z}\right| = \max\left(\left|\frac{r-\alpha}{r+\alpha}\right|,\left|\frac{r-\beta}{r+\beta}\right|\right) \ . \tag{6.2.22}$$

When $r=\sqrt{\alpha\beta}$, we have,

$$\left|\frac{r-\alpha}{r+\alpha}\right| = \left|\frac{r-\beta}{r+\beta}\right| = \frac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\alpha}+\sqrt{\beta}} \ . \tag{6.2.23}$$

For $0<r<\sqrt{\alpha\beta}$, we obtain,

$$\begin{aligned} \left|\frac{r-\beta}{r+\beta}\right| - \frac{(\sqrt{\beta}-\sqrt{\alpha})}{(\sqrt{\alpha}+\sqrt{\beta})} &= \frac{2\sqrt{\beta}(\sqrt{\alpha\beta}-r)}{(r+\beta)(\sqrt{\alpha}+\sqrt{\beta})} \\ &> 0 \ , \end{aligned}$$

i.e. $$\left|\frac{r-\beta}{r+\beta}\right| > \frac{(\sqrt{\beta}-\sqrt{\alpha})}{(\sqrt{\alpha}+\sqrt{\beta})} \quad . \tag{6.2.24}$$

Similarly, for $\sqrt{\alpha\beta}<r$, we get,

$$\left|\frac{r-\alpha}{r+\alpha}\right| - \frac{(\sqrt{\beta}-\sqrt{\alpha})}{(\sqrt{\alpha}+\sqrt{\beta})} = \frac{2\sqrt{\alpha}(r-\sqrt{\alpha\beta})}{(r+\alpha)(\sqrt{\alpha}+\sqrt{\beta})}$$
$$> 0 \; ,$$

i.e., $$\left|\frac{r-\alpha}{r+\alpha}\right| > \frac{(\sqrt{\beta}-\sqrt{\alpha})}{(\sqrt{\alpha}+\sqrt{\beta})} \quad . \tag{6.2.25}$$

Hence, using (6.2.21)-(6.2.25) we deduce that the bound $\phi(\alpha,\beta;r)$ for $\rho(M(r))$ is minimised when $r=\hat{r}=\sqrt{\alpha\beta}$ and $\rho(M(\hat{r}))\leq\phi(\alpha,\beta;\hat{r}) = (\frac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\alpha}+\sqrt{\beta}})^2$.

For an efficient implementation of the AGE algorithm, it is essential to vary the acceleration parameters r_p from iteration to iteration - *the non-stationary case*. This will result in a substantial improvement in the rate of convergence of the AGE method when the Peaceman-Rachford variant is employed. The Peaceman-Rachford formula (6.2.11) will now become

$$(G_1+r_{p+1}I)\underline{u}^{(p+\frac{1}{2})} = (r_{p+1}I-G_2)\underline{u}^{(p)}+\underline{f}$$

and $$(G_2+r_{p+1}I)\underline{u}^{(p+1)} = (r_{p+1}I-G_1)\underline{u}^{(p+\frac{1}{2})} + \underline{f} \; , \; p\geq 0 \; . \tag{6.2.26}$$

The best values of r_p can be ascertained provided G_1 and G_2 are *commutative* - a property which is not possessed by our model problem. However, these matrices commute if the boundary conditions are *periodic* and of order 4, that is, when the conditions (6.2.1a) are replaced by

$$u(0,t) = u(1,t) \; , \; \frac{\partial u}{\partial x}(0,t) = \frac{\partial u}{\partial x}(1,t) \; . \tag{6.2.27}$$

For the general application of (6.2.26), ($r_p>0$) we assume first of all that the positive definite matrices G_1 and G_2 commute. Thus, G_1 and G_2 have a common set of (orthonormal) eigenvectors. Let $(\mu_j,\underline{v}_j)_{j=1}^m$ and $(\eta_j,\underline{v}_j)_{j=1}^m$ be the eigensystem of G_1 and G_2 respectively.

For p iterations of (6.2.26), relation (6.2.13) yields, for $1 \leqslant j \leqslant m$,

$$\left(\prod_{i=1}^{p} M(r_i)\right)\underline{v}_j = \left\{\prod_{i=1}^{p}\left(\frac{r_i-\mu_j}{r_i+\mu_j}\right)\left(\frac{r_i-\eta_j}{r_i+\eta_j}\right)\right\}\underline{v}_j \;. \qquad (6.2.28)$$

It follows that $\prod_{i=1}^{p} M(r_i)$ is symmetric and therefore,

$$\begin{aligned} \left\|\prod_{i=1}^{p} M(r_i)\right\|_2 &= \rho\left(\prod_{i=1}^{p} M(r_i)\right) \\ &= \max_{1\leqslant j\leqslant m} \prod_{i=1}^{p}\left|\frac{r_i-\mu_j}{r_i+\mu_j}\right|\left|\frac{r_i-\eta_j}{r_i+\eta_j}\right| < 1 \end{aligned} \qquad (6.2.29)$$

and convergence of the iterative process is achieved. Now, it is clear that

$$\begin{aligned} \max_{1\leqslant j\leqslant m} \prod_{i=1}^{p}\left|\frac{r_i-\mu_j}{r_i+\mu_j}\right| \left|\frac{r_i-\eta_j}{r_i+\eta_j}\right| &\leqslant \max_{1\leqslant j\leqslant m} \prod_{i=1}^{p}\left|\frac{r_i-\mu_j}{r_i+\mu_j}\right| \max_{1\leqslant j\leqslant m} \prod_{i=1}^{p}\left|\frac{r_i-\eta_j}{r_i+\eta_j}\right| . \\ &\leqslant \left\{\max_{\alpha\leqslant z\leqslant\beta} \prod_{i=1}^{p}\left|\frac{r_i-z}{r_i+z}\right|\right\}^2 \end{aligned} \qquad (6.2.30)$$

where we have used the bounds for the eigenvalues given by (6.2.20).

Hence,

$$\rho\left(\prod_{i=1}^{p} M(r_i)\right) \leqslant \max_{\alpha\leqslant z\leqslant\beta} |R_p(z)|^2 = \phi(\alpha,\beta;r_1,\ldots,r_p) \qquad (6.2.31)$$

where $R_p(z) = \prod_{i=1}^{p} \frac{(r_i-z)}{(r_i+z)}$. The difficulty of determining the optimum parameters by minimising $\phi(\alpha,\beta;r_1,\ldots,r_p)$ has led a number of authors to devise alternative sequences. For example, the parameters used by Peaceman and Rachford (1955), are,

$$\hat{r}_j = \beta\left(\frac{\alpha}{\beta}\right)^{(2j-1)/(2p)} \;, \; j=1,2,\ldots,p \qquad (6.2.32)$$

from which we obtain the result,

$$\rho\left(\prod_{i=1}^{p} M(r_i)\right) \leqslant \left\{\frac{1-\left(\frac{\alpha}{\beta}\right)^{1/(2p)}}{1+\left(\frac{\alpha}{\beta}\right)^{1/(2p)}}\right\}^2 \;. \qquad (6.2.33)$$

Wachpress (1966), on the other hand, solved the minimisation problem analytically in terms of elliptic functions and arrived at the result

$$\hat{r}_j = \beta\,\left(\frac{\alpha}{\beta}\right)^{(j-1)/(p-1)}\ ,\ p\geq 2\ ,\quad j=1,2,\ldots,p. \qquad (6.2.34)$$

It must be noted that the requirement that G_1 and G_2 be commutative can be very restrictive indeed. However, in far more general situations we can expect convergence of the AGE iteration for a *fixed acceleration parameter* without the condition $G_1G_2=G_2G_1$.

We conclude this section by considering the case when we have an *odd number of intervals* (corresponding to an even number of internal points, i.e. m even) on $0\leq x\leq 1$. The coefficient matrix A will be split as in (6.2.5) but G_1 and G_2 now take the form,

$$G_1 = \left[\begin{array}{cc|cc|c|cc|cc}
a/2 & b & & & & & & & \\
c & a/2 & & & & & & & \\ \hline
 & & a/2 & b & & & & & \\
 & & c & a/2 & & & & & \\ \hline
 & & & & \ddots & & O & & \\ \hline
 & & & O & & a/2 & b & & \\
 & & & & & c & a/2 & & \\ \hline
 & & & & & & & a/2 & b \\
 & & & & & & & c & a/2
\end{array}\right]_{(m\times m)} \qquad (6.2.35)$$

and

$$G_2 = \left(\begin{array}{c|c|c|c|c|c}
a/2 & & & & & \\ \hline
 & \begin{matrix} a/2 & b \\ c & a/2 \end{matrix} & & & & \\ \hline
 & & \begin{matrix} a/2 & b \\ c & a/2 \end{matrix} & & \bigcirc & \\ \hline
 & & & \ddots & & \\ \hline
 & & \bigcirc & & \begin{matrix} a/2 & b \\ c & a/2 \end{matrix} & \\ \hline
 & & & & & a/2
\end{array}\right)_{(m\times m)} . \qquad (6.2.36)$$

All the preceding conditions and arguments regarding convergence and the choice of the optimal acceleration parameter for both the stationary and non-stationary cases remain valid.

6.3 VARIANTS OF THE AGE SCHEME AND ITS COMPUTATION

Many variants of the basic Peaceman-Rachford scheme can be proposed. For example, we have, on modifying the second stage of (6.2.26) (the non-stationary case),

$$\begin{aligned} &(G_1+r_{p+1}I)\underline{u}^{(p+\frac{1}{2})} = (r_{p+1}I-G_2)\underline{u}^{(p)} + \underline{f} \\ \text{and}\quad &(G_2+r_{p+1}I)\underline{u}^{(p+1)} = (G_2-(1-\omega)r_{p+1}I)\underline{u}^{(p)}+(2-\omega)r_{p+1}\underline{u}^{(p+\frac{1}{2})} \end{aligned} \qquad (6.3.1)$$

where ω is a parameter. For $\omega=0$, we have the Peaceman-Rachford scheme (6.2.26) and for $\omega=1$, we obtain the scheme due to Douglas and Rachford (1956). For G_1 and G_2 symmetric and positive definite and with a *fixed acceleration parameter* $r>0$, the resulting *generalised AGE scheme* is convergent for any $0\leq\omega\leq 2$. As we shall see in a subsequent chapter, a natural extension of the AGE algorithm is to implement it on higher dimensional boundary value problems using the Douglas-Rachford variant.

For the purpose of computation, we shall now attempt to derive equations that are satisfied at each intermediate (half-time) level. For *the Peaceman-Rachford variant*, in particular and with fixed parameter r, the *AGE method* can be applied to determine $\underline{u}^{(p+\frac{1}{2})}$ and $\underline{u}^{(p+1)}$ implicitly by,

$$\begin{aligned} &(G_1+rI)\underline{u}^{(p+\frac{1}{2})} = (rI-G_2)\underline{u}^{(p)} + \underline{f} \\ \text{and}\quad &(G_2+rI)\underline{u}^{(p+1)} = (rI-G_1)\underline{u}^{(p+\frac{1}{2})} + \underline{f} \end{aligned} \qquad (6.3.2)$$

or explicitly by,

$$\left.\begin{aligned} &\underline{u}^{(p+\frac{1}{2})} = (G_1+rI)^{-1}\{(rI-G_2)\underline{u}^{(p)}+\underline{f}\} \\ \text{and}\quad &\underline{u}^{(p+1)} = (G_2+rI)^{-1}\{(rI-G_1)\underline{u}^{(p+\frac{1}{2})} + \underline{f}\}. \end{aligned}\right\} \qquad (6.3.3)$$

If we assume *m to be odd (even number of intervals)* and if we write

$$\hat{G} = \begin{pmatrix} r_2 & b \\ c & r_2 \end{pmatrix} \qquad (6.3.4)$$

where $$r_2 = r + \frac{a}{2} \tag{6.3.5}$$

then from (6.2.6) and (6.2.7), we have,

$$(G_1+rI) = \left(\begin{array}{c|c|c|c|c} r_2 & & & & \\ \hline & \hat{G} & & & \\ \hline & & \hat{G} & & \\ \hline & & & \begin{array}{cc} \ddots & O \\ O & \ddots \end{array} & \\ \hline & & & & \hat{G} \end{array}\right)_{(m\times m)} \tag{6.3.6}$$

and

$$(G_2+rI) = \left(\begin{array}{c|c|c|c|c} \hat{G} & & & & \\ \hline & \hat{G} & & & \\ \hline & & \begin{array}{cc} \ddots & O \\ O & \ddots \end{array} & & \\ \hline & & & \hat{G} & \\ \hline & & & & r_2 \end{array}\right)_{(m\times m)} . \tag{6.3.7}$$

It is clear that (G_1+rI) and (G_2+rI) are block diagonal matrices. All the diagonal elements except the first (or the last for (G_2+rI)) are (2×2) submatrices. Therefore, (G_1+rI) and (G_2+rI) can be easily inverted by merely inverting their block diagonal entries. Hence,

$$(G_1+rI)^{-1} = \left(\begin{array}{c|c|c|c|c} \frac{1}{r_2} & & & & \\ \hline & \hat{G}^{-1} & & & \\ \hline & & \hat{G}^{-1} & & O \\ \hline & & & \ddots & \\ \hline & & O & & \hat{G}^{-1} \end{array}\right)_{(m\times m)} \tag{6.3.8}$$

$$= \frac{1}{\Delta}\left(\begin{array}{c|cc|c|cc} \Delta/r_2 & & & & & \\ \hline & r_2 & -b & & & \\ & -c & r_2 & & & \\ \hline & & & \ddots & & O \\ \hline & & O & & r_2 & -b \\ & & & & -c & r_2 \end{array}\right)_{(m\times m)} \tag{6.3.9}$$

where $\Delta = r_2^2 - bc$. (6.3.10)

Similarly, we obtain,

$$(G_2+rI)^{-1} = \frac{1}{\Delta}\left(\begin{array}{cc|cc|c|cc|c} r_2 & -b & & & & & & \\ -c & r_2 & & & & & & \\ \hline & & r_2 & -b & & & & \\ & & -c & r_2 & & O & & \\ \hline & & & & \ddots & & & \\ \hline & & O & & & r_2 & -b & \\ & & & & & -c & r_2 & \\ \hline & & & & & & & \Delta/r_2 \end{array}\right)_{(m\times m)} . \tag{6.3.11}$$

From (6.3.3), $\underline{u}^{(p+\frac{1}{2})}$ and $\underline{u}^{(p+1)}$ are given by,

$$\begin{pmatrix} u_1^{(p+\frac{1}{2})} \\ u_2^{(p+\frac{1}{2})} \\ u_3^{(p+\frac{1}{2})} \\ \vdots \\ \vdots \\ u_{m-1}^{(p+\frac{1}{2})} \\ u_m^{(p+\frac{1}{2})} \end{pmatrix} = \frac{1}{\Delta}\left(\begin{array}{c|cc|c|cc} \Delta/r_2 & & & & & \\ \hline & r_2 & -b & & & \\ & -c & r_2 & & & \\ \hline & & & \ddots & & O \\ \hline & & O & & & \\ & & & & r_2 & -b \\ & & & & -c & r_2 \end{array}\right) \begin{pmatrix} r_1u_1^{(p)} - bu_2^{(p)} + f_1 \\ -cu_1^{(p)} + r_1u_2^{(p)} + f_2 \\ r_1u_3^{(p)} - bu_4^{(p)} + f_3 \\ -cu_3^{(p)} + r_1u_4^{(p)} + f_4 \\ \vdots \\ r_1u_{m-2}^{(p)} - bu_{m-1}^{(p)} + f_{m-2} \\ -cu_{m-2}^{(p)} + r_1u_{m-1}^{(p)} + f_{m-1} \\ r_1u_m^{(p)} + f_m \end{pmatrix} \tag{6.3.12}$$

and

$$\begin{pmatrix} u_1^{(p+1)} \\ u_2^{(p+1)} \\ u_3^{(p+1)} \\ \vdots \\ u_{m-2}^{(p+1)} \\ u_{m-1}^{(p+1)} \\ u_m^{(p+1)} \end{pmatrix} = \frac{1}{\Delta} \left(\begin{array}{cc|cc|c|cc|c} r_2 & -b & & & & & & \\ -c & r_2 & & & & & & \\ \hline & & r_2 & -b & & & & \\ & & -c & r_2 & & \Theta & & \\ \hline & & & & \ddots & & & \\ \hline & & O & & & r_2 & -b & \\ & & & & & -c & r_2 & \\ \hline & & & & & & & \Delta/r_2 \end{array}\right) \begin{pmatrix} r_1u_1^{(p+\frac{1}{2})}+f_1 \\ r_1u_2^{(p+\frac{1}{2})}-bu_3^{(p+\frac{1}{2})}+f_2 \\ -cu_2^{(p+\frac{1}{2})}+r_1u_3^{(p+\frac{1}{2})}+f_3 \\ r_1u_4^{(p+\frac{1}{2})}-bu_5^{(p+\frac{1}{2})}+f_4 \\ -cu_4^{(p+\frac{1}{2})}+r_1u_5^{(p+\frac{1}{2})}+f_5 \\ \vdots \\ r_1u_{m-1}^{(p+\frac{1}{2})}-bu_m^{(p+\frac{1}{2})}+f_{m-1} \\ -cu_{m-1}^{(p+\frac{1}{2})}+r_1u_m^{(p+\frac{1}{2})}+f_m \end{pmatrix} \quad (6.3.13)$$

where $r_1=r-a/2$, $r_2=r+a/2$ and $\Delta=r_2^2-bc$. (6.3.14)

The *alternating implicit* nature of the (2×2) groups in the equations (6.3.2) is shown in Figure (6.3.1) where the implicit/explicit values are given on the forward/backward levels for sweeps on the (p+½)th and (p+1)th levels.

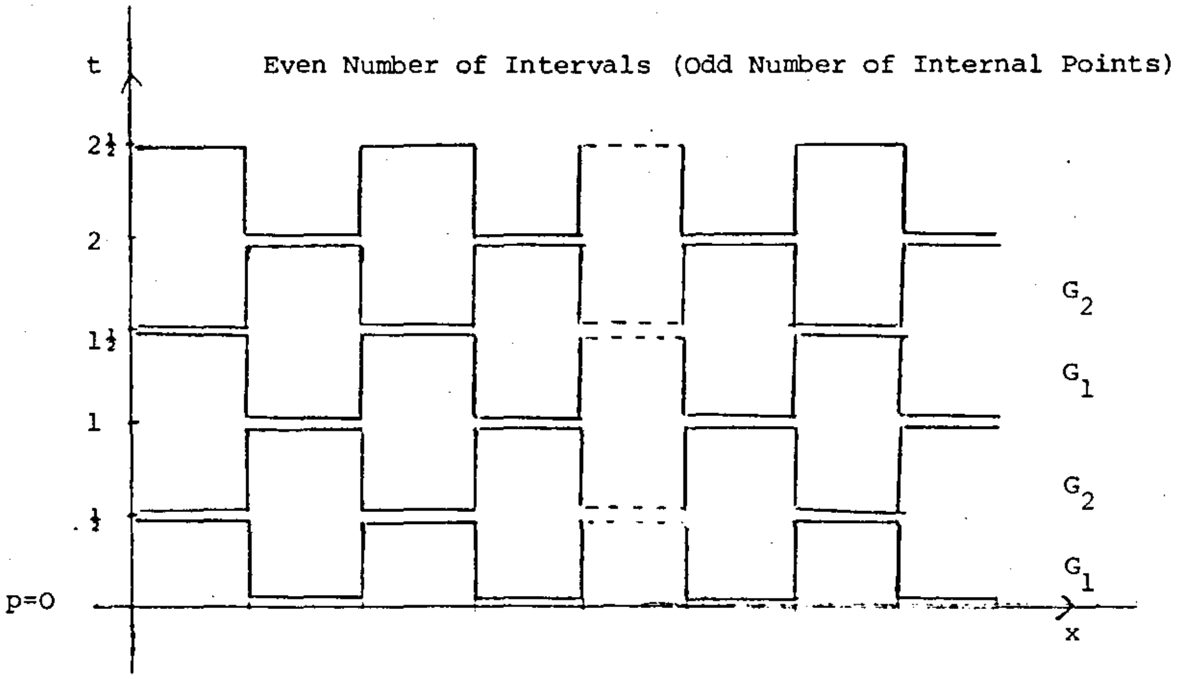


FIGURE 6.3.1: The AGE Method (Implicit)

The corresponding explicit expressions for the AGE equations are obtained by carrying out the multiplications in (6.3.12) and (6.3.13). Thus, we have,

(i) at level $(p+\frac{1}{2})$

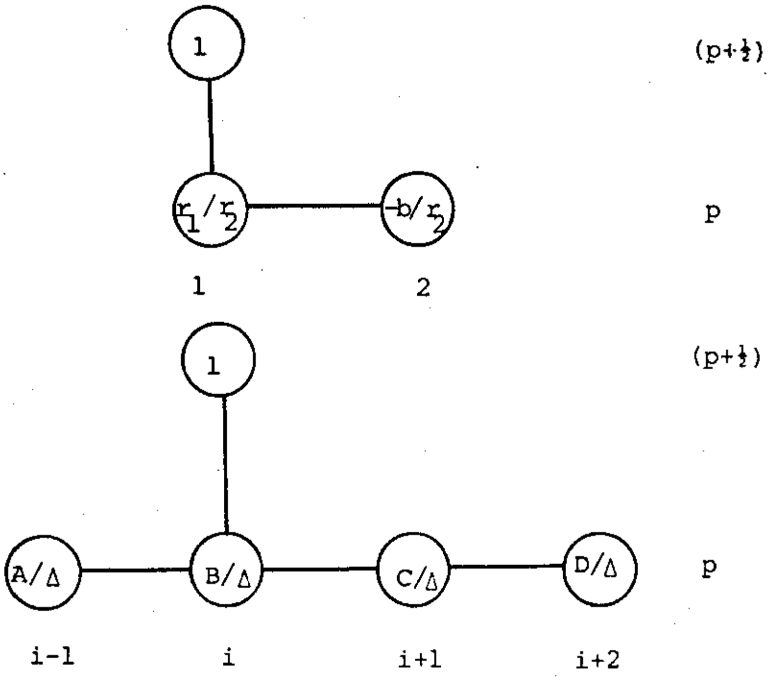
$$\left.\begin{aligned} u_1^{(p+\frac{1}{2})} &= (r_1u_1^{(p)}-bu_2^{(p)}+f_1)/r_2 \;, \\ \left.\begin{aligned} u_i^{(p+\frac{1}{2})} &= (Au_{i-1}^{(p)}+Bu_i^{(p)}+Cu_{i+1}^{(p)}+Du_{i+2}^{(p)}+E_i)/\Delta \\ \text{and}\quad u_{i+1}^{(p+\frac{1}{2})} &= (\tilde{A}u_{i-1}^{(p)}+\tilde{B}u_i^{(p)}+\tilde{C}u_{i+1}^{(p)}+\tilde{D}u_{i+2}^{(p)}+\tilde{E}_i)/\Delta \end{aligned}\right\} & i=2,4,\ldots,m-1 \end{aligned}\right\} \quad (6.3.15)$$

where,

$$\left.\begin{aligned} &A=-cr_2,\ B=r_1r_2,\ C=-br_1,\ E_i=r_2f_i-bf_{i+1},\ D=\begin{cases}0 & \text{for } i=m-1\\ b^2 & \text{otherwise}\end{cases} \\ &\text{and } \tilde{A}=c^2,\ \tilde{B}=-cr_1,\ \tilde{C}=r_1r_2,\ \tilde{E}_i=r_2f_{i+1}-cf_i,\ \tilde{D}=\begin{cases}0 \text{ for } i=m-1\\ br_1 \text{ otherwise}\end{cases} \end{aligned}\right\}$$

(6.3.15a)

with the following computational molecules (Figure 6.3.2),



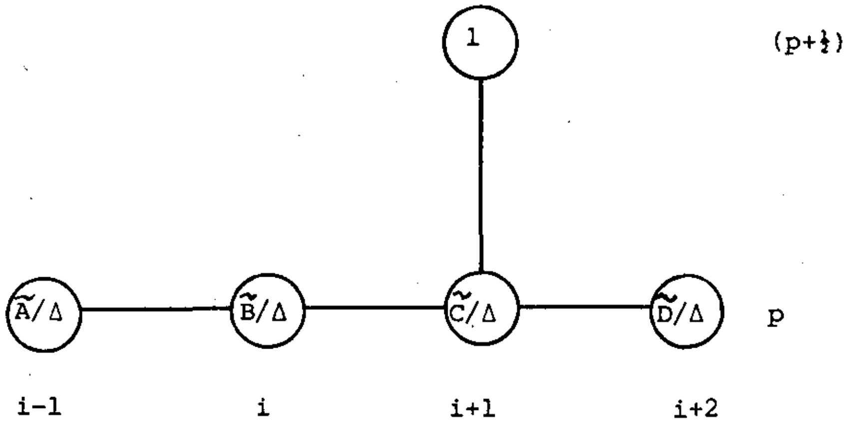


FIGURE 6.3.2: The AGE Method (Explicit) at level (p+½)

(ii) at level (p+1)

$$\left.\begin{array}{l}\left.\begin{array}{l} u_i^{(p+1)} = (Pu_{i-1}^{(p+\frac{1}{2})}+Qu_i^{(p+\frac{1}{2})}+Ru_{i+1}^{(p+\frac{1}{2})}+Su_{i+2}^{(p+\frac{1}{2})}+T_i)/\Delta \\ u_{i+1}^{(p+1)} = (\tilde{P}u_{i-1}^{(p+\frac{1}{2})}+\tilde{Q}u_i^{(p+\frac{1}{2})}+\tilde{R}u_{i+1}^{(p+\frac{1}{2})}+\tilde{S}u_{i+2}^{(p+\frac{1}{2})}+\tilde{T}_i)/\Delta \end{array}\right\} \\ \qquad\qquad\qquad\qquad i=1,3,\ldots,m-2 \\ \text{and} \quad u_m^{(p+1)} = (-cu_{m-1}^{(p+\frac{1}{2})}+r_1u_m^{(p+\frac{1}{2})}+f_m)/r_2 \end{array}\right\} \qquad (6.3.16)$$

where,

$$\left.\begin{array}{l} P = \begin{cases} 0 \text{ for } i=1 \\ -cr_2 \text{ for } i\neq 1 \end{cases}, \; Q=r_1r_2, \; R=-br_1, \; S=b^2, \; T_i=r_2f_i-bf_{i+1} \\ \text{and} \quad \tilde{P} = \begin{cases} 0 \text{ for } i=1 \\ c^2 \text{ for } i\neq 1 \end{cases}, \; \tilde{Q} = -cr_1, \; \tilde{R} = Q = r_1r_2, \; \tilde{S} = -br_2, \\ \qquad\qquad\qquad \tilde{T}_i = -cf_i+r_2f_{i+1} \end{array}\right\} \quad (6.3.16a)$$

with its computational molecules given by Figure (6.3.3).

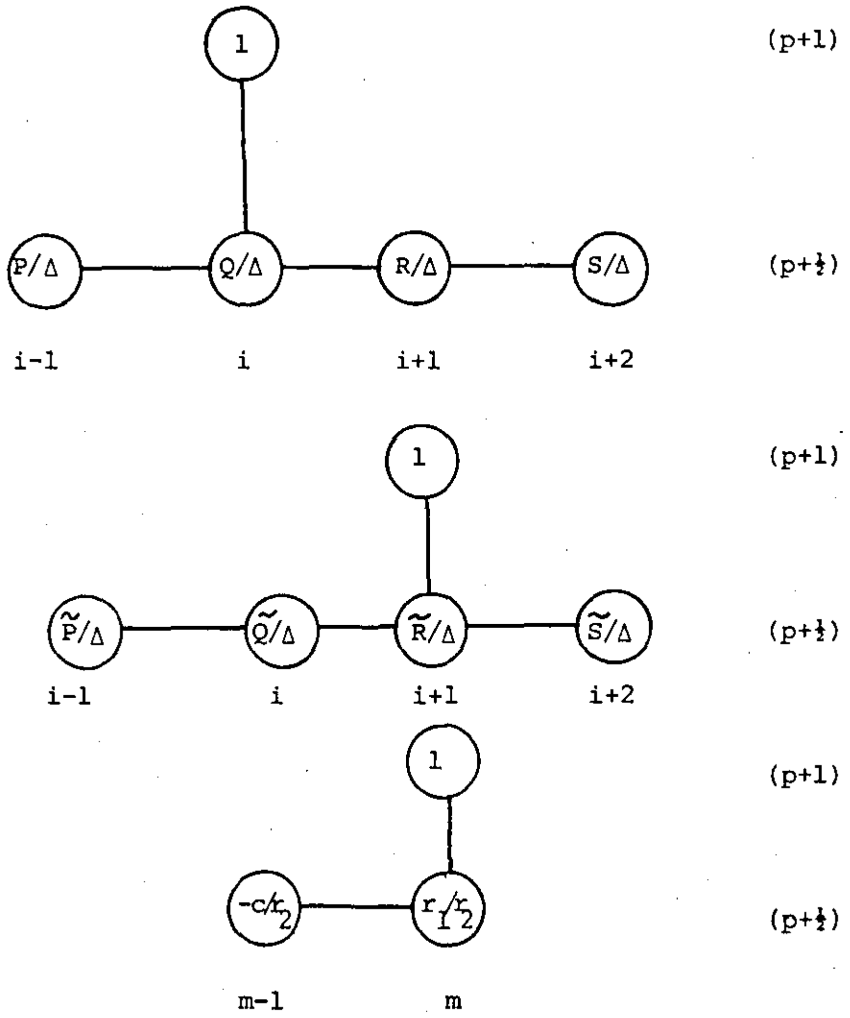


FIGURE 6.3.3: The AGE Method (Explicit) at Level (p+1)

For the generalised AGE Scheme (6.3.1) with fixed acceleration parameter r, the relevant equations at level (p+½) remain the same as in (6.3.15/15a). The equations at level (p+1) are, however, now replaced by,

$$\begin{bmatrix} u_1^{(p+1)} \\ u_2^{(p+1)} \\ u_3^{(p+1)} \\ \vdots \\ u_{m-2}^{(p+1)} \\ u_{m-1}^{(p+1)} \\ u_m^{(p+1)} \end{bmatrix} = \frac{1}{\Delta}\left[\begin{array}{cc|cc|c|cc|c} r_2 & -b & & & & & & \\ -c & r_2 & & & & & & \\ \hline & & r_2 & -b & & & & \\ & & -c & r_2 & & & & \\ \hline & & & & \ddots & & O & \\ \hline & & & & & r_2 & -b & \\ & & O & & & -c & r_2 & \\ \hline & & & & & & & \Delta/r_2 \end{array}\right] \begin{bmatrix} r_3u_1^{(p)}+bu_2^{(p)}+r_4u_1^{(p+\frac{1}{2})} \\ cu_1^{(p)}+r_3u_2^{(p)}+r_4u_2^{(p+\frac{1}{2})} \\ r_3u_3^{(p)}+bu_4^{(p)}+r_4u_3^{(p+\frac{1}{2})} \\ cu_3^{(p)}+r_3u_4^{(p)}+r_4u_4^{(p+\frac{1}{2})} \\ \vdots \\ r_3u_{m-2}^{(p)}+bu_{m-1}^{(p)}+r_4u_{m-2}^{(p+\frac{1}{2})} \\ cu_{m-2}^{(p)}+r_3u_{m-1}^{(p)}+r_4u_{m-1}^{(p+\frac{1}{2})} \\ r_3u_m^{(p)}+r_4u_m^{(p+\frac{1}{2})} \end{bmatrix}, \qquad (6.3.17)$$

where r_1, r_2 and Δ are given by (6.3.14) and $r_3 = \frac{a}{2}-(1-\omega)r$ and $r_4=(2-\omega)r$. This leads to

$$\left.\begin{array}{l} \left.\begin{array}{l} u_i^{(p+1)} = (Pu_i^{(p+\frac{1}{2})}+Qu_{i+1}^{(p+\frac{1}{2})}+Ru_i^{(p)}+Su_{i+1}^{(p)})/\Delta \\ u_{i+1}^{(p+1)} = (\tilde{P}u_i^{(p+\frac{1}{2})}+\tilde{Q}u_{i+1}^{(p+\frac{1}{2})}+\tilde{R}u_i^{(p)}+\tilde{S}u_{i+1}^{(p)})/\Delta \end{array}\right\} i=1,3,\ldots,m-2 \\ \text{and} \quad u_m^{(p+1)} = (r_3u_m^{(p)}+r_4u_m^{(p+\frac{1}{2})})/r_2 \end{array}\right\} \qquad (6.3.18)$$

where,

$$\left.\begin{array}{l} P = r_2r_4, \quad Q = -br_4, \quad R = r_2r_3-bc, \quad S = b(r_2-r_3) \\ \text{and} \quad \tilde{P} = -cr_4, \quad \tilde{Q} = P = r_2r_4, \quad \tilde{R} = c(r_2-r_3) \text{ and } \tilde{S} = r_2r_3-bc \end{array}\right\} \qquad (6.3.19)$$

and the computational molecules are given by Figure (6.3.4).

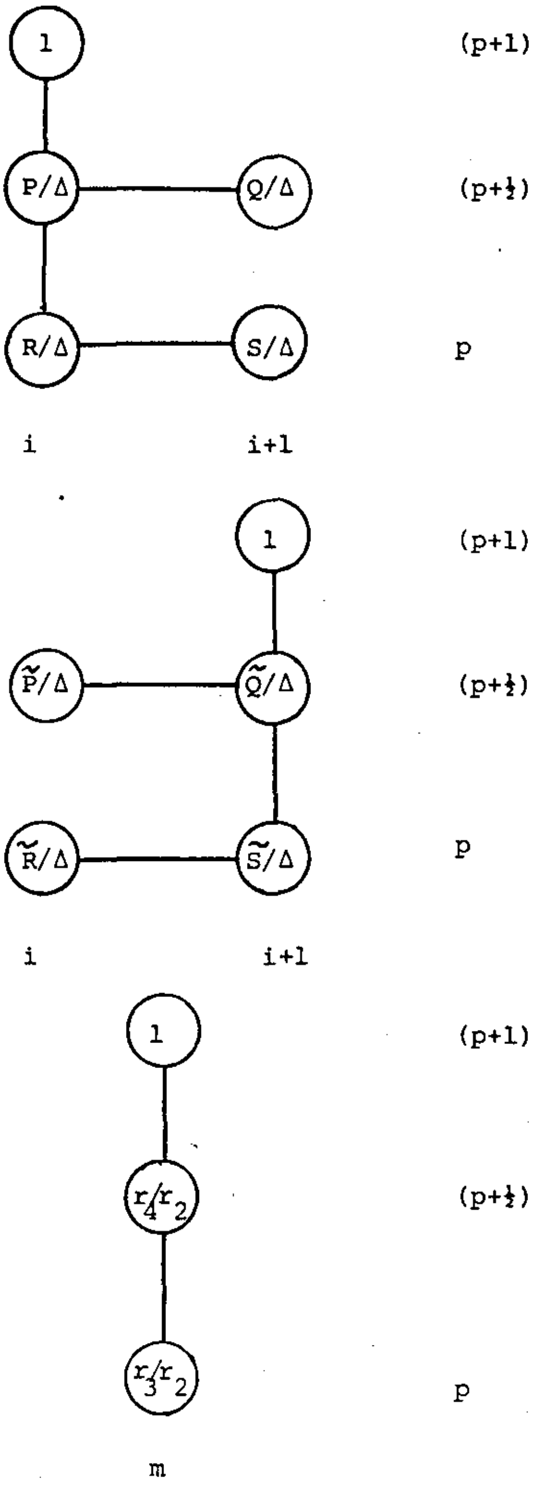


FIGURE 6.3.4: The generalised AGE method (Explicit) at level (p+1)

Finally, let us now consider the case when m *is even (corresponding to an odd number of intervals)*. We shall then have from (6.2.35) and (6.2.36),

$$(G_1+rI)^{-1}=\left(\begin{array}{c|c|c|c}
\hat{G}^{-1} & & & \\ \hline
 & \hat{G}^{-1} & & \\ \hline
 & & \begin{array}{cc} \ddots & O \\ O & \ddots \end{array} & \\ \hline
 & & & \hat{G}^{-1}
\end{array}\right)_{(m\times m)} \tag{6.3.20}$$

$$= \frac{1}{\Delta}\left(\begin{array}{cc|cc|c|cc}
r_2 & -b & & & & & \\
-c & r_2 & & & & & \\ \hline
 & & r_2 & -b & & & \\
 & & -c & r_2 & & & \\ \hline
 & & & & \begin{array}{cc} \ddots & O \\ O & \ddots \end{array} & & \\ \hline
 & & & & & r_2 & -b \\
 & & & & & -c & r_2
\end{array}\right)_{(m\times m)} \tag{6.3.21}$$

and

$$(G_2+rI)^{-1}=\left(\begin{array}{c|c|c|c|c}
1/r_2 & & & & \\ \hline
 & \hat{G}^{-1} & & & \\ \hline
 & & \begin{array}{cc} \ddots & O \\ O & \ddots \end{array} & & \\ \hline
 & & & \hat{G}^{-1} & \\ \hline
 & & & & 1/r_2
\end{array}\right)_{(m\times m)} \tag{6.3.22}$$

The Peaceman-Rachford variant in its implicit form (6.3.2) can be pictorially represented by Figure (6.3.5).

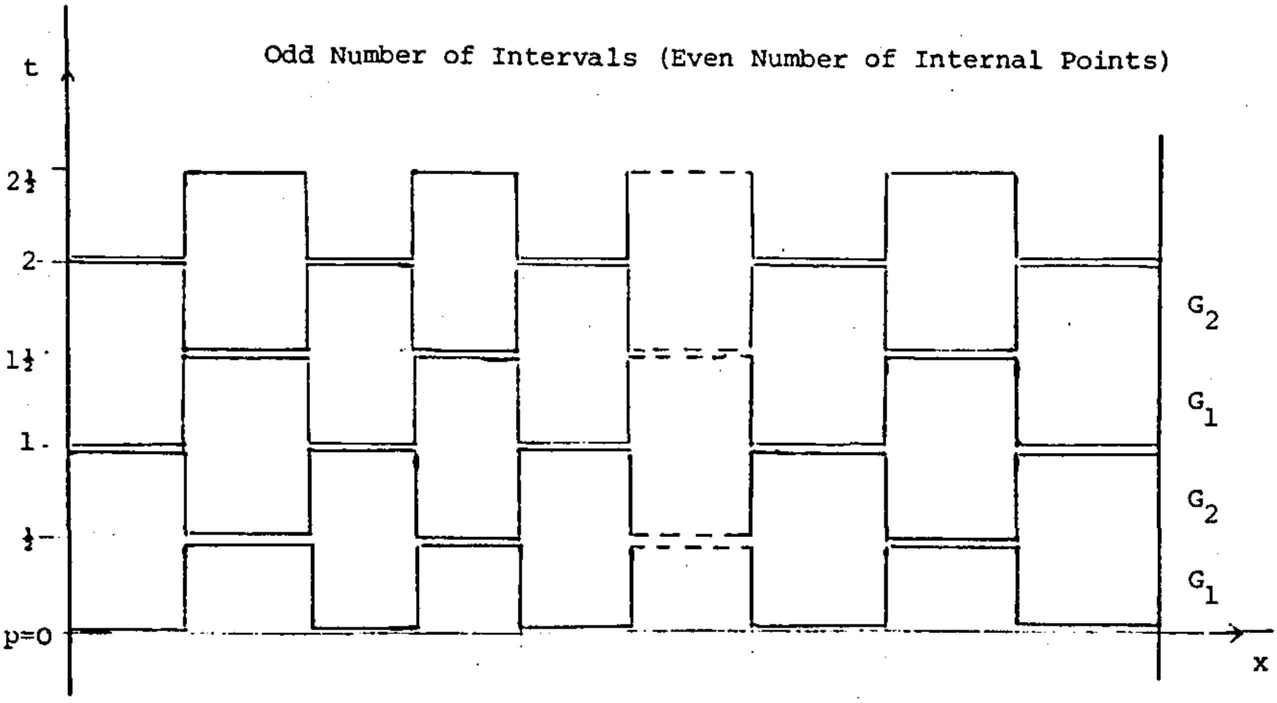


FIGURE 6.3.5: The AGE Method (Implicit)

By means of equations (6.2.35)-(6.2.36), (6.3.3) and (6.3.20)-(6.3.22) we obtain the following explicit expressions for the AGE scheme:

(i) at level (p+½)

$$\left.\begin{aligned} u_i^{(p+\frac{1}{2})} &= (Au_{i-1}^{(p)}+Bu_i^{(p)}+Cu_{i+1}^{(p)}+Du_{i+2}^{(p)}+E_i)/\Delta \\ \text{and}\quad u_{i+1}^{(p+\frac{1}{2})} &= (\tilde{A}u_{i-1}^{(p)}+\tilde{B}u_i^{(p)}+\tilde{C}u_{i+1}^{(p)}+\tilde{D}u_{i+2}^{(p)}+\tilde{E}_i)/\Delta \end{aligned}\right\} i=1,3,\ldots,m-1 \qquad (6.3.23)$$

where

$$\left.\begin{aligned} A = \begin{cases} 0 & \text{, for } i=1 \\ -cr_2, & \text{otherwise} \end{cases}, \; B = r_1r_2,\; C = -br_1,\; D = \begin{cases} 0, & \text{for } i=m-1 \\ b^2, & \text{otherwise} \end{cases} \\ E_i = r_2f_i-bf_{i+1} \end{aligned}\right|$$

and
$$\left.\begin{array}{l} \tilde{A} = \begin{cases} 0, \text{ for } i=1 \\ c^2, \text{otherwise} \end{cases}, \; \tilde{B} = -cr_1, \; \tilde{C} = r_1 r_2, \; \tilde{D} = \begin{cases} 0, \text{ for } i=m-1 \\ -br_2, \text{otherwise} \end{cases} \\ \qquad\qquad\qquad\qquad \tilde{E}_i = -cf_i + r_2 f_{i+1} \end{array}\right\} \qquad (6.3.23a)$$

(ii) at level (p+1)

$$\left.\begin{array}{ll} u_1^{(p+1)} = (r_1 u_1^{(p+\frac{1}{2})} - bu_2^{(p+\frac{1}{2})} + f_1)/r_2 \,, & \\ u_i^{(p+1)} = (Pu_{i-1}^{(p+\frac{1}{2})} + Qu_i^{(p+\frac{1}{2})} + Ru_{i+1}^{(p+\frac{1}{2})} + Su_{i+2}^{(p+\frac{1}{2})} + T_i)/\Delta, & \\ u_{i+1}^{(p+1)} = (\tilde{P}u_{i-1}^{(p+\frac{1}{2})} + \tilde{Q}u_i^{(p+\frac{1}{2})} + \tilde{R}u_{i+1}^{(p+\frac{1}{2})} + \tilde{S}u_{i+2}^{(p+\frac{1}{2})} + \tilde{T}_i)/\Delta, & i=2,4,\ldots,m-2 \\ \text{and} \quad u_m^{(p+1)} = (-cu_{m-1}^{(p+\frac{1}{2})} + r_1 u_m^{(p+\frac{1}{2})} + f_m)/r_2 & \end{array}\right\} \qquad (6.3.24)$$

where

$$\left.\begin{array}{l} P = -cr_2, \; Q = r_1 r_2, \; R = -br_1, \; S = b^2, \; T_i = r_2 f_i - bf_{i+1} \\ \text{and} \quad \tilde{P} = c^2, \; \tilde{Q} = -cr_1, \; \tilde{R} = Q = r_1 r_2, \; \tilde{S} = -br_2, \; \tilde{T}_i = -cf_i + r_2 f_{i+1} \;. \end{array}\right\} \qquad (6.3.24a)$$

For *the generalised AGE scheme* (6.3.1) with fixed acceleration parameter r, the same equations in (6.3.23) still apply for level (p+½). At level (p+1), we have the following equations,

$$\left.\begin{array}{l} u_1^{(p+1)} = (r_3 u_1^{(p)} + r_4 u_1^{(p+\frac{1}{2})})/r_2 \,, \\ \left.\begin{array}{l} u_i^{(p+1)} = (Pu_i^{(p)} + Qu_{i+1}^{(p)} + Ru_i^{(p+\frac{1}{2})} + Su_{i+1}^{(p+\frac{1}{2})})/\Delta, \\ u_{i+1}^{(p+1)} = (\tilde{P}u_i^{(p)} + \tilde{Q}u_{i+1}^{(p)} + \tilde{R}u_i^{(p+\frac{1}{2})} + \tilde{S}u_i^{(p+\frac{1}{2})})/\Delta, \end{array}\right\} i=2,4,\ldots,m-2 \\ u_m^{(p+1)} = (r_3 u_m^{(p)} + r_4 u_m^{(p+\frac{1}{2})})/r_2 \,, \end{array}\right\} \qquad (6.3.25)$$

where,

$$\left.\begin{array}{l} P = r_2 r_3 - bc, \; Q = b(r_2 - r_3), \; R = r_2 r_4, \; S = -br_4 \,, \\ \text{and} \quad \tilde{P} = c(r_2 - r_3), \; \tilde{Q} = r_2 r_3 - bc, \; \widetilde{R} = -cr_4, \; S = r_2 r_4 \;. \end{array}\right\} \qquad (6.3.25a)$$

The AGE algorithm is completed *explicitly* by using the required equations at levels (p+½) and (p+1) in *alternate sweeps* along all the points in the interval (0,1) until a specified convergence criterion is satisfied.

6.4 THE AGE METHOD TO SOLVE SECOND ORDER PARABOLIC EQUATIONS WITH PERIODIC BOUNDARY CONDITIONS

Consider the following second order parabolic equation,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + k(x,t)\,,\ 0\leq x\leq 1,\ 0<t\leq T\ , \tag{6.4.1}$$

subject to the initial condition,

$$U(x,0) = f(x)\ ,\ 0<x<1\ ,$$

and the boundary condition, (6.4.1a)

$$U(0,t) = U(1,t)\ ,\ \frac{\partial U}{\partial x}(0,t) = \frac{\partial U}{\partial x}(1,t)\ .$$

As in Section 6.2, a weighted approximation to the differential equation (6.4.1) at the point $(x_i,t_{j+\frac{1}{2}})$ is

$$-\lambda\theta u_{i-1,j+1}+(1+2\lambda\theta)u_{i,j+1}-\lambda\theta u_{i+1,j+1} = \lambda(1-\theta)u_{i-1,j}+(1-2\lambda(1-\theta))u_{ij}$$
$$+\lambda(1-\theta)u_{i+1,j}+\Delta t k_{i,j+\frac{1}{2}},\ \text{for } i=1,2,\ldots,m. \tag{6.4.2}$$

When written in the matrix form (6.2.3) and taking account of (6.4.1a) we obtain the system (6.2.4) where,

$$c = -\lambda\theta\ ,\ a = (1+2\lambda\theta),\ b = -\lambda\theta;$$

$$f_1 = (1-2\lambda(1-\theta))u_{1j}+\lambda(1-\theta)u_{2j}+\lambda(1-\theta)u_{mj}+\Delta t k_{1,j+\frac{1}{2}}$$

$$f_i = \lambda(1-\theta)u_{i-1,j}+(1-2\lambda(1-\theta))u_{ij}+\lambda(1-\theta)u_{i+1,j}+\Delta t k_{i,j+\frac{1}{2}},$$
$$i=2,\ldots,m-1,$$

$$f_m = \lambda(1-\theta)u_{1j}+\lambda(1-\theta)u_{m-1,j}+(1-2\lambda(1-\theta))u_{mj}+\Delta t k_{m,j+\frac{1}{2}}\ .$$

To implement the AGE algorithm, we split the coefficient matrix A as in (6.2.5) with A given by,

$$A = \begin{bmatrix} a & b & & & & c \\ c & a & b & & & \\ & c & a & b & O & \\ & & & \ddots & & \\ & O & & & \ddots & \\ & & & c & a & b \\ b & & & & c & a \end{bmatrix} \tag{6.4.3}$$

and to ascertain the forms of G_1 and G_2, we treat two different cases of m as before.

(a) *m even* (even number of points since $u_0=u_m$ at every level with $x_0=0$ and $x_m=1$). We have,

$$G_1 = \left(\begin{array}{cc|cc|c|cc} \frac{a}{2} & b & & & & & \\ c & \frac{a}{2} & & & & & \\ \hline & & \frac{a}{2} & b & & O & \\ & & c & \frac{a}{2} & & & \\ \hline & & & & \ddots & & \\ \hline & & O & & & \frac{a}{2} & b \\ & & & & & c & \frac{a}{2} \end{array}\right)_{(m\times m)}, \tag{6.4.4}$$

$$G_2 = \left(\begin{array}{c|cc|c|cc|c} \frac{a}{2} & & & & & & c \\ \hline & \frac{a}{2} & b & & & & \\ & c & \frac{a}{2} & & O & & \\ \hline & & & \ddots & & & \\ \hline & & O & & \frac{a}{2} & b & \\ & & & & c & \frac{a}{2} & \\ \hline b & & & & & & a/2 \end{array}\right)_{(m\times m)} \tag{6.4.5}$$

and

$$(G_1+rI)^{-1}=\frac{1}{\Delta}\left(\begin{array}{cc|cc|c|cc} r_2 & -b & & & & & \\ -c & r_2 & & & & & \\ \hline & & r_2 & -b & & O & \\ & & -c & r_2 & & & \\ \hline & & & & \ddots & & \\ \hline & & O & & & r_2 & -b \\ & & & & & -c & r_2 \end{array}\right)_{(m\times m)} \tag{6.4.6}$$

and

$$(G_2+rI)^{-1}=\frac{1}{\Delta}\left[\begin{array}{c|cc|c|cc|c}
r_2 & & & & & & -c \\
\hline
 & r_2 & -b & & & & \\
 & -c & r_2 & & & & \\
\hline
 & & & \ddots & O & & \\
\hline
 & & & O & r_2 & -b & \\
 & & & & -c & r_2 & \\
\hline
-b & & & & & & r_2
\end{array}\right]_{(m\times m)} . \qquad (6.4.7)$$

Hence, using (6.3.1), (6.4.4)-(6.4.7) we obtain the following equations for the computation of *the generalised AGE scheme:*

(i) at level $(p+\frac{1}{2})$

$$\begin{bmatrix} u_1^{(p+\frac{1}{2})} \\ u_2^{(p+\frac{1}{2})} \\ u_3^{(p+\frac{1}{2})} \\ \vdots \\ \vdots \\ u_{m-1}^{(p+\frac{1}{2})} \\ u_m^{(p+\frac{1}{2})} \end{bmatrix} = \frac{1}{\Delta}\left[\begin{array}{cc|cc|c|cc}
r_2 & -b & & & & & \\
-c & r_2 & & & & & \\
\hline
 & & r_2 & -b & & & \\
 & & -c & r_2 & & O & \\
\hline
 & & & & \ddots & & \\
 & & O & & & & \\
\hline
 & & & & & r_2 & -b \\
 & & & & & -c & r_2
\end{array}\right]\begin{bmatrix} r_1u_1^{(p)}-cu_m^{(p)}+f_1 \\ r_1u_2^{(p)}-bu_3^{(p)}+f_2 \\ -cu_2^{(p)}+r_1u_3^{(p)}+f_3 \\ r_1u_4^{(p)}-bu_5^{(p)}+f_4 \\ -cu_4^{(p)}+r_1u_5^{(p)}+f_5 \\ \vdots \\ r_1u_{m-2}^{(p)}-bu_{m-1}^{(p)}+f_{m-2} \\ -cu_{m-2}^{(p)}+r_1u_{m-1}^{(p)}+f_{m-1} \\ -bu_1^{(p)}+r_1u_m^{(p)}+f_m \end{bmatrix}$$

(6.4.8)

which leads to:

$$\left.\begin{array}{l}
\left.\begin{array}{ll}
u_i^{(p+\frac{1}{2})} = (Au_{i-1}^{(p)}+Bu_i^{(p)}+Cu_{i+1}^{(p)}+Du_{i+2}^{(p)}+E_i)/\Delta, & \\
\text{and } u_{i+1}^{(p+\frac{1}{2})} = (\tilde{A}u_{i-1}^{(p)}+\tilde{B}u_i^{(p)}+\tilde{C}u_{i+1}^{(p)}+\tilde{D}u_{i+2}^{(p)}+\tilde{E}_i)/\Delta, &
\end{array}\right\} i=1,3,\ldots,m-1 \\
\text{with } \quad u_0^{(p)} = u_m^{(p)} \text{ and } u_{m+1}^{(p)} = u_1^{(p)}
\end{array}\right\} \qquad (6.4.9)$$

where,

$$\left.\begin{array}{ll} & A = -cr_2, \quad B = r_1r_2, \; C = -br_1, \; D = b^2, \; E_i = r_2f_i-bf_{i+1} \\ \text{and} & \tilde{A} = c^2, \; \tilde{B} = -cr_1, \; \tilde{C} = r_1r_2, \; \tilde{D} = -br_2, \; \tilde{E}_i = -cf_i+r_2f_{i+1}. \end{array}\right\} \quad (6.4.9a)$$

(ii) at level (p+1)

$$\begin{pmatrix} u_1^{(p+1)} \\ u_2^{(p+1)} \\ \vdots \\ \vdots \\ u_{m-1}^{(p+1)} \\ u_m^{(p+1)} \end{pmatrix} = \frac{1}{\Delta} \left(\begin{array}{c|cc|c|cc|c} r_2 & & & & & & -c \\ \hline & r_2 & -b & & & & \\ & -c & r_2 & & & & \\ \hline & & & \ddots & & \bigcirc & \\ \hline & & & \bigcirc & r_2 & -b & \\ & & & & -c & r_2 & \\ \hline -b & & & & & & r_2 \end{array}\right) \begin{pmatrix} r_3u_1^{(p)}+cu_m^{(p)}+r_4u_1^{(p+\frac{1}{2})} \\ r_3u_2^{(p)}+bu_3^{(p)}+r_4u_2^{(p+\frac{1}{2})} \\ cu_2^{(p)}+r_3u_3^{(p)}+r_4u_3^{(p+\frac{1}{2})} \\ r_3u_4^{(p)}+bu_5^{(p)}+r_4u_4^{(p+\frac{1}{2})} \\ cu_4^{(p)}+r_3u_5^{(p)}+r_4u_5^{(p+\frac{1}{2})} \\ \vdots \\ r_3u_{m-2}^{(p)}+bu_{m-1}^{(p)}+r_4u_{m-2}^{(p+\frac{1}{2})} \\ cu_{m-2}^{(p)}+r_3u_{m-1}^{(p)}+r_4u_{m-1}^{(p+\frac{1}{2})} \\ bu_1^{(p)}+r_3u_m^{(p)}+r_4u_m^{(p+\frac{1}{2})} \end{pmatrix}$$

(6.4.10)

This gives,

$$\left.\begin{array}{ll} & \left.\begin{array}{l} u_i^{(p+1)} = (Pu_{i-1}^{(p)}+Qu_i^{(p)}+Ru_{i-1}^{(p+\frac{1}{2})}+Su_i^{(p+\frac{1}{2})})/\Delta \\ u_{i+1}^{(p+1)} = (\tilde{P}u_{i+1}^{(p)}+\tilde{Q}u_{i+2}^{(p)}+\tilde{R}u_{i+1}^{(p+\frac{1}{2})}+\tilde{S}u_{i+2}^{(p+\frac{1}{2})})/\Delta \end{array}\right\} i=1,3,\ldots,m-1 \\ \text{with} & u_0 = u_m \text{ and } u_1 = u_{m+1} \text{ at both alternating levels} \end{array}\right\} \quad (6.4.11)$$

where,

$$\left.\begin{array}{ll} & P = c(r_2-r_3), \; Q = r_2r_3-bc, \quad R = -cr_4, \; S = r_2r_4 \\ \text{and} & \tilde{P} = r_2r_3-bc, \; \tilde{Q} = b(r_2-r_3), \; \tilde{R} = r_2r_4 \text{ and } \tilde{S} = -br_4. \end{array}\right\} \quad (6.4.11a)$$

(b) m odd (odd number of points)

We have,

$$G_1 = \left[\begin{array}{c|cc|cc|c|cc}
a/2 & & & & & & & \\ \hline
 & a/2 & b & & & & & \\
 & c & a/2 & & & & & \\ \hline
 & & & a/2 & b & & \bigcirc & \\
 & & & c & a/2 & & & \\ \hline
 & & & & & \ddots & & \\ \hline
 & & & \bigcirc & & & a/2 & b \\
 & & & & & & c & a/2
\end{array}\right]_{(m\times m)}, \tag{6.4.12}$$

$$G_2 = \left[\begin{array}{cc|cc|c|c}
a/2 & b & & & & c \\
c & a/2 & & & & \\ \hline
 & & a/2 & b & & \\
 & & c & a/2 & & \\ \hline
 & & & & \ddots \quad \bigcirc & \\
 & & & \bigcirc & & \\ \hline
b & & & & & a/2
\end{array}\right]_{(m\times m)} \tag{6.4.13}$$

and

$$(G_1+rI)^{-1} = \frac{1}{\Delta}\left[\begin{array}{c|cc|cc|c|cc}
\Delta/r_2 & & & & & & & \\ \hline
 & r_2 & -b & & & & & \\
 & -c & r_2 & & & & & \\ \hline
 & & & r_2 & -b & & & \\
 & & & -c & r_2 & & & \\ \hline
 & & & & & \ddots & \bigcirc & \\
 & & & & & \bigcirc & & \\ \hline
 & & & & & & r_2 & -b \\
 & & & & & & -c & r_2
\end{array}\right] \tag{6.4.14}$$

and

$$(G_2+rI)^{-1}=\frac{1}{\Delta(\Delta_1)}\left(\begin{array}{cc|cc|c|cc|c}
r_2\Delta & -b\Delta & & & & & & -c\Delta \\
-c\Delta & \Delta^2/r_2 & & & & & & c^2\Delta/r_2 \\
\hline
 & & r_2\Delta_1 & -b\Delta_1 & & & & \\
 & & -c\Delta_1 & r_2\Delta_1 & & & & \\
\hline
 & & & & \ddots & 0 & & \\
\hline
 & & & & 0 & r_2\Delta_1 & -b\Delta_1 & \\
 & & & & & -c\Delta_1 & r_2\Delta_1 & \\
\hline
-b\Delta & b^2\Delta/r_2 & & & & & & \Delta^2/r_2
\end{array}\right)$$

(6.4.15)

where $\Delta_1 = r_2^2-2bc$. (6.4.16)

We shall now derive the AGE equations using (6.3.1) and (6.4.12)-(6.4.15):

(i) level (p+½)

$$\begin{pmatrix} u_1^{(p+\frac{1}{2})} \\ u_2^{(p+\frac{1}{2})} \\ \vdots \\ \vdots \\ u_{m-1}^{(p+\frac{1}{2})} \\ u_m^{(p+\frac{1}{2})} \end{pmatrix} = \frac{1}{\Delta}\left(\begin{array}{c|cc|c|cc}
\Delta/r_2 & & & & & \\
\hline
 & r_2 & -b & & & \\
 & -c & r_2 & & & \\
\hline
 & & & \begin{matrix} \ddots & 0 \\ 0 & \ddots \end{matrix} & & \\
\hline
 & & & & r_2 & -b \\
 & & & & -c & r_2
\end{array}\right)\begin{pmatrix} r_1u_1^{(p)}-bu_2^{(p)}-cu_m^{(p)}+f_1 \\ -cu_1^{(p)}+r_1u_2^{(p)}+f_2 \\ r_1u_3^{(p)}-bu_4^{(p)}+f_3 \\ -cu_3^{(p)}+r_1u_4^{(p)}+f_4 \\ \vdots \\ r_1u_{m-2}^{(p)}-bu_{m-1}^{(p)}+f_{m-2} \\ -cu_{m-2}^{(p)}+r_1u_{m-1}^{(p)}+f_{m-1} \\ r_1u_m^{(p)}+f_m-bu_1^{(p)} \end{pmatrix}$$

(6.4.17)

Hence, we obtain

$$\left.\begin{array}{ll}
u_1^{(p+\frac{1}{2})} = (r_1u_1^{(p)}-bu_2^{(p)}-cu_m^{(p)}+f_1)/r_2 & \\
\left.\begin{array}{l}
u_i^{(p+\frac{1}{2})} = (Au_{i-1}^{(p)}+Bu_i^{(p)}+Cu_{i+1}^{(p)}+Du_{i+2}^{(p)}+E_i)/\Delta \\
\text{and}\quad u_{i+1}^{(p+\frac{1}{2})} = (\tilde{A}u_{i-1}^{(p)}+\tilde{B}u_i^{(p)}+\tilde{C}u_{i+1}^{(p)}+\tilde{D}u_{i+2}^{(p)}+\tilde{E}_i)/\Delta
\end{array}\right\} & i=2,4,\ldots,m-3,m-1
\end{array}\right\}$$

(6.4.17a)

with $u_1^{(p)} = u_{m+1}^{(p)}$,

where,

$$\left.\begin{aligned} &A = -cr_2,\ B = r_1r_2,\ \ C = -br_1,\ \ D = b^2,\ \ E_i = r_2f_i - bf_{i+1} \\ \text{and}\quad &\tilde{A} = c^2,\ \ \tilde{B} = -cr_1,\ \ \tilde{C} = r_1r_2,\ \ \tilde{D} = -br_2 \text{ and } \tilde{E}_i = -cf_i + r_2f_{i+1} \end{aligned}\right\} \tag{6.4.17b}$$

(ii) level (p+1)

$$\begin{pmatrix} u_1^{(p+1)} \\ u_2^{(p+1)} \\ \vdots \\ \\ \\ u_{m-1}^{(p+1)} \\ u_m^{(p+1)} \end{pmatrix} = \frac{1}{\Delta(\Delta_1)} \left(\begin{array}{cc|cc|c|cc|c} r_2\Delta & -b\Delta & & & & & & -c\Delta \\ -c\Delta & \Delta^2/r_2 & & & & & & c^2\Delta/r_2 \\ \hline & & r_2\Delta_1 & -b\Delta_1 & & & & \\ & & -c\Delta_1 & r_2\Delta_1 & & & & \\ \hline & & & & \ddots & & O & \\ \hline & & & & & r_2\Delta_1 & -b\Delta_1 & \\ & & O & & & -c\Delta_1 & r_2\Delta_1 & \\ \hline -b\Delta & b^2\Delta/r_2 & & & & & & \Delta^2/r_2 \end{array}\right) \begin{pmatrix} r_3u_1^{(p)} + bu_2^{(p)} + cu_m^{(p)} + r_4u_1^{(p+\frac{1}{2})} \\ cu_1^{(p)} + r_3u_2^{(p)} + r_4u_2^{(p+\frac{1}{2})} \\ r_3u_3^{(p)} + bu_4^{(p)} + r_4u_3^{(p+\frac{1}{2})} \\ cu_3^{(p)} + r_3u_4^{(p)} + r_4u_4^{(p+\frac{1}{2})} \\ \vdots \\ r_3u_{m-2}^{(p)} + bu_{m-1}^{(p)} + r_4u_{m-2}^{(p+\frac{1}{2})} \\ cu_{m-2}^{(p)} + r_3u_{m-1}^{(p)} + r_4u_{m-1}^{(p+\frac{1}{2})} \\ bu_1^{(p)} + r_3u_m^{(p)} + r_4u_m^{(p+\frac{1}{2})} \end{pmatrix} \tag{6.4.18}$$

which leads to,

$$\left.\begin{aligned} u_1^{(p+1)} &= (P_1u_1^{(p)} + P_2u_2^{(p)} + P_3u_m^{(p)} + P_4u_1^{(p+\frac{1}{2})} + P_5u_2^{(p+\frac{1}{2})} + P_6u_m^{(p+\frac{1}{2})})/\Delta_1 \\ u_2^{(p+1)} &= (Q_1u_1^{(p)} + Q_2u_2^{(p)} + Q_3u_m^{(p)} + Q_4u_1^{(p+\frac{1}{2})} + Q_5u_2^{(p+\frac{1}{2})} + Q_6u_m^{(p+\frac{1}{2})})/\Delta_1 \end{aligned}\right|$$

$$\left.\begin{aligned}
\left.\begin{aligned}
u_i^{(p+1)} &= (Pu_i^{(p)}+Qu_{i+1}^{(p)}+Ru_i^{(p+\frac{1}{2})}+Su_{i+1}^{(p+\frac{1}{2})})/\Delta \\
u_{i+1}^{(p+1)} &= (\tilde{P}u_i^{(p)}+\tilde{Q}u_{i+1}^{(p)}+\tilde{R}u_i^{(p+\frac{1}{2})}+\tilde{S}u_{i+1}^{(p+\frac{1}{2})})/\Delta
\end{aligned}\right\} i=3,5,\ldots,m-2 \\
u_m^{(p+1)} = (R_1u_1^{(p)}+R_2u_2^{(p)}+R_3u_m^{(p)}+R_4u_1^{(p+\frac{1}{2})}+R_5u_2^{(p+\frac{1}{2})}+R_6u_m^{(p+\frac{1}{2})})/\Delta_1
\end{aligned}\right\} \tag{6.4.19}$$

where,

$$\left.\begin{aligned}
&P_1 = r_2r_3-2bc,\ P_2 = b(r_2-r_3),\ P_3 = c(r_2-r_3),\ P_4 = r_2r_4\ , \\
&\qquad\qquad\qquad\qquad P_5 = -br_4\ ,\ P_6 = -cr_4\ , \\
&Q_1 = c(r_2-r_3),\ Q_2 = -bc+r_3\Delta/r_2,\ Q_3 = -c^2(r_2-r_3)/r_2,\ Q_4 = -cr_4, \\
&\qquad\qquad\qquad\qquad Q_5 = r_4\Delta/r_2,\ Q_6 = r_4c^2/r_2\ , \\
&P = r_2r_3-bc,\ Q = b(r_2-r_3),\ R = r_2r_4,\ S = -br_4 \\
&\tilde{P} = c(r_2-r_3),\ \tilde{Q} = r_2r_3-bc,\ \tilde{R} = -cr_4,\ \tilde{S} = r_2r_4 \\
&R_1 = b(r_2-r_3),\ R_2 = b^2(r_3-r_2)/r_2,\ R_3 = -bc+r_3\Delta/r_2,\ R_4 = -br_4, \\
&\qquad\qquad\qquad R_5 = b^2r_4/r_2\ ,\quad R_6 = \Delta r_4/r_2\ .
\end{aligned}\right\} \tag{6.4.19a}$$

The iterative process is continued for each alternate sweep until convergence is reached.

6.5 THE AGE METHOD TO SOLVE THE DIFFUSION-CONVECTION EQUATION

Consider the following diffusion-convection equation (cf. (3.11.1)),

$$\frac{\partial U}{\partial t} = \varepsilon \frac{\partial^2 U}{\partial x^2} - k \frac{\partial U}{\partial x} , \qquad (6.5.1)$$

with its Dirichlet boundary conditions at x=0,1 specified by (6.2.1a) At the point $(x_i, t_{j+\frac{1}{2}})$, the derivatives in the equation (6.5.1) are approximated by finite differences given by (5.2.1)-(5.2.3). Thus, we are led to the following generalised formula,

$$\begin{aligned}&-(E\theta_2+K\alpha_3)u_{i-1,j+1}+[1+E(\theta_1+\theta_2)-K(\alpha_1-\alpha_3)]u_{i,j+1}-(E\theta_1-K\alpha_1)u_{i+1,j+1}\\ &= (E\theta_4+K\alpha_2)u_{i-1,j}+[1-E(\theta_3+\theta_4)+K(\alpha_4-\alpha_2)]u_{ij}+(E\theta_3-K\alpha_4)u_{i+1,j},\\ &\qquad\qquad i=1,2,\ldots,m,\end{aligned} \qquad (6.5.2)$$

where,

$$\lambda = \Delta t/(\Delta x)^2, \quad E = \varepsilon\lambda \ , \ K = \tfrac{1}{2}k\lambda\Delta x \ . \qquad (6.5.2a)$$

Different choices of the weighting parameters α,θ lead to a variety of finite-difference schemes. As examples, we have,

(a) $\theta_1=\theta_2=0$, $\alpha_1=\alpha_3=0$, $\theta_3=\theta_4=1$, $\alpha_2=\alpha_4=1$. Equation (6.5.2) reduces to

$$u_{i,j+1} = (E+K)u_{i-1,j}+(1-2E)u_{ij}+(E-K)u_{i+1,j} , \qquad (6.5.3)$$

which is *the classical explicit* scheme with $O([\Delta x]^2+\Delta t)$ accuracy.

(b) $\theta_1=\theta_2=1$, $\alpha_1=\alpha_3=1$, $\theta_3=\theta_4=0$, $\alpha_2=\alpha_4=0$ gives the following *fully implicit* scheme,

$$-(E+K)u_{i-1,j+1}+(1+2E)u_{i,j+1}-(E-K)u_{i+1,j+1} = u_{ij} \qquad (6.5.4)$$

with $O([\Delta x]^2+\Delta t)$ accuracy.

(c) $\theta_1=\theta_2=\theta_3=\theta_4=\frac{1}{2}$, $\alpha_2=\alpha_3=1$, $\alpha_1=\alpha_4=0$ gives

$$\begin{aligned}-\tfrac{1}{2}(E+2K)u_{i-1,j+1}+(1+E+K)u_{i,j+1}-\tfrac{1}{2}Eu_{i+1,j+1} &= \tfrac{1}{2}(E+2K)u_{i-1,j}+\\ &(1-E-K)u_{ij}+\tfrac{1}{2}Eu_{i+1,j} ,\end{aligned} \qquad (6.5.5)$$

which is *the Crank-Nicolson scheme with upwinding* with $O(\Delta x+[\Delta t]^2)$ accuracy.

(d) $\theta_1=\theta_2=\theta_3=\theta_4=\frac{1}{2}$, $\alpha_1=\alpha_2=\alpha_3=\alpha_4=\frac{1}{2}$ gives,

$$-\tfrac{1}{2}(E+K)u_{i-1,j+1}+(1+E)u_{i,j+1}-\tfrac{1}{2}(E-K)u_{i+1,j+1} = \tfrac{1}{2}(E+K)u_{i-1,j}$$

$$+(1-E)u_{ij}+\tfrac{1}{2}(E-K)u_{i+1,j}, \qquad (6.5.6)$$

which is *the Crank-Nicolson* scheme with $O([\Delta x]^2+[\Delta t]^2)$ accuracy.

The generalised finite-difference equation (6.5.2) generates a tridiagonal system of linear equations of the form (6.2.3) and (6.2.4) where,

$$c = -(E\theta_2+K\alpha_3),\quad a = 1+E(\theta_1+\theta_2)-K(\alpha_1-\alpha_3),\ b = -(E\theta_1-K\alpha_1);$$

$$f_1 = eu_{0j}-cu_{0,j+1}+fu_{1j}+gu_{2j}\ ,$$

$$f_i = eu_{i-1,j}+fu_{ij}+gu_{i+1,j},\ i=2,3,\ldots,m-1,$$

$$f_m = eu_{m-1,j}+fu_{mj}+gu_{m+1,j}-bu_{m+1,j+1}\ ,$$

with $e = E\theta_4+K\alpha_2$, $f = 1-E(\theta_3+\theta_4)+K(\alpha_4-\alpha_2)$, $g = E\theta_3-K\alpha_4$.

The AGE algorithm is applied to the system (6.5.2) or (6.2.4) with the usual splitting of the coefficient matrix A resulting in exactly the same equations for computation as was found in Section 6.3.

6.6 THE AGE METHOD TO SOLVE A FIRST ORDER HYPERBOLIC EQUATION

A first order hyperbolic equation takes the general form,

$$\frac{\partial U}{\partial t} = -\frac{\partial U}{\partial x} + k(x,t),\ 0\leq x\leq 1,\ 0<t\leq T\ . \qquad (6.6.1)$$

The Dirichlet boundary conditions of (6.2.1a) may be specified on the above hyperbolic problem. Following the arguments in Section 2.9, equation (6.6.1) is replaced at the point $(x_i, t_{j+\theta})$ by the difference analogue (cf. equation (2.9.8)),

$$-\tfrac{1}{2}\lambda\theta u_{i-1,j+1}+u_{i,j+1}+\tfrac{1}{2}\lambda\theta u_{i+1,j+1} = \tfrac{1}{2}\lambda(1-\theta)u_{i-1,j}+u_{ij}-\tfrac{1}{2}\lambda(1-\theta)u_{i+1,j}$$
$$+\Delta t k_{i,j+\theta},\ i=1,2,\ldots,m. \qquad (6.6.2)$$

Again, this tridiagonal system of linear equations can be presented in the matrix form (6.2.3/4) where,

$$c = -\tfrac{1}{2}\lambda\theta,\quad a = 1,\ b = \tfrac{1}{2}\lambda\theta\ ;$$

$$f_1 = u_{1j}+\tfrac{1}{2}\lambda(1-\theta)(u_{0j}-u_{2j})+\tfrac{1}{2}\lambda\theta u_{0,j+1}+\Delta t k_{1,j+\theta}, \qquad (6.6.2a)$$

$$f_i = \tfrac{1}{2}\lambda(1-\theta)(u_{i-1,j}-u_{i+1,j})+u_{ij}+\Delta t k_{i,j+\theta},\ i=2,3,\ldots,m-2,m-1,$$

and $$f_m = u_{mj}+\tfrac{1}{2}\lambda(1-\theta)(u_{m-1,j}-u_{m+1,j})-\tfrac{1}{2}\lambda\theta u_{m+1,j+1}+\Delta t k_{m,j+\theta}.$$

We note that equation (6.6.2) corresponds to the explicit (classical), fully implicit (centred-in-distance, backward-in-time) and the Crank-Nicolson type (centred-in-distance, centred-in-time) formula when θ=0,1 and $\tfrac{1}{2}$ with accuracies to the order of $O(\Delta x+\Delta t)$, $O([\Delta x]^2+\Delta t)$ and $O([\Delta x]^2+[\Delta t]^2)$ respectively.

When the AGE algorithm is implemented on the above system, we will arrive at exactly the same form of equations along the lines $(p+\tfrac{1}{2})$ and $(p+1)$ as those that we derived for the heat equation (6.2.1). These explicit equations are given by (6.3.15)-(6.3.16), (6.3.18)-(6.3.19) and (6.3.23)-(6.3.25).

It must be mentioned, however, that the constituent matrices G_1 and G_2 are not symmetric since from (6.6.2a) $c\neq b$. In fact both matrices (without loss of generality, we take the case m odd only) have one real and (m-1) complex eigenvalues given by $\frac{1}{2}, \frac{1}{2}(1+\lambda\theta i_c)$ (of multiplicity $\frac{1}{2}(m-1)$) and $\frac{1}{2}(1-\lambda\theta i_c)$ (also of multiplicity $\frac{1}{2}(m-1)$) where $i_c=\sqrt{-1}$. The same analysis to prove Theorem 6.1 cannot, therefore, be used to establish the condition (6.2.15) for convergence. However, an attempt made to employ the inequality (6.2.16) turned out to be unsuccessful. We have,

$$(rI-G_1)(G_1+rI)^{-1} = \frac{1}{\Delta}\left(\begin{array}{c|cc|c|cc} \Delta r_1/r_2 & & & & & \\ \hline & (r_1r_2+bc) & -b(r_1+r_2) & & & \\ & -c(r_1+r_2) & (r_1r_2+bc) & & \multicolumn{2}{c}{O} \\ \hline & & & \ddots & & \\ \hline & & O & & (r_1r_2+bc) & -b(r_1+r_2) \\ & & & & -c(r_1+r_2) & (r_1r_2+bc) \end{array}\right) \tag{6.6.3}$$

and similarly,

$$(rI-G_2)(G_2+rI)^{-1} = \frac{1}{\Delta}\left(\begin{array}{cc|c|cc|c} (r_1r_2+bc) & -b(r_1+r_2) & & & & \\ -c(r_1+r_2) & r_1r_2+bc & & & & \\ \hline & & \ddots & & O & \\ & O & & & & \\ \hline & & & (r_1r_2+bc) & -b(r_1+r_2) & \\ & & & -c(r_1+r_2) & (r_1r_2+bc) & \\ \hline & & & & & \frac{\Delta r_1}{r_2} \end{array}\right) \tag{6.6.4}$$

By virtue of (6.6.2a), we obtain,

$$r_1 = r-\tfrac{1}{2},\ r_2 = r+\tfrac{1}{2},\ r_1r_2+bc = r^2-\tfrac{1}{4}(1+\lambda^2\theta^2)\ ,$$
$$-b(r_1+r_2) = -\lambda\theta r,\ -c(r_1+r_2) = \lambda\theta r \text{ and } \Delta = (r+\tfrac{1}{2})^2+\tfrac{1}{4}\lambda^2\theta^2. \tag{6.6.5}$$

Hence by taking the L_∞ norm on inequality (6.2.16), we find that

$$||(rI-G_1)(G_1+rI)^{-1}||_\infty = ||(rI-G_2)(G_2+rI)^{-1}||_\infty$$

$$= \max\left\{\frac{|r-\frac{1}{2}|}{(r+\frac{1}{2})}\ ,\ \frac{|r^2-1/4(1+\lambda^2\theta^2)|+\lambda\theta r}{(r+\frac{1}{2})^2\ +1/4\lambda^2\theta^2}\right\},$$

$$\text{for } \lambda, r>0 \text{ and } 0\leq\theta\leq 1. \tag{6.6.6}$$

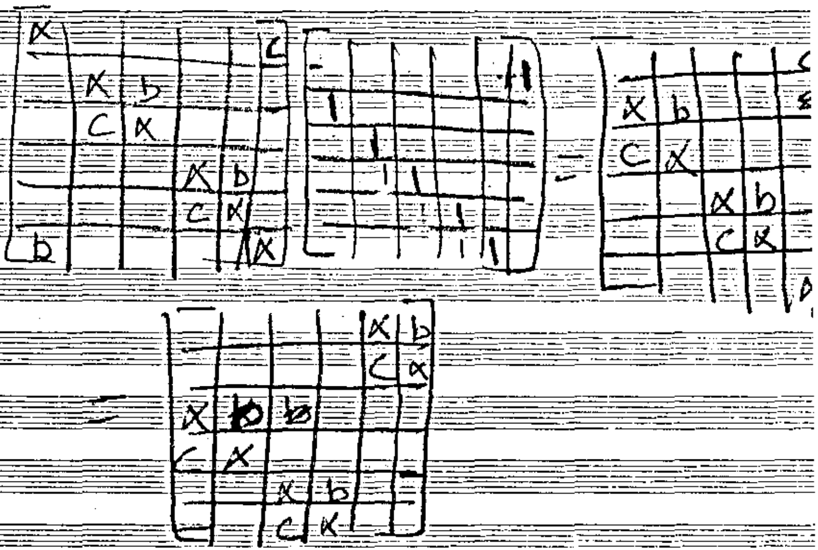
If, for example, we prescribe $\lambda=4$, $r=1$, and $\theta=1$, we get,

$$||(rI-G_1)(G_1+rI)^{-1}||_\infty = ||(rI-G_2)(G_2+rI)^{-1}||_\infty$$

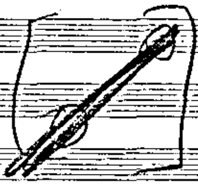
$$= 1.16$$

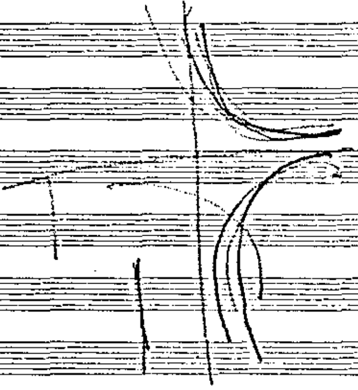
which leads to $||\tilde{M}(r)||_\infty > 1$ (6.6.7)

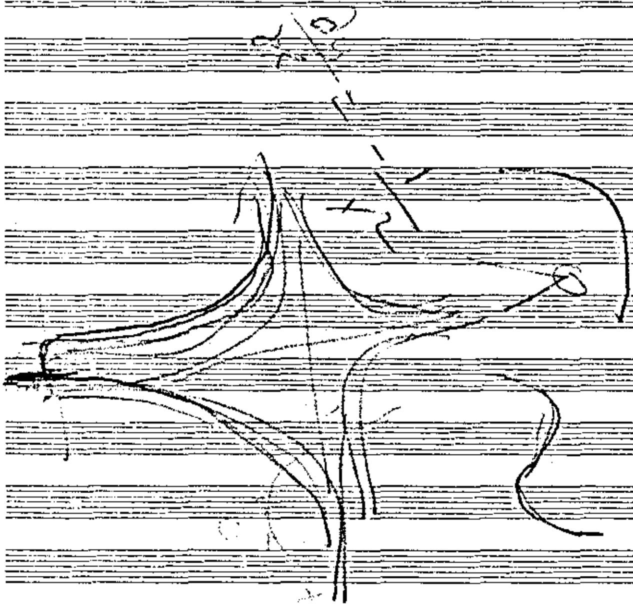
and this clearly does not satisfy the *sufficient* condition for convergence (i.e. $||\tilde{M}(r)||<1$). The failure of this test does not necessarily imply non-convergence of the AGE iterative process. It serves to confirm the theoretical difficulty that arises in dealing with unsymmetric matrices. A direct derivation of the eigenvalues of the AGE iteration matrix therefore becomes necessary and this can be very cumbersome if not impossible. An alternative to this analytical approach is to evaluate them numerically by means of, for example, *the power method* to obtain the dominant eigenvalue. This would enable us to show that $\rho(M(r)<1$.

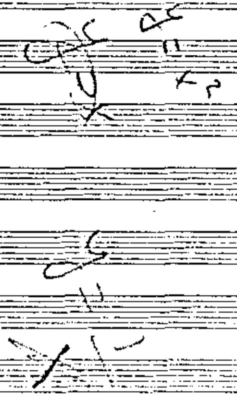


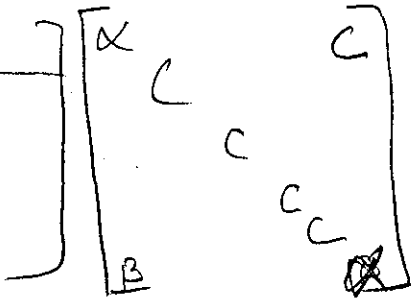
$$\begin{bmatrix} & 3\\ 2 & \end{bmatrix}\begin{bmatrix} & 1\\ 1 & \end{bmatrix} = \begin{bmatrix}3 & 2\\ & 2\end{bmatrix}\begin{bmatrix} & 2\\ 3 & \end{bmatrix}$$

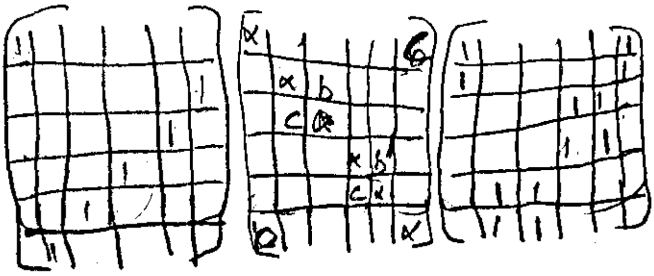


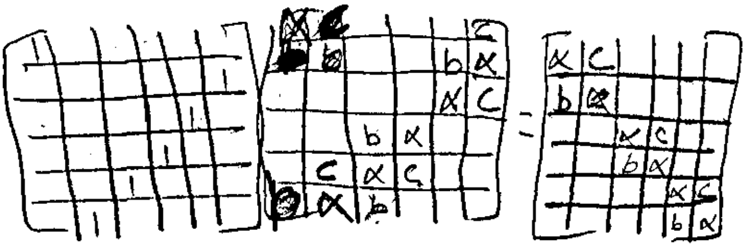


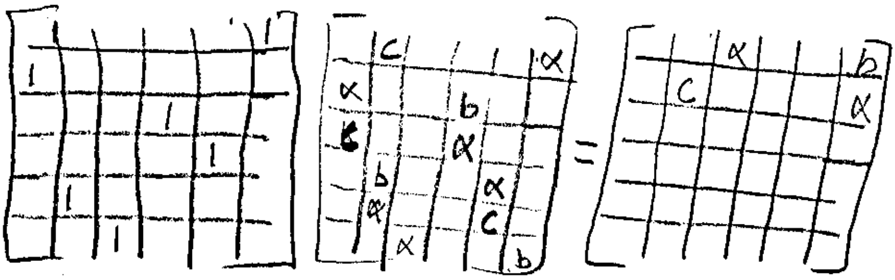


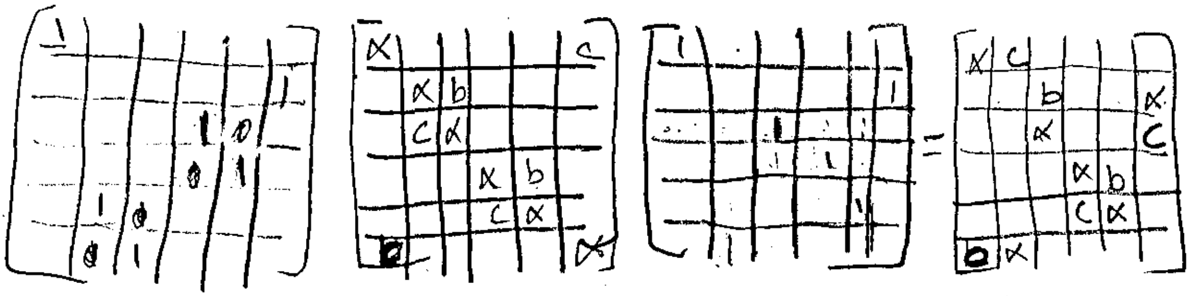


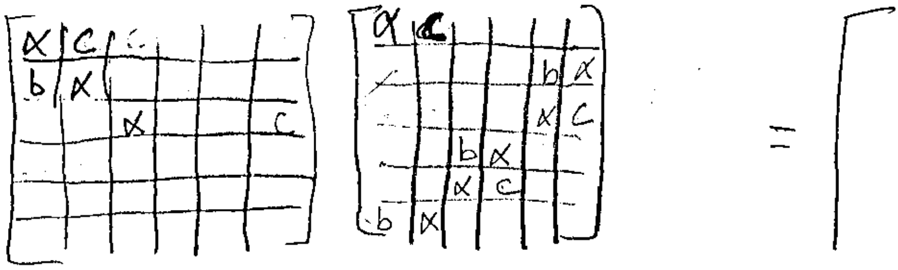












6.7 THE AGE METHOD TO SOLVE A SECOND ORDER HYPERBOLIC (WAVE) EQUATION

(a) *The Wave Equation with Dirichlet Boundary Conditions*

We seek the solution to the wave equation,

$$\frac{\partial^2 U}{\partial x^2} = \frac{\partial^2 U}{\partial t^2}\ , \quad 0\leqslant x\leqslant 1,\ 0\leqslant t<T\ , \tag{6.7.1}$$

subject to the following auxiliary conditions,

$$\left.\begin{aligned} U(x,0) &= f(x)\ ,\\ \frac{\partial U}{\partial t}(x,0) &= g(x)\ , \end{aligned}\right\} \quad 0\leqslant x\leqslant 1 \tag{6.7.1a}$$

$$\left.\begin{aligned} U(0,t) &= h(t)\ ,\\ U(1,t) &= k(t)\ . \end{aligned}\right\} \quad 0\leqslant t<T \tag{6.7.1b}$$

Following our discussion of Section 2.13, the general three-level implicit formula approximating (6.7.1) at the point $(i\Delta x,j\Delta t)$ is,

$$\frac{1}{(\Delta t)^2}\,\delta_t^2 u_{ij} = \frac{1}{(\Delta x)^2}[\alpha\delta_x^2 u_{i,j+1}+(1-2\alpha)\delta_x^2 u_{ij}+\alpha\delta_x^2 u_{i,j-1}] \tag{6.7.2}$$

where α, a weighting factor takes the values of $\alpha\geqslant 1/4$ for stability. The truncation error is $O([\Delta x]^2+[\Delta t]^2)$ for $\alpha=1/4$ and $\frac{1}{2}$. On expanding (6.7.2) we obtain,

$$\begin{aligned} -\alpha\lambda^2 u_{i-1,j+1}+(1+2\alpha\lambda^2)u_{i,j+1}-\alpha\lambda^2 u_{i+1,j+1} &= (1-2\alpha)\lambda^2 u_{i-1,j}+2(1-(1-2\alpha)\lambda^2)\\ &\quad u_{ij}+(1-2\alpha)\lambda^2 u_{i+1,j}+\alpha\lambda^2 u_{i-1,j-1}-(1+2\alpha\lambda^2)u_{i,j-1}+\alpha\lambda^2 u_{i+1,j-1},\\ &\qquad i=1,2,\ldots,m \end{aligned} \tag{6.7.3}$$

which gives a tridiagonal system of equations that can be displayed in the matrix form (6.2.3/4) where,

$$c = -\alpha\lambda^2,\quad a = 1+2\alpha\lambda^2,\ b = -\alpha\lambda^2,\ \lambda = \frac{\Delta t}{\Delta x}\ ;$$

$$\begin{aligned} f_1 = {} & [2(1-(1-2\alpha)\lambda^2)u_{1j}+(1-2\alpha)\lambda^2 u_{2j}]+[-(1+2\alpha\lambda^2)u_{1,j-1}+\alpha\lambda^2 u_{2,j-1}]+\\ & \alpha\lambda^2[u_{0,j+1}+u_{0,j-1}]+(1-2\alpha)\lambda^2 u_{0j}\ ; \end{aligned}$$

$$\begin{aligned} f_i = {} & (1-2\alpha)\lambda^2 u_{i-1,j}+2(1-(1-2\alpha)\lambda^2)u_{ij}+(1-2\alpha)\lambda^2 u_{i+1,j}+\alpha\lambda^2 u_{i-1,j-1}-\\ & (1+2\alpha\lambda^2)u_{i,j-1}+\alpha\lambda^2 u_{i+1,j-1},\ i=2,3,\ldots,m-1\ ; \end{aligned}$$

and $f_m = [(1-2\alpha)\lambda^2 u_{m-1,j} + 2(1-(1-2\alpha)\lambda^2)u_{mj}] + [\alpha\lambda^2 u_{m-1,j-1} - (1+2\alpha\lambda^2)u_{m,j-1}]$

$$+\alpha\lambda^2[u_{m+1,j+1}+u_{m+1,j-1}] + (1-2\alpha)\lambda^2 u_{m+1,j} \ .$$

The u values on the first time level are given by the initial condition. Values on the second time level are obtained from applying the forward difference approximation to first order, at t=0,

$$\frac{\partial U}{\partial t}(x_i,0) \simeq \frac{u_{i,1}-u_{i,0}}{\Delta t} \qquad (6.7.3a)$$

$$= g_i \text{ from } (6.7.1a),$$

giving $u_{i1}=u_{i0}+\Delta t g_i$. Solutions on the third and subsequent time levels are generated iteratively by applying the AGE algorithm on lines (p+½) and (p+1) and utilising the same explicit equations used by the first order hyperbolic equation of (6.6.1), i.e. the formulae (6.3.15)-(6.3.16), (6.3.18)-(6.3.19) and (6.3.23)-(6.3.25) for both cases of even and odd number of intervals.

(b) The Wave Equation with Derivative Boundary Condition

We will again solve the wave equation (6.7.1) together with the initial conditions (6.7.1a) but now the boundary conditions (6.7.1b) are replaced by,

$$\frac{\partial U}{\partial x}(0,t) = h(t) \ ,$$

and $$U(1,t) = k(t) \ . \qquad (6.7.4)$$

A central difference approximation is used to represent the boundary condition at x=0, i.e.,

$$\frac{u_{1j}-u_{-1,j}}{2\Delta x} = h_j$$

giving $$u_{-1,j} = u_{1j} - 2\Delta x h_j \ . \qquad (6.7.5)$$

Similarly we have,

$$u_{-1,j-1} = u_{1,j-1} - 2\Delta x h_{j-1} \ , \qquad (6.7.5a)$$

and $$u_{-1,j+1} = u_{1,j+1} - 2\Delta x h_{j+1} , \tag{6.7.5b}$$

Now, returning to equation (6.7.3) and putting i=0, we get,

$$-\alpha\lambda^2 u_{-1,j+1}+(1-2\alpha\lambda^2)u_{0,j+1}-\alpha\lambda^2 u_{1,j+1} = (1-2\alpha)\lambda^2 u_{-1,j}+2(1-(1-2\alpha)\lambda^2)u_{0,j}$$
$$+(1-2\alpha)\lambda^2 u_{1,j}+\alpha\lambda^2 u_{-1,j-1}-(1+2\alpha\lambda^2)u_{0,j-1}+\alpha\lambda^2 u_{1,j-1} . \tag{6.7.6}$$

The insertion of the calculated values (6.7.5)-(6.7.5b) into equation (6.7.6) at the fictitious points (-1,j-1), (-1,j) and (-1,j+1) yields,

$$(1+2\alpha\lambda^2)u_{0,j+1}-2\alpha\lambda^2 u_{1,j+1} = 2(1-(1-2\alpha)\lambda^2)u_{0j}+2(1-2\alpha)\lambda^2 u_{1j}-(1+2\alpha\lambda^2)$$
$$u_{0,j-1}+2\alpha\lambda^2 u_{1,j-1}-2(\Delta x)\alpha\lambda^2 h_{j-1}-2(\Delta x)(1-2\alpha)\lambda^2 h_j-$$
$$2(\Delta x)\alpha\lambda^2 h_{j+1} . \tag{6.7.7}$$

Hence when the tridiagonal system of equations incorporating (6.7.7) and (6.7.3) (for i=2,3,...,m-1,m) is written in matrix form, we have,

$$A\underline{u} = \underline{f} , \tag{6.7.8}$$

where,

$$A = \begin{pmatrix} a & b_1 & & & & \\ c & a & b & & & \\ & c & a & b & & 0 \\ & & \ddots & \ddots & \ddots & \\ & 0 & & c & a & b \\ & & & & c & a \end{pmatrix}_{(m+1)\times(m+1)} , \tag{6.7.9}$$

$\underline{u} = (u_0,u_1,\ldots,u_{m-1},u_m)^T$, (the unknown values at level (j+1))

$\underline{f} = (f_0,f_1,\ldots,f_{m-1},f_m)^T$, (the known values at the preceding levels)

and $c = -\alpha\lambda^2$, $a = 1+2\alpha\lambda^2$, $b = -\alpha\lambda^2$, $b_1 = -2\alpha\lambda^2$.

As in (a), the solutions on the first time level are obtained from the initial condition and the values on the second level from (6.7.3a).

To be in conformity with previous convention, we shall take the values of u in the order of $u_1, u_2, \ldots, u_M$ where $M=m+1$. Hence, u_1 corresponds with the value of u at $x=0$, u_M with u at $x=(M-1)\Delta x=m\Delta x$ and u_{M+1} with u at $x=M\Delta x=(m+1)\Delta x=1$ on the line segment $0 \leq x \leq 1$. We then have, from (6.7.8),

$$\begin{pmatrix} a & b_1 & & & & & \\ c & a & b & & & O & \\ & c & a & b & & & \\ & & \ddots & \ddots & \ddots & & \\ & O & & & c & a & b \\ & & & & & c & a \end{pmatrix}_{M\times M} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{M-1} \\ u_M \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{M-1} \\ f_M \end{pmatrix} \qquad (6.7.10)$$

with

$$f_1 = 2(1-(1-2\alpha)\lambda^2)u_{1j}+2(1-2\alpha)\lambda^2 u_{2j}-(1+2\alpha\lambda^2)u_{1,j-1}+2\alpha\lambda^2 u_{2,j-1}-2(\Delta x)\alpha\lambda^2$$
$$h_{j-1}-2(\Delta x)(1-2\alpha)\lambda^2 h_j-2(\Delta x)\alpha\lambda^2 h_{j+1} ;$$

$$f_i = (1-2\alpha)\lambda^2 u_{i-1,j}+2(1-(1-2\alpha)\lambda^2)u_{ij}+(1-2\alpha)\lambda^2 u_{i+1,j}+\alpha\lambda^2 u_{i-1,j-1}-(1+2\alpha\lambda^2)$$
$$u_{i,j-1}+\alpha\lambda^2 u_{i+1,j+1}, \quad i=2,3,\ldots,M-1;$$

and $$f_M = [(1-2\alpha)\lambda^2 u_{M-1,j}+2(1-(1-2\alpha)\lambda^2)u_{Mj}]+[\alpha\lambda^2 u_{M-1,j-1}-(1+2\alpha\lambda^2)u_{M,j-1}]$$
$$+\alpha\lambda^2[u_{M+1,j+1}+u_{M+1,j-1}]+(1-2\alpha)\lambda^2 u_{M+1,j} .$$

If we assume *M odd*, then with the usual splitting $A=G_1+G_2$, G_1 takes exactly the same form (with m replaced by M) as in (6.2.6). The other constituent matrix G_2, however, is given by,

$$G_2 = \left[\begin{array}{cc|cc|c|cc|c} a/2 & b_1 & & & & & & \\ c & a/2 & & & & & & \\ \hline & & a/2 & b & & & & \\ & & c & a/2 & & & & \\ \hline & & & & \ddots & & 0 & \\ \hline & & 0 & & & a/2 & b & \\ & & & & & c & a/2 & \\ \hline & & & & & & & a/2 \end{array}\right]_{M\times M} . \qquad (6.7.11)$$

Similarly $(G_1+rI)^{-1}$ is given by (6.3.8/9) whilst $(G_2+rI)^{-1}$ turns out to be,

$$(G_2+rI)^{-1} = \frac{1}{\Delta}\left[\begin{array}{cc|cc|c|cc|c} \Delta r_2/\Delta_2 & -\Delta b_1/\Delta_2 & & & & & & \\ -\Delta c/\Delta_2 & \Delta r_2/\Delta_2 & & & & & & \\ \hline & & r_2 & -b & & & & \\ & & -c & r_2 & & & & \\ \hline & & & & \ddots & & 0 & \\ \hline & & 0 & & & r_2 & -b & \\ & & & & & -c & r_2 & \\ \hline & & & & & & & \Delta/r_2 \end{array}\right]_{M\times M} \qquad (6.7.12)$$

where $\Delta_2 = r_2^2-b_1c$. (6.7.13)

It is clear that the required explicit equations for the computation of the AGE scheme at level $(p+\frac{1}{2})$ remain the same as (6.3.15/15a) except the first equation which is replaced by $u_1^{(p+\frac{1}{2})}=(r_1u_1^{(p)}-b_1u_2^{(p)}+f_1)/r_2$.

We arrive at the following AGE equations at level (p+1):

$$\left.\begin{aligned} u_1^{(p+1)} &= [r_2(r_3u_1^{(p)}+b_1u_2^{(p)}+r_4u_1^{(p+\frac{1}{2})})-b_1(cu_1^{(p)}+r_3u_2^{(p)}+r_4u_2^{(p+\frac{1}{2})})]/\Delta_2 \\ u_2^{(p+1)} &= [-c(r_3u_1^{(p)}+b_1u_2^{(p)}+r_4u_1^{(p+\frac{1}{2})})+r_2(cu_1^{(p)}+r_3u_2^{(p)}+r_4u_2^{(p+\frac{1}{2})})]/\Delta_2 \end{aligned}\right|$$

$$\left.\begin{aligned} u_i^{(p+1)} &= (Pu_i^{(p+\frac{1}{2})}+Qu_{i+1}^{(p+\frac{1}{2})}+Ru_i^{(p)}+Su_{i+1}^{(p)})/\Delta \\ u_{i+1}^{(p+1)} &= (\tilde{P}u_i^{(p+\frac{1}{2})}+\tilde{Q}u_{i+1}^{(p+\frac{1}{2})}+\tilde{R}u_i^{(p)}+\tilde{S}u_{i+1}^{(p)})/\Delta \qquad i=3,4,\ldots,M-2 \\ \text{and}\quad u_M^{(p+1)} &= (r_3u_M^{(p)}+r_4u_M^{(p+\frac{1}{2})})/r_2 \end{aligned}\right\} \quad (6.7.14)$$

where P, $\tilde{P}$, Q, $\tilde{Q}$, R, $\tilde{R}$, S and $\tilde{S}$ are given by (6.3.19).

If *M is even*, we find that,

$$G_1 = \left[\begin{array}{cc:cc:c:cc:cc} a/2 & b_1 & & & & & & & \\ c & a/2 & & & & & & & \\ \hdashline & & a/2 & b & & & & & \\ & & c & a/2 & & & & & \\ \hdashline & & & & \ddots & & O & & \\ \hdashline & & & & & a/2 & b & & \\ & & O & & & c & a/2 & & \\ \hdashline & & & & & & & a/2 & b \\ & & & & & & & c & a/2 \end{array}\right]_{(M\times M)} \quad (6.7.15)$$

with G_2 as in (6.2.36),

$$(G_1+rI)^{-1} = \frac{1}{\Delta}\left[\begin{array}{cc:cc:c:cc:cc} \Delta r_2/\Delta_2 & -\Delta b_1/\Delta_2 & & & & & & & \\ -\Delta c/\Delta_2 & r_2/\Delta_2 & & & & & & & \\ \hdashline & & r_2 & -b & & & & & \\ & & -c & r_2 & & & O & & \\ \hdashline & & & & \ddots & & & & \\ \hdashline & & & & & r_2 & -b & & \\ & & O & & & -c & r_2 & & \\ \hdashline & & & & & & & r_2 & -b \\ & & & & & & & -c & r_2 \end{array}\right] \quad (6.7.16)$$

and $(G_2+rI)^{-1}$ as in (6.3.22). Hence, the explicit expressions for the generated AGE scheme at level $(p+\frac{1}{2})$ are given by,

$$\left.\begin{aligned}
u_1^{(p+\frac{1}{2})} &= (r_2(r_1u_1^{(p)}+f_1)-b_1(r_1u_2^{(p)}-bu_3^{(p)}+f_2))/\Delta_2\\
u_2^{(p+\frac{1}{2})} &= (-c(r_1u_1^{(p)}+f_1)+r_2(r_1u_2^{(p)}-bu_3^{(p)}+f_2))/\Delta_2\\
&\left.\begin{aligned}
u_i^{(p+\frac{1}{2})} &= (Au_{i-1}^{(p)}+Bu_i^{(p)}+Cu_{i+1}^{(p)}+Du_{i+2}^{(p)}+E_i)/\Delta\\
\text{and}\quad u_{i+1}^{(p+\frac{1}{2})} &= (\widetilde{A}u_{i-1}^{(p)}+\widetilde{B}u_i^{(p)}+\widetilde{C}u_{i+1}^{(p)}+\widetilde{D}u_{i+2}^{(p)}+\widetilde{E}_i)/\Delta
\end{aligned}\right\}i=3,5,\ldots,M-1
\end{aligned}\right\}\quad(6.7.17)$$

where $A,\widetilde{A},B,\widetilde{B},C,\widetilde{C},D,\widetilde{D}$, E_i and $\widetilde{E}_i$ are obtained from (6.3.23a). The same equations in (6.3.24) remain valid for use at level (p+1).

6.8 THE AGE METHOD TO SOLVE NON-LINEAR PARABOLIC AND HYPERBOLIC EQUATIONS

The concept of the AGE method is now extended to a variety of *non-linear* problems.

(i) Solving the equation: $\frac{\partial U}{\partial t} = \frac{\partial^2 U^n}{\partial x^2}$ *, $n \geq 2$.*

We shall now consider implementing the AGE algorithm on the following non-linear parabolic problem of the form,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U^n}{\partial x^2}\ ,\ n \geq 2\ , \tag{6.8.1}$$

given the initial condition,

$$U(x,0) = f(x),\ 0<x<1\ ,$$

and the boundary conditions,

$$\left.\begin{aligned} U(0,t) &= g(t)\ , \\ \text{and} \quad U(1,t) &= h(t)\ . \end{aligned}\right\} 0<t\leq T\ . \tag{6.8.1a}$$

The equation (6.8.1) is approximated at the grid points by finite-difference schemes and we shall adopt the approach of Richtmyer and Lee to linearise them which result in tridiagonal systems of equations as before.

(a) Richtmyer's linearisation

The non-linear equation (6.8.1) is approximated by the implicit weighted average difference formula,

$$\frac{1}{\Delta t}(u_{i,j+1}-u_{ij}) = \frac{1}{(\Delta x)^2}[\theta\delta_x^2(u^n_{i,j+1})+(1-\theta)\delta_x^2(u^n_{ij})] \tag{6.8.2}$$

with $i=1,2,\ldots,m$.

As in the linear case $n=1$, the above formula corresponds to the fully implicit, the Crank-Nicolson and the Douglas schemes when $\theta=1,\frac{1}{2}$ and $(6\lambda-1)/12\lambda$ respectively with $\lambda = \frac{\Delta t}{(\Delta x)^2}$. By resorting to the Taylor series expansion of $u^n_{i,j+1}$ about the point (x_i,t_j) we have,

$$\begin{aligned} u^{n}_{i,j+1} &= u^{n}_{ij}+(\Delta t)\,\frac{\partial u^{n}_{ij}}{\partial t}+\ldots \\ &= u^{n}_{ij}\;+(\Delta t)\,\frac{\partial u^{n}_{ij}}{\partial u_{ij}}\;\frac{\partial u_{ij}}{\partial t}+\ldots \quad . \end{aligned}$$

Hence to terms of order n, the approximation,

$$u^{n}_{i,j+1} = u^{n}_{ij} + nu^{n-1}_{ij}(u_{i,j+1}-u_{ij}) \;, \qquad (6.8.3)$$

replaces the non-linear unknown $u^{n}_{i,j+1}$ by an approximation which is linear in $u_{i,j+1}$. Now, if we let,

$$w_{i} = u_{i,j+1}-u_{ij} \;, \qquad (6.8.4)$$

then using (6.8.2) and (6.8.3) we obtain,

$$\begin{aligned} \frac{1}{\Delta t}\,w_{i} &= \frac{1}{(\Delta x)^{2}}[\theta\delta^{2}_{x}(u^{n}_{ij}+nu^{n-1}_{ij}w_{i})+(1-\theta)\,\delta^{2}_{x}u^{n}_{ij}] \\ &= \frac{1}{(\Delta x)^{2}}[n\theta\delta^{2}_{x}u^{n-1}_{ij}w_{i}+\delta^{2}_{x}u^{n}_{ij}] \\ &= \frac{1}{(\Delta x)^{2}}[n\theta(u^{n-1}_{i-1,j}w_{i-1}-2u^{n-1}_{ij}w_{i}+u^{n-1}_{i+1,j}w_{i+1})+(u^{n}_{i-1,j}-2u^{n}_{ij}+u^{n}_{i+1,j})] \end{aligned}$$

$$(6.8.5)$$

which gives a set of linear equations for the w_{i}. Now the system of equations (6.8.5) can be written in the more compact matrix form as *(with n=2)*,

$$A\underline{w} = \underline{f} \;, \qquad (6.8.6)$$

which is solved for $\underline{w}$ and by means of (6.8.4), the solution at time level (j+1) is given by,

$$\underline{u}_{j+1} = \underline{w} + \underline{u}_{j} \quad . \qquad (6.8.7)$$

The coefficient matrix A takes the following tridiagonal form,

$$A = \begin{bmatrix} a_{1j} & b_{1j} & & & \\ c_{2j} & a_{2j} & b_{2j} & & \bigcirc \\ & \ddots & \ddots & \ddots & \\ \bigcirc & & c_{m-1,j} & a_{m-1,j} & b_{m-1,j} \\ & & & c_{mj} & a_{mj} \end{bmatrix}_{(m\times m)} \qquad (6.8.8)$$

where $a_{ij} = 1+4\lambda\theta u_{ij}$, $i=1,2,\ldots,m;$

$b_{ij} = -2\lambda\theta u_{i+1,j}$, $i=1,2,\ldots,m-1;$

and $c_{ij} = -2\lambda\theta u_{i-1,j}$, $i=2,3,\ldots,m.$

The entries of the right hand side vector of (6.8.6) are given by

$$f_1 = (-2\lambda u_{1j})u_{1j}+(\lambda u_{2j})u_{2j}+\lambda u_{0j}[u_{0j}+2\theta(u_{0,j+1}-u_{0j})];$$

$$f_i = (\lambda u_{i-1,j})u_{i-1,j}+(-2\lambda u_{ij})u_{ij}+(\lambda u_{i+1,j})u_{i+1,j},\ i=2,3,\ldots,m-1;$$

and $$f_m = (\lambda u_{m-1,j})u_{m-1,j}+(-2\lambda u_{mj})u_{mj}+\lambda u_{m+1,j}[u_{m+1,j}+2\theta(u_{m+1,j+1}-u_{m+1,j})].$$

If we assume *m odd* then when A is split into the sum of the matrices G_1 and G_2, these constituent matrices take the forms,

$$G_1 = \left[\begin{array}{c|cc|cc|c|cc} a_{1j}/2 & & & & & & & \\ \hline & a_{2j}/2 & b_{2j} & & & & & \\ & c_{3j} & a_{3j}/2 & & & & & \\ \hline & & & a_{4j}/2 & b_{4j} & & \bigcirc & \\ & & & c_{5j} & a_{5j}/2 & & & \\ \hline & & & & & \ddots & & \\ \hline & & \bigcirc & & & & a_{m-1,j}/2 & b_{m-1,j} \\ & & & & & & c_{m,j} & a_{m,j}/2 \end{array}\right]_{(m\times m)} \qquad (6.8.9)$$

and

$$G_2 = \left[\begin{array}{cc|cc|c|cc|c}
a_{1j}/2 & b_{1j} & & & & & & \\
c_{2j} & a_{2j}/2 & & & & & & \\ \hline
 & & a_{3j}/2 & b_{3j} & & & & \\
 & & c_{4j} & a_{4j}/2 & & & & \\ \hline
 & & & & \ddots & & O & \\ \hline
 & & & O & & a_{m-2,j}/2 & b_{m-2,j} & \\
 & & & & & c_{m-1,j} & a_{m-1,j}/2 & \\ \hline
 & & & & & & & a_{mj}/2
\end{array}\right]_{(m\times m)} .$$

(6.8.10)

If we define

$$G^{(i)} = \begin{bmatrix} \frac{1}{2}a_{2i,j}+r & b_{2i,j} \\ c_{2i+1,j} & \frac{1}{2}a_{2i+1,j}+r \end{bmatrix} \tag{6.8.11}$$

then $\alpha_i = |G^{(i)}|$

$$= (\tfrac{1}{2}a_{2i,j}+r)(\tfrac{1}{2}a_{2i+1,j}+r)-b_{2i,j}c_{2i+1,j} , \tag{6.8.12}$$

and,

$$(G_1+rI)^{-1} = \left[\begin{array}{c|c|c|c|c}
\frac{1}{(r+\frac{1}{2}a_{1j})} & & & & \\ \hline
 & \hat{G}^{(1)} & & & \\ \hline
 & & \hat{G}^{(2)} & & O \\ \hline
 & & & \ddots & \\ \hline
 & & O & & \hat{G}^{(m-1)/2}
\end{array}\right]_{(m\times m)} , \tag{6.8.13}$$

where, $\hat{G}^{(i)} = (G^{(i)})^{-1}$

$$= \frac{1}{\alpha_i} \begin{bmatrix} \frac{1}{2}a_{2i+1,j}+r & -b_{2i,j} \\ -c_{2i+1,j} & \frac{1}{2}a_{2i,j}+r \end{bmatrix} , \; i=1,2,\ldots,\tfrac{1}{2}(m-1). \tag{6.8.14}$$

Similarly, we have,

$$(G_2+rI)^{-1} = \left[\begin{array}{c:c:c:c:c}
\hat{\hat{G}}^{(1)} & & & & \\ \hdashline
 & \hat{\hat{G}}^{(2)} & & & \\ \hdashline
 & & \ddots & \bigcirc & \\ \hdashline
 & \bigcirc & & \hat{\hat{G}}^{(m-1)/2} & \\ \hdashline
 & & & & 1/(\tfrac{1}{2}a_{mj}+r)
\end{array}\right]_{(m\times m)} \qquad (6.8.15)$$

where

$$\hat{\hat{G}}^{(i)} = \frac{1}{\hat{\alpha}_i}\begin{bmatrix} \tfrac{1}{2}a_{2i,j}+r & -b_{2i-1,j} \\ -c_{2i,j} & \tfrac{1}{2}a_{2i-1,j}+r \end{bmatrix} \qquad (6.8.16)$$

with $\hat{\alpha}_i = (\tfrac{1}{2}a_{2i,j}+r)(\tfrac{1}{2}a_{2i-1,j}+r)-b_{2i-1,j}c_{2i,j},\ i=1,2,\ldots,\tfrac{1}{2}(m-1).$

(6.8.17)

We, therefore, arrive at the following equations for the computation of the solution of the non-linear problem using the generalised AGE scheme:

(1) at level (iterate) $(p+\frac{1}{2})$

$$\left.\begin{aligned}
w_1^{(p+\frac{1}{2})} &= (\bar{s}_1 w_1^{(p)} - b_1 w_2^{(p)} + f_1)/\bar{r}_1 \\
w_i^{(p+\frac{1}{2})} &= (A_i w_{i-1}^{(p)} + B_i w_i^{(p)} + C_i w_{i+1}^{(p)} + D_i w_{i+2}^{(p)} + E_i)/\alpha_{i/2} \\
w_{i+1}^{(p+\frac{1}{2})} &= (\tilde{A}_i w_{i-1}^{(p)} + \tilde{B}_i w_i^{(p)} + \tilde{C}_i w_{i+1}^{(p)} + \tilde{D}_i w_{i+2}^{(p)} + \tilde{E}_i)/\alpha_{i/2}
\end{aligned}\right\} i=2,4,\ldots,m-1 \qquad (6.8.18)$$

where,

$$\left.\begin{aligned}
&A_i = -c_i\bar{r}_{i+1},\ B_i = \bar{r}_{i+1}\bar{s}_i,\ C_i = -b_i\bar{s}_{i+1},\ D_i = \begin{cases} 0, & \text{for } i=m-1 \\ b_ib_{i+1}, & \text{otherwise} \end{cases} \\
&\qquad E_i = \bar{r}_{i+1}f_i - b_if_{i+1}, \\
&\tilde{A}_i = c_ic_{i+1},\ \tilde{B}_i = -c_{i+1}\bar{s}_i,\ \tilde{C}_i = \bar{r}_i\bar{s}_{i+1},\ \tilde{D}_i = \begin{cases} 0, & \text{for } i=m-1 \\ -b_{i+1}\bar{r}_i, & \text{otherwise} \end{cases} \\
&\qquad \tilde{E}_i = -c_{i+1}f_i + \bar{r}_if_{i+1}, \\
&\text{with}\quad \bar{r}_i = r+\tfrac{1}{2}a_i \quad \text{and} \quad \bar{s}_i = r-\tfrac{1}{2}a_i,\quad i=1,2,\ldots,m.
\end{aligned}\right\}$$

(6.8.18a)

(2) *at level (iterate) (p+1)*

$$\left.\begin{array}{l} \left.\begin{array}{l} w_i^{(p+1)} = (P_i w_i^{(p)} + Q_i w_{i+1}^{(p)} + R_i w_i^{(p+\frac{1}{2})} + S_i w_{i+1}^{(p+\frac{1}{2})})/\hat{\alpha}_{(i+1)/2} \\ w_{i+1}^{(p+1)} = (\tilde{P}_i w_i^{(p)} + \tilde{Q}_i w_{i+1}^{(p)} + \tilde{R}_i w_i^{(p+\frac{1}{2})} + \tilde{S}_i w_{i+1}^{(p+\frac{1}{2})})/\hat{\alpha}_{(i+1)/2} \end{array}\right\} i=1,3,\ldots,m-2 \\ w_m^{(p+1)} = (\bar{q}_m w_m^{(p)} + d w_m^{(p+\frac{1}{2})})/\bar{r}_m \,, \end{array}\right\} \tag{6.8.19}$$

where,

$$\left.\begin{array}{l} P_i = \bar{r}_{i+1}\bar{q}_i - b_i c_{i+1},\ Q_i = b_i(\bar{r}_{i+1} - \bar{q}_{i+1}),\ R_i = \bar{r}_{i+1}d,\ S_i = -b_i d \\ \text{and}\ \ \tilde{P}_i = c_{i+1}(\bar{r}_i - \bar{q}_i),\ \tilde{Q}_i = -b_i c_{i+1} + \bar{r}_i \bar{q}_{i+1},\ \tilde{R}_i = -c_{i+1}d,\ \tilde{S}_i = \bar{r}_i d \\ \text{with} \\ \bar{r}_i = r + \frac{1}{2}a_i,\ \bar{q}_i = \frac{1}{2}a_i - (1-\omega)r \text{ and } d = (2-\omega)r,\ i=1,2,\ldots,m. \end{array}\right\} \tag{6.8.19a}$$

If on the other hand, we assume that *m even* then we obtain the following forms of the constituent matrices:

$$G_1 = \left(\begin{array}{cc|cc|c|cc} a_{1j}/2 & b_{1j} & & & & & \\ c_{2j} & a_{2j}/2 & & & & & \\ \hline & & a_{3j}/2 & b_{3j} & & & \\ & & c_{4j} & a_{4j}/2 & & \bigcirc & \\ \hline & & & & \ddots & & \\ \hline & & \bigcirc & & & a_{m-1,j}/2 & b_{m-1,j} \\ & & & & & c_{m\ \ ,j} & a_{mj}/2 \end{array}\right)_{(m\times m)} \tag{6.8.20}$$

$$G_2 = \left(\begin{array}{c|cc|c|cc|c} a_{1j}/2 & & & & & & \\ \hline & a_{2j}/2 & b_{2j} & & & & \\ & c_{3j} & a_{3j}/2 & & \bigcirc & & \\ \hline & & & \ddots & & & \\ \hline & & \bigcirc & & a_{m-2,j}/2 & b_{m-2,j} & \\ & & & & c_{m-1,j} & a_{m-1,j}/2 & \\ \hline & & & & & & a_{mj}/2 \end{array}\right)_{(m\times m)} \tag{6.8.21}$$

$$(G_1+rI)^{-1} = \left(\begin{array}{c|c|c|c} \hat{\hat{G}}^{(1)} & & & \\ \hline & \hat{\hat{G}}^{(2)} & & O \\ \hline & & \ddots & \\ \hline & O & & \hat{\hat{G}}^{(m/2)} \end{array}\right)_{(m\times m)} \qquad (6.8.22)$$

and

$$(G_2+rI)^{-1} = \left(\begin{array}{c|c|c|c|c} \frac{1}{(r+\frac{1}{2}a_{1j})} & & & & \\ \hline & \hat{G}^{(1)} & & O & \\ \hline & & \ddots & & \\ \hline & & O & \hat{G}^{\left(\frac{m-2}{2}\right)} & \\ \hline & & & & \frac{1}{(r+\frac{1}{2}a_{mj})} \end{array}\right)_{(m\times m)} \qquad (6.8.23)$$

where $\hat{G}^{(i)}$ and $\hat{\hat{G}}^{(i)}$ are given by (6.8.14) and (6.8.16) respectively. The AGE equations can be derived along similar lines as before and are given by,

(1) at level $(p+\frac{1}{2})$

$$\left.\begin{aligned} w_i^{(p+\frac{1}{2})} &= (A_i w_{i-1}^{(p)} + B_i w_i^{(p)} + C_i w_{i+1}^{(p)} + D_i w_{i+2}^{(p)} + E_i)/\hat{\alpha}_{(i+1)/2} \\ w_{i+1}^{(p+\frac{1}{2})} &= (\tilde{A}_i w_{i-1}^{(p)} + \tilde{B}_i w_i^{(p)} + \tilde{C}_i w_{i+1}^{(p)} + \tilde{D}_i w_{i+2}^{(p)} + \tilde{E}_i)/\hat{\alpha}_{(i+1)/2} \end{aligned}\right\} i=1,3,\ldots,m-1 \qquad (6.8.24)$$

where,

$$\left.\begin{aligned} A_i &= \begin{cases} 0, \text{ for } i=1 \\ -c_i\bar{r}_{i+1}, \text{otherwise} \end{cases}, \quad B_i = \bar{r}_{i+1}\bar{s}_i, \quad C_i = -b_i\bar{s}_{i+1}, \\ D_i &= \begin{cases} 0, \text{ for } i=m-1 \\ b_ib_{i+1}, \text{otherwise} \end{cases}, \quad E_i = \bar{r}_{i+1}f_i - b_if_{i+1}, \\ \tilde{A}_i &= \begin{cases} 0, \text{ for } i=1 \\ c_ic_{i+1}, \text{otherwise} \end{cases}, \quad \tilde{B}_i = -c_{i+1}\bar{s}_i, \quad \tilde{C}_i = \bar{r}_i\bar{s}_{i+1}, \\ \tilde{D}_i &= \begin{cases} 0, \text{ for } i=m-1 \\ -\bar{r}_ib_{i+1}, \text{otherwise} \end{cases}, \quad \tilde{E}_i = -c_{i+1}f_i + \bar{r}_if_{i+1} \end{aligned}\right\} \qquad (6.8.24a)$$

(2) at level (p+1)

$$\left.\begin{array}{l} w_1^{(p+1)} = (\bar{q}_1 w_1^{(p)} + dw_1^{(p+\frac{1}{2})})/\bar{r}_1 \\ \left.\begin{array}{l} w_i^{(p+1)} = (P_i w_i^{(p)} + Q_i w_{i+1}^{(p)} + R_i w_i^{(p+\frac{1}{2})} + S_i w_{i+1}^{(p+\frac{1}{2})})/\alpha_{i/2} \\ w_{i+1}^{(p+1)} = (\widetilde{P}_i w_i^{(p)} + \widetilde{Q}_i w_{i+1}^{(p)} + \widetilde{R}_i w_i^{(p+\frac{1}{2})} + \widetilde{S}_i w_{i+1}^{(p+\frac{1}{2})})/\alpha_{i/2} \end{array}\right\} i=2,4,\ldots,m-2 \\ w_m^{(p+1)} = (\bar{q}_m w_m^{(p)} + dw_m^{(p+\frac{1}{2})})/\bar{r}_m \ , \end{array}\right\} \qquad (6.8.25)$$

where

$$\left.\begin{array}{l} P_i = (\bar{r}_{i+1}\bar{q}_i - b_i c_{i+1}),\ Q_i = b_i(\bar{r}_{i+1} - \bar{q}_{i+1}),\ R_i = d\bar{r}_{i+1},\ S_i = -db_i \ , \\ \widetilde{P}_i = c_{i+1}(\bar{r}_i - \bar{q}_i),\ \widetilde{Q}_i = \bar{r}_i\bar{q}_{i+1} - c_{i+1}b_i,\ \widetilde{R}_i = -dc_{i+1},\ \widetilde{S}_i = d\bar{r}_i \ . \end{array}\right\} \qquad (6.8.25a)$$

The iterative process is continued until the convergence requirement is met.

(b) Lee's three-level linearisation

Lees (1966) considered the non-linear equation,

$$b(U)\frac{\partial U}{\partial t} = \frac{\partial}{\partial x}\{a(U)\frac{\partial U}{\partial x}\} \ , \ a(U)>0,\ b(U)>0 \ , \qquad (6.8.26)$$

and investigated a difference scheme that:

(i) achieved *linearity* in the unknowns $u_{i,j+1}$ by evaluating all coefficients of $u_{i,j+1}$ at a time level of known solution values,

(ii) preserved *stability* by averaging u_{ij} over three time levels, and

(iii) maintained *accuracy* by using central-difference approximations.

Since,

$$\begin{aligned}(\frac{\partial U}{\partial x})_{i,j} &\simeq \frac{1}{\Delta x}(U_{i+\frac{1}{2},j} - U_{i-\frac{1}{2},j}) \\ &= \frac{1}{\Delta x}\,\delta_x U_{ij}\end{aligned}$$

then a central-difference approximation to (6.8.26) is given by

$$\begin{aligned} b(u_{i,j})\,\frac{1}{2\Delta t}(u_{i,j+1} - u_{i,j-1}) &= \frac{1}{\Delta x}\,\delta_x\{a(u_{ij})\,\frac{1}{\Delta x}\,\delta_x u_{ij}\} \\ &= \frac{1}{(\Delta x)^2}\,\delta_x\{a(u_{ij})\,\delta_x\}u_{ij} \ , \qquad (6.8.27)\end{aligned}$$

which reduces to the Richardson formula (3.3.1) when $a(u)=b(u)=1$ and is therefore unconditionally unstable. However, in *the linear constant coefficient case* (see Mitchell and Griffiths (1980), pages 89-92), unconditional stability is obtained by replacing $\delta_x^2 u_{i,j}$ by $\frac{1}{3}\delta_x^2(u_{i,j+1}+u_{ij}+u_{i,j-1})$. Following this procedure, equation (6.8.27) is rewritten as,

$$b(u_{ij})(u_{i,j+1}-u_{i,j-1}) = 2\lambda[a(u_{i+\frac{1}{2},j})(u_{i+1,j}-u_{ij})-a(u_{i-\frac{1}{2},j})(u_{ij}-u_{i-1,j})]$$

and then $u_{i+1,j}$, u_{ij} and $u_{i-1,j}$ are replaced by,

$$\frac{1}{3}(u_{i+1,j+1}+u_{i+1,j}+u_{i+1,j-1}),\ \frac{1}{3}(u_{i,j+1}+u_{ij}+u_{i,j-1}) \text{ and } \frac{1}{3}(u_{i-1,j+1}+u_{i-1,j}+u_{i-1,j-1})$$

respectively. Furthermore, since $u_{i\pm\frac{1}{2},j}$ do not fall on the grid points, we replace $a(u_{i+\frac{1}{2},j})$ and $a(u_{i-\frac{1}{2},j})$ by $a(\frac{u_{i+1,j}+u_{ij}}{2})$ and $a(\frac{u_{ij}+u_{i-1,j}}{2})$ respectively. This leads to *the linearised three-level formula*,

$$\begin{aligned} b(u_{ij})(u_{i,j+1}-u_{i,j-1}) = \tfrac{2}{3}\lambda[\beta^+\{(u_{i+1,j+1}-u_{i,j+1})+(u_{i+1,j}-u_{ij})+ \\ (u_{i+1,j-1}-u_{i,j-1})\}-\beta^-\{(u_{i,j+1}-u_{i-1,j+1})+(u_{ij}-u_{i-1,j}) \\ +(u_{i,j-1}-u_{i-1,j-1})\}] \ , \end{aligned} \qquad (6.8.28)$$

where $$\beta^+ = a(\frac{u_{i+1,j}+u_{ij}}{2}) \text{ and } \beta^- = a(\frac{u_{ij}+u_{i-1,j}}{2}) \ . \qquad (6.8.29)$$

Lees (1966) proved the convergence result for (6.8.28) by showing that for sufficiently small values of Δx and Δt,

$$\max_{i,j} |U_{ij}-u_{ij}| \leq K((\Delta x)^2+(\Delta t)^2) \ ,$$

where K is a constant. For this method to be applied to (6.8.1), it is necessary to write the equation in the *self-adjoint form* as,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U^n}{\partial x^2} \ , \quad n\geq 2 \ ,$$

$$= \frac{\partial}{\partial x}(nU^{n-1} \frac{\partial U}{\partial x}) .$$

On comparing this equation with (6.8.26), we find that for the particular value of n=2, a(U)=2U and b(U)=1 and from (6.8.29),

$$\beta^{+} = u_{i+1,j}+u_{ij} \ , \ \beta^{-} = u_{ij}+u_{i-1,j} \ . \qquad (6.8.30)$$

Hence the formula (6.8.28) becomes,

$$-\frac{2}{3}\lambda\beta^{-}u_{i-1,j+1}+(1+\frac{2}{3}\lambda(\beta^{+}+\beta^{-}))u_{i,j+1}-\frac{2}{3}\lambda\beta^{+}u_{i+1,j+1}$$

$$= \frac{2}{3}\lambda\beta^{-}u_{i-1,j} - \frac{2}{3}\lambda(\beta^{+}+\beta^{-})u_{ij} + \frac{2}{3}\lambda\beta^{+}u_{i+1,j}$$

$$+\frac{2}{3}\lambda\beta^{-}u_{i-1,j-1}+(1-\frac{2}{3}\lambda(\beta^{+}+\beta^{-}))u_{i,j-1} + \frac{2}{3}\lambda\beta^{+}u_{i+1,j-1}, \text{ for } i=1,2,\ldots,m \qquad (6.8.31)$$

which is a tridiagonal system of equations that can be written in the matrix form (6.8.6) (with $\underline{u}$ replacing $\underline{w}$) where A takes the form (6.8.8) and,

$$a_{ij} = 1 + \frac{2}{3}\lambda(\beta^{+}+\beta^{-})$$

$$= 1 + \frac{2}{3}\lambda(u_{i-1,j}+2u_{ij}+u_{i+1,j}), \ i=1,2,\ldots,m;$$

$$b_{ij} = -\frac{2}{3}\lambda\beta^{+}$$

$$= -\frac{2}{3}\lambda(u_{i+1,j}+u_{ij}), \ i=1,2,\ldots,m-1;$$

and $$c_{ij} = -\frac{2}{3}\lambda\beta^{-}u_{i-1,j+1}$$

$$= -\frac{2}{3}\lambda(u_{ij}+u_{i-1,j}), \ i=2,3,\ldots,m.$$

The components of the right hand side vector $\underline{f}$ are given by,

$$f_1 = \frac{2}{3}\lambda[(u_{1j}+u_{0j})(u_{0,j-1}+u_{0,j}+u_{0,j+1})+(u_{2j}+u_{1j})(u_{2,j-1}+u_{2j})-$$

$$(u_{0j}+2u_{1j}+u_{2j})u_{1j}] + [1-\frac{2}{3}\lambda(u_{0j}+2u_{1j}+u_{2j})]u_{1,j-1}$$

$$f_i = \frac{2}{3}\lambda[\beta^{+}(u_{i+1,j}+u_{i+1,j-1})+\beta^{-}(u_{i-1,j}+u_{i-1,j-1})-(\beta^{+}+\beta^{-})u_{ij}]+$$

$$[1-\frac{2}{3}\lambda(\beta^{+}+\beta^{-})]u_{i,j-1}, \text{ for } i= \ 2,3,\ldots,m-1,$$

and

$$f_m = \frac{2}{3}\lambda[(u_{m+1,j}+u_{mj})(u_{m+1,j-1}+u_{m+1,j}+u_{m+1,j+1})+(u_{mj}+u_{m-1,j})(u_{m-1,j-1}+ u_{m-1,j})-(u_{m+1,j}+2u_{mj}+u_{m-1,j})u_{mj}]+[1-\frac{2}{3}\lambda(u_{m+1,j}+2u_{mj}+u_{m-1,j})]u_{m,j-1}.$$

When the AGE procedure is implemented on the above tridiagonal system of equations, we will arrive at the same computational formulae (with $\underline{w}$ replaced by $\underline{u}$) at the $(p+\frac{1}{2})$ and $(p+1)$ iterates as that for Richtmyer's linearisation. This implies that equations (6.8.18)-(6.8.19) (for the case m odd) and the equations (6.8.24)-(6.8.25) (when m is even) will be used for our iterative process.

(ii) Solving Burgers' Equation: $\varepsilon \frac{\partial^2 U}{\partial x^2} = \frac{\partial U}{\partial t} + U\frac{\partial U}{\partial x}$, $\varepsilon>0$.

The *general non-linear parabolic equation* for initial boundary value problems is given by,

$$\frac{\partial U}{\partial t} = \phi(x,t,U,\frac{\partial U}{\partial x}, \frac{\partial^2 U}{\partial x^2}), \quad 0<x<1, \quad 0<t\leq T , \tag{6.8.32}$$

subject to smooth initial and boundary conditions. This problem is well posed in the region (see, for example, Friedman (1964)) if

$$\frac{\partial \phi}{\partial U_{xx}} \geq a > 0 . \tag{6.8.33}$$

If this holds, then the implicit relation (6.8.32) may be solved for $\frac{\partial^2 U}{\partial x^2}$. Thus, we assume the partial differential equation to have the form,

$$\frac{\partial^2 U}{\partial x^2} = \psi(x,t,U,\frac{\partial U}{\partial x}, \frac{\partial U}{\partial t}) , \tag{6.8.34}$$

where the properly posed requirement is

$$\frac{\partial \psi}{\partial U_t} \geq a > 0 . \tag{6.8.35}$$

In some instances, (6.8.34) may be written in the quasi-linear form,

$$\frac{\partial^2 U}{\partial x^2} + f(x,t,U)\frac{\partial U}{\partial x} + g(x,t,U) = p(x,t,U)\frac{\partial U}{\partial t} . \tag{6.8.36}$$

(Specifically Burgers' equation $\varepsilon \frac{\partial^2 U}{\partial x^2} = \frac{\partial U}{\partial t} + U\frac{\partial U}{\partial x}$ is of this form).

(a) The fully implicit form

At the point (i,j+1), (6.8.36) can be approximated by the formula,

$$\frac{1}{(\Delta x)^2}\,\delta_x^2 u_{i,j+1} + \frac{1}{2\Delta x}\,f[i\Delta x,(j+1)\Delta t,u_{ij}]\mu\delta_x u_{i,j+1} + g[i\Delta x,(j+1)\Delta t,u_{ij}] = p[i\Delta x,(j+1)\Delta t,u_{ij}]\,\frac{(u_{i,j+1}-u_{ij})}{\Delta t} \qquad (6.8.37)$$

which contains $u_{i,j+1}$ only linearly and where the difference operators δ and μ are defined by,

$$\delta y_n = y_{n+\frac{1}{2}} - y_{n-\frac{1}{2}} \quad \text{(central)}$$

and

$$\mu y_n = \tfrac{1}{2}[y_{n+\frac{1}{2}}+y_{n-\frac{1}{2}}] \quad \text{(averaging)}\ .$$

Thus, the algebraic problem is *linear and tridiagonal* at each time step. For Burgers' equation, we have the analogue,

$$\frac{\varepsilon}{(\Delta x)^2}\,\delta_x^2 u_{i,j+1} + \frac{1}{2\Delta x}[-u_{ij}]\mu\delta_x u_{i,j+1} = \frac{u_{i,j+1}-u_{ij}}{\Delta t}$$

which leads to

$$-\lambda(\varepsilon + \frac{\Delta x}{4}\,u_{ij})u_{i-1,j+1}+(1+2\varepsilon\lambda)u_{i,j+1}-\lambda(\varepsilon-\frac{\Delta x}{4}\,u_{ij})u_{i+1,j+1} = u_{ij},$$
$$i=1,\ldots,m. \qquad (6.8.38)$$

(b) The Crank-Nicolson form

The application of the Crank-Nicolson concept to the equation (6.8.36) gives,

$$\frac{1}{2(\Delta x)^2}\,\delta_x^2(u_{i,j+1}+u_{ij}) + \frac{1}{2\Delta x}\,f[i\Delta x,(j+\tfrac{1}{2})\Delta t,\tfrac{1}{2}(u_{i,j+1}+u_{ij})\mu\delta_x(u_{i,j+1}+u_{ij})]+g[i\Delta x,(j+\tfrac{1}{2})\Delta t,\tfrac{1}{2}(u_{i,j+1}+u_{ij})] = p[i\Delta x,(j+\tfrac{1}{2})\Delta t,\tfrac{1}{2}(u_{i,j+1}+u_{ij})]\,\frac{(u_{i,j+1}-u_{ij})}{\Delta t}\,, \qquad (6.8.39)$$

which are *non-linear*. For Burgers' equation, these simplify to

$$u_{i,j+1}-u_{ij} = \frac{\varepsilon\lambda}{2}[(u_{i-1,j}-2u_{ij}+u_{i+1,j})+(u_{i-1,j+1}-2u_{i,j+1}+u_{i+1,j+1})] - \frac{\lambda\Delta x}{4}[(u_{i+1,j}-u_{i-1,j})\alpha_{ij}+(u_{i+1,j+1}-u_{i-1,j+1})\beta_{ij}] \qquad (6.8.40)$$

where $\alpha_{ij}=\beta_{ij}=(u_{i,j+1}+u_{ij})/2$. This equation, however, can be *linearised* if we replace α_{ij} by $u_{i,j+1}$ and β_{ij} by u_{ij}. Thus, (6.8.40) becomes,

$$-(\frac{\varepsilon\lambda}{2} + \frac{\Delta x}{4}\,\lambda u_{ij})u_{i-1,j+1}+(1+\varepsilon\lambda+\frac{\Delta x}{4}\lambda(u_{i+1,j}-u_{i-1,j}))u_{i,j+1}-(\frac{\varepsilon\lambda}{2} - \frac{\Delta x}{4}\lambda u_{ij}) u_{i+1,j+1} = \frac{\varepsilon\lambda}{2}u_{i-1,j}+(1-\varepsilon\lambda)u_{ij} + \frac{\varepsilon\lambda}{2}u_{i+1,j},\ i=1,2,\ldots,m\,. \qquad (6.8.41)$$

(c) The Predictor-Corrector form

Non-linear algebraic equations can be avoided if two-step methods called the predictor-corrector methods are used. The 'predictor' gives a first approximation to the solution and the 'corrector' is used repeatedly, if necessary, to provide the final result. If ψ of (6.8.34) assumes the form,

$$\psi = f_1(x,t,U)\frac{\partial U}{\partial t} + f_2(x,t,U)\frac{\partial U}{\partial x} + f_3(x,t,U) \tag{6.8.42a}$$

or

$$\psi = g_1(x,t,U,\frac{\partial U}{\partial x})\frac{\partial U}{\partial t} + g_2(x,t,U,\frac{\partial U}{\partial x}) \tag{6.8.42b}$$

then a predictor-corrector modification of the Crank-Nicolson procedure is possible so that the resulting algebraic problem is linear. The class of equation (6.8.42a) includes Burgers' equation and if ψ is of the form (6.8.42a), then one possibility of the predictor is

$$\frac{1}{(\Delta x)^2}\,\delta_x^2 u_{i,j+\frac{1}{2}} = \psi[i\Delta x,(j+\tfrac{1}{2})\Delta t,u_{ij},\frac{1}{\Delta x}\,\mu\delta_x u_{ij},\frac{2}{\Delta t}(u_{i,j+\frac{1}{2}}-u_{ij})] \tag{6.8.43}$$

for i=1,2,...,m followed by the corrector,

$$\frac{1}{2(\Delta x)^2}\,\delta_x^2[u_{i,j+1}+u_{ij}] = \psi[i\Delta x,(j+\tfrac{1}{2})\Delta t,u_{i,j+\frac{1}{2}},\frac{1}{2\Delta x}\mu\delta_x(u_{i,j+1}+u_{ij}),\frac{1}{\Delta t}(u_{i,j+1}-u_{ij})]\ . \tag{6.8.44}$$

For Burgers' equation, the corresponding predictor-corrector (P-C) pair are given by,

$$P:\varepsilon\lambda u_{i-1,j+\frac{1}{2}}-2(1+\varepsilon\lambda)u_{i,j+\frac{1}{2}}+\varepsilon\lambda u_{i+1,j+\frac{1}{2}} = [\tfrac{1}{2}\lambda(\Delta x)(u_{i+1,j}-u_{i-1,j})-2]u_{ij} \tag{6.8.45}$$

and

$$C: -\frac{\lambda}{2}(\varepsilon+\frac{\Delta x}{2}u_{i,j+\frac{1}{2}})u_{i-1,j+1}+(1+\varepsilon\lambda)u_{i,j+1}-\frac{\lambda}{2}(\varepsilon-\frac{\Delta x}{2}u_{i,j+\frac{1}{2}})u_{i+1,j+1}$$

$$= \frac{\varepsilon\lambda}{2}(u_{i-1,j}-2u_{ij}+u_{i+1,j})+u_{ij}-\frac{\lambda\Delta x}{4}\,u_{i,j+\frac{1}{2}}(u_{i+1,j}-u_{i-1,j})\ . \tag{6.8.46}$$

The above predictor-corrector formulae are known to have second-order accuracy in both space and time (Douglas and Jones (1963)). Note that (6.8.43) is a backward difference equation. One may also use the following modified Crank-Nicolson predictor,

$$\frac{1}{2(\Delta x)^2}\delta_x^2(u_{i,j+\frac{1}{2}}+u_{ij}) = \psi[i\Delta x,(j+\tfrac{1}{2})\Delta t,u_{ij}, \frac{1}{\Delta x}\mu\delta_x u_{ij}, \frac{2}{\Delta t}(u_{i,j+\frac{1}{2}}-u_{ij})]. \tag{6.8.46a}$$

While the procedure (6.8.43) and (6.8.44) leads to a set of linear algebraic equations to solve for ψ when it is of the form (6.8.42a), it does not for ψ of the form (6.8.42b). If (6.8.44) is replaced by,

$$\frac{1}{2(\Delta x)^2}\delta_x^2(u_{i,j+1}+u_{ij}) = \psi[i\Delta x,(j+\tfrac{1}{2})\Delta t,u_{i,j+\frac{1}{2}},\frac{1}{\Delta x}\mu\delta_x u_{i,j+\frac{1}{2}}, \frac{1}{\Delta t}(u_{i,j+1}-u_{ij})] , \tag{6.8.47}$$

the system (6.8.43) and (6.8.47) does produce the desired linear algebraic equations with a local truncation error $O((\Delta x)^2)+O((\Delta t)^{3/2})$.

All of the equations (6.8.38), (6.8.41) and (6.8.46) generate, as before, tridiagonal systems of the form,

$$A\underline{u} = \underline{f} , \tag{6.8.47a}$$

where A takes the same form as (6.8.8), $\underline{u}=(u_{1,j+1},u_{2,j+1},\ldots,u_{m,j+1})^T$ and $\underline{f}=(f_1,f_2,\ldots,f_m)^T$. Hence, we have,

(1) for the implicit form (6.8.38)

$$a_{i,j} = 1+2\varepsilon\lambda , \quad i=1,2,\ldots,m$$

$$b_{i,j} = -\lambda(\varepsilon - \frac{\Delta x}{4}u_{ij}), \; i=1,2,\ldots,m-1$$

$$c_{i,j} = -\lambda(\varepsilon + \frac{\Delta x}{4}u_{i,j}), \; i=2,3,\ldots,m$$

$$f_1 = \lambda(\varepsilon + \frac{\Delta x}{4}u_{1j})u_{0,j+1}+u_{1j} ,$$

$$f_i = u_{ij} , \quad i=2,3,\ldots,m-1$$

and

$$f_m = \lambda(\varepsilon - \frac{\Delta x}{4}u_{mj})u_{m+1,j+1} + u_{mj} ;$$

(2) for the Crank-Nicolson form (6.8.41)

$$a_{i,j} = 1 + \lambda[\varepsilon + \frac{\Delta x}{4}(u_{i+1,j}-u_{i-1,j})], \quad i=1,2,\ldots,m$$

$$b_{i,j} = -\frac{\lambda}{2}(\varepsilon - \frac{\Delta x}{2} u_{ij}), \quad i=1,2,\ldots,m-1$$

$$c_{i,j} = -\frac{\lambda}{2}(\varepsilon + \frac{\Delta x}{2}u_{ij}), \quad i=2,3,\ldots,m$$

$$f_1 = (1-\varepsilon\lambda)u_{1j}+ \frac{\lambda}{2}[\varepsilon(u_{0j}+u_{2j})+(\varepsilon + \frac{\Delta x}{2}u_{1j})u_{0,j+1}]$$

$$f_i = \frac{\varepsilon\lambda}{2}(u_{i-1,j}+u_{i+1,j})+(1-\varepsilon\lambda)u_{ij} \ , \quad i=2,3,\ldots,m-1$$

and

$$f_m = (1-\varepsilon\lambda)u_{mj} + \frac{\lambda}{2}[\varepsilon(u_{m-1,j}+u_{m+1,j})+(\varepsilon - \frac{\Delta x}{2} u_{mj})u_{m+1,j+1}] \ ;$$

(3) for the predictor-corrector form (6.8.46)

$$c_{i,j} = -\frac{\lambda}{2}(\varepsilon + \frac{\Delta x}{2}u_{i,j+\frac{1}{2}}), \quad i=1,2,\ldots,m$$

$$a_{i,j} = 1+\varepsilon\lambda, \quad i=1,2,\ldots,m-1$$

$$b_{i,j} = -\frac{\lambda}{2}(\varepsilon - \frac{\Delta x}{2} u_{i,j+\frac{1}{2}}), \quad i=2,3,\ldots,m$$

$$f_1 = \frac{\lambda}{2}[\varepsilon(u_{0j}-2u_{1j}+u_{2j})+(\varepsilon+ \frac{\Delta x}{2}u_{1,j+\frac{1}{2}})u_{0,j+1} - \frac{\Delta x}{2} u_{1,j+\frac{1}{2}}(u_{2j}-u_{0j})]+u_{1j}$$

$$f_i = \frac{\lambda}{2}[\varepsilon(u_{i-1,j}-2u_{ij}+u_{i+1,j})- \frac{\Delta x}{2}u_{i,j+\frac{1}{2}}(u_{i+1,j}-u_{i-1,j})]+u_{ij}, \quad i=2,3,\ldots,m-1$$

and

$$f_m = \frac{\lambda}{2}[\varepsilon(u_{m-1,j}-2u_{mj}+u_{m+1,j})+(\varepsilon- \frac{\Delta x}{2} u_{m,j+\frac{1}{2}})u_{m+1,j+1} - \frac{\Delta x}{2}u_{m,j+\frac{1}{2}}(u_{m+1,j}-u_{m-1,j})]+u_{mj} \ .$$

When the AGE algorithm is implemented, our iterative process will require the same equations for computation as in (6.8.18) and (6.8.19) (for the case, m odd) and (6.8.24) and (6.8.25) (for the case, m even) with $\underline{w}$ replaced by $\underline{u}$. We note, however, that for the predictor-corrector form, the solutions at the *predictor* stage are obtained using the Thomas elimination algorithm.

(iii) A Non-Linear Example for the Reaction-Diffusion Equation

We shall now consider the following one-dimensional *reaction-diffusion equation* taken from Ramos (1985),

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + S \ , \tag{6.8.48}$$

where $S=U^2(1-U)$ and $-\infty<x<\infty$ and $t\geq 0$. This equation has an exact travelling wave solution given by,

$$U(x,t) = 1/(1+\exp(V(x-Vt))) \ , \tag{6.8.49}$$

where, $$U(-\infty,t) = 1, \ U(\infty,t) = 0 \ , \tag{6.8.50}$$

and V is the *steady-state wave speed* which is equal to $1/\sqrt{2}$. In our numerical experiments, however, equation (6.8.48) was solved in a truncated domain $-50\leq x\leq 400$, where the locations of the boundaries were selected so that they did not influence the wave propagation. In the truncated domain, the following initial condition was used,

$$U(x,0) = 1/(1+\exp(Vx)) \ , \tag{6.8.51}$$

i.e. the initial condition corresponds to the exact solution.

Various schemes in the GE and AGE class of methods are now developed to solve (6.8.48).

(a) GE schemes involving an explicit evaluation of the source term:GE-EXP

Following the same argument as in Section 5.2, a generalised approximation to (6.8.48) at the point $(x_i,t_{j+\frac{1}{2}})$ is

$$\frac{(u_{i,j+1}-u_{ij})}{\Delta t} = \frac{1}{(\Delta x)^2}\{\theta_1\delta_x u_{i+\frac{1}{2},j+1}-\theta_2\delta_x u_{i-\frac{1}{2},j+1}+\theta_3\delta_x u_{i+\frac{1}{2},j}-\theta_4\delta_x u_{i-\frac{1}{2},j}\} + s_{ij}. \tag{6.8.52}$$

By letting $\theta_1=\theta_4=1$ and $\theta_2=\theta_3=0$ in (6.8.52), we obtain the following asymmetric LR approximation,

$$-\lambda u_{i+1,j+1}+(1+\lambda)u_{i,j+1} = \lambda u_{i-1,j}+(1-\lambda)u_{ij}+\Delta t \ s_{ij} \tag{6.8.53}$$

where $\lambda = \frac{\Delta t}{(\Delta x)^2}$. If we choose $\theta_2 = \theta_3 = 1$, and $\theta_1 = \theta_4 = 0$ we arrive at the following RL formula,

$$(1+\lambda)u_{i,j+1} - \lambda u_{i-1,j+1} = \lambda u_{i+1,j} + (1-\lambda)u_{ij} + \Delta t\ s_{ij}$$

or equivalently, at the point $(i+1,j+\frac{1}{2})$,

$$-\lambda u_{i,j+1} + (1+\lambda)u_{i+1,j+1} = (1-\lambda)u_{i+1,j} + \lambda u_{i+2,j} + \Delta t\ s_{i+1,j}. \tag{6.8.54}$$

When we couple equations (6.8.53) and (6.8.54), we obtain,

$$\begin{bmatrix} (1+\lambda) & -\lambda \\ -\lambda & (1+\lambda) \end{bmatrix} \begin{bmatrix} u_{i,j+1} \\ u_{i+1,j+1} \end{bmatrix} = \begin{bmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{bmatrix} \begin{bmatrix} u_{ij} \\ u_{i+1,j} \end{bmatrix} + \begin{bmatrix} \lambda u_{i-1,j} + \Delta t s_{ij} \\ \lambda u_{i+2,j} + \Delta t s_{i+1,j} \end{bmatrix},$$

i.e.,

$$A\underline{u}_{j+1} = B\underline{u}_j + \hat{\underline{u}}_j ,$$

or

$$\underline{u}_{j+1} = A^{-1}(B\underline{u}_j + \hat{\underline{u}}_j)$$

giving

$$\begin{bmatrix} u_{i,j+1} \\ \\ u_{i+1,j+1} \end{bmatrix} = \frac{1}{(1+2\lambda)} \begin{bmatrix} \lambda(1+\lambda)u_{i-1,j} + (1-\lambda^2)u_{ij} + \lambda(1-\lambda)u_{i+1,j} + \lambda^2 u_{i+2,j} + \\ \Delta t[(1+\lambda)s_{ij} + \lambda s_{i+1,j}] \\ \lambda^2 u_{i-1,j} + \lambda(1-\lambda)u_{ij} + (1-\lambda^2)u_{i+1,j} + \lambda(1+\lambda)u_{i+2,j} \\ +\Delta t\,[\lambda s_{ij} + (1+\lambda)s_{i+1,j}] \end{bmatrix}$$

Letting $a_1 = \frac{\lambda(1+\lambda)}{(1+2\lambda)}$, $a_2 = \frac{(1-\lambda^2)}{(1+2\lambda)}$, $a_3 = \frac{\lambda(1-\lambda)}{(1+2\lambda)}$,

$a_4 = \frac{\lambda^2}{(1+2\lambda)}$, $a_5 = \frac{\Delta t(1+\lambda)}{(1+2\lambda)}$ and $a_6 = \frac{\Delta t\lambda}{(1+2\lambda)}$

we have the following set of explicit equations defining the GE schemes at the general points,

$$u_{i,j+1} = a_1 u_{i-1,j} + a_2 u_{ij} + a_3 u_{i+1,j} + a_4 u_{i+2,j} + a_5 s_{ij} + a_6 s_{i+1,j} , \tag{6.8.55}$$

and

and $$u_{i+1,j+1} = a_4u_{i-1,j}+a_3u_{ij}+a_2u_{i+1,j}+a_1u_{i+2,j}+a_6s_{ij}+a_5s_{i+1,j}. \tag{6.8.56}$$

For the ungrouped point at the right end boundary, we have using (6.8.53)

$$u_{m-1,j+1} = \frac{1}{(1+\lambda)}(\lambda(u_{m-2,j}+u_{m,j+1})+(1-\lambda)u_{m-1,j}+\Delta ts_{m-1,j})$$

i.e.,

$$u_{m-1,j+1} = a_7(u_{m-2,j}+u_{m,j+1})+a_8u_{m-1,j}+a_9s_{m-1,j} , \tag{6.8.57}$$

where

$$a_7 = \lambda/(1+\lambda), \quad a_8 = (1-\lambda)/(1+\lambda), \quad a_9 = \Delta t/(1+\lambda) .$$

Similarly, using (6.8.54), the ungrouped point at the left end boundary is given by,

$$u_{1,j+1} = a_8u_{1j}+a_7(u_{2j}+u_{0,j+1})+a_9s_{1j} . \tag{6.8.58}$$

With equations (6.8.55)-(6.8.58) we can then construct the alternating schemes of (S)AGE/(D)AGE-EXP schemes based on their constituent GER and GEL formulae.

(b) GE methods employing the predictor-corrector technique: GE-PC

The formulation is the same as in iii(a) except that the source term S is now approximated by $(s_{i,j}+\bar{s}_i)/2$ where $\bar{s}_i=S(\bar{u}_i)$ and the value $\bar{u}_i$ in the predictor step is determined by the solution of the system of differential equations (the explicit method of lines)

$$\frac{du_i}{dt} = \frac{1}{(\Delta x)^2}[u_{i+1,j}-2u_{i,j}+u_{i-1,j}]+s_i, \quad i=1,2,\ldots,m-1. \tag{6.8.59}$$

(The system (6.8.59) is derived by discretizing the spatial derivatives in (6.8.48) and keeping the time as a continuous variable. The diffusion terms are evaluated at the previous time step). These equations are solved by means of *an explicit, fourth order accurate Runge-Kutta method.*

The solution is then employed in the corrector step by the appropriate GE schemes whose set of explicit equations could be similarly derived as in iii(a) and given by,

$$u_{i,j+1} = a_1u_{i-1,j}+a_2u_{ij}+a_3u_{i+1,j}+a_4u_{i+2,j}+a_5(s_{ij}+\bar{s}_i)/2+a_6(s_{i+1,j}+\bar{s}_{i+1})/2 \tag{6.8.60}$$

and

$$u_{i+1,j+1} = a_4u_{i-1,j}+a_3u_{ij}+a_2u_{i+1,j}+a_1u_{i+2,j}+a_6(s_{ij}+\bar{s}_i)/2+ a_5(s_{i+1,j}+\bar{s}_{i+1})/2\ . \tag{6.8.61}$$

In the same way, the ungrouped point at the right end boundary is given as

$$u_{m-1,j+1} = a_7(u_{m-2,j}+u_{m,j+1})+a_8u_{m-1,j}+a_9(s_{m-1,j}+\bar{s}_{m-1})/2 \tag{6.8.62}$$

and at the left end boundary by,

$$u_{1,j+1} = a_8u_{1j}+a_7(u_{2j}+u_{0,j+1})+a_9(s_{1j}+\bar{s}_1)/2\ . \tag{6.8.63}$$

Having found the basic equations for our GE schemes given by (6.8.60)-(6.8.63) the (S)AGE/(D)AGE-PC methods can be constructed.

(c) The AGE methods employing the predictor-corrector technique on implicit approximation: AGE-PC

As in iii(b), before applying the AGE algorithm on the corrector, the explicit method of lines is again employed to determine the values of u_i from the system of ordinary differential equations,

$$\frac{du_i}{dt} = \frac{1}{(\Delta x)^2}[u_{i+1,j}-2u_{ij}+u_{i-1,j}]+s_i,\ i=1,2,\ldots,m$$

which is solved by means of an explicit, fourth-order accurate, Runge-Kutta method. The solution $\bar{u}_i$ say, is then employed in the following implicit approximation to (6.8.48) in the corrector step,

$$\frac{(u_{i,j+1}-u_{ij})}{\Delta t} = \frac{1}{2(\Delta x)^2}[(u_{i+1,j}-2u_{ij}+u_{i-1,j})+(u_{i+1,j+1}-2u_{i,j+1}+u_{i-1,j+1})]+(s_{ij}+\bar{s}_i)/2 ,$$

where $\bar{s}_i = S(\bar{u}_i)$. This C-N like scheme can be written in tridiagonal matrix form as,

$$-\lambda u_{i-1,j+1}+(2+2\lambda)u_{i,j+1}-\lambda u_{i+1,j+1} = 2u_{ij}+\lambda(u_{i+1,j}-2u_{ij}+u_{i-1,j}) +\Delta t(s_{ij}+\bar{s}_i) \qquad (6.8.64)$$

or

$$A\underline{u} = \underline{f} ,$$

where A takes the same form as in (6.2.3), $\underline{u}=(u_{1,j+1},u_{2,j+1},\ldots,u_{m,j+1})^T$ and $\underline{f}=(f_1,f_2,\ldots,f_m)^T$

with $c = b = -\lambda \; , \; a = 2(1+\lambda)$

$$f_1 = 2u_{1j}+\lambda(u_{0j}+u_{0,j+1}+u_{2j}-2u_{1j})+\Delta t(s_{1j}+\bar{s}_1) ,$$

$$f_i = 2u_{ij}+\lambda(u_{i+1,j}-2u_{ij}+u_{i-1,j})+\Delta t(s_{ij}+\bar{s}_i), \; i=2,3,\ldots,m-1$$

and

$$f_m = 2u_{mj}+\lambda(u_{m+1,j}+u_{m+1,j+1}-2u_{mj}+u_{m-1,j})+\Delta t(s_{mj}+\bar{s}_m).$$

The AGE equations required to solve the system (6.8.64) are given by (6.3.15) and (6.3.18) for the case m odd and (6.3.23) and (6.3.25) for the case m even.

(d) The AGE methods employing time linearisation techniques on implicit approximation: AGE-TL

Four different time linearisation schemes can be derived as approximations to the differential equation (6.8.48).

(1) First time linearisation scheme (1TL)

The familiar fully implicit approximation to (6.8.48) is given by

$$u_{i,j+1} = u_{ij}+\lambda(u_{i-1,j+1}-2u_{i,j+1}+u_{i+1,j+1})+\Delta t s_{i,j+1}, \; i=1,2,\ldots,m. \qquad (6.8.65)$$

The source term $s_{i,j+1}$ can be linearised around the previous time step by means of Taylor's series as

$$s_{i,j+1} = s_{ij} + (\frac{\partial S}{\partial U})_{ij}(u_{i,j+1}-u_{ij}) \ . \qquad (6.8.66)$$

The substitution of the expression in (6.8.66) into (6.8.65) yields the following equation,

$$-\lambda u_{i-1,j+1}+(1+2\lambda-\Delta t(\frac{\partial S}{\partial U})_{ij})u_{i,j+1}-\lambda u_{i+1,j+1} = u_{ij}+\Delta t(s_{ij}-(\frac{\partial S}{\partial U})_{ij}u_{ij}). \qquad (6.8.67)$$

(2) Second time linearisation scheme (2TL)

By applying the Crank-Nicolson concept to the differential equation (6.8.48), we obtain the following approximation,

$$u_{i,j+1} = u_{ij}+\lambda(u_{i+1,j+1}-2u_{i,j+1}+u_{i-1,j+1})/2+\lambda(u_{i+1,j}-2u_{ij}+u_{i-1,j})/2$$
$$+\Delta t(s_{ij}+s_{i,j+1})/2 \text{ for } i=1,2,\ldots,m. \qquad (6.8.68)$$

The substitution of (6.8.66) into (6.8.68) leads to,

$$-\lambda u_{i-1,j+1}+(2+2\lambda-\Delta t(\frac{\partial S}{\partial U})_{ij})u_{i,j+1}-\lambda u_{i+1,j+1} = 2u_{ij}+\lambda(u_{i+1,j}-2u_{ij}+$$
$$u_{i-1,j})+\Delta t(2s_{ij}-(\frac{\partial S}{\partial U})_{ij}u_{ij}) \ . \qquad (6.8.69)$$

(3) Third time linearisation scheme (3TL)

This technique employs a first-order accurate time discretisation and a three-point compact formula for the diffusion terms (Kopal (1961)) and can be written as,

$$u_{i,j+1} = u_{ij} + \lambda \frac{\delta_x^2}{(1+\delta_x^2/12)} u_{i,j+1}+\Delta t s_{i,j+1}, \ i=1,2,\ldots,m$$

or,

$$(1+\frac{\delta_x^2}{12})u_{i,j+1} = (1+\frac{\delta_x^2}{12})u_{ij} +\lambda\delta_x^2u_{i,j+1}+\Delta t(1+\frac{\delta_x^2}{12})s_{i,j+1} \qquad (6.8.70)$$

where δ denotes the usual central difference operator. By substituting (6.8.66) into equation (6.8.70) we obtain,

$$[-\lambda+\frac{1}{12}\{1-\Delta t(\frac{\partial S}{\partial U})_{i-1,j}\}]u_{i-1,j+1}+[2\lambda+\frac{5}{6}\{1-\Delta t(\frac{\partial S}{\partial U})_{i,j}\}]u_{i,j+1}+[-\lambda+\frac{1}{12}$$

$$\{1-\Delta t(\frac{\partial S}{\partial U})_{i+1,j}\}]u_{i+1,j+1}=\frac{1}{12}[1-\Delta t(\frac{\partial S}{\partial U})_{i-1,j}]u_{i-1,j}+\frac{5}{6}[1-\Delta t(\frac{\partial S}{\partial U})_{ij}]u_{ij}+$$

$$\frac{1}{12}[1-\Delta t(\frac{\partial S}{\partial U})_{i+1,j}]u_{i+1,j}+\frac{1}{6}\Delta t[5s_{ij}+\tfrac{1}{2}(s_{i-1,j}+s_{i+1,j})], \quad (6.8.71)$$

$$i=1,2,\dots,m.$$

(4) Fourth time linearisation scheme (4TL)

This scheme employs a second-order accurate time discretisation and a three-point compact formula for the diffusion terms (Kopal (1961)) which is fourth-order accurate in space. The finite-difference form of the 4TL method can be written as,

$$u_{i,j+1}=u_{ij}+\lambda\frac{\delta_x^2}{(1+\delta_x^2/12)}[u_{ij}+u_{i,j+1}]/2+\Delta t[s_{ij}+s_{i,j+1}]/2,$$

$$i=1,2,\dots,m$$

or

$$(2+\frac{\delta_x^2}{6})u_{i,j+1}=(2+\frac{\delta_x^2}{6})u_{ij}+\lambda\,\delta_x^2(u_{ij}+u_{i,j+1})+\Delta t(1+\frac{\delta_x^2}{12})(s_{ij}+s_{i,j+1}). \quad (6.8.72)$$

The substitution of (6.8.66) into (6.8.72) yields

$$\{\frac{1}{6}[1-\frac{\Delta t}{2}(\frac{\partial S}{\partial U})_{i-1,j}]-\lambda\}u_{i-1,j+1}+\{\frac{5}{6}[2-\Delta t(\frac{\partial S}{\partial U})_{i,j}]+2\lambda\}u_{i,j+1}+\{\frac{1}{6}[1-\frac{\Delta t}{2}(\frac{\partial S}{\partial U})_{i+1,j}]$$

$$-\lambda\}u_{i+1,j+1}=\{\frac{1}{6}[1-\frac{\Delta t}{2}(\frac{\partial S}{\partial U})_{i-1,j}]+\lambda\}u_{i-1,j}+\{\frac{5}{6}[2-\Delta t(\frac{\partial S}{\partial U})_{ij}]-2\lambda\}u_{ij}$$

$$+\{\frac{1}{6}[1-\frac{\Delta t}{2}(\frac{\partial S}{\partial U})_{i+1,j}]+\lambda\}u_{i+1,j}+\frac{\Delta t}{6}[S_{i-1,j}+10S_{ij}+S_{i+1,j}]. \quad (6.8.73)$$

We note that equations (6.8.67), (6.8.69), (6.8.71) and (6.8.73) generate tridiagonal systems of the form (6.8.47a). Hence, we have,

(1) for 1TL (6.8.67)

$c_{ij}=-\lambda$, $i=2,3,\dots,m$

$a_{ij}=1+2\lambda-\Delta t(\frac{\partial S}{\partial U})_{ij}$, $i=1,2,\dots,m$

$b_{ij}=-\lambda$, $i=1,2,\dots,m-1$

$f_1=u_{1j}+\Delta t(s_{1j}-(\frac{\partial S}{\partial U})_{1,j}u_{1j})+\lambda u_{0,j+1}$

$$f_i = u_{ij}+\Delta t\,(s_{ij}-(\frac{\partial S}{\partial U})_{ij}u_{ij}), \quad i=2,3,\ldots,m-1$$

and $$f_m = u_{mj}+\Delta t\,(s_{mj}-(\frac{\partial S}{\partial U})_{mj}u_{mj})+\lambda u_{m+1,j+1} ;$$

(2) for 2TL (6.8.69)

$$c_{ij} = -\lambda , \quad i=2,3,\ldots,m$$

$$a_{ij} = 2+2\lambda-\Delta t\,(\frac{\partial S}{\partial U})_{i,j} , \quad i=1,2,\ldots,m$$

$$b_{ij} = -\lambda, \quad i=1,2,\ldots,m-1$$

$$f_1 = 2u_{1j}+\lambda(u_{2j}-2u_{1j})+\Delta t\,(2s_{1j}-(\frac{\partial S}{\partial U})_{1j}u_{1j})+\lambda(u_{0j}+u_{0,j+1})$$

$$f_i = 2u_{ij}+\lambda(u_{i-1,j}-2u_{ij}+u_{i+1,j})+\Delta t\,(2s_{ij}-(\frac{\partial S}{\partial U})_{ij}u_{ij}), \quad i=2,\ldots,m-1$$

and $$f_m = 2u_{mj}+\lambda(u_{m-1,j}-2u_{mj})+\Delta t\,(2s_{mj}-(\frac{\partial S}{\partial U})_{mj}u_{mj})+\lambda(u_{m+1,j}+u_{m+1,j+1});$$

(3) for 3TL (6.8.71)

$$c_{ij} = -\lambda+\frac{1}{12}[1-\Delta t\,(\frac{\partial S}{\partial U})_{i-1,j}] , \quad i=2,3,\ldots,m$$

$$a_{ij} = 2\lambda+\frac{5}{6}[1-\Delta t\,(\frac{\partial S}{\partial U})_{ij}] , \quad i=1,2,\ldots,m$$

$$b_{ij} = -\lambda+\frac{1}{12}[1-\Delta t\,(\frac{\partial S}{\partial U})_{i+1,j}] , \quad i=1,2,\ldots,m-1$$

$$f_1 = \frac{1}{12}[1-\Delta t\,(\frac{\partial S}{\partial U})_{0,j}]\,[u_{0j}-u_{0,j+1}]+\lambda u_{0,j+1}+\frac{5}{6}[1-\Delta t\,(\frac{\partial S}{\partial U})_{1,j}]u_{1j}$$
$$+\frac{1}{6}\,\Delta t\,[5s_{1j}+\tfrac{1}{2}(s_{0j}+s_{2j})]$$

$$f_i = \frac{1}{12}[1-\Delta t\,(\frac{\partial S}{\partial U})_{i-1,j}]u_{i-1,j}+\frac{5}{6}[1-\Delta t\,(\frac{\partial S}{\partial U})_{ij}+\frac{1}{12}[1-\Delta t\,(\frac{\partial S}{\partial U})_{i+1,j}]u_{i+1,j}$$
$$+\frac{1}{6}\Delta t\,[5s_{ij}+\tfrac{1}{2}(s_{i-1,j}+s_{i+1,j})] , \quad i=2,\ldots,m-1,$$

and $$f_m = \frac{1}{12}[1-\Delta t\,(\frac{\partial S}{\partial U})_{m+1,j}]\,[u_{m+1,j}-u_{m+1,j+1}]+\lambda u_{m+1,j+1}+\frac{1}{12}[1-\Delta t\,(\frac{\partial S}{\partial U})_{m-1,j}]$$
$$u_{m-1,j}+\frac{1}{6}\{5\,[1-\Delta t\,(\frac{\partial S}{\partial U})_{mj}]u_{mj}+\Delta t\,[5s_{mj}+\tfrac{1}{2}(s_{m-1,j}+s_{m+1,j})]] ;$$

(4) for 4TL (6.8.73)

$$c_{ij} = \frac{1}{6}[1-\frac{\Delta t}{2}(\frac{\partial S}{\partial U})_{i-1,j}]-\lambda , \quad i=2,3,\ldots,m,$$

$$a_{ij} = \frac{5}{6}[2-\Delta t(\frac{\partial S}{\partial U})_{ij}]+2\lambda, \quad i=1,2,\ldots,m$$

$$b_{ij} = \frac{1}{6}[1- \frac{\Delta t}{2}(\frac{\partial S}{\partial U})_{i+1,j}]-\lambda, \quad i=1,2,\ldots,m-1$$

$$f_1 = \frac{1}{6}[1- \frac{\Delta t}{2}(\frac{\partial S}{\partial U})_{0,j}](u_{0j}-u_{0,j+1})+\lambda(u_{0j}+u_{0,j+1})+[\frac{5}{6}(2-\Delta t(\frac{\partial S}{\partial U})_{1j})-$$
$$2\lambda]u_{1j} + [\frac{1}{6}(1 - \frac{\Delta t}{2}(\frac{\partial S}{\partial U})_{2,j})+\lambda]u_{2j} + \frac{\Delta t}{6}[s_{0j}+10s_{1j}+s_{2j}]$$

$$f_i = [\frac{1}{6}(1- \frac{\Delta t}{2}(\frac{\partial S}{\partial U})_{i-1,j})+\lambda]u_{i-1,j}+[\frac{5}{6}(2-\Delta t(\frac{\partial S}{\partial U})_{ij})-2\lambda]u_{ij} +$$
$$[\frac{1}{6}(1- \frac{\Delta t}{2}(\frac{\partial S}{\partial U})_{i+1,j})+\lambda]u_{i+1,j} + \frac{\Delta t}{6}(s_{i-1,j}+10s_{ij}+s_{i+1,j}),$$
$$i=2,3,\ldots,m-1$$

and
$$f_m = \frac{1}{6}[1- \frac{\Delta t}{2}(\frac{\partial S}{\partial U})_{m+1,j}](u_{m+1,j}-u_{m+1,j+1})+\lambda(u_{m+1,j}+u_{m+1,j+1})$$
$$+[\frac{1}{6}(1- \frac{\Delta t}{2}(\frac{\partial S}{\partial U})_{m-1,j}+\lambda]u_{m-1,j} + [\frac{5}{6}(2-\Delta t(\frac{\partial S}{\partial U})_{m,j})-2\lambda]u_{mj}$$
$$+ \frac{\Delta t}{6}(s_{m-1,j}+10s_{mj}+s_{m+1,j}) .$$

Again we find that the equations governing the convergence of the AGE iterative process that utilises each of the above time linearisation schemes are given by (6.8.18) and (6.8.19) (for the case m odd) and (6.8.24) and (6.8.25) (for the case m even) with $\underline{w}$ replaced by $\underline{u}$.

(iv) Solving the Non-Linear First Order Hyperbolic (Convection) Equation

A non-linear first order hyperbolic equation may take the following form,

$$\frac{\partial U}{\partial t} = -U\frac{\partial U}{\partial x} . \qquad (6.8.74)$$

By resorting to the generalised difference formula developed in Section 2.9, an approximation to (6.8.74) at the point (i,j+θ) is given by,

$$\frac{(u_{i,j+1}-u_{ij})}{\Delta t} = -u_{i,j+\theta}\{\frac{1}{\Delta x}[(1-w)u_{i+1,j+1}+(2w-1)u_{i,j+1}-$$
$$wu_{i-1,j+1}]+(1-\theta)[(1-w)u_{i+1,j}+(2w-1)u_{ij}-wu_{i-1,j}]\},$$

where $0\leq w, \theta\leq 1$. for $i=1,2,\ldots,m$. (6.8.75) *(a) The implicit, centred-in-distance, backward-in-time scheme (CDBT)*

If we choose $w=\frac{1}{2}$ and $\theta=1$ in (6.8.75) we obtain the following formula,

$$\frac{u_{i,j+1}-u_{ij}}{\Delta t} = -\frac{\alpha_{i,j+1}}{2\Delta x}[u_{i+1,j+1}-u_{i-1,j+1}] \, , \qquad (6.8.76)$$

where $\alpha_{i,j+1} = u_{i,j+1}$. Equation (6.8.76) is linearised by replacing $\alpha_{i,j+1}$ by $u^{(p)}_{i,j+1}$ (the old value in our iterative process) leading to the tridiagonal system of equations,

$$(-\frac{\lambda}{2}u^{(p)}_{i,j+1})u_{i-1,j+1}+u_{i,j+1}+(\frac{\lambda}{2}\, u^{(p)}_{i,j+1})u_{i+1,j+1} = u_{ij}, \; i=1,2,\ldots,m; \qquad (6.8.77)$$

where $\lambda=\Delta t/\Delta x$, the mesh ratio.

(b) The centred-in-distance, centred-in-time scheme (CDCT)

The choice of $w=\frac{1}{2}$ and $\theta=\frac{1}{2}$ in (6.8.75) gives the difference analogue,

$$\frac{u_{i,j+1}-u_{ij}}{\Delta t} = \frac{u_{i,j+\frac{1}{2}}}{4\Delta x}[(u_{i+1,j+1}-u_{i-1,j+1})+(u_{i+1,j}-u_{i-1,j})]. \qquad (6.8.78)$$

Since $u_{i,j+\frac{1}{2}}$ does not fall on the grid point, it can be replaced by $(u_{i,j}+u_{i,j+1})/2$ and (6.8.78) becomes,

$$\frac{(u_{i,j+1}-u_{ij})}{\Delta t} = -\frac{1}{4\Delta x}[\alpha_{ij}(u_{i+1,j+1}-u_{i-1,j+1})+\beta_{ij}(u_{i+1,j}-u_{i-1,j})]$$

which is non-linear with $\alpha_{ij}=\beta_{ij}=(u_{ij}+u_{i,j+1})/2$. It can, however, be linearised by letting instead $\alpha_{ij}=u_{ij}$ and $\beta_{ij}=u_{i,j+1}$ and consequently we get the tridiagonal system of equations,

$$(-\frac{\lambda}{4}u_{ij})u_{i-1,j+1}+(1+\frac{\lambda}{4}(u_{i+1,j}-u_{i-1,j}))u_{i,j+1}+(\frac{\lambda}{4}u_{ij})u_{i+1,j+1} = u_{ij} \qquad (6.8.79)$$

As previously, the tridiagonal systems (6.8.77) and (6.8.79) are of the form (6.8.47a) and therefore we have

(1) for CDBT (6.8.77)

$$c_{ij} = -\frac{\lambda}{2}u^{(p)}_{i,j+1} \quad , \; i=2,3,\ldots,m$$

$$a_{ij} = 1, \quad i=1,2,\ldots,m$$

$$b_{ij} = \frac{\lambda}{2}u^{(p)}_{i,j+1}, \quad i=1,2,\ldots,m-1$$

$$f_1 = u_{1j} + (\frac{\lambda}{2}u^{(p)}_{1,j+1})u_{0,j+1}$$

$$f_i = u_{ij}, \; i=2,3,\ldots,m-1,$$

$$f_m = u_{mj}-(\frac{\lambda}{2}\,u^{(p)}_{m,j+1})u_{m+1,j+1} \;;$$

(2) for CDCT (6.8.79)

$$c_{ij} = -\frac{\lambda}{4}\,u_{ij}, \; i=2,3,\ldots,m$$

$$a_{ij} = 1 + \frac{\lambda}{4}(u_{i+1,j}-u_{i-1,j}), \; i=1,2,\ldots,m$$

$$b_{ij} = \frac{\lambda}{4}u_{ij} \; , \; i=1,2,\ldots,m-1,$$

$$f_1 = u_{1j}\;(1 + \frac{\lambda}{4}\,u_{0,j+1})$$

$$f_i = u_{ij} \; , \quad i=2,\ldots,m-1,$$

and $$f_m = u_{mj}(1 - \frac{\lambda}{4}\,u_{m+1,j+1}).$$

The AGE iterative process is completed when convergence is reached using the equations (6.8.18) and (6.8.19) (if m odd) and (6.8.24) and (6.8.25) (if m even) with $\underline{w}$ replaced by $\underline{u}$.

6.9 NUMERICAL EXPERIMENTS AND COMPARATIVE RESULTS

A number of experiments were conducted to demonstrate the application of the AGE algorithm on parabolic and hyperbolic problems and where appropriate the solutions were compared with that of the GE class of methods either given by Evans and Abdullah (1983) or obtained by the author. In most cases, both the Peaceman-Rachford (PR) and the Douglas-Rachford (DR) variants were employed for the implementation of the AGE scheme and the acceleration parameter r was chosen so as to provide the most rapid convergence. Unless otherwise stated, the *convergence criterion* was taken as $eps=10^{-4}$.

EXPERIMENT 1

We considered the following problem taken from Saulev (1964),

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} , \quad 0\leq x\leq 1 , \tag{6.9.1}$$

subject to the initial condition,

$$U(x,0) = 4x(1-x) , \quad 0\leq x\leq 1$$

and the boundary conditions

$$U(0,t) = U(1,t) = 0 , \quad t\geq 0 . \tag{6.9.1a}$$

The exact solution is given by,

$$U(x,t) = \frac{32}{\pi^3} \sum_{k=1,(2)}^{\infty} \frac{1}{k^3} e^{-\pi^2 k^2 t} \sin(k\pi x) . \tag{6.9.2}$$

Tables 6.9.1-6.9.3 provide a comparison of the accuracy of the methods under consideration in terms of the absolute errors at the appropriate grid points for various values of λ. It is very clear that in the AGE class of methods, the Douglas formula (AGE-DG) employing the PR variant is the most accurate in comparison with the fully implicit formula (AGE-IMP) as well as the Crank-Nicolson scheme (AGE-CN). This is to be expected since DG is second-order accurate in space and fourth-order accurate in time whereas CN and IMP have accuracies to the order of $O[(\Delta x)^2+(\Delta t)^2]$ and $O[(\Delta x)^2+\Delta t]$ respectively. By the same reasoning, the PR variant which we know is second-order accurate in both space and time is expected to produce a better solution and takes smaller iteration than the DR variant whose truncation error is $O((\Delta x)^2+\Delta t)$. It is also apparent that (D)AGE and AGE-CN have comparable accuracies.

To indicate the efficiency of the AGE iterative methods, it is necessary to consider *the computational complexity* for each iteration as well as the number of iterations required for convergence. One way of estimating this computational complexity is to count the number of arithmetic operations performed on each mesh line (time level) where there are m internal points. Thus, using (6.2.4a), (6.3.15/15a) and (6.3.18/19) (with m odd) we find that to complete one iteration of the generalised AGE scheme, the amount of work done is given by the table below:

Scheme	Number of multiplications	Number of additions
AGE-IMP	$\frac{(19m-1)}{2}+31$	6m+19
AGE-CN	$\frac{(23m-1)}{2}+31$	8m+20
AGE-DG	$\frac{(23m-1)}{2}+45$	8m+23

TABLE 6.9.4

In comparison, *the Thomas algorithm* requires approximately (11m-3) multiplications and (7m-3) subtractions to solve the CN scheme directly. We infer from Table 6.9.4 that for large m, the number of multiplications incurred in the implementation of the AGE algorithm for CN or DG are only slightly more than that required by the Thomas algorithm. On the average only two iterations are needed by the PR variant for convergence and it can therefore be concluded that the AGE scheme has merits as an alternative iterative method with respect to stability, accuracy, efficiency and rate of convergence.

EXPERIMENT 2

In this experiment, we attempted to solve the following heat conduction problem with periodic boundary conditions,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial t^2} + 10(1-x)xt \ , \qquad (6.9.3)$$

subject to the initial condition,

$$U(x,0) = x(1-x) \ , \qquad (6.9.3a)$$

and the boundary conditions,

$$U(0,t) = U(1,t) \ , \ \frac{\partial U}{\partial x}(0,t) = \frac{\partial U}{\partial x}(1,t) \ . \qquad (6.9.3b)$$

The above problem has the exact solution given by,

$$U(x,t) = \frac{(1+5t^2)}{6} - \frac{5}{8}\sum_{k=1}^{\infty} \frac{\cos(2k\pi x)}{(k\pi)^6}\{4k^2\pi^2 t - 1 + e^{-4k^2\pi^2 t}\}$$

$$- \sum_{k=1}^{\infty} \frac{e^{-4k^2\pi^2 t}}{k^2\pi^2} \cos(2k\pi x) \ . \qquad (6.9.4)$$

The AGE solutions at the grid points on selected time rows and for various time steps are displayed in Tables 6.9.5 and 6.9.6 for both the PR and DR variants. It is observed that except for λ=0.1 (at t=0.1), an examination of all the averages of the percentage errors indicates that the AGE solutions are more accurate than the (D)AGE values derived by Abdullah (1983). While the number of iterations for the PR variant remains fairly constant (as well as the percentage errors), the iteration count for the DR variant tends to become large with increasing λ.

EXPERIMENT 3

This experiment involved the solution of the diffusion-convection equation,

$$\frac{\partial U}{\partial t} = \varepsilon \frac{\partial^2 U}{\partial x^2} - k \frac{\partial U}{\partial x} \;, \qquad (6.9.5)$$

with the initial condition,

$$U(x,0) = 0,\; 0<x<1 \;, \qquad (6.9.5a)$$

and the Dirichlet boundary conditions,

$$\left.\begin{aligned} U(0,t) &= 0 \;, \\ U(1,t) &= 1. \end{aligned}\right\} \; t\geq 0 \;. \qquad (6.9.5b)$$

The coefficients k and ε assumed the same values of 1. The exact solution is given by,

$$U(x,t) = \frac{(e^{kx/\varepsilon}-1)}{(e^{k/\varepsilon}-1)} + 2\sum_{n=1}^{\infty} \frac{(-1)^n\, n\pi}{(n\pi)^2+(k/2\varepsilon)^2}\, e^{\frac{k}{2\varepsilon}(x-1)} \sin(n\pi x)\, e^{-[(n\pi)^2\varepsilon+\frac{k^2}{4\varepsilon}]t} \;. \qquad (6.9.6)$$

From Abdullah (1983) the solutions of (6.9.5) by means of the Crank-Nicolson scheme with upwinding (CNU) and the (D)AGE scheme were included. These were then compared with the AGE scheme employing the fully implicit formula (AGE-IMP) as well as the methods of Crank-Nicolson (AGE-CN) and Crank-Nicolson with upwinding (AGE-CNU).

The absolute errors of the numerical solutions of these schemes are shown in Tables 6.9.7-6.9.10. It must be remembered that the CNU scheme is only first-order accurate in space and second-order accurate in time. This explains the poor accuracy of CNU and AGE-CNU even when compared with the fully implicit formula whose truncation error is $O((\Delta x)^2+\Delta t)$ since smaller values of Δt were taken in our experiments with $\Delta x=0.1$. For small λ, the (D)AGE process appears to have a slight edge on AGE-CN. For larger mesh ratios such as $\lambda=2.0$, however, the

AGE-CN method can be very competitive since it exhibits better accuracy and requires only one iteration for convergence and is therefore worthy of recommendation.

EXPERIMENT 4

The AGE algorithm was implemented on the same first-order hyperbolic equations of Section 4.13, i.e. Problem 1 and Problem 2 given by:

(a) Problem 1

$$\frac{\partial U}{\partial t}+\frac{\partial U}{\partial x}=0\ , \tag{6.9.7}$$

subject to,

$$U(x,0) = \cos x \tag{6.9.7a}$$

$$U(0,t) = \cos t$$

and $$U(1,t) = \cos(1-t)\ . \tag{6.9.7b}$$

The exact solution is given by,

$$U(x,t) = \cos(x-t)\ ; \tag{6.9.8}$$

and

(b) Problem 2

$$\frac{\partial U}{\partial t}+\frac{\partial U}{\partial x}=k(x,t)\ , \tag{6.9.9}$$

$$(k(x,t) = -2\sin(x-t)e^{-2t})$$

subject to,

$$U(x,0) = \sin x \tag{6.9.9a}$$

$$U(0,t) = -\sin(t)e^{-2t},$$

and $$U(1,t) = \sin(1-t)e^{-2t}\ . \tag{6.9.9b}$$

The exact solution is given by,

$$U(x,t) = \sin(x-t)e^{-2t}\ . \tag{6.9.10}$$

It is known that the truncation errors of the fully implicit scheme and the Crank-Nicolson type scheme (the centred-in-distance, centred-in-time formula) are $O((\Delta x)^2+\Delta t)$ and $O((\Delta x)^2+(\Delta t)^2)$ respectively. The accuracies of the AGE method utilising these schemes (AGE-IMP and AGE-CN) are depicted in Tables 6.9.11-6.9.14 for both problems using $\lambda=0.5$ and $\lambda=1$. They are then compared with the results derived from the GE class of methods and other well-known schemes given in Tables 4.13.1-4.13.4. It is interesting to note that the CN method (using the Thomas algorithm) and the AGE-CN scheme (employing the PR variant) exhibit the same order of accuracy although the latter with 3 to 4 iterations requires more computational load than the former. Furthermore, for Problem 1, these schemes evidently emerge as the next most accurate method of solution after the Lax-Wendroff formula. For Problem 2, however, the (D)AGE process appears to be more favourable. As expected, the DR variant of the AGE class of methods produces a slightly less accurate solution and entails two to three times more iterations than the corresponding PR formula.

EXPERIMENT 5

This experiment dealt with the following problems involving the one-dimensional wave equation:

(a) Problem 1 (with Dirichlet boundary conditions)

$$\frac{\partial^2 U}{\partial x^2} = \frac{\partial^2 U}{\partial t^2} \quad , \tag{6.9.11}$$

subject to the initial conditions,

$$\left.\begin{aligned} U(x,0) &= \frac{1}{8}\sin(\pi x) \quad , \\ \frac{\partial U}{\partial t}(x,0) &= 0 \quad , \end{aligned}\right. \tag{6.9.11a}$$

and the boundary conditions,

$$U(0,t) = U(1,t) = 0 \ . \tag{6.9.11b}$$

The exact solution is given by,

$$U(x,t) = \frac{1}{8}\sin(\pi x)\cos(\pi t) \; ; \qquad (6.9.12)$$

and

(b) Problem 2 (with derivative boundary conditions)

$$\frac{\partial^2 U}{\partial x^2} = \frac{\partial^2 U}{\partial t^2} \;, \qquad (6.9.13)$$

subject to the initial conditions,

$$\begin{aligned} U(x,0) &= 100x^2 \;, \\ \frac{\partial U}{\partial t}(x,0) &= 200x \qquad (6.9.13a) \end{aligned}$$

and the boundary conditions,

$$\frac{\partial U}{\partial x}(0,t) = 200t$$

and $$U(1,t) = 100(1+t)^2 \;. \qquad (6.9.13b)$$

The exact solution is of the form,

$$U(x,t) = 100(x+t)^2 \;. \qquad (6.9.14)$$

In applying the AGE algorithm to the general equations (6.7.3), we chose $\alpha=1/4$ to correspond with the implicit method (2.13.15) which as we know from Section 2.15 has second-order accuracy in both space and time. In Table 6.9.15 is shown the absolute errors of the AGE solutions for Problem 1 for $\lambda=0.5$ and $\lambda=1.0$. As solutions by means of the GE class of methods are available, the results of this particular experiment are compared with that of Table 4.13.8. It is immediately evident that while the (S)AGE-LW combination is favoured in *applying the GE technique* on the wave equation it is, however, slightly less accurate than the AGE scheme which unlike the former has the additional advantage of being unconditionally stable. The iteration number necessary for convergence is also found to be considerably small. In Table 6.9.16 is presented the percentage errors of the AGE solutions to the wave equation for Problem 2 for the values of λ progressing from 0.1 to 2.0.

Unfortunately, no solutions from other schemes have been worked out to provide a comparison for this particular example.

EXPERIMENT 6

This experiment was concerned with the solution of the non-linear problem,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U^2}{\partial x^2} \qquad (6.9.15)$$

with the exact solution (Abdullah (1983)),

$$(2U-3)+\ell n(U-\tfrac{1}{2}) = 2(2t-x) \ . \qquad (6.9.16)$$

The appropriate boundary data were given to satisfy the above exact solution.

The analytical solution was obtained iteratively by means of the Newton-Raphson method through the formula,

$$U^{(p+1)} = \tfrac{1}{2}+[\frac{3}{2} - \tfrac{1}{2}\ell n(U^{(p)}-\tfrac{1}{2})-x+2t](1-\frac{1}{2U^{(p)}}) ,$$
$$p=1,2,3,... \qquad (6.9.17)$$

and the initial guesses were taken as,

$$U^{(0)} = 1 + \frac{4-e^{2(x-2t)}}{4+2e^{2(x-2t)}} \ . \qquad (6.9.17a)$$

The AGE solutions using the fully implicit, Crank-Nicolson and Douglas formulae which were linearised using the Ritchmyer's method (AGE-IMP(RCHM),AGE-CN(RCHM) and AGE-DG(RCHM)) are shown in Tables 6.9.17-6.9.18. They are compared with the results obtained from the linearised schemes of Crank-Nicolson (using the Richtmyer's method) and Lee and as well as from the (D)AGE formulae. The linearised schemes of Crank-Nicolson (CN) and Lee employed the Thomas algorithm as a method of solution to the resulting tridiagonal system of equations. As the (D)AGE scheme had to be solved iteratively, convergence of the iterative process was considered for every group of 2 points and the figures in brackets in the tables indicate the number of iterations required.

It is observed that in the AGE class of methods, the AGE-LEE scheme provides the most accurate solutions for both $\lambda=0.05$ and $\lambda=0.1$. The approximate amount of arithmetic involved in the computation of the AGE solutions per iteration is given in Table 6.9.19 below.

Method \ Operation	Number of multiplications	Number of additions
AGE-IMP/CN/DG	41m-16	24m+1
AGE-LEE	50m-21	45m-20

TABLE 6.9.19

To reduce the storage requirement to a minimum, the entries of the coefficient matrix A and the right-hand side vector $\underline{f}$ of (6.8.6) were generated rather than stored at each of the $(p+\frac{1}{2})$ and $(p+1)$ iterates. Consequently as is expected for a non-linear problem, the computational complexity of the AGE schemes can be quite substantial.

It is also apparent that the (D)AGE, CN and LEE schemes possess solutions close to the analytical ones. However, it must be mentioned that the (D)AGE scheme requires 3 iterations for *every group of 2 points* while the AGE methods require only 1 or 2 iterations over the *whole mesh line*.

EXPERIMENT 7

The following Burger's equation was considered,

$$\frac{\partial U}{\partial t} = \varepsilon \frac{\partial^2 U}{\partial x^2} - U\frac{\partial U}{\partial x} \qquad (6.9.18)$$

and the initial and boundary conditions were prescribed so as to satisfy the exact solution. Two problems were solved using the AGE algorithm.

Problem 1

This problem has the exact solution (Madsen and Sincovec (1976)),

$$U(x,t) = (0.1e^{-A} + 0.5e^{-B} + e^{-C})/(e^{-A} + e^{-B} + e^{-C}),\ 0\leq x\leq 1,\ t\geq 0$$

where, (6.9.19)

$$\left.\begin{aligned} A &= \frac{0.05}{\varepsilon}(x-0.5+4.95t)\ , \\ B &= \frac{0.25}{\varepsilon}(x-0.5+0.75t)\ , \\ \text{and}\quad C &= \frac{0.5}{\varepsilon}(x-0.375)\ ; \end{aligned}\right\} \qquad (6.9.19a)$$

and

Problem 2

The exact solution to this problem is given by (Cole (1951)),

$$U(x,t) = \frac{2\pi \sum_{k=1}^{\infty} k\ A_k\ \sin(k\pi x)e^{(-\varepsilon k^2\pi^2 t)}}{\frac{1}{\varepsilon}\ \{A_0 + \sum_{k=1}^{\infty} A_k \cos(k\pi x)e^{(-\varepsilon k^2\pi^2 t)}\}} \qquad (6.9.20)$$

where,

$$\left.\begin{aligned} A_k &= 2\int_0^1 \cos(k\pi x)\exp\{-\frac{1}{2\varepsilon}\int_0^x F(x')dx'\}dx,\ (k=1,2,3,\ldots) \\ A_0 &= \int_0^1 \exp\ \{-\frac{1}{2\varepsilon}\int_0^x F(x')dx'\}dx \end{aligned}\right\} \qquad (6.9.20a)$$

and

(i) $F(x) = 4x(1-x)$, (6.9.20b)

(ii) $F(x) = \sin(\pi x)$. (6.9.20c)

Comparative results for Problem 1 using the (D)AGE and AGE schemes for $\lambda=1.0$ and $\varepsilon=0.1$ and $\varepsilon=1.0$ at $t=1.0$ are given in Table 6.9.20. The figures in brackets for the (D)AGE scheme indicate the number of iterations required for convergence at every group of two points. The numerical solutions to Burgers' equation for the same problem obtained by means of the various difference schemes at different time levels for smaller values of ε are presented in Tables 6.9.21-6.9.22. The solutions to Problem 2 are shown in Tables 6.9.23-6.9.24 where only the PR variant was employed for the generalised AGE method.

It can be inferred from Table 6.9.20 that the AGE-CN and (D)AGE processes exhibit comparable accuracies at the grid point. The number of iterations required by the AGE-CN (PR) scheme (about 3-4 iterations) can be further reduced by relaxing the convergence criterion which at 10^{-6} may be regarded as quite stringent. Similar features of accuracy of the (D)AGE and AGE-CN schemes are also observed in Tables 6.9.21-6.9.22 for a much smaller value ε=0.003 although now the solutions are at slight variance with the exact ones. However, a *large number of iterations* were required to achieve convergence of the (D)AGE procedure (Evans and Abdullah (1984)). The AGE-CN (PR) method, on the other hand, required only 2 iterations over the whole mesh line at t=0.1 and t=0.5. Hence, it can be said that the AGE-CN scheme is just as competitive, if not better, than the (D)AGE method.

For Problem 2 with F(x) given by (6.9.20b) we find from Table 6.9.23 that while the (D)AGE and AGE-CN schemes have about the same order of accuracy, the AGE-IMP method appears to be more accurate. The same conclusion can be drawn from Table 6.9.24 with F(x) given by (6.9.20c). In this particular case, the explicit scheme (Caldwell and Smith (1982)) demonstrates that it provides a better solution. In general, it is not expected of these two schemes to perform better. However, the inaccuracies in the solutions for small values of ε seems to be a difficulty experienced by most finite difference methods.

Finally, as an indication of efficiency of the AGE algorithm, we present below estimates of the computational complexity per iteration of the relevant difference schemes employed to solve the above Burger's equation:

Operations / Schemes	Number of multiplications	Number of additions
AGE-IMP	24m	$\frac{(39m-1)}{2}-3$
AGE-CN	41m-16	37m-19

TABLE 6.9.25

where we have assumed that m is odd.

EXPERIMENT 8

This experiment dealt with the calculation of the propagation of a one-dimensional wave governed by the reaction-diffusion equation,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + S \ , \quad -\infty<x<\infty \ , \quad t\geq 0 \ , \qquad (6.9.21)$$

where $$S = U^2(1-U) \ . \qquad (6.9.21a)$$

As was explained in Section 6.8(iii), for practical purposes, the above domain was truncated to $-50\leq x\leq 400$, $t\geq 0$ and the following initial and boundary conditions were prescribed,

$$U(x,0) = 1/(1+e^{Vx}) \ , \qquad (6.9.21b)$$

$$U(-50,t) = 1 \ ,$$

and $$U(400,t) = 0 \ , \qquad (6.9.21c)$$

where $V= 1/\sqrt{2}$, the steady-state wave speed. The exact travelling wave solution is given by,

$$U(x,t) = 1/(1+e^{V(x-Vt)}) \ . \qquad (6.9.22)$$

The numerical solutions to the reaction-diffusion problem were obtained using the GE-PC, AGE-PC and AGE-TL schemes and comparisons among these methods were presented in terms of the absolute errors of the solutions.

The mesh ratio was taken as $\lambda=0.2$ and the calculations were performed with different spatial grids and time steps.

An examination of the absolute errors in Tables 6.9.26-6.9.29 for solutions at different time levels with $\Delta x=1.0$ and $\Delta t=0.2$ shows that for small t, the AGE-4TL method is the most accurate followed by the AGE-PC, (D)AGE-PC, AGE-2TL, (S)AGE-PC, AGE-1TL, AGE-3TL, (S)AGE-EXP and (D)AGE-EXP schemes while for large t we have in decreasing order of accuracy, the AGE-4TL, AGE-2TL, AGE-PC, (D)AGE-PC, (S)AGE-PC, AGE-1TL, AGE-3TL, (D)AGE-EXP and (S)AGE-EXP schemes. The same conclusion can be drawn for finer grids with $\Delta x=0.5$ and $\Delta t=0.05$ (Tables 6.9.31-6.9.32). However, as t progresses, AGE-3TL becomes more accurate than AGE-1TL. One possible explanation to the above observation is that the accuracy of the methods is very much dependent on *the time step* employed in the calculation. For the TL methods, for example, the non-linear reaction terms are expanded in Taylor series about the previous time and the 1TL, 2Tl, 3TL and 4TL formulae are known to have accuracies to the order of $O([\Delta x]^2+\Delta t)$, $O([\Delta x]^2+[\Delta t]^2)$, $O([\Delta x]^4+\Delta t)$ and $O([\Delta x]^4+[\Delta t]^2)$ respectively. The results presented in Tables 6.9.26-6.9.29 and Tables 6.9.31-6.9.32 therefore suggest that *temporal approximations play* a more dominant role than *spatial approximations* in determining the accuracy of the methods. A similar reasoning can also be applied to other methods employing the PC technique. For each grid point along the mesh line (time row), the Runge-Kutta method of $O([\Delta t]^5)$ accuracy is used which is then corrected by utilising the Crank-Nicolson formula which in turn is second-order accurate in both space and time. Thus, in the (S)AGE-PC and (D)AGE-PC methods, the utilisation of these high-order formulae coupled with the cancellation of errors resulting from the alternate use of the constituent GER and GEL schemes lead to accurate solutions of these schemes.

In Tables 6.9.30 and 6.9.33, we compare the accuracy of the various methods in terms of the computed wave speed V. For a steady state wave propagation problem, the following condition applies,

$$\frac{\partial U}{\partial t} + V \frac{\partial U}{\partial x} = 0 \ . \qquad (6.9.23)$$

By substituting $\frac{\partial U}{\partial t}$ into (6.9.21) and then integrating the result using (6.9.21c), V can be obtained as,

$$V = \int_{-50}^{400} U^2 (1-U) dx \ . \qquad (6.9.24)$$

In our experiment, the solutions worked out from each of the numerical methods were used along the whole mesh line and then the composite trapezium rule was employed to compute V which was then compared with the exact value $1/\sqrt{2}$ =0.7071058. It must be mentioned that Ramos (1985) employed the Thomas algorithm to solve tridiagonal systems of equations arising from the application of the relevant finite difference formulae. Since the values of V from the solutions of these methods are available, we are then able to compare the accuracy of the solutions of the (S)AGE, (D)AGE and AGE class of methods with those obtained from the application of the Thomas algorithm. In general, we find that the computed values of V from the AGE(PR) solutions (CV) are in close agreement with those of Ramos (CV-R). This can also be read from a comparison of the appropriate percentage errors in the computed values (PCV and PCV-R) as presented in Tables 6.9.30 and 6.9.33 for both coarse and fine grids.

An estimate of the computational effort involved by all the methods to execute the calculations per time row (in the case of the GE schemes) or per iteration (in the case of the AGE schemes) is given in the following table:

Operations Schemes		Number of multiplications	Number of additions
GE-EXP		12m+10	7m+1
GE-PC	P	38m	29
	C	18m+10	9m+1
	Total	56m+10	9m+30
AGE-PC	P	38m	29
	C	22m-23	$\frac{(33m-1)}{2}+6$
	Total	60m-23	$\frac{(33m-1)}{2}+35$
AGE-1TL		$\frac{(73m-1)}{2}+34$	31m-13
AGE-2TL		38m-12	$\frac{(63m-1)}{2}-11$
AGE-3TL		76m-12	52m-13
AGE-4TL		75m-10	55m-11

TABLE 6.9.34

The estimate of the amount of work done is based on an odd number (m odd) of internal points on the time row.

From the preceding discussion, we conclude that among all the methods worthy of recommendation, the AGE-4TL(PR) procedure provides the most accurate solution to the reaction-diffusion problem. As a high-order method , it is expected to yield more accurate results for a wide range of spatial and temporal step sizes. By contrast, the 3TL technique is the least competitive in the AGE class because it is not only less accurate but also requires more arithmetic work and iterations. Although the GE-PC and AGE-PC procedures result in highly accurate solutions for our particular example, the *explicit evaluations* of the diffusion terms at the predictor stage may lead to some stability restrictions on the time step size.

EXPERIMENT 9

The following non-linear first order hyperbolic (convection) equation was considered

$$\frac{\partial U}{\partial t} = -U\frac{\partial U}{\partial x} , \tag{6.9.25}$$

and two problems were solved using the AGE algorithm.

Problem 1

Equation (6.9.25) was solved subject to the auxiliary conditions (Casulli et al, 1984),

$$U(x,0) = 1-x, \quad 0 \leq x \leq 1 \; ; \tag{6.9.26}$$

$$U(0,t) = 1 \; , \quad 0 < t \leq 1 \; ,$$

and

$$U(1,t) = \begin{cases} 1 \text{ for } 1 \leq t \\ 0 \text{ for } t < 1 \, . \end{cases} \tag{6.9.26a}$$

The exact solution is given by,

$$U(x,t) = \begin{cases} 1 \text{ for } x \leq t \\ \frac{x-1}{t-1} \text{ for } t < x \leq 1 . \end{cases} \tag{6.9.27}$$

Problem 2

For this problem, the initial and boundary conditions take the form (Ames (1977)),

$$U(x,0) = x \; , \quad 0 \leq x \leq 1 \; ; \tag{6.9.28}$$

$$U(0,t) = 0 \; , \; t \geq 0 \; ,$$

and

$$U(1,t) = 1/(1+t) \; , \; t \geq 0 \; . \tag{6.9.28a}$$

This problem has the exact solution given by,

$$U(x,t) = x/(1+t) \; . \tag{6.9.29}$$

The solutions to both problems were obtained using the AGE algorithm which was implemented on the CDBT and CDCT formulae given by the linearised tridiagonal systems (6.8.77) and (6.8.79) respectively.

A number of computer runs were carried out using $\Delta x=0.05$ and with varying time steps to determine the accuracy of the AGE-CDBT and AGE-CDCT schemes. The results given in terms of the absolute errors are shown in Tables 6.9.35-6.9.37.

It is observed from Table 6.9.35 that for small Δt, the average of absolute errors on time level $t=1.0$ indicates that the AGE-CDBT scheme is more accurate than the AGE-CDCT procedure. When the mesh ratio is progressively increased (Δx being fixed) we find that the accuracy of the AGE-CDCT scheme begins to improve over the AGE-CDBT method. Since the truncation errors of the CDCT formula is $O([\Delta x]^2+[\Delta t]^2)$, it is therefore clear that the AGE-CDCT scheme produces a more accurate solution than AGE-CDBT whose truncation error is only $O([\Delta x]^2+\Delta t)$. The accuracy of the AGE-CDCT scheme is even more pronounced for the second example. It is immediately evident from Table 6.9.37 that despite having to satisfy the stringent convergence requirement of $eps=10^{-6}$, the AGE-CDCT(PR) scheme demonstrates that it requires only two iterations while at the same time it maintains its high order of accuracy for large time steps. Thus, although the computational complexity of the AGE-CDCT scheme may be slightly large as illustrated in Table 6.9.38, it offers great promise as an accurate, stable and efficient numerical procedure to solve non-linear hyperbolic problems.

Operations / Schemes	Number of multiplications	Number of additions
AGE-CDBT	22m+2	$\frac{29(m-1)}{2}+14$
AGE-CDCT	34m-16	26m-12

TABLE 6.9.38: Computational complexity per iteration (m odd)

$\lambda=0.1$, $t=0.05$, $\Delta t=0.001$, $\Delta x=0.1$, $r=0.5$, $eps=10^{-4}$

x Scheme		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Number of Iterations
GER		6.0×10^{-5}	1.6×10^{-3}	1.5×10^{-3}	2.3×10^{-3}	2.3×10^{-3}	2.3×10^{-3}	2.6×10^{-3}	1.6×10^{-3}	1.9×10^{-3}	-
GEL		1.9×10^{-3}	1.6×10^{-3}	2.6×10^{-3}	2.3×10^{-3}	2.3×10^{-3}	2.3×10^{-3}	1.5×10^{-3}	1.6×10^{-3}	6.0×10^{-5}	-
(S)AGE		9.0×10^{-4}	1.4×10^{-3}	2.0×10^{-3}	2.0×10^{-3}	2.1×10^{-3}	2.1×10^{-3}	1.9×10^{-3}	1.6×10^{-3}	8.0×10^{-4}	-
(D)AGE		9.0×10^{-4}	1.5×10^{-3}	1.9×10^{-3}	2.1×10^{-3}	2.1×10^{-3}	2.1×10^{-3}	1.9×10^{-3}	1.6×10^{-3}	9.0×10^{-4}	-
AGE-IMP	PR	1.5×10^{-3}	2.6×10^{-3}	3.3×10^{-3}	3.7×10^{-3}	3.8×10^{-3}	3.7×10^{-3}	3.3×10^{-3}	2.6×10^{-3}	1.5×10^{-3}	2
	DR	2.3×10^{-3}	4.2×10^{-3}	5.5×10^{-3}	6.2×10^{-3}	6.4×10^{-3}	6.2×10^{-3}	5.5×10^{-3}	4.2×10^{-3}	2.3×10^{-3}	6
AGE-CN	PR	9.1×10^{-4}	1.6×10^{-3}	2.1×10^{-3}	2.3×10^{-3}	2.4×10^{-3}	2.3×10^{-3}	2.1×10^{-3}	1.6×10^{-3}	9.1×10^{-4}	2
	DR	1.7×10^{-3}	3.2×10^{-3}	4.2×10^{-3}	4.8×10^{-3}	5.0×10^{-3}	4.8×10^{-3}	4.2×10^{-3}	3.2×10^{-3}	1.7×10^{-3}	6
AGE-DG	PR	1.7×10^{-6}	5.5×10^{-6}	1.2×10^{-5}	1.7×10^{-5}	2.0×10^{-5}	1.7×10^{-5}	1.1×10^{-5}	5.5×10^{-6}	1.6×10^{-6}	2
	DR	8.1×10^{-4}	1.5×10^{-3}	2.1×10^{-3}	2.4×10^{-3}	2.6×10^{-3}	2.4×10^{-3}	2.1×10^{-3}	1.5×10^{-3}	8.1×10^{-4}	6
EXACT SOLUTION		0.1950648	0.3707705	0.5098716	0.5989617	0.6296137	0.5989617	0.5098716	0.3707705	0.1950648	-

TABLE 6.9.1: The absolute errors of the numerical solutions to Problem (6.9.1)
(Parabolic problem with Dirichlet boundary conditions)

$\lambda=0.5$, $t=0.25$, $\Delta t=0.005$, $\Delta x=0.1$, $r=0.5$, $eps=10^{-4}$

Scheme x		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Number of Iterations
GER		1.0×10^{-4}	1.1×10^{-3}	1.1×10^{-3}	1.7×10^{-3}	1.8×10^{-3}	1.7×10^{-3}	1.9×10^{-3}	1.1×10^{-3}	1.2×10^{-3}	-
GEL		1.2×10^{-3}	1.1×10^{-3}	1.9×10^{-3}	1.7×10^{-3}	1.8×10^{-3}	1.7×10^{-3}	1.1×10^{-3}	1.1×10^{-3}	1.0×10^{-4}	-
(S)AGE		2.0×10^{-4}	1.6×10^{-3}	1.8×10^{-3}	2.5×10^{-3}	2.5×10^{-3}	2.3×10^{-3}	2.1×10^{-3}	1.0×10^{-3}	8.0×10^{-4}	-
(D)AGE		2.0×10^{-4}	7.0×10^{-4}	3.0×10^{-4}	7.0×10^{-4}	7.0×10^{-4}	5.0×10^{-4}	7.0×10^{-4}	3.0×10^{-4}	3.0×10^{-4}	-
AGE-IMP	PR	2.2×10^{-3}	4.2×10^{-3}	5.7×10^{-3}	6.7×10^{-3}	7.0×10^{-3}	6.7×10^{-3}	5.7×10^{-3}	4.1×10^{-3}	2.2×10^{-3}	3
	DR	2.6×10^{-3}	5.0×10^{-3}	6.8×10^{-3}	8.0×10^{-3}	8.4×10^{-3}	7.9×10^{-3}	6.8×10^{-3}	4.9×10^{-3}	2.7×10^{-3}	6
AGE-CN	PR	5.3×10^{-4}	1.0×10^{-3}	1.4×10^{-3}	1.6×10^{-3}	1.7×10^{-3}	1.6×10^{-3}	1.4×10^{-3}	1.0×10^{-3}	5.4×10^{-4}	2
	DR	9.7×10^{-4}	1.9×10^{-3}	2.5×10^{-3}	3.0×10^{-3}	3.1×10^{-3}	3.0×10^{-3}	2.5×10^{-3}	1.8×10^{-3}	9.9×10^{-4}	6
AGE-DG	PR	1.5×10^{-5}	2.5×10^{-5}	3.8×10^{-5}	4.4×10^{-5}	4.7×10^{-5}	4.5×10^{-5}	3.6×10^{-5}	2.8×10^{-5}	1.2×10^{-5}	2
	DR	4.2×10^{-4}	8.0×10^{-4}	1.1×10^{-3}	1.3×10^{-3}	1.3×10^{-3}	1.3×10^{-3}	1.1×10^{-3}	7.9×10^{-4}	4.2×10^{-4}	6
EXACT SOLUTION		0.0270461	0.0514447	0.0708075	0.0832392	0.0875229	0.0832392	0.0708075	0.0514447	0.0270461	-

TABLE 6.9.2: The absolute errors of the numerical solutions to Problem (6.9.1) (Parabolic problem with Dirichlet boundary conditions)

$\lambda=1.0$, $t=0.5$, $\Delta t=0.01$, $\Delta x=0.1$, $r=0.5$, $eps=10^{-4}$

x Scheme		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Number of Iterations
GER		9.0×10^{-6}	2.0×10^{-4}	2.0×10^{-4}	3.0×10^{-4}	3.0×10^{-4}	3.0×10^{-4}	4.0×10^{-4}	2.0×10^{-4}	2.0×10^{-4}	-
GEL		2.2×10^{-4}	2.1×10^{-4}	3.5×10^{-4}	3.3×10^{-4}	3.5×10^{-4}	3.3×10^{-4}	2.1×10^{-4}	2.1×10^{-4}	9×10^{-6}	-
(S)AGE		6.0×10^{-4}	1.5×10^{-3}	1.9×10^{-3}	2.4×10^{-3}	2.5×10^{-3}	2.3×10^{-3}	2.0×10^{-3}	1.3×10^{-3}	8.0×10^{-4}	-
(D)AGE		4.0×10^{-4}	3.0×10^{-4}	2.0×10^{-4}	3.0×10^{-4}	3.0×10^{-4}	2.4×10^{-4}	2.8×10^{-4}	7.0×10^{-5}	1.5×10^{-4}	-
AGE-IMP	PR	6.3×10^{-4}	1.3×10^{-3}	1.6×10^{-3}	2.0×10^{-3}	2.0×10^{-3}	1.9×10^{-3}	1.7×10^{-3}	1.2×10^{-3}	7.0×10^{-4}	2
	DR	8.1×10^{-4}	1.6×10^{-3}	2.1×10^{-3}	2.5×10^{-3}	2.6×10^{-3}	2.5×10^{-3}	2.2×10^{-3}	1.5×10^{-3}	9.1×10^{-4}	4
AGE-CN	PR	7.3×10^{-5}	1.5×10^{-4}	1.9×10^{-4}	2.3×10^{-4}	2.4×10^{-4}	2.2×10^{-4}	2.0×10^{-4}	1.4×10^{-4}	8.1×10^{-5}	2
	DR	3.0×10^{-4}	5.8×10^{-4}	7.7×10^{-4}	9.2×10^{-4}	9.6×10^{-4}	9.1×10^{-4}	7.9×10^{-4}	5.6×10^{-4}	3.1×10^{-4}	4
AGE-DG	PR	1.6×10^{-5}	2.7×10^{-5}	4.3×10^{-5}	4.9×10^{-5}	5.3×10^{-5}	5.1×10^{-5}	4.0×10^{-5}	3.1×10^{-5}	1.2×10^{-5}	2
	DR	2.0×10^{-4}	4.0×10^{-4}	5.4×10^{-4}	6.3×10^{-4}	6.6×10^{-4}	6.3×10^{-4}	5.4×10^{-4}	3.9×10^{-4}	2.2×10^{-4}	4
EXACT SOLUTION		0.0022936	0.0043628	0.006048	0.0070591	0.0074224	0.00705091	0.0060048	0.0043628	0.0022936	-

TABLE 6.9.3: The absolute errors of the numerical solutions to Problem (6.9.1)
(Parabolic problem with Dirichlet boundary conditions)

$\Delta x=0.1$, $r=0.5$, $eps=10^{-4}$ PR variant

Mesh Ratio λ	x / Scheme	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Average of Percentage Errors	Number of Iteration
0.1	(D)AGE	0.1693	0.1718	0.1745	0.1766	0.1774	0.1767	0.1747	0.1720	1.1696	0.1684	≈1%	-
Δt=0.001	AGE-IMP	0.1696	0.1720	0.1748	0.1768	0.1776	0.1768	0.1748	0.1720	1.1696	0.1687	1.004%	2
	AGE-CN	0.1697	0.1721	0.1747	0.1767	0.1774	0.1767	0.1747	0.1721	0.1697	0.1686	1.003%	2
t=0.1	AGE-DG	0.1699	0.1722	0.1746	0.1765	0.1772	0.1765	0.1746	0.1722	0.1699	0.1689	1.001%	2
	EXACT SOL.	0.1718	0.1739	0.1763	0.1781	0.1788	0.1781	0.1763	0.1739	0.1718	0.1709	-	-
0.5	(D)AGE	0.3556	0.3639	0.3724	0.3784	0.3808	0.3788	0.3730	0.3650	0.3567	0.3517	≈2%	-
Δt=0.005	AGE-IMP	0.3607	0.3682	0.3759	0.3814	0.3834	0.3814	0.3759	0.3682	0.3607	0.3568	1.003%	2
	AGE-CN	0.3607	0.3682	0.3759	0.3814	0.3834	0.3814	0.3759	0.3682	0.3607	0.3567	1.004%	2
t=0.5	AGE-DG	0.3607	0.3682	0.3759	0.3814	0.3834	0.3814	0.3759	0.3682	0.3607	0.3567	1.004%	2
	EXACT SOL.	0.3650	0.3720	0.3793	0.3846	0.3865	0.3846	0.3793	0.3720	0.3650	0.3618	-	-
1.0	(D)AGE	0.9526	0.9684	0.9841	0.9971	1.0028	0.9989	0.9866	0.9698	0.9525	0.9425	≈3%	-
Δt=0.01	AGE-IMP	0.9684	0.9837	0.9994	0.0108	1.0149	1.0108	0.9995	0.9837	0.9684	0.9603	1.002%	2
	AGE-CN	0.9683	0.9837	0.9995	1.0109	1.0150	1.0109	0.9995	0.9837	0.9683	0.9602	1.002%	2
t=1.0	AGE-DG	0.9683	0.9837	0.9995	1.0109	1.0150	1.0109	0.9995	0.9837	0.9683	0.9602	1.002%	2
	EXACT SOL.	0.9795	0.9937	1.0088	1.0197	1.0237	1.0197	1.0088	0.9937	0.9795	0.9729	-	-
1.5	(D)AGE	1.9832	1.9999	1.9769	1.9944	2.0403	2.0458	2.0097	1.9685	1.9350	1.9332	≈3%	-
Δt=0.015	AGE-IMP	1.9885	2.0117	2.0355	2.0528	2.0590	2.0528	2.0356	2.0117	1.9886	1.9764	1.001%	3
	AGE-CN	1.9884	2.0117	2.0356	2.0529	2.0592	2.0529	2.0356	2.0117	1.9884	1.9762	1.001%	2
t=1.5	AGE-DG	1.9884	2.0117	2.0356	2.0529	2.0592	2.0529	2.0356	2.0117	1.9884	1.9762	1.001%	2
	EXACT SOL.	2.0107	2.0322	2.0549	2.0715	2.0775	2.0715	2.0549	2.0322	2.0107	2.0007	-	-

TABLE 6.9.5: The numerical solutions to Problem (6.9.3)
(Parabolic problem with periodic boundary conditions)

$\Delta x=0.1$, $r=0.5$, $eps=10^{-4}$

DR variant

Mesh Ratio	x / Scheme	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Average of Percentage Errors	Number of Iterations
0.1 Δt=0.001 t=0.1	(D)AGE	0.1693	0.1718	0.1745	0.1766	0.1774	0.1767	0.1747	0.1720	11696	0.1684	1%	-
	AGE-IMP	0.1703	0.1738	0.1764	0.1775	0.1764	0.1737	0.1703	0.1672	0.1658	0.1672	1.802%	2
	AGE-CN	0.1703	0.1737	0.1763	0.1772	0.1763	0.1737	0.1703	0.1674	0.1660	0.1674	1.809%	2
	AGE-DG	0.1704	0.1735	0.1760	0.1769	0.1760	0.1735	0.1704	0.1676	0.1663	0.1676	1.828%	2
	EXACT SOL.	0.1718	0.1739	0.1763	0.1781	0.1788	0.1781	0.1763	0.1739	0.1718	0.1709	-	-
0.5 Δt=0.005 t=0.5	(D)AGE	0.3556	0.3639	0.3724	0.3784	0.3808	0.3788	0.3730	0.3650	0.3567	0.3517	2%	-
	AGE-IMP	0.3625	0.3701	0.3756	0.3776	0.3756	0.3701	0.3625	0.3550	0.3511	0.3550	2.528%	6
	AGE-CN	0.3625	0.3702	0.3757	0.3777	0.3757	0.3702	0.3625	0.3550	0.3510	0.3550	2.521%	6
	AGE-DG	0.3626	0.3702	0.3758	0.3778	0.3758	0.3702	0.3626	0.3550	0.3511	0.3550	2.508%	6
	EXACT SOL.	0.3650	0.3720	0.3793	0.3846	0.3865	0.3846	0.3793	0.3720	0.3650	0.3618	-	-
1.0 Δt=0.01 t=1.0	(D)AGE	0.9526	0.9684	0.9841	0.9971	1.0028	0.9989	0.9866	0.9698	0.9525	0.9425	3%	-
	AGE-IMP	0.9774	0.9931	1.0045	1.0086	1.0045	0.9932	0.9774	0.9621	0.9541	0.9621	1.632%	8
	AGE-CN	0.9773	0.9931	1.0045	1.0087	1.0045	0.9931	0.9773	0.9620	0.9539	0.9620	1.638%	8
	AGE-DG	0.9774	0.9932	1.0046	1.0087	1.0046	0.9932	0.9774	0.9620	0.9539	0.9620	1.632%	8
	EXACT SOL.	0.9795	0.9937	1.0088	1.0197	1.0237	1.0197	1.0088	0.9937	0.9795	0.9729	-	-
1.5 Δt=0.015 t=1.5	(D)AGE	1.9832	1.9999	1.9769	1.9944	2.0403	2.0458	2.0097	1.9685	1.9350	1.9332	3%	-
	AGE-IMP	2.0056	2.0294	2.0466	2.0528	2.0466	2.0294	2.0056	1.9824	1.9703	1.9824	1.302%	9
	AGE-CN	2.0052	2.0291	2.0464	2.0526	2.0464	2.0291	2.0052	1.9819	1.9697	1.9819	1.320%	9
	AGE-DG	2.0052	2.0292	2.0465	2.0527	2.0465	2.0292	2.0052	1.9820	1.9697	1.9820	1.316%	9
	EXACT SOL.	2.0107	2.0322	2.0549	2.0715	2.0775	2.0715	2.0549	2.0322	2.0107	2.0007	-	-

TABLE 6.9.6: The numerical solutions to Problem (6.9.3)

$k=1.0$, $\epsilon=1.0$, $\Delta t=0.001$, $\Delta x=0.1$, $\lambda=0.1$, $t=0.1$, $r=0.5$, $eps=5.0\times10^{-7}$

x Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Number of Iterations
CNU		2.4×10^{-3}	4.9×10^{-3}	7.3×10^{-3}	9.4×10^{-3}	1.1×10^{-2}	1.2×10^{-2}	1.15×10^{-2}	9.4×10^{-3}	5.6×10^{-3}	-
DAGE		1.6×10^{-4}	2.7×10^{-4}	2.8×10^{-4}	1.3×10^{-4}	9.0×10^{-5}	4.0×10^{-4}	6.1×10^{-4}	6.2×10^{-4}	4.3×10^{-4}	-
AGE-IMP	DR	1.3×10^{-4}	1.4×10^{-4}	5.4×10^{-5}	4.8×10^{-4}	1.04×10^{-3}	1.6×10^{-3}	1.9×10^{-3}	1.8×10^{-3}	1.1×10^{-3}	12
	PR	1.3×10^{-4}	1.5×10^{-4}	4.4×10^{-5}	4.7×10^{-4}	1.0×10^{-3}	1.6×10^{-3}	1.9×10^{-3}	1.8×10^{-3}	1.1×10^{-3}	3
AGE-CNU	DR	2.4×10^{-3}	4.9×10^{-3}	7.37×10^{-3}	9.6×10^{-3}	1.1×10^{-2}	1.2×10^{-2}	1.15×10^{-2}	9.4×10^{-3}	5.6×10^{-3}	12
	PR	2.4×10^{-3}	4.9×10^{-3}	7.38×10^{-3}	9.6×10^{-3}	1.1×10^{-2}	1.2×10^{-2}	1.15×10^{-2}	9.4×10^{-3}	5.6×10^{-3}	3
AGE-CN	DR	1.7×10^{-4}	2.8×10^{-4}	2.7×10^{-4}	9.1×10^{-5}	1.8×10^{-4}	5.1×10^{-4}	7.2×10^{-4}	7.2×10^{-4}	4.6×10^{-4}	12
	PR	1.8×10^{-4}	2.9×10^{-4}	2.8×10^{-4}	1.0×10^{-4}	1.6×10^{-4}	4.9×10^{-4}	7.0×10^{-4}	7.1×10^{-4}	4.5×10^{-4}	3
EXACT SOLUTION		0.01895	0.04370	0.07892	0.12982	0.20177	0.29986	0.42794	0.58805	0.77976	-

eps: convergence criterion

TABLE 6.9.7: The absolute errors of the numerical solutions to Problem (6.9.5) (Diffusion-convection problem)

$k=1.0$, $\epsilon=1.0$, $\Delta t=0.005$, $\Delta x=0.1$, $\lambda=0.5$, $t=0.5$, $r=0.5$, $eps=5.0\times10^{-7}$

x Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Number of Iterations
CNU		1.7×10^{-3}	3.2×10^{-3}	4.5×10^{-3}	5.4×10^{-3}	6.1×10^{-3}	6.2×10^{-3}	5.8×10^{-3}	4.7×10^{-3}	2.8×10^{-3}	-
DAGE		2.0×10^{-5}	6.0×10^{-5}	5.0×10^{-5}	8.0×10^{-5}	8.0×10^{-5}	8.0×10^{-5}	9.0×10^{-5}	6.0×10^{-5}	6.0×10^{-5}	-
AGE-IMP	DR	1.6×10^{-4}	3.22×10^{-4}	4.6×10^{-4}	5.7×10^{-4}	6.3×10^{-4}	6.3×10^{-4}	5.7×10^{-4}	4.4×10^{-4}	2.4×10^{-4}	9
	PR	1.6×10^{-4}	3.19×10^{-4}	4.6×10^{-4}	5.7×10^{-4}	6.3×10^{-4}	6.3×10^{-4}	5.7×10^{-4}	4.4×10^{-4}	2.5×10^{-4}	4
AGE-CNU	DR	1.7×10^{-3}	3.2×10^{-3}	4.4×10^{-3}	5.4×10^{-3}	6.04×10^{-3}	6.2×10^{-3}	5.8×10^{-3}	4.7×10^{-3}	2.8×10^{-3}	9
	PR	1.7×10^{-3}	3.2×10^{-3}	4.5×10^{-3}	5.4×10^{-3}	6.1×10^{-3}	6.2×10^{-3}	5.8×10^{-3}	4.7×10^{-3}	2.8×10^{-3}	3
AGE-CN	DR	5.8×10^{-5}	1.1×10^{-4}	1.6×10^{-4}	1.98×10^{-4}	2.2×10^{-4}	2.2×10^{-4}	2.0×10^{-4}	1.6×10^{-4}	9.3×10^{-5}	9
	PR	5.6×10^{-5}	1.1×10^{-4}	1.6×10^{-4}	1.9×10^{-4}	2.1×10^{-4}	2.2×10^{-4}	1.97×10^{-4}	1.6×10^{-4}	8.99×10^{-5}	3
EXACT SOLUTION.		0.06043	0.12730	0.20136	0.28345	0.37447	0.47539	0.58724	0.71114	0.84830	-

eps: convergence criterion

TABLE 6.9.8: The absolute errors of the numerical solutions to Problem (6.9.5) (Diffusion-convection problem)

$k=1.0,\ \varepsilon=1.0,\ \Delta t=0.01,\ \Delta x=0.1,\ \lambda=1.0,\ t=1.0,\ r=0.5,\ eps=5.0\times10^{-7}$

Method \ x		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Number of Iterations
CNU		1.5×10^{-3}	2.9×10^{-3}	4.1×10^{-3}	4.9×10^{-3}	5.5×10^{-3}	5.7×10^{-3}	5.3×10^{-3}	4.3×10^{-5}	2.6×10^{-5}	-
DAGE		2.7×10^{-5}	5.2×10^{-5}	7.2×10^{-5}	8.8×10^{-5}	9.8×10^{-5}	1.0×10^{-4}	9.4×10^{-5}	7.6×10^{-5}	4.6×10^{-5}	-
AGE-IMP	DR	3.1×10^{-5}	6.0×10^{-5}	8.4×10^{-5}	1.0×10^{-4}	1.1×10^{-4}	1.2×10^{-4}	1.1×10^{-4}	8.8×10^{-5}	5.3×10^{-5}	3
	PR	3.1×10^{-5}	5.9×10^{-5}	8.2×10^{-5}	1.0×10^{-4}	1.1×10^{-4}	1.1×10^{-4}	1.1×10^{-4}	8.6×10^{-5}	5.2×10^{-5}	2
AGE-CNU	DR	1.5×10^{-3}	2.9×10^{-3}	4.1×10^{-3}	4.97×10^{-3}	5.5×10^{-3}	5.7×10^{-3}	5.3×10^{-3}	4.3×10^{-3}	2.6×10^{-3}	3
	PR	1.5×10^{-3}	2.9×10^{-3}	4.1×10^{-3}	4.98×10^{-3}	5.5×10^{-3}	5.7×10^{-3}	5.3×10^{-3}	4.3×10^{-3}	2.6×10^{-3}	2
AGE-CN	DR	2.8×10^{-5}	5.4×10^{-5}	7.6×10^{-5}	9.3×10^{-5}	1.0×10^{-4}	1.1×10^{-4}	9.9×10^{-5}	8.0×10^{-5}	4.8×10^{-5}	3
	PR	2.7×10^{-5}	5.2×10^{-5}	7.3×10^{-5}	8.9×10^{-5}	9.9×10^{-5}	1.0×10^{-4}	9.6×10^{-5}	7.8×10^{-5}	4.7×10^{-5}	2
EXACT SOLUTION		0.06120	0.12884	0.20360	0.28621	0.37752	0.47843	0.58996	0.71322	0.84945	-

eps: convergence criterion

TABLE 6.9.9: The absolute errors of the numerical solutions to Problem (6.9.5) (Diffusion-convection problem)

$k=1.0$, $\varepsilon=1.0$, $\Delta t=0.020$, $\Delta x=0.1$, $\lambda=2.0$, $t=2.0$, $r=0.5$, $eps=5\times10^{-7}$

x Method	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Number of Iterations
CNU	1.5×10^{-3}	2.9×10^{-3}	4.1×10^{-3}	4.9×10^{-3}	5.5×10^{-3}	5.7×10^{-3}	5.3×10^{-3}	4.3×10^{-3}	2.6×10^{-3}	-
DAGE TWO-LEVEL	3.0×10^{-5}	5.0×10^{-5}	7.0×10^{-5}	9.0×10^{-5}	1.0×10^{-4}	1.0×10^{-4}	1.0×10^{-4}	8.0×10^{-5}	5.0×10^{-5}	-
AGE-IMP DR	2.7×10^{-5}	5.1×10^{-5}	7.2×10^{-5}	8.8×10^{-5}	9.8×10^{-5}	10×10^{-4}	9.4×10^{-5}	7.7×10^{-5}	4.6×10^{-5}	1
PR	2.7×10^{-5}	5.1×10^{-5}	7.2×10^{-5}	8.8×10^{-5}	9.8×10^{-5}	1.0×10^{-4}	9.4×10^{-5}	7.7×10^{-5}	4.6×10^{-5}	1
AGE-CNU DR	1.5×10^{-3}	2.9×10^{-3}	4.1×10^{-3}	4.97×10^{-3}	5.5×10^{-3}	5.7×10^{-3}	5.3×10^{-3}	4.3×10^{-3}	2.6×10^{-3}	1
PR	1.5×10^{-3}	2.9×10^{-3}	4.1×10^{-3}	4.97×10^{-3}	5.5×10^{-3}	5.7×10^{-3}	5.3×10^{-3}	4.3×10^{-3}	2.6×10^{-3}	1
AGE-CN DR	2.7×10^{-5}	5.2×10^{-5}	7.2×10^{-5}	8.8×10^{-5}	9.8×10^{-5}	1.0×10^{-4}	9.4×10^{-5}	7.7×10^{-5}	4.6×10^{-5}	1
PR	2.7×10^{-5}	5.1×10^{-5}	7.2×10^{-5}	8.8×10^{-5}	9.8×10^{-5}	1.0×10^{-4}	9.4×10^{-5}	7.7×10^{-5}	4.6×10^{-5}	1
EXACT SOLUTION	0.061207	0.128851	0.203610	0.28631	0.377541	0.478454	0.589980	0.713236	0.849455	-

eps: convergence criterion

TABLE 6.9.10: The absolute errors of the numerical solutions to Problem (6.9.5) (Diffusion-convection problem)

(a) $t=0.4$, $\lambda=0.5$, $\Delta t=0.05$, $\Delta x=0.1$, $r=0.5$, $eps=10^{-4}$

x Scheme		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter atio
AGE-IMP	PR	2.46×10^{-3}	4.94×10^{-3}	7.16×10^{-3}	8.18×10^{-3}	9.79×10^{-3}	8.39×10^{-3}	1.18×10^{-2}	5.53×10^{-3}	1.53×10^{-2}	8.17×10^{-3}	3
	DR	2.31×10^{-3}	4.92×10^{-3}	7.04×10^{-3}	8.24×10^{-3}	9.93×10^{-3}	8.53×10^{-3}	1.24×10^{-2}	5.67×10^{-3}	1.61×10^{-2}	8.35×10^{-3}	9
AGE-CN	PR	5.81×10^{-5}	7.19×10^{-5}	4.16×10^{-5}	5.25×10^{-6}	9.75×10^{-5}	9.82×10^{-5}	3.44×10^{-4}	5.77×10^{-5}	7.15×10^{-4}	1.65×10^{-4}	3
	DR	1.52×10^{-4}	1.29×10^{-4}	9.54×10^{-5}	4.6×10^{-5}	2.23×10^{-4}	2.42×10^{-4}	7.57×10^{-4}	1.75×10^{-4}	1.51×10^{-3}	3.7×10^{-4}	9
EXACT SOLUTION		0.9553365	0.9800666	0.9950042	1.0	0.9950042	0.9800666	0.9553365	0.9210610	0.8775826	-	-

(b) $t=1.0$, $\lambda=0.5$, $\Delta t=0.05$, $\Delta x=0.1$, $r=0.5$, $eps=10^{-4}$

x Scheme		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter-ation
AGE-IMP	PR	5.63×10^{-3}	5.39×10^{-4}	1.23×10^{-2}	6.04×10^{-4}	2.21×10^{-2}	2.51×10^{-4}	3.3×10^{-2}	1.59×10^{-4}	4.19×10^{-2}	1.29×10^{-2}	4
	DR	5.43×10^{-3}	1.08×10^{-3}	1.19×10^{-3}	1.64×10^{-3}	2.18×10^{-3}	1.46×10^{-3}	3.28×10^{-2}	6.82×10^{-4}	4.19×10^{-2}	1.32×10^{-2}	9
AGE-CN	PR	1.16×10^{-6}	4.58×10^{-4}	4.99×10^{-5}	8.29×10^{-4}	7.27×10^{-6}	9.04×10^{-4}	4.42×10^{-6}	5.75×10^{-4}	3.49×10^{-5}	3.18×10^{-4}	3
	DR	7.29×10^{-5}	9.64×10^{-4}	2.28×10^{-4}	1.9×10^{-3}	1.29×10^{-4}	2.19×10^{-3}	4.4×10^{-5}	1.45×10^{-3}	1.9×10^{-4}	7.96×10^{-4}	9
EXACT SOLUTION		0.6216100	0.6967067	0.7648422	0.8253356	0.8775826	0.9210610	0.9553365	0.9800666	0.9950042	-	-

TABLE 6.9.11: The absolute errors of the AGE solutions to hyperbolic Problem 1 (6.9.7)

(a) t=0.8, λ=1.0, Δt=0.1, Δx=0.1, r=0.5, eps=10^{-4}

x Scheme		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iteratio
AGE-IMP	PR	8.23×10^{-3}	4.36×10^{-3}	2.06×10^{-2}	8.89×10^{-3}	3.67×10^{-2}	1.02×10^{-2}	5.35×10^{-2}	6.73×10^{-3}	6.77×10^{-2}	2.41×10^{-2}	5
	DR	7.92×10^{-3}	4.27×10^{-3}	2.03×10^{-2}	8.68×10^{-3}	3.65×10^{-2}	9.88×10^{-3}	5.34×10^{-2}	6.5×10^{-3}	6.78×10^{-2}	2.39×10^{-2}	12
AGE-CN	PR	1.1×10^{-4}	3.63×10^{-4}	2.11×10^{-4}	6.66×10^{-4}	4.013×10^{-5}	8.18×10^{-4}	4.81×10^{-4}	5.72×10^{-4}	7.29×10^{-4}	4.43×10^{-4}	4
	DR	2.97×10^{-4}	5.47×10^{-4}	4.73×10^{-4}	1.01×10^{-3}	5.85×10^{-5}	1.19×10^{-3}	6.94×10^{-4}	8.31×10^{-4}	1.2×10^{-3}	7.0×10^{-4}	10
EXACT SOLUTION		0.7648422	0.8253356	0.8775826	0.921061	0.9553365	0.9800666	0.9950042	1.0	0.9950042	-	-

(b) t=2.0, λ=1.0, Δt=0.1, Δx=0.1, r=0.5, eps=10^{-4}

x Scheme		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter- atio
AGE-IMP	PR	7.57×10^{-2}	1.18×10^{-2}	7.09×10^{-2}	1.81×10^{-2}	6.67×10^{-2}	1.79×10^{-2}	6.47×10^{-2}	1.15×10^{-2}	6.61×10^{-2}	4.48×10^{-2}	6
	DR	7.51×10^{-2}	1.15×10^{-2}	7.01×10^{-2}	1.76×10^{-2}	6.57×10^{-2}	1.75×10^{-2}	6.35×10^{-2}	1.12×10^{-2}	6.49×10^{-2}	4.41×10^{-2}	14
AGE-CN	PR	1.6×10^{-4}	4.24×10^{-4}	8.72×10^{-4}	5.26×10^{-4}	1.92×10^{-3}	3.88×10^{-4}	2.99×10^{-3}	1.28×10^{-4}	3.75×10^{-3}	1.24×10^{-3}	4
	DR	1.41×10^{-4}	7.73×10^{-4}	1.25×10^{-3}	1.01×10^{-3}	2.78×10^{-3}	7.92×10^{-4}	4.18×10^{-3}	3.53×10^{-4}	5.07×10^{-3}	1.82×10^{-3}	11
EXACT SOLUTION		-0.3232896	-0.2272021	-0.1288445	-0.2919952	0.0707372	0.1699671	0.2674988	0.3623578	0.4535961	-	-

TABLE 6.9.12: The absolute errors of the AGE solutions to hyperbolic Problem 1 (6.9.7)

(a) t=0.4, λ=0.5, Δt=0.05, Δx=0.1, r=0.5, eps=10^{-4}

x Scheme		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter-ations
AGE-IMP	PR	4.42×10^{-3}	1.04×10^{-2}	1.78×10^{-2}	2.1×10^{-2}	2.97×10^{-2}	2.3×10^{-2}	4.27×10^{-2}	1.38×10^{-2}	6.04×10^{-2}	2.48×10^{-2}	4
	DR	4.36×10^{-3}	1.05×10^{-2}	1.8×10^{-2}	2.13×10^{-2}	3.01×10^{-2}	2.33×10^{-2}	4.34×10^{-2}	1.41×10^{-2}	6.13×10^{-2}	2.52×10^{-2}	10
AGE-CN	PR	6.89×10^{-5}	2.08×10^{-4}	3.58×10^{-4}	4.56×10^{-4}	5.62×10^{-4}	5.04×10^{-4}	7.86×10^{-4}	2.9×10^{-4}	1.17×10^{-3}	4.9×10^{-4}	3
	DR	8.27×10^{-5}	3.32×10^{-4}	5.5×10^{-4}	7.39×10^{-4}	9.26×10^{-4}	8.56×10^{-4}	1.36×10^{-3}	5.56×10^{-4}	2.04×10^{-3}	8.26×10^{-4}	9
EXACT SOLUTION		-0.1327158	-0.0890597	-0.044858	0	0.044858	0.0892679	0.1327858	0.1749769	0.2154198	-	-

(b) t=1.0, λ=0.5, Δt=0.05, Δx=0.1, r=0.5, eps=10^{-4}

x Scheme		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter-ations
AGE-IMP	PR	1.85×10^{-2}	1.45×10^{-2}	2.82×10^{-2}	2.6×10^{-2}	4.37×10^{-2}	2.73×10^{-2}	5.94×10^{-2}	1.65×10^{-2}	7.04×10^{-2}	3.38×10^{-2}	3
	DR	1.83×10^{-2}	1.47×10^{-2}	2.81×10^{-2}	2.64×10^{-2}	4.43×10^{-2}	2.77×10^{-2}	6.08×10^{-2}	1.68×10^{-2}	7.23×10^{-2}	3.44×10^{-2}	7
AGE-CN	PR	2.22×10^{-4}	2.63×10^{-4}	4.99×10^{-4}	5.97×10^{-4}	8.46×10^{-4}	6.47×10^{-4}	1.17×10^{-3}	3.34×10^{-4}	1.44×10^{-3}	6.69×10^{-4}	3
	DR	1.89×10^{-4}	6.38×10^{-4}	6.11×10^{-4}	1.2×10^{-3}	1.43×10^{-3}	1.25×10^{-3}	2.35×10^{-3}	6.76×10^{-4}	3.14×10^{-3}	1.28×10^{-3}	7
EXACT SOLUTION		-0.1060118	-0.0970836	-0.0871854	-0.0764160	-0.0648832	-0.052702	-0.03999431	-0.026887	-0.013511	-	-

TABLE 6.9.13: The absolute errors of the AGE solutions to hyperbolic Problem 2 (6.9.9)

(a) $t=0.8$, $\lambda=1.0$, $\Delta t=0.1$, $\Delta x=0.1$, $r=0.5$, $eps=10^{-4}$

Scheme \ x		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter- ations
AGE-IMP	PR	2.1×10^{-2}	1.1×10^{-2}	4.1×10^{-2}	1.62×10^{-2}	7.85×10^{-2}	1.71×10^{-2}	1.2×10^{-1}	1.27×10^{-2}	1.48×10^{-1}	5.17×10^{-2}	5
	DR	2.07×10^{-2}	1.1×10^{-2}	4.1×10^{-2}	1.62×10^{-2}	7.89×10^{-2}	1.71×10^{-2}	1.2×10^{-1}	1.27×10^{-2}	1.49×10^{-1}	5.19×10^{-2}	10
AGE-CN	PR	9.97×10^{-5}	2.97×10^{-4}	3.27×10^{-4}	7.44×10^{-4}	1.26×10^{-3}	1.04×10^{-3}	2.35×10^{-3}	6.98×10^{-4}	2.74×10^{-3}	1.06×10^{-3}	3
	DR	3.41×10^{-5}	3.81×10^{-4}	3.52×10^{-4}	8.43×10^{-4}	1.55×10^{-3}	1.09×10^{-3}	2.97×10^{-3}	7.22×10^{-4}	3.59×10^{-3}	1.28×10^{-3}	9
EXACT SOLUTION		-0.1300653	-0.113993	-0.0967943	-0.0786222	-0.0596645	-0.0401106	-0.02015602	0	0.020156	-	-

(b) $t=2.0$, $\lambda=1.0$, $\Delta t=0.1$, $\Delta x=0.1$, $r=0.5$, $eps=10^{-4}$

Scheme \ x		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter- ations
AGE-IMP	PR	1.19×10^{-1}	1.18×10^{-2}	1.06×10^{-1}	1.88×10^{-2}	8.62×10^{-2}	1.85×10^{-2}	6.64×10^{-2}	1.14×10^{-2}	5.43×10^{-2}	5.48×10^{-2}	4
	DR	1.2×10^{-1}	1.16×10^{-2}	1.07×10^{-1}	1.85×10^{-2}	8.59×10^{-2}	1.83×10^{-2}	6.56×10^{-2}	1.12×10^{-2}	5.32×10^{-2}	5.46×10^{-2}	9
AGE-CN	PR	2.05×10^{-3}	1.05×10^{-3}	2.01×10^{-3}	1.52×10^{-3}	1.28×10^{-3}	1.33×10^{-3}	4.64×10^{-4}	6.61×10^{-4}	2.11×10^{-4}	1.18×10^{-3}	3
	DR	2.89×10^{-3}	9.6×10^{-4}	2.45×10^{-3}	1.37×10^{-3}	1.11×10^{-3}	1.2×10^{-3}	3.03×10^{-4}	5.95×10^{-4}	9.0×10^{-4}	1.31×10^{-3}	6
EXACT SOLUTION		-0.0173321	-0.0178366	-0.018163	-0.0183078	-0.0182698	-0.0180491	-0.0176482	-0.0170709	-0.016323	-	-

TABLE 6.9.14: The absolute errors of the AGE solutions to hyperbolic Problem 2 (6.9.9)

x Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter- ation
(a) $t=1.0$, $\lambda=0.5$, $\Delta t=0.05$, $\Delta x=0.1$, $r=0.5$, $eps=10^{-4}$												
AGE	PR	6.43×10^{-5}	1.22×10^{-4}	1.69×10^{-4}	1.98×10^{-4}	2.08×10^{-4}	1.98×10^{-4}	1.69×10^{-4}	1.22×10^{-4}	6.44×10^{-5}	1.46×10^{-4}	2
	DR	1.86×10^{-3}	3.54×10^{-3}	4.87×10^{-3}	5.72×10^{-3}	6.02×10^{-3}	5.72×10^{-3}	4.87×10^{-3}	3.54×10^{-3}	1.86×10^{-3}	4.22×10^{-3}	6
EXACT SOLUTION		-0.0386272	-0.0734732	-0.1011271	-0.1188821	-0.125000	-0.1188827	-0.1011271	-0.0734732	-0.0386272	-	-
(b) $t=2.0$, $\lambda=1.0$, $\Delta t=0.1$, $\Delta x=0.1$, $r=0.5$, $eps=10^{-4}$												
AGE	PR	5.6×10^{-4}	1.06×10^{-3}	1.47×10^{-3}	1.73×10^{-3}	1.81×10^{-3}	1.73×10^{-3}	1.47×10^{-3}	1.07×10^{-3}	5.59×10^{-4}	1.27×10^{-3}	3
	DR	1.47×10^{-3}	2.8×10^{-3}	3.83×10^{-3}	4.52×10^{-3}	4.74×10^{-3}	4.5×10^{-3}	3.85×10^{-3}	2.78×10^{-3}	1.47×10^{-3}	3.33×10^{-3}	8
EXACT SOLUTION		0.0386272	0.0734732	0.1011271	0.1188821	0.125000	0.1188821	0.1011271	0.0734732	0.0386272	-	-

TABLE 6.9.15: The absolute errors of the AGE solutions to the wave equation for Problem 1 (6.9.11)

x Method		0	0.12	0.24	0.36	0.48	0.6	0.72	0.84	0.96	Average of all percentage errors	No. of iter-ation
(a) $t=0.4$, $\lambda=0.1$, $\Delta t=0.004$, $\Delta x=0.04$, $r=0.5$, $eps=10^{-4}$												
AGE	PR	9.99×10^{-1}	5.92×10^{-1}	3.91×10^{-1}	2.77×10^{-1}	2.07×10^{-1}	1.55×10^{-1}	8.87×10^{-2}	4.15×10^{-2}	8.57×10^{-3}	2.86×10^{-1}	2
	DR	1.49×10	9.17×10^{-1}	6.64×10^{-1}	5.29×10^{-1}	4.42×10^{-1}	3.73×10^{-1}	2.64×10^{-1}	1.47×10^{-1}	3.57×10^{-2}	5.16×10^{-1}	14
EXACT SOLUTION		16	27.04	40.96	57.76	77.44	100	125.44	153.76	184.96	-	-
(b) $t=2.0$, $\lambda=0.5$, $\Delta t=0.02$, $\Delta x=0.04$, $r=0.5$, $eps=10^{-4}$												
AGE	PR	1.46×10^{-3}	7.53×10^{-4}	5.58×10^{-4}	1.32×10^{-4}	1.34×10^{-3}	2.05×10^{-3}	2.61×10^{-3}	3.79×10^{-3}	2.93×10^{-3}	1.51×10^{-3}	4
	DR	2.6×10^{-2}	2.35×10^{-2}	2.17×10^{-2}	1.83×10^{-2}	1.38×10^{-2}	1.39×10^{-2}	5.82×10^{-3}	8.64×10^{-3}	4.14×10^{-3}	1.53×10^{-2}	17
EXACT SOLUTION		400	449.44	501.76	556.96	615.04	676	739.84	806.56	876.16	-	-
(c) $t=8.0$, $\lambda=2.0$, $\Delta t=0.08$, $\Delta x=0.04$, $r=0.5$, $eps=10^{-4}$												
AGE	PR	3.12×10^{-3}	9.96×10^{-4}	2.97×10^{-3}	3.33×10^{-3}	1.75×10^{-3}	7.36×10^{-3}	8.77×10^{-3}	4.91×10^{-3}	8.35×10^{-4}	3.53×10^{-3}	14
	DR	3.11×10^{-3}	1.01×10^{-3}	2.97×10^{-3}	3.33×10^{-3}	1.75×10^{-3}	7.37×10^{-3}	8.77×10^{-3}	4.91×10^{-3}	8.35×10^{-4}	3.53×10^{-3}	33
EXACT SOLUTION		6400	6593.44	6789.76	6988.96	7191.04	7396	7603.84	7814.56	8028.16	-	-

TABLE 6.9.16: The percentage errors of the AGE solutions to the wave equation for Problem 2 (6.9.13)

t=0.05, Δx=0.1, λ=0.05, eps=10^{-4}

x Scheme		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iterations
(D)AGE		1.5000$_{(3)}$	1.43409	1.36976$_{(3)}$	1.30713	1.24630$_{(3)}$	1.18740	1.13056$_{(3)}$	1.07590	1.02356$_{(3)}$	NA	-
AGE-IMP(RCHM)	PR	1.49914	1.43226	1.36766	1.30452	1.24381	1.18481	1.12845	1.07411	1.02267	1.92×10^{-3}	1
	DR	1.49830	1.43049	1.36561	1.30200	1.24141	1.18233	1.12643	1.07240	1.02183	3.77×10^{-3}	1
AGE-CN(RCHM)	PR	1.49957	1.43318	1.36868	1.30580	1.24502	1.18608	1.12947	1.07498	1.02309	9.82×10^{-4}	1
	DR	1.49914	1.43227	1.36762	1.30450	1.24375	1.18478	1.12839	1.07408	1.02264	1.95×10^{-3}	1
CN		1.50000	1.43409	1.36976	1.30713	1.24630	1.18740	1.13056	1.07590	1.02355	NA	-
AGE-DG(RCHM)	PR	1.50098	1.43618	1.37241	1.31033	1.24954	1.19069	1.13337	1.07820	1.02478	2.42×10^{-3}	1
	DR	1.50204	1.43841	1.37527	1.31381	1.25308	1.19430	1.13647	1.08075	1.02613	5.06×10^{-3}	1
AGE-LEE	PR	1.500098	1.43430	1.37011	1.30753	1.24677	1.18785	1.13099	1.07619	1.02370	3.14×10^{-4}	2
	DR	1.49909	1.43254	1.36779	1.30485	1.24394	1.18505	1.12845	1.07415	1.02247	1.82×10^{-3}	3
LEE		1.50000	1.43409	1.36976	1.30713	1.24630	1.18740	1.13056	1.07590	1.02355	NA	-
EXACT SOLUTION		1.50000	1.43409	1.36974	1.30713	1.24631	1.18741	1.13057	1.07590	1.02355	-	-

TABLE 6.9.17: Numerical solutions to non-linear problem (6.9.15)

t=0.1, Δx=0.1, λ=0.1, eps=10^{-4}

x Scheme		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iterations
(D)AGE		1.56738 (3)	1.50001	1.43407 (3)	1.36976	1.30710 (3)	1.24630	1.18738 (3)	1.13056	1.07588 (3)	NA	-
AGE-IMP(RCHM)	PR	1.56514	1.49530	1.42859	1.36286	1.30041	1.23928	1.18162	1.12569	1.07351	5.13×10^{-3}	1
	DR	1.56300	1.49078	1.42334	1.35629	1.29402	1.23263	1.17614	1.12110	1.07125	9.99×10^{-3}	1
AGE-CN(RCHM)	PR	1.56627	1.49766	1.43126	1.36625	1.30364	1.24270	1.18438	1.12805	1.07462	2.63×10^{-3}	1
	DR	1.56517	1.49536	1.42848	1.36280	1.30021	1.23915	1.18141	1.12558	1.07337	5.22×10^{-3}	1
CN		1.56739	1.5000	1.43408	1.36976	1.30712	1.24630	1.18740	1.13056	1.07590	NA	-
AGE-DG(RCHM)	PR	1.56809	1.50148	1.43599	1.37210	1.30953	1.24874	1.18950	1.13227	1.07680	1.77×10^{-3}	1
	DR	1.56882	1.50302	1.43797	1.37452	1.31201	1.25127	1.19168	1.13404	1.07774	3.62×10^{-3}	1
AGE-LEE	PR	1.5675	1.50025	1.43450	1.37024	1.30769	1.24684	1.18792	1.13092	1.07608	3.78×10^{-4}	2
	DR	1.56691	1.49920	1.43309	1.36859	1.30591	1.24507	1.18631	1.12964	1.07531	9.45×10^{-4}	4
LEE		1.56739	1.50000	1.43408	1.36976	1.30712	1.24630	1.18740	1.13056	1.07590	NA	-
EXACT SOLUTION		1.56739	1.50000	1.43409	1.36974	1.30713	1.24631	1.18741	1.13057	1.07590	-	-

TABLE 6.9.18: Numerical solutions to non-linear problem (6.9.15)

$\varepsilon=1.0$, $\lambda=1.0$, $\Delta t=0.01$, $\Delta x=0.1$, $t=1.0$, eps=10^{-6}

x Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	No.of iterations
DAGE		$7.0\times10^{-8}_{(2)}$	2.5×10^{-7}	$3.7\times10^{-7}_{(2)}$	2.9×10^{-7}	$6.5\times10^{-7}_{(2)}$	3.5×10^{-7}	$6.1\times10^{-7}_{(2)}$	9.6×10^{-7}	$1.22\times10^{-6}_{(2)}$	-
AGE-IMP	PR	8.4×10^{-4}	1.5×10^{-3}	1.99×10^{-3}	2.3×10^{-3}	2.4×10^{-3}	2.3×10^{-3}	2.0×10^{-3}	1.6×10^{-3}	8.8×10^{-4}	5
	DR	8.5×10^{-4}	1.5×10^{-3}	1.99×10^{-3}	2.3×10^{-3}	2.4×10^{-3}	2.3×10^{-3}	2.0×10^{-3}	1.6×10^{-3}	8.8×10^{-4}	10
AGE-CN	PR	8.2×10^{-8}	1.1×10^{-7}	1.8×10^{-7}	1.4×10^{-7}	2.2×10^{-7}	1.1×10^{-7}	2.9×10^{-7}	1.1×10^{-7}	2.7×10^{-7}	4
	DR	1.8×10^{-6}	2.9×10^{-6}	3.9×10^{-6}	4.3×10^{-6}	4.5×10^{-6}	4.2×10^{-6}	3.8×10^{-6}	2.5×10^{-6}	2.4×10^{-6}	10
EXACT SOLUTION		0.58918851	0.58233803	0.5747863	0.56861393	0.56174757	0.55488319	0.54802440	0.54117482	0.53433803	-

$\varepsilon=0.1$, $\lambda=1.0$, $\Delta t=0.01$, $\Delta x=0.1$, $t=1.0$, eps=10^{-6}

x Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	No.of iteratio
DAGE		$1.83\times10^{-4}_{(3)}$	3.26×10^{-4}	$3.88\times10^{-4}_{(3)}$	2.39×10^{-4}	$9.6\times10^{-5}_{(3)}$	6.49×10^{-4}	$1.19\times10^{-3}_{(3)}$	1.51×10^{-3}	$1.22\times10^{-3}_{(3)}$	-
AGE-IMP	PR	2.9×10^{-2}	6.1×10^{-2}	9.4×10^{-2}	1.236×10^{-1}	1.4498×10^{-1}	1.542×10^{-1}	1.4769×10^{-1}	1.2229×10^{-1}	7.485×10^{-2}	5
	DR	2.9×10^{-2}	6.1×10^{-2}	9.4×10^{-2}	1.236×10^{-1}	1.450×10^{-1}	1.543×10^{-1}	1.4777×10^{-1}	1.2235×10^{-1}	7.489×10^{-2}	14
AGE-CN	PR	1.62×10^{-4}	2.78×10^{-4}	2.83×10^{-4}	1.07×10^{-4}	2.8×10^{-4}	8.3×10^{-4}	1.38×10^{-3}	1.64×10^{-3}	1.30×10^{-3}	3
	DR	1.60×10^{-4}	2.66×10^{-4}	2.59×10^{-4}	7.17×10^{-5}	3.3×10^{-4}	8.94×10^{-4}	1.45×10^{-3}	1.71×10^{-3}	1.34×10^{-3}	13
EXACT SOLUTION		0.932745	0.911271	0.883314	0.847514	0.802758	0.748601	0.685736	0.616304	0.543775	-

TABLE 6.9.20: The absolute errors of the numerical solutions to Burgers' equation for Problem 1 (6.9.19)

$\varepsilon=0.003$, $\lambda=1.0$, $\Delta x=0.01$, $\Delta t=0.0001$, $t=0.1$, $eps=10^{-6}$

Method x	DAGE	PR		DR		EXACT SOLUTION
		AGE-IMP	AGE-CN	AGE-IMP	AGE-CN	
0.05	1.000000	1.000000	1.000000	0.999999	0.999999	1.000000
0.10	1.000000	0.999999	0.999999	0.999999	0.999999	1.000000
0.15	1.000000	0.999999	0.999999	0.999999	0.999999	1.000000
0.20	0.999993	0.999976	0.999994	0.999976	0.999994	0.999985
0.25	0.999572	0.998482	0.999572	0.998475	0.999569	0.999037
0.30	0.952565	0.930082	0.952562	0.929798	0.952191	0.944636
0.35	0.557343	0.593433	0.557344	0.593049	0.557074	0.555361
0.40	0.501467	0.503739	0.501468	0.503713	0.501458	0.500894
0.45	0.498438	0.499422	0.498438	0.499421	0.498435	0.498093
0.50	0.454613	0.484010	0.454613	0.483994	0.454527	0.452319
0.55	0.181832	0.325894	0.181833	0.325743	0.181710	0.183443
0.60	0.103999	0.129626	0.103999	0.129561	0.103992	0.103726
0.65	0.100148	0.101435	0.100148	0.101430	0.100148	0.100134
0.70	0.100005	0.100053	0.100005	0.100053	0.100005	0.100004
0.75	0.100000	0.100002	0.100000	0.100002	0.100000	0.100000
0.80	0.100000	0.100000	0.100000	0.100000	0.100000	0.100000
0.85	0.100000	0.100000	0.100000	0.100000	0.100000	0.100000
0.90	0.100000	0.100000	0.100000	0.100000	0.100000	0.100000
0.95	0.100000	0.100000	0.100000	0.100000	0.099999	0.100000
Number of iterations	-	2	2	9	10	-

TABLE 6.9.21: The numerical solutions to Burgers' equation for Problem 1 (6.9.19)

$\varepsilon=0.003$, $\lambda=1.0$, $\Delta x=0.01$, $\Delta t=0.0001$, $t=0.5$, $eps=10^{-6}$

Method / x	DAGE	PR		DR		EXACT SOLUTION
		AGE-IMP	AGE-CN	AGE-IMP	AGE-CN	
0.05	1.0000	1.000000	1.000000	1.000000	1.000000	1.0000
0.10	1.0000	0.999999	1.000000	0.999999	1.000000	1.0000
0.15	1.0000	0.999999	1.000000	0.999999	1.000000	1.0000
0.20	1.0000	0.999999	1.000000	0.999999	0.999999	1.0000
0.25	1.0000	0.999999	1.000000	0.999999	0.999999	1.0000
0.30	1.0000	0.999995	0.999999	0.999995	0.999999	1.0000
0.35	1.0000	0.999755	0.999999	0.999746	0.999999	1.0000
0.40	1.0000	0.992646	0.999999	0.992410	0.999999	1.0000
0.45	1.0000	0.894929	0.999999	0.892547	0.999999	1.0000
0.50	1.0000	0.620463	1.000001	0.618006	1.000001	0.9999
0.55	1.0000	0.492275	1.000020	0.491687	1.000028	0.9999
0.60	0.9552	0.360375	0.955208	0.359512	0.953063	0.9413
0.65	0.3362	0.173939	0.336164	0.173327	0.334005	0.3410
0.70	0.1145	0.109650	0.114505	0.109531	0.114373	0.1138
0.75	0.1006	0.100808	0.100642	0.100795	0.100636	0.1005
0.80	0.1000	0.100049	0.100027	0.100048	0.100026	0.1000
0.85	0.1000	0.100002	0.100001	0.100002	0.100001	0.1000
0.90	0.1000	0.100000	0.100000	0.100000	0.100000	0.1000
0.95	0.1000	0.100000	0.100000	0.100000	0.100000	0.1000
Number of iter-ations	-	2	2	8	10	-

TABLE 6.9.22: The numerical solution to Burgers' equation for Problem 1 (6.9.19)

Case (i) $F(x)=4x(1-x)$ with $\varepsilon=0.01$, $\Delta t=0.01$, $\Delta x=0.25$, $r=0.5$, $eps=10^{-6}$

x=0.25	Exact	Implicit	Explicit	DAGE	PR	
t					AGE-IMP	AGE-CN
0.01	0.7492	0.7346	0.7342	0.7342	0.7417	0.7344
0.05	0.7460	0.6766	0.6748	0.6755	0.7098	0.6757
0.10	0.7420	0.6122	0.6087	0.6104	0.6724	0.6104
0.15	0.7380	0.5562	0.5512	0.5537	0.6376	0.5537
0.20	0.7340	NA	0.5016	0.5046	0.6054	0.5047
0.25	0.7300	NA	NA	0.4652	0.5755	0.4625

x=0.5	Exact	Implicit	Explicit	DAGE	PR	
t					AGE-IMP	AGE-CN
0.01	0.9992	0.9986	0.9992	0.9989	0.9991	0.9989
0.05	0.9960	0.9873	0.9901	0.9887	0.9938	0.9887
0.10	0.9920	0.9611	0.9662	0.9636	0.9841	0.9636
0.15	0.9880	0.9233	0.9299	0.9263	0.9712	0.9265
0.20	0.9840	NA	0.8835	0.8796	0.9554	0.8797
0.25	0.9800	NA	NA	0.8256	0.9370	0.8254

x=0.75	Exact	Implicit	Explicit	DAGE	PR	
t					AGE-IMP	AGE-CN
0.01	0.7492	0.7644	0.7642	0.7645	0.7567	0.7643
0.05	0.7460	0.8241	0.8232	0.8240	0.7837	0.8237
0.10	0.7420	0.9027	0.9012	0.9020	0.8181	0.9019
0.15	0.7380	0.9843	0.9828	0.9832	0.8531	0.9836
0.20	0.7340	NA	1.0065	1.0671	0.8886	1.0670
0.25	0.7300	NA	NA	1.1513	0.9242	1.1507

Method t	Number of Iterations	
	AGE-IMP	AGE-CN
0.01	2	3
0.05	2	3
0.10	2	3
0.15	2	3
0.20	2	3
0.25	2	3

TABLE 6.9.23: The numerical solution to Burgers' equation for Problem 2 (6.9.20)

Case (ii) $F(x)=\sin(\pi x)$ with $\varepsilon=1.0$, $\Delta t=0.01$, $\Delta x=0.25$, $r=0.5$, $eps=10^{-6}$

x=0.25 t	Exact	Implicit	Explicit	DAGE	PR	
					AGE-IMP	AGE-CN
0.01	0.6290	0.6377	0.6267	0.6259	0.6416	0.6327
0.05	0.4131	0.4339	0.4099	0.4168	0.4419	0.4224
0.10	0.2536	0.2768	0.2525	0.2639	0.2821	0.2648
0.15	0.1566	0.1784	0.1565	0.1681	0.1811	0.1676
0.20	0.0964	NA	0.0967	0.1047	0.1163	0.1059
0.25	0.0592	NA	NA	0.0651	0.0746	0.0668

x=0.5 t	Exact	Implicit	Explicit	DAGE	PR	
					AGE-IMP	AGE-CN
0.01	0.9057	0.9141	0.9063	0.9082	0.9142	0.9103
0.05	0.6096	0.6380	0.6100	0.6222	0.6386	0.6246
0.10	0.3716	0.4075	0.3729	0.3876	0.4080	0.3906
0.15	0.2268	0.2604	0.2281	0.2417	0.2607	0.2445
0.20	0.1385	NA	0.1395	0.1509	0.1666	0.1530
0.25	0.0 845	NA	NA	0.0942	0.1064	0.0957

x=0.75 t	Exact	Implicit	Explicit	DAGE	PR	
					AGE-IMP	AGE-CN
0.01	0.6524	0.6556	0.6550	0.6612	0.6514	0.6550
0.05	0.4502	0.4702	0.4556	0.4668	0.4617	0.4631
0.10	0.2726	0.3007	0.2762	0.2871	0.2952	0.2888
0.15	0.1644	0.1904	0.1663	0.1753	0.1877	0.1786
0.20	0.0994	NA	0.1006	0.1094	0.1193	0.1105
0.25	0.0603	NA	NA	0.0684	0.0759	0.0686

Method t	Number of Iterations	
	AGE-IMP	AGE-CN
0.01	4	4
0.05	4	4
0.10	4	4
0.15	4	3
0.20	4	3
0.25	4	3

TABLE 6.9.24: The numerical solutions to Burgers' equation for Problem 2 (6.9.20)

$t=3,\ \Delta x=1.0,\ \Delta t=0.2,\ \lambda=0.2,\ eps=10^{-4}$

Method \ x		-49	-33	-17	-1	15	207	339	Average of all absolute errors	Number of iterations
(S)AGE-EXP		7.11×10^{-14}	8.17×10^{-13}	7.96×10^{-8}	3.31×10^{-3}	2.45×10^{-6}	3.37×10^{-70}	3.01×10^{-123}	1.06×10^{-4}	-
(D)AGE-EXP		4.26×10^{-14}	6.47×10^{-13}	6.58×10^{-8}	1.82×10^{-3}	5.4×10^{-6}	7.34×10^{-70}	2.58×10^{-123}	1.12×10^{-4}	-
(S)AGE-PC		8.53×10^{-14}	1.67×10^{-12}	1.23×10^{-7}	1.64×10^{-3}	2.47×10^{-6}	3.37×10^{-70}	3.01×10^{-123}	3.65×10^{-5}	-
(D)AGE-PC		5.68×10^{-14}	1.88×10^{-12}	1.41×10^{-7}	2.06×10^{-4}	5.43×10^{-6}	7.34×10^{-70}	2.58×10^{-123}	3.62×10^{-5}	-
AGE-PC	PR	1.42×10^{-13}	8.81×10^{-13}	7.91×10^{-8}	9.2×10^{-4}	7.25×10^{-6}	9.84×10^{-70}	2.47×10^{-123}	3.52×10^{-5}	2
	DR	0	9.09×10^{-13}	8.75×10^{-8}	3.95×10^{-4}	6.58×10^{-6}	8.95×10^{-70}	2.53×10^{-123}	4.21×10^{-5}	8
AGE-1TL	PR	0	2.3×10^{-12}	1.84×10^{-7}	7.46×10^{-4}	1.76×10^{-5}	2.4×10^{-69}	1.5×10^{-123}	5.92×10^{-5}	3
	DR	5.68×10^{-14}	2.57×10^{-12}	1.94×10^{-7}	1.26×10^{-3}	1.66×10^{-5}	2.27×10^{-69}	1.6×10^{-123}	4.85×10^{-5}	8
AGE-2TL	PR	5.68×10^{-14}	1.11×10^{-12}	8.66×10^{-8}	1.08×10^{-3}	7.25×10^{-6}	9.84×10^{-70}	2.47×10^{-123}	3.64×10^{-5}	2
	DR	1.14×10^{-13}	1.53×10^{-12}	9.51×10^{-8}	5.57×10^{-4}	6.58×10^{-6}	8.95×10^{-70}	2.53×10^{-123}	4.17×10^{-5}	8
AGE-3TL	PR	2.42×10^{-13}	1.68×10^{-12}	9.6×10^{-8}	1.64×10^{-3}	8.95×10^{-6}	1.2×10^{-69}	2.32×10^{-123}	7.82×10^{-5}	5
	DR	3.27×10^{-13}	1.68×10^{-12}	9.98×10^{-8}	2.02×10^{-3}	8.27×10^{-6}	1.11×10^{-69}	2.38×10^{-123}	6.95×10^{-5}	11
AGE-4TL	PR	1.56×10^{-13}	1.85×10^{-13}	3.73×10^{-9}	1.9×10^{-4}	2.67×10^{-8}	4.18×10^{-72}	3.14×10^{-123}	3.17×10^{-6}	2
	DR	1.43×10^{-13}	5.26×10^{-13}	4.69×10^{-9}	3.43×10^{-4}	6.44×10^{-7}	8.65×10^{-71}	3.2×10^{-123}	1.63×10^{-5}	8
EXACT SOLUTION		1.0000	1.0000	0.9999987	0.9008854	1.11×10^{-4}	1.48×10^{-68}	1.32×10^{-122}	-	-

TABLE 6.9.26: The absolute errors of the numerical solutions to non-linear problem (6.9.21) (The reaction-diffusion equation)

t=12, $\Delta x=1.0$, $\Delta t=0.2$, $\lambda=0.2$, eps=10^{-4}

x Method		-49	-33	-17	-1	15	207	339	Average of all absolute errors	Number of iterations
(S)AGE-EXP		5.68×10^{-14}	1.28×10^{-13}	2.43×10^{-9}	6.77×10^{-7}	5.99×10^{-4}	1.8×10^{-67}	1.68×10^{-121}	4.28×10^{-4}	-
(D)AGE-EXP		8.53×10^{-14}	1.07×10^{-13}	3.03×10^{-9}	2.63×10^{-5}	6.43×10^{-4}	1.94×10^{-67}	1.65×10^{-121}	4.29×10^{-4}	-
(S)AGE-PC		7.11×10^{-14}	2.63×10^{-13}	5.24×10^{-9}	1.11×10^{-4}	9.91×10^{-5}	1.8×10^{-67}	1.68×10^{-121}	1.0×10^{-4}	-
(D)AGE-PC		8.53×10^{-14}	2.63×10^{-13}	4.49×10^{-9}	8.67×10^{-5}	4.74×10^{-5}	1.94×10^{-67}	1.65×10^{-121}	1.01×10^{-4}	-
AGE-PC	PR	1.42×10^{-14}	7.11×10^{-14}	2.78×10^{-9}	6.58×10^{-5}	5.48×10^{-4}	3.92×10^{-67}	2.43×10^{-122}	1.03×10^{-4}	2
	DR	0	2.84×10^{-14}	3.21×10^{-9}	9.6×10^{-5}	3.24×10^{-4}	3.53×10^{-67}	5.07×10^{-122}	1.7×10^{-4}	8
AGE-1TL	PR	0	2.56×10^{-13}	8.04×10^{-9}	1.22×10^{-4}	2.87×10^{-3}	1.1×10^{-66}	4.55×10^{-121}	2.92×10^{-4}	3
	DR	5.68×10^{-14}	3.55×10^{-13}	8.65×10^{-9}	1.54×10^{-4}	2.57×10^{-3}	1.03×10^{-66}	4.06×10^{-121}	2.22×10^{-4}	8
AGE-2TL	PR	5.68×10^{-14}	1.56×10^{-13}	3.05×10^{-9}	6.69×10^{-5}	5.53×10^{-4}	3.92×10^{-67}	2.43×10^{-122}	9.6×10^{-5}	2
	DR	1.14×10^{-13}	8.38×10^{-13}	3.48×10^{-9}	9.72×10^{-5}	3.29×10^{-4}	3.53×10^{-67}	5.07×10^{-122}	1.63×10^{-4}	8
AGE-3TL	PR	2.42×10^{-13}	5.83×10^{-13}	4.27×10^{-9}	5.68×10^{-5}	2.33×10^{-3}	4.86×10^{-67}	3.97×10^{-122}	3.67×10^{-4}	5
	DR	3.27×10^{-13}	9.66×10^{-13}	4.49×10^{-9}	7.81×10^{-5}	2.08×10^{-3}	4.47×10^{-67}	1.29×10^{-122}	3.05×10^{-4}	11
AGE-4TL	PR	1.14×10^{-13}	4.83×10^{-13}	1.49×10^{-10}	5.33×10^{-6}	1.22×10^{-5}	1.51×10^{-69}	2.91×10^{-121}	3.61×10^{-6}	2
	DR	1.42×10^{-13}	5.26×10^{-13}	4.69×10^{-9}	3.43×10^{-4}	6.44×10^{-7}	8.65×10^{-71}	3.2×10^{-123}	1.63×10^{-5}	8
EXACT SOLUTION		1.0000	1.0000	0.9999999	0.9987793	0.0098869	1.33×10^{-66}	1.19×10^{-120}	-	-

TABLE 6.9.27: The absolute errors of the numerical solutions to non-linear problem (6.9.21)
(The reaction-diffusion equation)

$t=21,\ \Delta x=1.0,\ \Delta t=0.2,\ \lambda=0.2,\ eps=10^{-4}$

x Method		-49	-33	-17	-1	15	207	339	Average of all absolute errors	Number of iteration
(S)AGE-EXP		7.11×10^{-14}	1.85×10^{-13}	2.66×10^{-11}	3.17×10^{-7}	6.97×10^{-2}	2.66×10^{-65}	9.04×10^{-120}	7.88×10^{-4}	-
(D)AGE-EXP		8.53×10^{-14}	1.85×10^{-13}	3.32×10^{-11}	7.13×10^{-9}	7.01×10^{-2}	3.21×10^{-65}	5.65×10^{-120}	7.87×10^{-4}	-
(S)AGE-PC		8.53×10^{-14}	1.85×10^{-13}	7.96×10^{-11}	2.73×10^{-6}	1.68×10^{-2}	2.66×10^{-65}	9.04×10^{-120}	1.88×10^{-4}	-
(D)AGE-PC		9.95×10^{-14}	1.85×10^{-13}	7.05×10^{-11}	2.43×10^{-6}	1.71×10^{-2}	3.21×10^{-65}	5.65×10^{-120}	1.87×10^{-4}	-
AGE-PC	PR	1.42×10^{-14}	2.84×10^{-14}	4.5×10^{-11}	1.79×10^{-6}	1.56×10^{-2}	6.84×10^{-65}	2.02×10^{-119}	1.64×10^{-4}	2
	DR	0	2.84×10^{-14}	5.39×10^{-11}	2.42×10^{-6}	2.56×10^{-2}	6.1×10^{-65}	1.52×10^{-119}	2.85×10^{-4}	8
AGE-1TL	PR	0	1.14×10^{-13}	1.26×10^{-10}	2.57×10^{-6}	4.41×10^{-2}	2.24×10^{-64}	1.25×10^{-118}	5.88×10^{-4}	3
	DR	5.68×10^{-14}	1.56×10^{-13}	1.4×10^{-10}	3.27×10^{-6}	3.39×10^{-2}	2.06×10^{-64}	1.14×10^{-118}	4.58×10^{-4}	8
AGE-2TL	PR	5.68×10^{-14}	2.7×10^{-13}	4.93×10^{-11}	1.87×10^{-6}	1.48×10^{-2}	6.84×10^{-65}	2.02×10^{-119}	1.54×10^{-4}	2
	DR	1.14×10^{-13}	3.98×10^{-13}	5.85×10^{-11}	2.51×10^{-6}	2.49×10^{-2}	6.1×10^{-65}	1.52×10^{-119}	2.76×10^{-4}	8
AGE-3TL	PR	2.42×10^{-13}	3.55×10^{-13}	6.37×10^{-11}	7.57×10^{-7}	5.64×10^{-2}	8.68×10^{-65}	3.27×10^{-119}	7.19×10^{-4}	5
	DR	3.27×10^{-13}	7.39×10^{-13}	6.91×10^{-11}	1.18×10^{-6}	4.76×10^{-2}	7.9×10^{-65}	2.74×10^{-119}	6.07×10^{-4}	11
AGE-4TL	PR	2.13×10^{-13}	4.97×10^{-13}	2.03×10^{-12}	9.6×10^{-8}	7.99×10^{-5}	2.37×10^{-67}	2.62×10^{-119}	3.54×10^{-6}	2
	DR	1.42×10^{-13}	9.95×10^{-13}	5.22×10^{-12}	4.75×10^{-7}	1.01×10^{-2}	4.82×10^{-66}	2.93×10^{-119}	1.25×10^{-4}	8
EXACT SOLUTION		1.0000	1.0000	0.9999999	0.9999864	0.4733711	1.198×10^{-64}	1.07×10^{-118}	-	-

TABLE 6.9.28: The absolute errors of the numerical solutions to non-linear problem (6.9.21)
(The reaction-diffusion equation)

$t=30$, $\Delta x=1.0$, $\Delta t=0.2$, $\lambda=0.2$, $eps=10^{-4}$

x Method		-49	-33	-17	-1	15	207	339	Average of all absolute errors	Number of iterations
(S)AGE-EXP		5.68×10^{-14}	1.85×10^{-13}	5.68×10^{-14}	1.24×10^{-8}	3.84×10^{-3}	3.84×10^{-63}	2.57×10^{-118}	1.15×10^{-3}	-
(D)AGE-EXP		7.11×10^{-14}	1.85×10^{-13}	1.28×10^{-13}	7.85×10^{-9}	3.8×10^{-3}	4.7×10^{-63}	8.2×10^{-118}	1.14×10^{-3}	-
(S)AGE-PC		7.11×10^{-14}	1.85×10^{-13}	1.42×10^{-12}	5.48×10^{-8}	1.3×10^{-3}	3.84×10^{-63}	2.57×10^{-118}	2.75×10^{-4}	-
(D)AGE-PC		7.11×10^{-14}	1.85×10^{-13}	1.3×10^{-12}	5.03×10^{-8}	1.25×10^{-3}	4.7×10^{-63}	8.2×10^{-118}	2.71×10^{-4}	-
AGE-PC	PR	1.42×10^{-14}	2.84×10^{-14}	5.68×10^{-13}	3.2×10^{-8}	9.05×10^{-4}	9.78×10^{-63}	4.27×10^{-117}	2.23×10^{-4}	2
	DR	0	2.84×10^{-14}	7.11×10^{-13}	4.29×10^{-8}	1.66×10^{-3}	8.63×10^{-63}	3.5×10^{-117}	3.99×10^{-4}	8
AGE-1TL	PR	0	1.14×10^{-13}	1.79×10^{-12}	4.2×10^{-8}	1.82×10^{-3}	3.78×10^{-62}	2.33×10^{-116}	8.94×10^{-4}	3
	DR	5.68×10^{-14}	1.42×10^{-13}	2.3×10^{-12}	5.42×10^{-8}	1.21×10^{-3}	3.43×10^{-62}	2.09×10^{-116}	7.08×10^{-4}	8
AGE-2TL	PR	5.68×10^{-14}	2.7×10^{-13}	9.24×10^{-13}	3.38×10^{-8}	8.54×10^{-4}	9.78×10^{-63}	4.27×10^{-117}	2.1×10^{-4}	2
	DR	1.14×10^{-13}	3.98×10^{-13}	1.28×10^{-12}	4.48×10^{-8}	1.61×10^{-3}	8.63×10^{-63}	3.5×10^{-117}	3.86×10^{-4}	8
AGE-3TL	PR	2.42×10^{-13}	3.84×10^{-13}	1.12×10^{-12}	9.21×10^{-9}	2.45×10^{-3}	1.27×10^{-62}	6.26×10^{-117}	1.07×10^{-3}	5
	DR	3.27×10^{-13}	7.39×10^{-13}	1.52×10^{-12}	1.6×10^{-8}	1.98×10^{-3}	1.15×10^{-62}	5.41×10^{-117}	9.14×10^{-4}	11
AGE-4TL	PR	1.42×10^{-13}	4.97×10^{-13}	2.98×10^{-13}	1.48×10^{-9}	3.54×10^{-5}	3.05×10^{-65}	2.37×10^{-117}	3.41×10^{-6}	2
	DR	2.56×10^{-13}	9.95×10^{-13}	6.54×10^{-13}	7.63×10^{-9}	6.8×10^{-4}	6.15×10^{-64}	2.76×10^{-117}	1.79×10^{-4}	8
EXACT SOLUTION		1.0000	1.0000	1.0000	0.9999985	0.9877919	1.08×10^{-62}	9.65×10^{-117}	-	-

TABLE 6.9.29: The absolute errors of the numerical solutions to non-linear problem (6.9.21) (The reaction-diffusion equation)

t	Method	(S)AGE-EXP	(D)AGE-EXP	(S)AGE-PC	(D)AGE-PC	AGE-PC		AGE-1TL		AGE-2TL		AGE-3TL		AGE-4TL	
						PR	DR	PR	DR	PR	DR	PR	DR	PR	DR
3	CV	0.6908973	0.6887024	0.7038429	0.7036834	0.7032153	0.7032711	0.7144151	0.714478	0.7030501	0.7031086	0.7178107	0.7175869	0.7066162	0.7066676
	CV-R	(0.6916288)		(0.7031970)		(0.7031970)		(0.714411)		(0.7030459		(0.7178736)		(0.7066156)	
	PCV-R(%)	(2.19)		(0.553)		(0.533)		(1.033)		(0.574)		(1.523)		(0.069)	
	PCV(%)	2.292	2.603	0.461	0.484	0.55	0.542	1.034	1.043	0.574	0.565	1.514	1.482	0.0692	0.062
12	CV	0.6889379	0.6890496	0.7026014	0.7027083	0.7035437	0.703606	0.7203569	0.7203964	0.7032528	0.7033159	0.7234951	0.7232126	0.7065187	0.7065909
	CV-R	(0.6870319)		(0.7035245)		(0.7035245)		(0.7203591)		(0.7032573)		(0.7235846)		(0.7065183)	
	PCV-R(%)	(2.84)		(0.506)		(0.506)		(1.874)		(0.544)		(2.330)		(0.083)	
	PCV(%)	2.569	2.554	0.637	0.622	0.504	0.495	1.874	1.88	0.545	0.536	2.318	2.278	0.083	0.073
21	CV	0.6892023	0.6892408	0.7027116	0.7027591	0.7038907	0.7039569	0.7211176	0.7211574	0.7035924	0.7036592	0.7239050	0.7236222	0.7065232	0.706597
	CV-R	(0.6869790)		(0.7038702)		(0.7038702)		(0.7211216)		(0.7035991)		(0.7239955)		(0.7065230)	
	PCV-R(%)	(2.846)		(0.458)		(0.458)		(1.982)		(0.496)		(2.389)		(0.0824)	
	PCV(%)	2.532	2.526	0.621	0.615	0.455	0.445	1.982	1.987	0.497	0.487	2.376	2.336	0.0824	0.072
30	CV	0.6892778	0.689422	0.7027429	0.7028983	0.7039567	0.7040239	0.7212402	0.7212807	0.7036572	0.7037251	0.7239644	0.7236825	0.7065240	0.706598
	CV-R	(0.6869822)		(0.7039361)		(0.7039361)		(0.7212447)		(0.7036646)		(0.7240550)		(0.7065237)	
	PCV-R(%)	(2.846)		(0.448)		(0.448)		(2.00)		(0.487)		(2.397)		(0.0823)	
	PCV(%)	2.521	2.501	0.617	0.595	0.445	0.436	1.999	2.005	0.488	0.478	2.384	2.344	0.0823	0.0718

continued...

CV: computed wave speed
CV-R: computed wave speed from Ramos
PCV: percentage error in CV
PCV-R: percentage error in CV-R
Exact $V=1/\sqrt{2}=0.7071058$

TABLE 6.9.30: The computed wave speeds and their percentage errors

$\Delta x=0.5$, $\Delta t=0.05$, $\lambda=0.2$, $eps=10^{-4}$, $t=3$

x Method		-49	-33	-17	-1	15	207	399	Average of all absolute errors	Number of Iterations
AGE-PC	PR	2.84×10^{-14}	2.27×10^{-13}	2.04×10^{-8}	2.27×10^{-4}	1.74×10^{-6}	2.34×10^{-70}	2.97×10^{-123}	8.79×10^{-6}	2
	DR	2.84×10^{-14}	2.13×10^{-13}	5.3×10^{-8}	1.89×10^{-3}	8.79×10^{-7}	1.15×10^{-70}	3.21×10^{-123}	7.69×10^{-5}	6
AGE-1TL	PR	2.27×10^{-13}	1.07×10^{-12}	4.7×10^{-8}	2.6×10^{-4}	3.87×10^{-6}	5.2×10^{-70}	2.78×10^{-123}	1.46×10^{-5}	2
	DR	3.84×10^{-13}	2.06×10^{-12}	8.0×10^{-8}	2.38×10^{-3}	1.16×10^{-6}	1.59×10^{-70}	3.02×10^{-123}	6.57×10^{-5}	6
AGE-2TL	PR	4.0×10^{-13}	4.97×10^{-13}	2.12×10^{-8}	2.35×10^{-4}	1.74×10^{-6}	2.34×10^{-70}	2.97×10^{-123}	8.76×10^{-6}	2
	DR	4.41×10^{-13}	1.49×10^{-12}	5.38×10^{-8}	1.88×10^{-3}	8.79×10^{-7}	1.15×10^{-70}	3.21×10^{-123}	7.67×10^{-5}	6
AGE-3TL	PR	4.26×10^{-13}	1.99×10^{-12}	2.7×10^{-8}	6.67×10^{-4}	1.83×10^{-6}	2.45×10^{-70}	2.97×10^{-123}	1.72×10^{-5}	4
	DR	3.27×10^{-13}	2.57×10^{-12}	5.19×10^{-8}	2.44×10^{-3}	4.97×10^{-7}	6.59×10^{-71}	3.18×10^{-123}	6.03×10^{-5}	8
AGE-4TL	PR	1.42×10^{-13}	8.1×10^{-13}	2.45×10^{-10}	1.25×10^{-5}	1.8×10^{-9}	2.83×10^{-73}	3.13×10^{-123}	2.05×10^{-7}	2
	DR	8.53×10^{-14}	2.07×10^{-12}	3.2×10^{-8}	2.1×10^{-3}	2.54×10^{-6}	3.39×10^{-70}	3.36×10^{-123}	7.34×10^{-5}	6
EXACT SOLUTION		1.0000	1.0000	0.9999987	0.9008854	1.11×10^{-4}	1.48×10^{-68}	1.32×10^{-122}	-	-

continued.....

$\Delta x=0.5$, $\Delta t=0.05$, $\lambda=0.2$, $eps=10^{-4}$, $t=12$

Method \ x		-49	-33	-17	-1	15	207	399	Average of all absolute errors	Number of Iterations
AGE-PC	PR	2.84×10^{-14}	1.71×10^{-13}	7.12×10^{-10}	1.73×10^{-5}	1.39×10^{-4}	8.63×10^{-68}	2.31×10^{-121}	2.55×10^{-5}	2
	DR	2.84×10^{-14}	2.84×10^{-14}	2.27×10^{-9}	1.39×10^{-4}	7.36×10^{-4}	4.1×10^{-68}	3.17×10^{-121}	3.16×10^{-4}	6
AGE-1TL	PR	2.7×10^{-13}	1.04×10^{-12}	1.83×10^{-9}	3.26×10^{-5}	6.79×10^{-4}	1.97×10^{-67}	1.56×10^{-121}	7.0×10^{-5}	2
	DR	3.84×10^{-13}	2.09×10^{-12}	3.49×10^{-9}	1.55×10^{-4}	2.47×10^{-4}	5.76×10^{-68}	2.51×10^{-121}	2.27×10^{-4}	6
AGE-2TL	PR	3.98×10^{-13}	1.85×10^{-13}	7.4×10^{-10}	1.75×10^{-5}	1.41×10^{-4}	8.63×10^{-68}	2.31×10^{-121}	2.46×10^{-5}	2
	DR	4.41×10^{-13}	2.4×10^{-12}	2.3×10^{-9}	1.39×10^{-4}	7.35×10^{-4}	4.1×10^{-68}	3.17×10^{-121}	3.15×10^{-4}	6
AGE-3TL	PR	4.26×10^{-13}	3.17×10^{-12}	1.12×10^{-9}	2.5×10^{-5}	4.47×10^{-4}	9.01×10^{-68}	2.29×10^{-121}	6.14×10^{-5}	4
	DR	3.27×10^{-13}	3.95×10^{-12}	2.35×10^{-9}	1.28×10^{-4}	3.7×10^{-4}	2.45×10^{-68}	3.06×10^{-121}	2.01×10^{-4}	8
AGE-4TL	PR	9.52×10^{-13}	1.17×10^{-12}	8.63×10^{-12}	3.62×10^{-7}	8.32×10^{-7}	1.02×10^{-70}	2.9×10^{-121}	2.35×10^{-7}	2
	DR	8.53×10^{-14}	3.5×10^{-12}	1.48×10^{-9}	1.2×10^{-4}	8.69×10^{-4}	1.18×10^{-67}	3.69×10^{-121}	2.93×10^{-4}	6
EXACT SOLUTION		1.0000	1.0000	0.9999999	0.9987793	0.0098869	1.33×10^{-66}	1.19×10^{-120}	-	-

TABLE 6.9.31: The absolute errors of the numerical solutions to non-linear problem (6.9.21)
(The reaction-diffusion equation)

$\Delta x=0.5$, $\Delta t=0.05$, $\lambda=0.2$, $eps=10^{-4}$, $t=21$

x Method		-49	-33	-17	-1	15	207	399	Average of all absolute errors	Number of Iterations
AGE-PC	PR	2.84×10^{-14}	0	1.07×10^{-11}	4.57×10^{-7}	3.84×10^{-3}	1.39×10^{-65}	1.66×10^{-119}	4.1×10^{-5}	2
	DR	2.84×10^{-14}	1.28×10^{-13}	4.24×10^{-11}	2.96×10^{-6}	4.43×10^{-2}	6.38×10^{-66}	3.04×10^{-119}	5.52×10^{-4}	6
AGE-1TL	PR	2.7×10^{-13}	1.21×10^{-12}	2.73×10^{-11}	6.61×10^{-7}	1.0×10^{-2}	3.27×10^{-65}	3.93×10^{-120}	1.41×10^{-4}	2
	DR	3.84×10^{-13}	2.26×10^{-12}	6.26×10^{-11}	3.20×10^{-6}	3.0×10^{-2}	9.21×10^{-66}	1.98×10^{-119}	3.76×10^{-4}	6
AGE-2TL	PR	3.98×10^{-13}	3.55×10^{-13}	1.2×10^{-11}	4.66×10^{-7}	3.71×10^{-3}	1.39×10^{-65}	1.66×10^{-119}	3.96×10^{-5}	2
	DR	4.41×10^{-13}	2.57×10^{-12}	4.44×10^{-11}	2.97×10^{-6}	4.41×10^{-2}	6.38×10^{-66}	3.04×10^{-119}	5.51×10^{-4}	6
AGE-3TL	PR	4.26×10^{-13}	3.34×10^{-12}	1.82×10^{-11}	3.88×10^{-7}	9.98×10^{-3}	1.46×10^{-65}	1.62×10^{-119}	1.25×10^{-4}	4
	DR	3.27×10^{-13}	4.12×10^{-12}	4.4×10^{-11}	2.47×10^{-6}	2.6×10^{-2}	3.85×10^{-66}	2.87×10^{-119}	3.33×10^{-4}	8
AGE-4TL	PR	9.52×10^{-13}	1.34×10^{-12}	8.95×10^{-13}	6.54×10^{-9}	1.6×10^{-6}	1.6×10^{-68}	2.6×10^{-119}	2.42×10^{-7}	2
	DR	8.53×10^{-14}	3.67×10^{-12}	3.14×10^{-11}	2.42×10^{-6}	4.04×10^{-2}	1.79×10^{-65}	3.82×10^{-119}	5.13×10^{-4}	6
EXACT SOLUTION		1.0000	1.0000	0.9999999	0.9999864	0.4733711	1.198×10^{-64}	1.07×10^{-118}	-	-

continued......

$\Delta x=0.5$, $\Delta t=0.05$, $\lambda=0.2$, $eps=10^{-4}$, $t=30$

x Method		-49	-33	-17	-1	15	207	399	Average of all absolute errors	Number of Iterations
AGE-PC	PR	2.84×10^{-14}	0	2.7×10^{-13}	7.99×10^{-9}	2.26×10^{-4}	1.83×10^{-63}	1.1×10^{-117}	5.59×10^{-5}	2
	DR	2.84×10^{-14}	1.28×10^{-13}	3.69×10^{-13}	4.99×10^{-8}	3.43×10^{-3}	8.11×10^{-64}	2.9×10^{-117}	7.87×10^{-4}	6
AGE-1TL	PR	2.7×10^{-13}	1.21×10^{-12}	9.09×10^{-13}	1.05×10^{-8}	4.59×10^{-4}	4.44×10^{-63}	6.59×10^{-118}	2.15×10^{-4}	2
	DR	3.84×10^{-13}	2.26×10^{-12}	2.53×10^{-12}	5.31×10^{-8}	2.58×10^{-3}	1.20×10^{-63}	1.53×10^{-117}	5.23×10^{-4}	6
AGE-2TL	PR	3.98×10^{-13}	3.55×10^{-13}	7.82×10^{-13}	8.15×10^{-9}	2.2×10^{-4}	1.83×10^{-63}	1.1×10^{-117}	5.38×10^{-5}	2
	DR	4.41×10^{-13}	2.57×10^{-12}	2.47×10^{-12}	5.01×10^{-8}	3.43×10^{-3}	8.11×10^{-64}	2.9×10^{-117}	7.85×10^{-4}	6
AGE-3TL	PR	4.26×10^{-13}	3.34×10^{-12}	2.81×10^{-12}	5.4×10^{-9}	3.97×10^{-4}	1.92×10^{-63}	1.05×10^{-117}	1.9×10^{-4}	4
	DR	3.27×10^{-13}	4.12×10^{-12}	3.5×10^{-12}	4.0×10^{-8}	2.25×10^{-3}	4.92×10^{-64}	2.68×10^{-117}	4.65×10^{-4}	8
AGE-4TL	PR	9.52×10^{-13}	1.34×10^{-12}	1.41×10^{-12}	9.99×10^{-11}	2.69×10^{-6}	2.06×10^{-66}	2.35×10^{-117}	2.51×10^{-7}	2
	DR	8.53×10^{-4}	3.67×10^{-12}	2.7×10^{-12}	3.98×10^{-8}	3.16×10^{-3}	2.23×10^{-63}	3.86×10^{-117}	7.33×10^{-4}	6
EXACT SOLUTION		1.0000	1.0000	1.0000	0.9999985	0.9877919	1.08×10^{-62}	9.65×10^{-117}	-	-

TABLE 6.9.32: The absolute errors of the numerical solutions to non-linear problem (6.9.21)
(The reaction-diffusion equation)

$\Delta x=0.5$, $\Delta t=0.05$, $\lambda=0.2$, $eps=10^{-4}$

t	Method	AGE-PC		AGE-1TL		AGE-2TL		AGE-3TL		AGE-4TL	
		PR	DR	PR	DR	PR	DR	PR	DR	PR	DR
3	CV	0.7061664	0.7062216	0.7090783	0.7091131	0.7061763	0.7062321	0.7098810	0.7096623	0.7070767	0.7071058
	CV-R	(0.7061863)		(0.7090694)		(0.7061753)		(0.7099601)		(0.7070766)	
	PCV-R(%)	(0.13)		(2.78)		(0.132)		(0.404)		(0.00413)	
	PCV(%)	0.133	0.125	0.279	0.284	0.131	0.124	0.392	0.362	0.00412	0.00246
12	CV	0.7062323	0.7062913	0.7105140	0.7105653	0.7062357	0.706295	0.7112576	0.7109959	0.7070700	0.7071058
	CV-R	(0.7062578)		(0.7105382)		(0.7062386)		(0.7113624)		(0.7070711)	
	PCV-R(%)	(0.120)		(0.485)		(0.123)		(0.602)		(0.0049)	
	PCV(%)	0.124	0.115	0.482	0.489	0.123	0.115	0.587	0.55	0.00505	0.00458
21	CV	0.7063150	0.706377	0.7106940	0.7107515	0.7063178	0.7063801	0.7113578	0.7110972	0.7070703	0.7071058
	CV-R	(0.7063419)		(0.7107250)		(0.7063221)		(0.7114640)		(0.7070724)	
	PCV-R(%)	(0.108)		(0.512)		(0.111)		(0.616)		(0.00472)	
	PCV(%)	0.112	0.103	0.507	0.516	0.111	0.103	0.601	0.564	0.00502	0.00484
30	CV	0.7063307	0.7063937	0.7107242	0.7107833	0.7063334	0.7063966	0.7113730	0.7111128	0.7071058	0.7071058
	CV-R	(0.7063589)		(0.7107582)		(0.7063389)		(0.7114793)		(0.7070734)	
	PCV-R(%)	(0.106)		(0.517)		(0.108)		(0.619)		(0.00458)	
	PCV(%)	0.11	0.101	0.512	0.52	0.109	0.1	0.603	0.567	0.00501	0.00488

CV: computed wave speed

CV-R: computed wave speed from Ramos

PCV: percentage error in CV

PCV-R: percentage error in CV-R

Exact $V=1/\sqrt{2}=0.7071058$

TABLE 6.9.33: The computed wave speeds and their percentage errors

$t=1.0,\ \Delta x=0.05,\ \Delta t=0.025,\ \lambda=0.5,\ eps=10^{-6}$

x Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter-ations
AGE-CDBT	PR	1.21×10^{-6}	5.25×10^{-6}	3.66×10^{-5}	1.2×10^{-4}	2.19×10^{-4}	1.55×10^{-4}	2.15×10^{-3}	7.97×10^{-4}	3.04×10^{-2}	1.21×10^{-2}	5
	DR	1.19×10^{-6}	5.24×10^{-6}	3.65×10^{-5}	1.21×10^{-4}	2.18×10^{-4}	1.54×10^{-4}	2.15×10^{-3}	7.93×10^{-4}	3.04×10^{-2}	1.21×10^{-2}	16
AGE-CDCT	PR	9.9×10^{-4}	1.89×10^{-3}	1.88×10^{-3}	8.25×10^{-4}	6.55×10^{-3}	7.06×10^{-3}	1.22×10^{-2}	1.63×10^{-2}	1.3×10^{-1}	1.64×10^{-2}	4
	DR	9.9×10^{-4}	1.89×10^{-3}	1.88×10^{-3}	8.23×10^{-4}	6.55×10^{-3}	7.07×10^{-3}	1.22×10^{-2}	1.64×10^{-2}	1.3×10^{-1}	1.64×10^{-2}	15
EXACT SOLUTION		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	-	-

$t=1.0,\ \Delta x=0.05,\ \Delta t=0.05,\ \lambda=1.0,\ eps=10^{-6}$

x Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter-ations
AGE-CDBT	PR	4.62×10^{-7}	2.43×10^{-7}	5.49×10^{-6}	2.22×10^{-5}	1.55×10^{-4}	2.27×10^{-4}	3.44×10^{-3}	2.1×10^{-2}	6.77×10^{-2}	3.25×10^{-2}	8
	DR	1.84×10^{-7}	8.19×10^{-7}	5.82×10^{-6}	2.28×10^{-5}	1.53×10^{-4}	2.24×10^{-4}	3.45×10^{-3}	2.11×10^{-2}	6.77×10^{-2}	3.25×10^{-2}	20
AGE-CDCT	PR	7.42×10^{-4}	2.22×10^{-3}	4.22×10^{-3}	3.69×10^{-3}	4.28×10^{-3}	1.27×10^{-2}	8.78×10^{-3}	3.7×10^{-2}	2.2×10^{-1}	2.93×10^{-2}	6
	DR	7.42×10^{-4}	2.22×10^{-3}	4.22×10^{-3}	3.69×10^{-3}	4.28×10^{-3}	1.27×10^{-2}	8.77×10^{-3}	3.7×10^{-2}	2.2×10^{-1}	2.93×10^{-2}	18
EXACT SOLUTION		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	-	-

continued..........

$t=1.0$, $\Delta x=0.05$, $\Delta t=0.1$, $\lambda=2.0$, $eps=10^{-6}$

x Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter-ations
AGE-CDBT	PR	3.07×10^{-5}	8.7×10^{-5}	2.61×10^{-4}	5.14×10^{-4}	1.81×10^{-4}	4.91×10^{-3}	1.98×10^{-2}	5.51×10^{-2}	1.21×10^{-1}	6.22×10^{-2}	17
	DR	3.08×10^{-5}	8.91×10^{-5}	2.65×10^{-4}	5.18×10^{-4}	1.77×10^{-4}	4.91×10^{-3}	1.98×10^{-2}	5.51×10^{-2}	1.21×10^{-1}	6.22×10^{-2}	34
AGE-CDCT	PR	2.11×10^{-3}	5.16×10^{-3}	4.95×10^{-3}	6.32×10^{-3}	1.38×10^{-2}	1.48×10^{-2}	2.47×10^{-2}	1.27×10^{-1}	3.75×10^{-1}	6.03×10^{-2}	10
	DR	2.11×10^{-3}	5.16×10^{-3}	4.95×10^{-3}	6.32×10^{-3}	1.38×10^{-2}	1.48×10^{-2}	2.47×10^{-2}	1.27×10^{-1}	3.75×10^{-1}	6.03×10^{-2}	24
EXACT SOLUTION		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	-	-

TABLE 6.9.35: The absolute errors of the numerical solutions to equation (6.9.25) for Problem 1 (Non-linear first order hyperbolic (convection) equation)

t=1.0, $\Delta x=0.05$, $\Delta t=0.025$, $\lambda=0.5$, eps=10^{-6}

x Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter-ations
AGE-CDBT	PR	1.07×10^{-8}	1.69×10^{-6}	6.38×10^{-5}	8.59×10^{-4}	4.69×10^{-3}	1.04×10^{-2}	9.73×10^{-3}	7.61×10^{-3}	7.19×10^{-3}	4.41×10^{-3}	3
	DR	1.56×10^{-6}	4.86×10^{-6}	6.87×10^{-5}	8.66×10^{-4}	4.7×10^{-3}	1.04×10^{-2}	9.73×10^{-3}	7.62×10^{-3}	7.19×10^{-3}	4.41×10^{-3}	13
AGE-CDCT	PR	1.12×10^{-11}	2.0×10^{-10}	1.32×10^{-9}	3.63×10^{-9}	7.37×10^{-10}	4.72×10^{-9}	5.56×10^{-9}	2.65×10^{-9}	2.22×10^{-9}	3.3×10^{-9}	3
	DR	1.76×10^{-6}	3.53×10^{-6}	5.33×10^{-6}	7.11×10^{-6}	8.46×10^{-6}	8.17×10^{-6}	6.29×10^{-6}	4.72×10^{-6}	3.29×10^{-6}	9.38×10^{-6}	13
EXACT SOLUTION		0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	-	-

t=1.0, $\Delta x=0.05$, $\Delta t=0.05$, $\lambda=1.0$, eps=10^{-6}

x Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter-ations
AGE-CDBT	PR	1.65×10^{-7}	1.45×10^{-5}	3.17×10^{-4}	2.68×10^{-3}	1.02×10^{-2}	1.91×10^{-2}	1.94×10^{-2}	1.6×10^{-2}	1.44×10^{-2}	8.87×10^{-3}	4
	DR	5.82×10^{-7}	1.54×10^{-5}	3.19×10^{-4}	2.68×10^{-3}	1.02×10^{-2}	1.91×10^{-2}	1.94×10^{-2}	1.6×10^{-2}	1.44×10^{-2}	8.87×10^{-3}	15
AGE-CDCT	PR	3.09×10^{-11}	7.17×10^{-10}	3.73×10^{-9}	8.38×10^{-10}	7.65×10^{-9}	9.69×10^{-9}	9.78×10^{-9}	9.4×10^{-9}	3.2×10^{-9}	5.31×10^{-9}	4
	DR	7.46×10^{-7}	1.53×10^{-6}	2.37×10^{-6}	3.23×10^{-6}	3.78×10^{-6}	3.53×10^{-6}	2.82×10^{-6}	2.07×10^{-6}	1.37×10^{-6}	4.26×10^{-6}	14
EXACT SOLUTION		0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	-	-

continued.........

$t=1.0$, $\Delta x=0.05$, $\Delta t=0.1$, $\lambda=2.0$, $eps=10^{-6}$

x Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter-ations
AGE-CDBT	PR	4.78×10^{-6}	1.79×10^{-4}	1.85×10^{-3}	8.52×10^{-3}	2.17×10^{-2}	3.44×10^{-2}	3.78×10^{-2}	3.38×10^{-2}	2.95×10^{-2}	1.8×10^{-2}	6
	DR	4.81×10^{-6}	1.79×10^{-4}	1.85×10^{-3}	8.52×10^{-3}	2.17×10^{-2}	3.44×10^{-2}	3.78×10^{-2}	3.38×10^{-2}	2.95×10^{-2}	1.8×10^{-2}	18
AGE-CDCT	PR	4.68×10^{-11}	5.76×10^{-10}	1.07×10^{-9}	7.76×10^{-9}	9.13×10^{-9}	4.75×10^{-9}	2.11×10^{-8}	3.15×10^{-8}	1.19×10^{-8}	8.92×10^{-9}	5
	DR	1.92×10^{-7}	4.36×10^{-7}	7.78×10^{-7}	1.12×10^{-6}	1.11×10^{-6}	7.7×10^{-7}	6.43×10^{-7}	5.89×10^{-7}	2.98×10^{-7}	1.5×10^{-6}	16
EXACT SOLUTION		0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	-	-

TABLE 6.9.36: The absolute errors of the numerical solutions to equation (6.9.25) for Problem 2 (Non-linear first order hyperbolic (convection) equation)

$t=5$, $\Delta x=0.1$, $\Delta t=0.05$, $\lambda=0.5$, $eps=10^{-6}$

x Method	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter-ations
AGE-CDBT PR	5.93×10^{-4}	3.34×10^{-3}	6.11×10^{-3}	4.87×10^{-3}	1.31×10^{-3}	2.48×10^{-3}	2.74×10^{-3}	8.61×10^{-4}	1.33×10^{-3}	2.63×10^{-3}	3
AGE-CDBT DR	5.89×10^{-4}	3.35×10^{-3}	6.1×10^{-3}	4.89×10^{-3}	1.29×10^{-3}	2.49×10^{-3}	2.7×10^{-3}	8.68×10^{-4}	1.27×10^{-3}	2.62×10^{-3}	11
AGE-CDCT PR	1.11×10^{-10}	2.18×10^{-10}	4.55×10^{-11}	1.06×10^{-9}	9.55×10^{-10}	6.45×10^{-10}	1.64×10^{-10}	2.1×10^{-9}	3.16×10^{-9}	1.1×10^{-9}	2
AGE-CDCT DR	3.26×10^{-6}	5.95×10^{-6}	1.17×10^{-5}	8.13×10^{-6}	2.39×10^{-5}	8.26×10^{-6}	3.8×10^{-5}	5.61×10^{-6}	5.49×10^{-5}	1.77×10^{-5}	11
EXACT SOLUTION	0.0166699	0.0333393	0.05000	0.0666667	0.0833333	0.1000	0.1166667	0.1333333	0.1500	-	-

$t=10$, $\Delta x=0.1$, $\Delta t=0.1$, $\lambda=1.0$, $eps=10^{-6}$

x Method	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter-ations
AGE-CDBT PR	2.24×10^{-3}	7.02×10^{-3}	3.51×10^{-3}	3.81×10^{-4}	2.34×10^{-3}	1.49×10^{-3}	5.29×10^{-4}	2.09×10^{-3}	2.53×10^{-4}	2.21×10^{-3}	2
AGE-CDBT DR	2.23×10^{-3}	7.03×10^{-3}	3.5×10^{-3}	3.9×10^{-4}	2.32×10^{-3}	1.5×10^{-3}	4.92×10^{-4}	2.09×10^{-3}	2.01×10^{-4}	2.2×10^{-3}	10
AGE-CDCT PR	1.43×10^{-10}	1.48×10^{-10}	5.8×10^{-11}	1.59×10^{-9}	7.64×10^{-10}	2.03×10^{-10}	2.67×10^{-9}	2.71×10^{-9}	3.07×10^{-9}	1.26×10^{-9}	2
AGE-CDCT DR	3.15×10^{-6}	5.16×10^{-6}	1.17×10^{-5}	7.57×10^{-6}	2.24×10^{-5}	7.77×10^{-6}	3.55×10^{-5}	5.81×10^{-6}	5.1×10^{-5}	1.67×10^{-5}	10
EXACT SOLUTION	0.0090909	0.0181818	0.0272273	0.0363636	0.0454545	0.0545455	0.0636364	0.0727273	0.0818181	-	-

continued.........

$t=20,\ \Delta x=0.1,\ \Delta t=0.2,\ \lambda=2.0,\ eps=10^{-6}$

x Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Average of all absolute errors	No. of iter-ations
AGE-CDBT	PR	3.29×10^{-3}	5.18×10^{-3}	1.81×10^{-3}	2.76×10^{-3}	2.4×10^{-4}	1.43×10^{-4}	7.23×10^{-4}	2.26×10^{-4}	6.29×10^{-4}	1.67×10^{-3}	2
	DR	3.28×10^{-3}	5.19×10^{-3}	1.81×10^{-3}	2.77×10^{-3}	2.23×10^{-4}	1.49×10^{-4}	6.91×10^{-4}	2.22×10^{-4}	5.8×10^{-4}	1.66×10^{-3}	9
AGE-CDCT	PR	5.59×10^{-11}	4.11×10^{-10}	6.25×10^{-10}	3.53×10^{-9}	2.79×10^{-9}	3.19×10^{-9}	2.21×10^{-9}	1.77×10^{-9}	3.38×10^{-9}	2.0×10^{-9}	2
	DR	3.08×10^{-6}	4.8×10^{-6}	1.12×10^{-5}	7.3×10^{-6}	2.15×10^{-5}	7.44×10^{-6}	3.42×10^{-5}	5.34×10^{-6}	4.88×10^{-5}	1.6×10^{-5}	9
EXACT SOLUTION		0.0047619	0.0095238	0.0142857	0.0190476	0.0238095	0.0285714	0.0333333	0.0380952	0.0428571	-	-

TABLE 6.9.37: The absolute errors of the numerical solutions to equation (6.9.25) for Problem 2 (Non-linear first order hyperbolic (convection) equation)

CHAPTER SEVEN

THE ALTERNATING GROUP EXPLICIT (AGE) ITERATIVE METHOD TO SOLVE HIGHER DIMENSIONAL PARABOLIC PROBLEMS

7.1 INTRODUCTION

The AGE method can be readily extended to higher space dimensions. To ensure unconditional stability, the Douglas-Rachford (DR) variant is used instead of the Peaceman-Rachford (PR) formula (cf. Section 3.18). In two space dimensions, for example, the specific problem we are considering is the heat equation,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + h(x,y,t), \quad (x,y,t) \in R\times(0,T] \qquad (7.1.1)$$

with the initial condition,

$$U(x,y,0) = F(x,y), \quad (x,y,t) \in R\times\{0\} \quad , \qquad (7.1.1a)$$

and U(x,y,t) is specified on the boundary of R, ∂R by

$$U(x,y,t) = G(x,y,t), \quad (x,y,t) \in \partial R\times(0,T] \quad , \qquad (7.1.1b)$$

where for simplicity we assume that the region R of the xy-plane is a rectangle. Similarly, the three-dimensional heat equation is given by,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} + h(x,y,z,t) \quad , \qquad (7.1.2)$$

with the initial and boundary conditions specified on R which is now a cube. Based on the AGE concept for the one-dimensional case, the formulation for higher dimensional problems can be done in very much the same way by employing the *fractional splitting* strategy introduced by Yanenko (1971).

7.2 THE AGE METHOD TO SOLVE TWO-DIMENSIONAL PARABOLIC PROBLEMS

Consider the two-dimensional heat equation (7.1.1) with the auxiliary conditions (7.1.1a) and (7.1.1b). The region R is a rectangle defined by

$$R = \{(x,y): 0 \leq x \leq L,\ 0 \leq y \leq M\}.$$

At the point $P(x_i,y_j,t_k)$ in the solution domain, the value of $U(x,y,t)$ is denoted by $U_{i,j,k}$ where $x_i=i\Delta x$, $y_j=j\Delta y$ for $0 \leq i \leq (m+1)$, $0 \leq j \leq (n+1)$ and $\Delta x=L/(m+1)$, $\Delta y=M/(n+1)$. The increment in the time t, Δt is chosen such that $t_k=k\Delta t$ for $k=0,1,2,\ldots$ For simplicity of presentation, we assume that m and n are chosen so that $\Delta x=\Delta y$ and consequently the mesh ratio is defined by $\lambda=\Delta t/(\Delta x)^2$. Analogous to the heat equation in one space dimension, a weighted finite-difference approximation to (7.1.1) at the point $(i,j,k+\frac{1}{2})$ is given by (with $0 \leq \theta \leq 1$)

$$\frac{\Delta_t u_{i,j,k}}{\Delta t} = \frac{1}{(\Delta x)^2}\{\theta(\delta_x^2+\delta_y^2)u_{i,j,k+1}+(1-\theta)(\delta_x^2+\delta_y^2)u_{i,j,k}\}+h_{i,j,k+\frac{1}{2}} \tag{7.2.1}$$

which leads to the *five-point formula*

$$-\lambda\theta u_{i-1,j,k+1}+(1+4\lambda\theta)u_{i,j,k+1}-\lambda\theta u_{i+1,j,k+1}-\lambda\theta u_{i,j-1,k+1}-\lambda\theta u_{i,j+1,k+1}$$

$$= \lambda(1-\theta)u_{i-1,j,k}+(1-4\lambda(1-\theta))u_{i,j,k}+\lambda(1-\theta)u_{i+1,j,k}+\lambda(1-\theta)u_{i,j-1,k}+$$

$$\lambda(1-\theta)u_{i,j+1,k}+\Delta t h_{i,j,k+\frac{1}{2}}\ , \tag{7.2.2}$$

$$\text{for } i=1,2,\ldots,m;\quad j=1,2,\ldots,n.$$

We note that when θ takes the values 0, $\frac{1}{2}$ and 1, we obtain the classical explicit, the Crank-Nicolson and the fully implicit schemes whose truncation errors are $O([\Delta x]^2+\Delta t)$, $O([\Delta x]^2+[\Delta t]^2)$ and $O([\Delta x]^2+\Delta t)$ respectively. The explicit scheme is stable only for $\lambda \leq 1/4$ (if $\Delta x \neq \Delta y$, we need $\Delta t/[(\Delta x)^2+(\Delta y)^2] \leq \frac{1}{8}$). The fully implicit and the

Crank-Nicolson schemes are, however, unconditionally stable. The computational molecules of these schemes are presented in Figures 7.2.1-7.2.3 below:

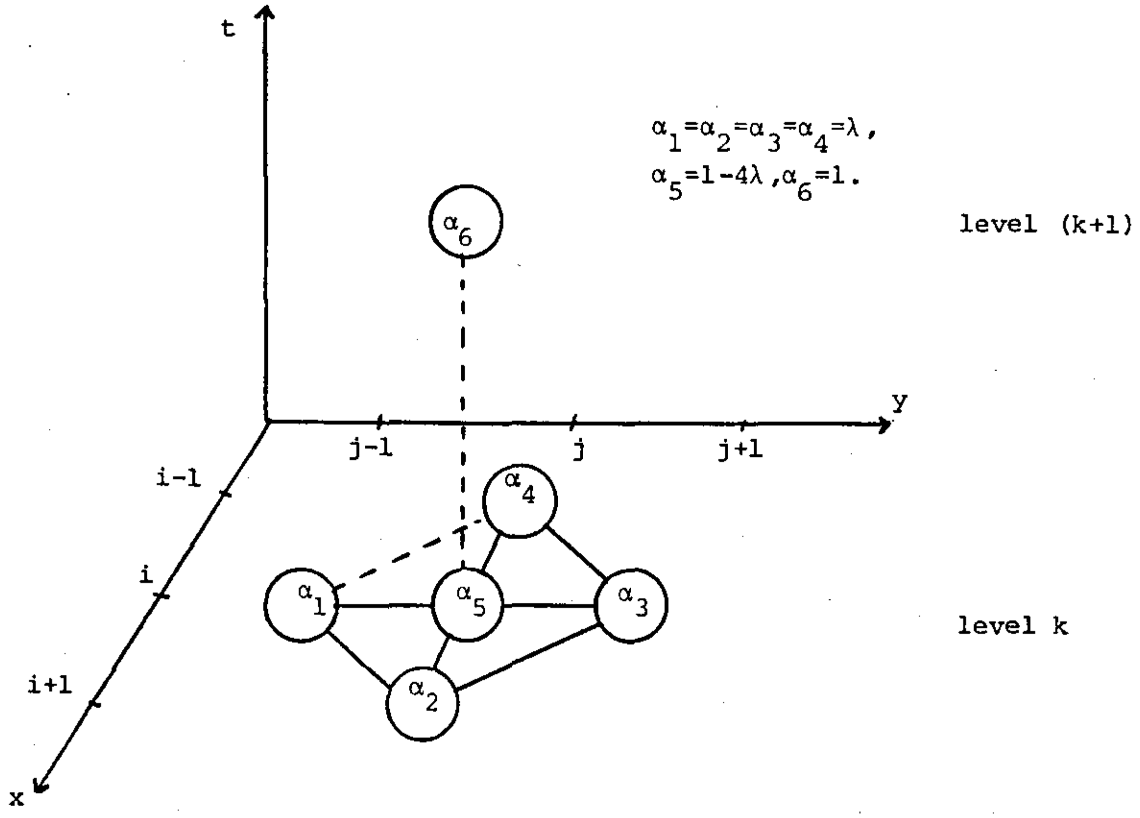


FIGURE 7.2.1: The classical explicit scheme

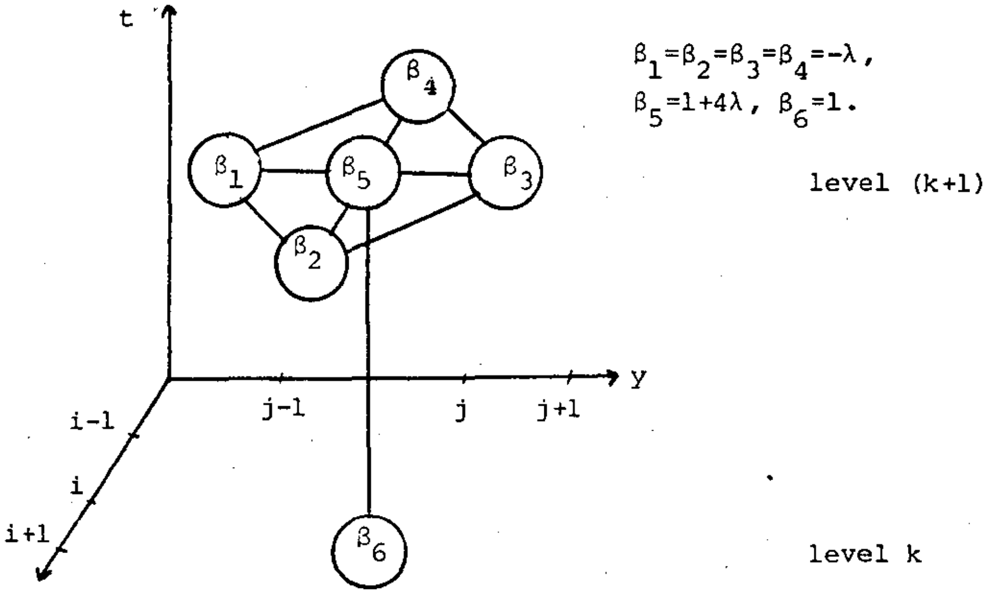


FIGURE 7.2.2: The fully implicit scheme

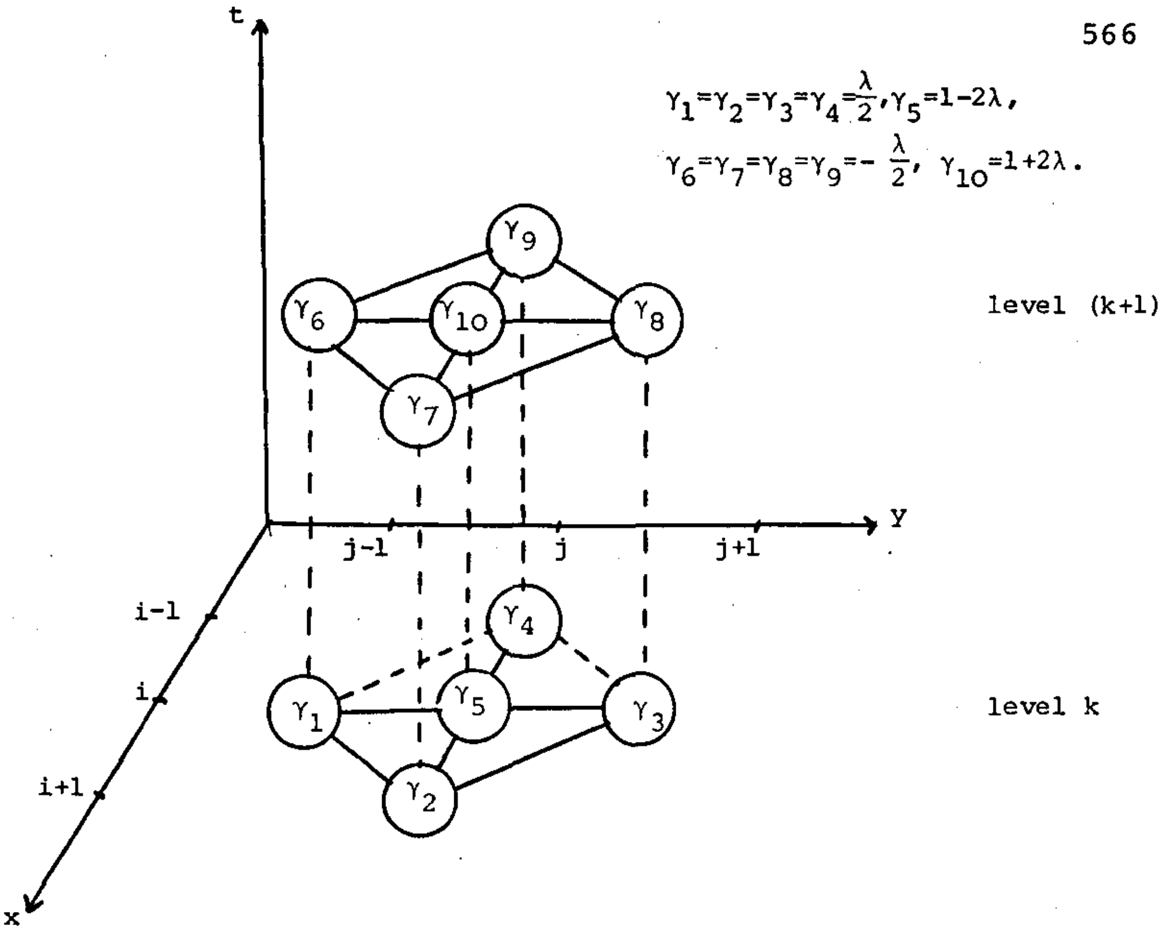


FIGURE 7.2.3: The Crank-Nicolson scheme

The weighted finite-difference equations (7.2.2) can be expressed in the more compact matrix form as

$$A\underline{u}_{(r)}^{[k+1]} = B\underline{u}_{(r)}^{[k]} + \underline{b} + \underline{g} \quad , \tag{7.2.3}$$

$$= \underline{f} \tag{7.2.4}$$

where

$\underline{u}_{(r)}^{[k]}$ are the known u-values at time level k and

$\underline{u}_{(r)} = (\underline{u}_1,\underline{u}_2,\ldots,\underline{u}_n)^T$ with $\underline{u}_j=(u_{1j},u_{2j},\ldots,u_{mj})^T$, $j=1,2,\ldots,n$.

Thus, the mn internal mesh points on the rectangular grid system R are ordered *row-wise* as shown in Figure 7.2.4 below:

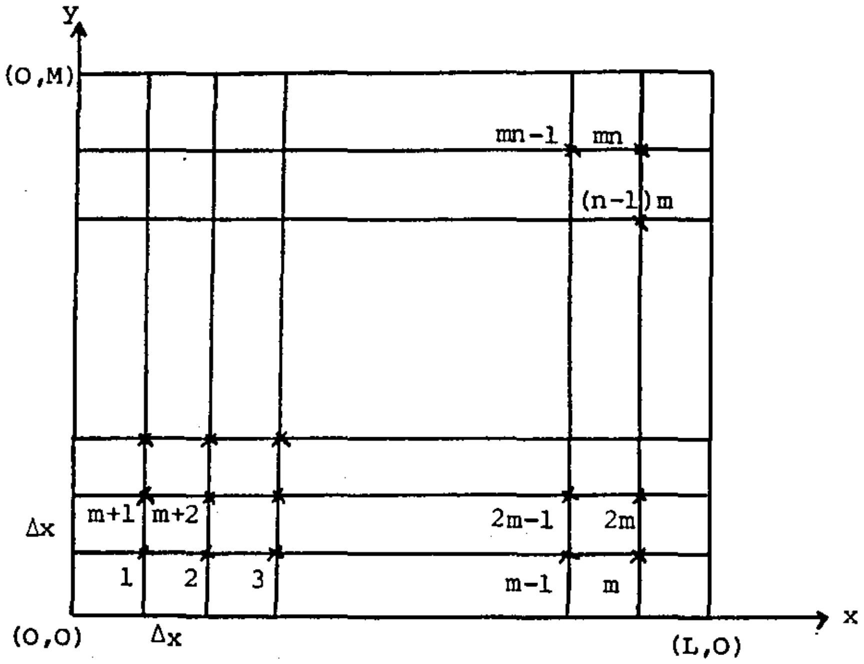


FIGURE 7.2.4

The vector $\underline{b}$ consists of the boundary values where,

$$\underline{b} = (\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_n)^T \text{ with}$$

$$\underline{b}_1 = (\lambda(1-\theta)[u_{0,1,k}+u_{1,0,k}]+\lambda\theta[u_{0,1,k+1}+u_{1,0,k+1}], \lambda(1-\theta)u_{2,0,k}$$

$$+\lambda\theta u_{2,0,k+1}, \ldots, \lambda(1-\theta)u_{m-1,0,k}+\lambda\theta u_{m-1,0,k+1},$$

$$\lambda(1-\theta)[u_{m,0,k}+u_{m+1,1,k}]+\lambda\theta[u_{m,0,k+1}+u_{m+1,1,k+1}])^T \; ;$$

$$\underline{b}_j = (\lambda(1-\theta)u_{0,j,k}+\lambda\theta u_{0,j,k+1}, 0, \ldots, 0, \lambda(1-\theta)u_{m+1,j,k}+\lambda\theta u_{m+1,j,k+1})^T$$

$$\text{for } j=2,3,\ldots,n-1;$$

and
$$\underline{b}_n = (\lambda(1-\theta)[u_{0,n,k}+u_{1,n+1,k}]+\lambda\theta[u_{0,n,k+1}+u_{1,n+1,k+1}],$$

$$\lambda(1-\theta)u_{2,n+1,k}+\lambda\theta u_{2,n+1,k+1}, \ldots, \lambda(1-\theta)u_{m-1,n+1,k}$$

$$+\lambda\theta u_{m-1,n+1,k+1}, \lambda(1-\theta)[u_{m,n+1,k}+u_{m+1,n,k}] +$$

$$\lambda\theta[u_{m,n+1,k+1}+u_{m+1,n,k+1}])^T$$

and the vector $\underline{g}$ contains the source term of (7.2.2) given by

$$\underline{g} = (\underline{g}_1, \underline{g}_2, \ldots, \underline{g}_n)^T \text{ with}$$

$$\underline{g}_j = (g_{1j}, g_{2j}, \ldots, g_{mj})^T$$

$$= \Delta t(h_{1,j,k+\frac{1}{2}}, h_{2,j,k+\frac{1}{2}}, \ldots, h_{m,j,k+\frac{1}{2}})^T \text{ for } j=1,2,\ldots,n.$$

The coefficient matrix A in (7.2.3) takes the block tridiagonal form as,

$$A = \left(\begin{array}{c|c|c|c|c|c|c}
A_1 & A_2 & & & & & \\ \hline
A_2 & A_1 & A_2 & & & & \\ \hline
 & A_2 & A_1 & A_2 & & \bigcirc & \\ \hline
 & & \ddots & \ddots & \ddots & & \\ \hline
 & \bigcirc & & & & & \\
 & & & & A_2 & A_1 & A_2 \\ \hline
 & & & & & A_2 & A_1
\end{array}\right)_{(mn\times mn)} \qquad (7.2.5)$$

with

$$A_1 = \begin{pmatrix}
c & a_1 & & & \\
a_1 & c & a_1 & & \bigcirc \\
 & \ddots & \ddots & \ddots & \\
\bigcirc & & a_1 & c & a_1 \\
 & & & a_1 & c
\end{pmatrix}_{(m\times m)} \qquad (7.2.5a)$$

and $A_2 = \text{diag}(a_1)$ of order $(m\times m)$ (7.2.5b)

where

$$c = 1+4\lambda\theta \quad \text{and} \quad a_1 = -\lambda\theta. \qquad (7.2.5c)$$

Similarly, the matrix B is of the form,

$$B = \left(\begin{array}{c|c|c|c|c|c|c}
B_1 & B_2 & & & & & \\ \hline
B_2 & B_1 & B_2 & & & & \\ \hline
 & B_2 & B_1 & B_2 & & \bigcirc & \\ \hline
 & & \ddots & \ddots & \ddots & & \\ \hline
 & & \bigcirc & & B_2 & B_1 & B_2 \\ \hline
 & & & & & B_2 & B_1
\end{array}\right)_{(mn\times mn)} \qquad (7.2.6)$$

with

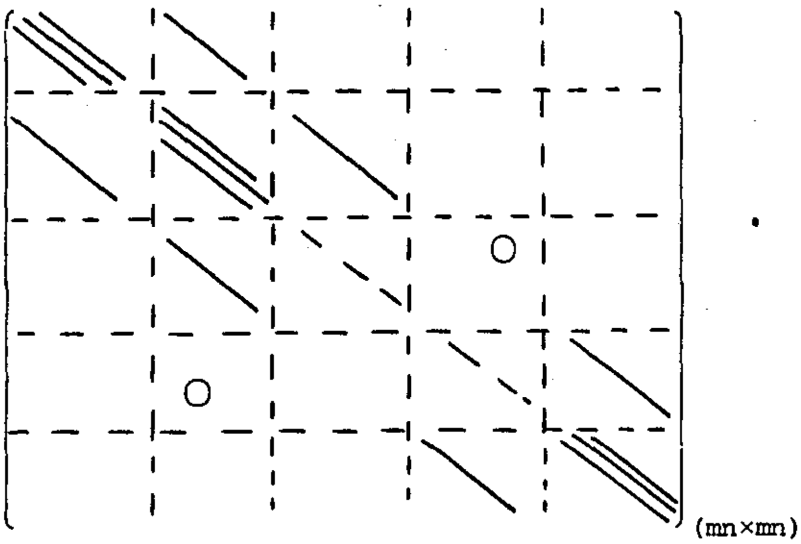
$$B_1 = \begin{bmatrix} d & e_1 & & & \\ e_1 & d & e_1 & & O \\ & \ddots & \ddots & \ddots & \\ & O & e_1 & d & e_1 \\ & & & e_1 & d \end{bmatrix}_{(m\times m)} \qquad (7.2.6a)$$

and $B_2 = \text{diag}(e_1)$ of order $(m\times m)$, (7.2.6b)

where $d = 1-4\lambda(1-\theta)$ and $e_1 = \lambda(1-\theta)$. (7.2.6c)

We observe from (7.2.5) that the matrix A is of the form,

A =

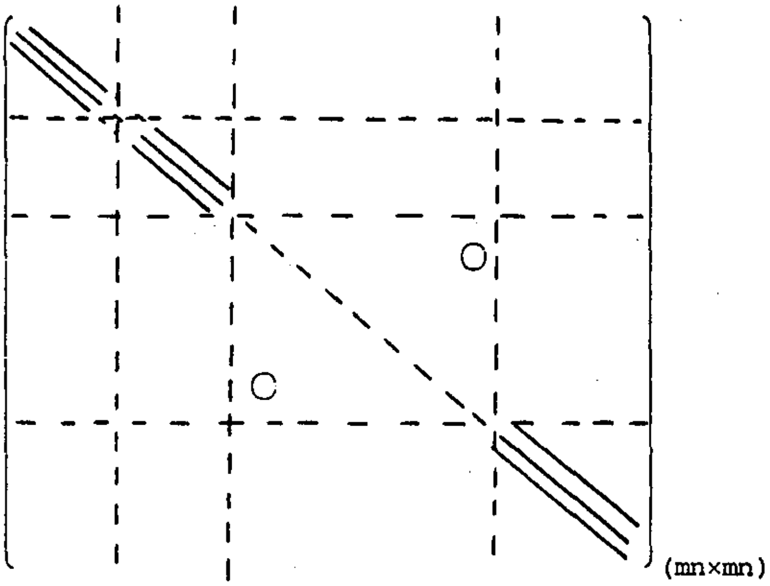


If we split A into the sum of its constituent symmetric and positive definite matrices G_1, G_2, G_3 and G_4, we have,

$$A = G_1 + G_2 + G_3 + G_4 , \qquad (7.2.7)$$

where,

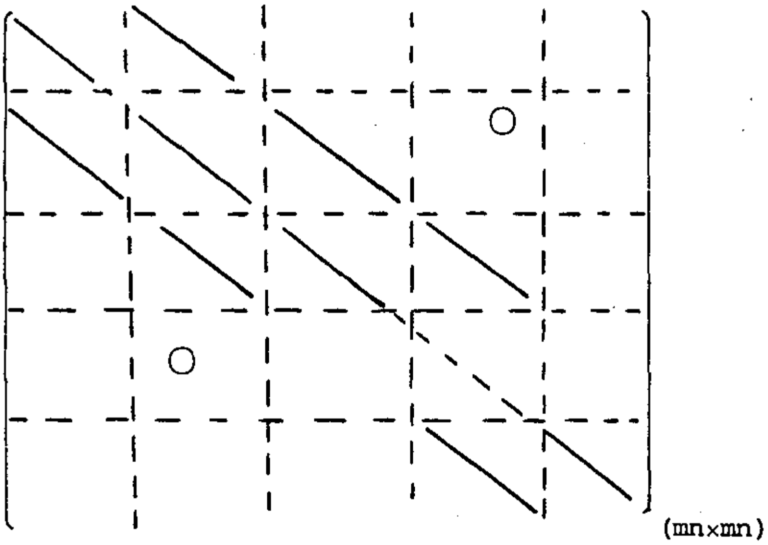
$G_1 + G_2 =$



with $\text{diag}(G_1+G_2) = \frac{1}{2}\text{diag}(A)$

and

$G_3 + G_4 =$



with $\text{diag}(G_3+G_4) = \frac{1}{2}\text{diag}(A)$.

In particular, if we let,

$$\hat{A}_1 = \begin{pmatrix} c/2 & a_1 & & & \\ a_1 & c/2 & a_1 & & O \\ & \ddots & \ddots & \ddots & \\ O & & a_1 & c/2 & a_1 \\ & & & a_1 & c/2 \end{pmatrix}_{(m\times m)} \tag{7.2.8}$$

and $A_4 = \text{diag}(c/2)$ of order $(m\times m)$, we have

$$G_1+G_2=\left[\begin{array}{c|c|ccc|c} \hat{A}_1 & & & & & \\ \hline & \hat{A}_1 & & & & \\ \hline & & \ddots & & O & \\ & & & \ddots & & \\ & & O & & \ddots & \\ \hline & & & & & \hat{A}_1 \end{array}\right]_{(mn\times mn)} \tag{7.2.9}$$

and

$$G_3+G_4=\left[\begin{array}{c|c|ccc|c|c} A_4 & A_2 & & & & & \\ \hline A_2 & A_4 & A_2 & & & & \\ \hline & \ddots & \ddots & \ddots & & O & \\ & & \ddots & \ddots & \ddots & & \\ & O & & \ddots & \ddots & \ddots & \\ \hline & & & & A_2 & A_4 & A_2 \\ \hline & & & & & A_2 & A_4 \end{array}\right]_{(mn\times mn)} . \tag{7.2.10}$$

The Douglas-Rachford formula for the AGE fractional scheme then takes the form,

$$\left.\begin{aligned} (G_1+rI)\underline{u}_{(r)}^{(p+1/4)} &= (rI-G_1-2G_2-2G_3-2G_4)\underline{u}_{(r)}^{(p)}+2\underline{f} \ , \\ (G_2+rI)\underline{u}_{(r)}^{(p+1/2)} &= G_2\underline{u}_{(r)}^{(p)}+r\underline{u}_{(r)}^{(p+1/4)} \ , \\ (G_3+rI)\underline{u}_{(r)}^{(p+3/4)} &= G_3\underline{u}_{(r)}^{(p)}+r\underline{u}_{(r)}^{(p+1/2)} \ , \\ (G_4+rI)\underline{u}_{(r)}^{(p+1)} &= G_4\underline{u}_{(r)}^{(p)}+r\underline{u}_{(r)}^{(p+3/4)} \ . \end{aligned}\right\} \tag{7.2.11}$$

We now consider the above iterative formulae at each of the four intermediate levels:

(i) At the first intermediate level (the $(p+1/4)^{th}$ iterate)

Since $A=G_1+G_2+G_3+G_4$, then using the first expression of (7.2.11) and (7.2.4) we obtain,

$$(G_1+rI)\underline{u}_{(r)}^{(p+1/4)} = ((rI+G_1)-2A)\underline{u}_{(r)}^{(p)}+2B\underline{u}_{(r)}^{[k]}+2(\underline{b}+\underline{g})$$

or
$$\underline{u}_{(r)}^{(p+1/4)} = (G_1+rI)^{-1}[((rI+G_1)-2A)\underline{u}_{(r)}^{(p)}+2B\underline{u}_{(r)}^{[k]}+2(\underline{b}+\underline{g})]. \tag{7.2.12}$$

Without loss of generality we assume that the *size of the matrix A is odd.* We find that

$$(rI+G_1) = \left(\begin{array}{c|c|c|c|c|c} C_1 & & & & & \\ \hline & C_2 & & & & \\ \hline & & C_1 & & & \\ \hline & & & \ddots \quad \text{O} & & \\ & & \text{O} & & & \\ \hline & & & & C_2 & \\ \hline & & & & & C_1 \end{array}\right)_{(mn\times mn)} \tag{7.2.13}$$

and

$$(rI+G_1)^{-1} = \left(\begin{array}{c|c|c|c|c|c} C_1^{-1} & & & & & \\ \hline & C_2^{-1} & & & & \\ \hline & & C_1^{-1} & & & \\ \hline & & & \ddots \quad \text{O} & & \\ & & \text{O} & & & \\ \hline & & & & C_2^{-1} & \\ \hline & & & & & C_1^{-1} \end{array}\right)_{(mn\times mn)} \tag{7.2.14}$$

where

$$C_1 = \left(\begin{array}{c|cc|cc|c|cc} r_1 & & & & & & & \\ \hline & r_1 & a_1 & & & & & \\ & a_1 & r_1 & & & & & \\ \hline & & & r_1 & a_1 & & \text{O} & \\ & & & a_1 & r_1 & & & \\ \hline & & & & & \ddots & & \\ \hline & & \text{O} & & & & r_1 & a_1 \\ & & & & & & a_1 & r_1 \end{array}\right)_{(m\times m)}, \tag{7.2.15}$$

$$C_2 = \left[\begin{array}{cc:cc:c:cc:c}
r_1 & a_1 & & & & & & \\
a_1 & r_1 & & & & & & \\ \hdashline
 & & r_1 & a_1 & & & & \\
 & & a_1 & r_1 & & \text{O} & & \\ \hdashline
 & & & & \ddots & & & \\ \hdashline
 & & & \text{O} & & r_1 & a_1 & \\
 & & & & & a_1 & r_1 & \\ \hdashline
 & & & & & & & r_1
\end{array}\right]_{(m\times m)} . \tag{7.2.16}$$

and

$$C_1^{-1} = \frac{1}{\Delta}\left[\begin{array}{c:cc:cc:c:cc}
\frac{\Delta}{r_1} & & & & & & & \\ \hdashline
 & r_1 & -a_1 & & & & & \\
 & -a_1 & r_1 & & & & & \\ \hdashline
 & & & r_1 & -a_1 & & \text{O} & \\
 & & & -a_1 & r_1 & & & \\ \hdashline
 & & & & & \ddots & & \\ \hdashline
 & & & \text{O} & & & r_1 & -a_1 \\
 & & & & & & -a_1 & r_1
\end{array}\right] , \tag{7.2.17}$$

$$C_2^{-1} = \frac{1}{\Delta}\left[\begin{array}{cc:cc:c:cc:c}
r_1 & -a_1 & & & & & & \\
-a_1 & r_1 & & & & & & \\ \hdashline
 & & r_1 & -a_1 & & & & \\
 & & -a_1 & r_1 & & \text{O} & & \\ \hdashline
 & & & & \ddots & & & \\ \hdashline
 & & & \text{O} & & r_1 & -a_1 & \\
 & & & & & -a_1 & r_1 & \\ \hdashline
 & & & & & & & \frac{\Delta}{r_1}
\end{array}\right] , \tag{7.2.18}$$

with $r_1 = r + \frac{c}{4}$ and $\Delta = (r_1+a_1)(r_1-a_1)$. (7.2.19)

By writing $D_1 = C_1-2A_1$, $D_2 = C_2-2A_1$, $E_1 = -2A_2$, $F_1 = 2B_1$ and $F_2 = 2B_2$ we get

$$((rI+G_1)-2A)\underline{u}^{(p)}_{(r)} + 2B\underline{u}^{[k]}_{(r)} + 2(\underline{b}+\underline{g})$$

$$= \begin{bmatrix} D_1 & E_1 & & & & & \\ E_1 & D_2 & E_1 & & & & \\ & E_1 & D_1 & E_1 & & \bigcirc & \\ & & \ddots & \ddots & \ddots & & \\ & \bigcirc & & & E_1 & D_2 & E_1 \\ & & & & & E_1 & D_1 \end{bmatrix} \begin{bmatrix} \underline{u}^{(p)}_{1(r)} \\ \underline{u}^{(p)}_{2(r)} \\ \underline{u}^{(p)}_{3(r)} \\ \vdots \\ \underline{u}^{(p)}_{n-1(r)} \\ \underline{u}^{(p)}_{n(r)} \end{bmatrix} +$$

$$\begin{bmatrix} F_1 & F_2 & & & & & \\ F_2 & F_1 & F_2 & & & & \\ & F_2 & F_1 & F_2 & & \bigcirc & \\ & & \ddots & \ddots & \ddots & & \\ & \bigcirc & & & F_2 & F_1 & F_2 \\ & & & & & F_2 & F_1 \end{bmatrix} \begin{bmatrix} \underline{u}^{[k]}_{1(r)} \\ \underline{u}^{[k]}_{2(r)} \\ \underline{u}^{[k]}_{3(r)} \\ \vdots \\ \underline{u}^{[k]}_{n-1(r)} \\ \underline{u}^{[k]}_{n(r)} \end{bmatrix} + 2 \begin{bmatrix} \underline{b}_1+\underline{g}_1 \\ \underline{b}_2+\underline{g}_2 \\ \underline{b}_3+\underline{g}_3 \\ \vdots \\ \underline{b}_{n-1}+\underline{g}_{n-1} \\ \underline{b}_n+\underline{g}_n \end{bmatrix} . \quad (7.2.20)$$

Hence, using (7.2.12) we obtain the following set of equations for the computation of the AGE algorithm at the $(p+1/4)^{th}$ iterate:

$$\underline{u}^{(p+1/4)}_{1(r)} = C_1^{-1}(D_1\underline{u}^{(p)}_{1(r)}+E_1\underline{u}^{(p)}_{2(r)}+F_1\underline{u}^{[k]}_{1(r)}+F_2\underline{u}^{[k]}_{2(r)}+2(\underline{b}_1+\underline{g}_1)); \quad (7.2.21a)$$

$$\underline{u}^{(p+1/4)}_{j(r)} = C_2^{-1}(E_1(\underline{u}^{(p)}_{j-1(r)}+\underline{u}^{(p)}_{j+1(r)})+D_2\underline{u}^{(p)}_{j(r)}+F_2(\underline{u}^{[k]}_{j-1(r)}+\underline{u}^{[k]}_{j+1(r)})+$$
$$F_1\underline{u}^{[k]}_{j(r)}+2(\underline{b}_j+\underline{g}_j)) \text{ , for } j=2,4,\ldots,n-1; \quad (7.2.21b)$$

$$\underline{u}^{(p+1/4)}_{j(r)} = C_1^{-1}(E_1(\underline{u}^{(p)}_{j-1(r)}+\underline{u}^{(p)}_{j+1(r)})+D_1\underline{u}^{(p)}_{j(r)}+F_2(\underline{u}^{[k]}_{j-1(r)}+\underline{u}^{[k]}_{j+1(r)})$$
$$F_1\underline{u}^{[k]}_{j(r)}+2(\underline{b}_j+\underline{g}_j)) \text{, for } j=3,5,\ldots,n-2 \text{ ;} \quad (7.2.21c)$$

$$\underline{u}^{(p+1/4)}_{n(r)} = C_1^{-1}(E_1\underline{u}^{(p)}_{n-1(r)}+D_1\underline{u}^{(p)}_{n(r)}+F_2\underline{u}^{[k]}_{n-1(r)}+F_1\underline{u}^{[k]}_{n(r)}+2(\underline{b}_n+\underline{g}_n)). \quad (7.2.21d)$$

Let $\alpha_1=r_1-2c$, $\alpha_2=-2a_1$, $\alpha_3=2d$ and $\alpha_4=2e_1$. When the above equations are written component-wise, we have,

(a) for (7.2.21a)

$$u_{11}^{(p+1/4)} = [\alpha_1u_{11}^{(p)}+\alpha_2(u_{21}^{(p)}+u_{12}^{(p)})+\alpha_3u_{11}^{[k]}+\alpha_4(u_{21}^{[k]}+u_{12}^{[k]})+2(b_{11}+g_{11})]/r_1$$

$$\left.\begin{aligned} u_{i1}^{(p+1/4)} &= [r_1v_i-a_1\bar{v}_i]/\Delta \\ u_{i+1,1}^{(p+1/4)} &= [-a_1v_i+r_1\bar{v}_i]/\Delta \end{aligned}\right\} \quad i=2,4,\ldots,m-1, \qquad (7.2.22a)$$

where

$$v_i = -a_1u_{i+1,1}^{(p)}+\alpha_1u_{i1}^{(p)}+\alpha_2(u_{i-1,1}^{(p)}+u_{i2}^{(p)})+\alpha_3u_{i1}^{[k]}+\alpha_4(u_{i-1,1}^{[k]}+u_{i+1,1}^{[k]}+ u_{i2}^{[k]})+2(g_{i1}+b_{i1})$$

and

$$\bar{v}_i = -a_1u_{i1}^{(p)}+\alpha_1u_{i+1,1}^{(p)}+\alpha_2(u_{i+2,1}^{(p)}+u_{i+1,2}^{(p)})+\alpha_3u_{i+1,1}^{[k]}+\alpha_4(u_{i1}^{[k]}+u_{i+2,1}^{[k]}+ u_{i+1,2}^{[k]})+2(b_{i+1,1}+g_{i+1,1})$$

with $u_{i1}=0$ for $i>m$;

(b) for (7.2.21b)

$$\left.\begin{aligned} u_{i,j}^{(p+1/4)} &= [r_1v_{ij}-a_1\bar{v}_{ij}]/\Delta \\ u_{i+1,j}^{(p+1/4)} &= [-a_1v_{ij}+r_1\bar{v}_{ij}]/\Delta \end{aligned}\right\} j=2,4,\ldots,n-1,\ i=1,3,\ldots,m-2,$$

$$\begin{aligned} u_{mj}^{(p+1/4)} = {} & [\alpha_1u_{mj}^{(p)}+\alpha_2(u_{m-1,j}^{(p)}+u_{m,j-1}^{(p)}+u_{m,j+1}^{(p)})+\alpha_3u_{mj}^{[k]}+\alpha_4(u_{m-1,j}^{[k]} \\ & +u_{m,j-1}^{[k]}+u_{m,j+1}^{[k]})+2(b_{mj}+g_{mj})]/r_1,\ j=2,4,\ldots,n-1, \end{aligned} \qquad (7.2.22b)$$

where,

$$v_{ij} = \alpha_1u_{ij}^{(p)}+\alpha_2(u_{i,j-1}^{(p)}+u_{i,j+1}^{(p)}+u_{i-1,j}^{(p)})+\alpha_3u_{ij}^{[k]}+\alpha_4(u_{i,j-1}^{[k]}+u_{i,j+1}^{[k]}+ u_{i+1,j}^{[k]}+u_{i-1,j}^{[k]})-a_1u_{i+1,j}^{(p)}+2(b_{i+1,j}+g_{i+1,j})$$

and

$$\bar{v}_{ij} = \alpha_1u_{i+1,j}^{(p)}+\alpha_2(u_{i+1,j-1}^{(p)}+u_{i+1,j+1}^{(p)}+u_{i+2,j}^{(p)})+\alpha_3u_{i+1,j}^{[k]}+\alpha_4(u_{ij}^{[k]}+ u_{i+1,j-1}^{[k]}+u_{i+1,j+1}^{[k]}+u_{i+2,j}^{[k]})-a_1u_{ij}^{(p)}+2(b_{i+1,j}+g_{i+1,j})$$

with $u_{0j}=0$.

(c) for (7.2.21c)

$$u_{1j}^{(p+1/4)} = [\alpha_1 u_{1j}^{(p)} + \alpha_2(u_{1,j-1}^{(p)} + u_{1,j+1}^{(p)} + u_{2j}^{(p)}) + \alpha_3 u_{1j}^{[k]} + \alpha_4(u_{1,j-1}^{[k]} + u_{1,j+1}^{[k]} + u_{2,j}^{[k]}) + 2(b_{1j} + g_{1j})]/r_1 \ , \ j=3,5,\ldots,n-2, \tag{7.2.22c}$$

$$\left.\begin{aligned} u_{ij}^{(p+1/4)} &= [r_1 w_{ij} - a_1 \overline{w}_{ij}]/\Delta \ , \\ u_{i+1,j}^{(p+1/4)} &= [-a_1 w_{ij} + r_1 \overline{w}_{ij}]/\Delta \ , \end{aligned}\right\} \text{for } j=3,5,\ldots,n-2;\ i=2,4,\ldots,m-1,$$

where,

$$w_{ij} = \alpha_1 u_{ij}^{(p)} + \alpha_2(u_{i-1,j}^{(p)} + u_{i,j-1}^{(p)} + u_{i,j+1}^{(p)}) + \alpha_3 u_{ij}^{[k]}$$
$$+\alpha_4(u_{i-1,j}^{[k]} + u_{i,j-1}^{[k]} + u_{i,j+1}^{[k]} + u_{i+1,j}^{[k]}) - a_1 u_{i+1,j}^{(p)} + 2(b_{ij} + g_{ij})$$

and

$$\overline{w}_{ij} = \alpha_1 u_{i+1,j}^{(p)} + \alpha_2(u_{i+1,j-1}^{(p)} + u_{i+1,j+1}^{(p)} + u_{i+2,j}^{(p)}) + \alpha_3 u_{i+1,j}^{[k]}$$
$$+\alpha_4(u_{ij}^{[k]} + u_{i+1,j-1}^{[k]} + u_{i+1,j+1}^{[k]} + u_{i+2,j}^{[k]}) - a_1 u_{ij}^{(p)} + 2(b_{i+1,j} + g_{i+1,j})$$

with $u_{ij}=0$ for $i>m$,

(d) for (7.2.21d)

$$u_{1n}^{(p+1/4)} = [\alpha_1 u_{1n}^{(p)} + \alpha_2(u_{1,n-1}^{(p)} + u_{2,n}^{(p)}) + \alpha_3 u_{1n}^{[k]} + \alpha_4(u_{1,n-1}^{[k]} + u_{2,n}^{[k]}) + 2(b_{1n} + g_{1n})]/r_1$$

$$\left.\begin{aligned} u_{i,n}^{(p+1/4)} &= [r_1 z_i - a_1 \overline{z}_i]/\Delta \\ u_{i+1,n}^{(p+1/4)} &= [-a_1 z_i + r_1 \overline{z}_i]/\Delta \end{aligned}\right\} i=2,4,\ldots,m-1 \ , \tag{7.2.22d}$$

where

$$z_i = \alpha_1 u_{i,n}^{(p)} + \alpha_2(u_{i-1,n}^{(p)} + u_{i,n-1}^{(p)}) + \alpha_3 u_{i,n}^{[k]} + \alpha_4(u_{i-1,n}^{[k]} + u_{i,n-1}^{[k]} + u_{i+1,n}^{[k]})$$
$$-a_1 u_{i+1,n}^{(p)} + 2(b_{i,n} + g_{i,n})$$

and

$$\overline{z}_i = \alpha_1 u_{i+1,n}^{(p)} + \alpha_2(u_{i+1,n-1}^{(p)} + u_{i+2,n}^{(p)}) + \alpha_3 u_{i+1,n}^{[k]} + \alpha_4(u_{i,n}^{[k]} + u_{i+1,n-1}^{[k]} +$$
$$u_{i+2,n}^{[k]}) - a_1 u_{i,n}^{(p)} + 2(b_{i+1,n} + g_{i+1,n})$$

with $u_{i,n}=0$ for $i>m$.

(ii) At the second intermediate level (the $(p+\frac{1}{2})^{th}$ iterate)

From the second equation of (7.2.11) we have

$$\underline{u}_{(r)}^{(p+\frac{1}{2})} = (G_2+rI)^{-1}[G_2\underline{u}_{(r)}^{(p)} + r\underline{u}_{(r)}^{(p+1/4)}] \; . \tag{7.2.23}$$

Let $\hat{C}_1 \equiv C_1$ as in (7.2.15) with the diagonal elements r_1 replaced by $c/4$ and $\hat{C}_2 \equiv C_2$ as in (7.2.16) with the diagonal elements r_1 replaced by $c/4$.

We also have,

$$G_2+rI = \left(\begin{array}{c|c|c|ccc|c|c} C_2 & & & & & & & \\ \hline & C_1 & & & & & & \\ \hline & & C_2 & & & & & \\ \hline & & & \ddots & & \bigcirc & & \\ & & & & \ddots & & & \\ & & & \bigcirc & & \ddots & & \\ \hline & & & & & & C_1 & \\ \hline & & & & & & & C_2 \end{array}\right)_{(mn\times mn)} \tag{7.2.24}$$

and

$$(G_2+rI)^{-1} = \left(\begin{array}{c|c|c|ccc|c|c} C_2^{-1} & & & & & & & \\ \hline & C_1^{-1} & & & & & & \\ \hline & & C_2^{-1} & & & & & \\ \hline & & & \ddots & & \bigcirc & & \\ & & & & \ddots & & & \\ & & & \bigcirc & & \ddots & & \\ \hline & & & & & & C_1^{-1} & \\ \hline & & & & & & & C_2^{-1} \end{array}\right)_{(mn\times mn)} \; . \tag{7.2.25}$$

Hence, equation (7.2.23) yields,

$$\begin{bmatrix} \underline{u}_{1(r)}^{(p+\frac{1}{2})} \\ \underline{u}_{2(r)}^{(p+\frac{1}{2})} \\ \vdots \\ \underline{u}_{n-1(r)}^{(p+\frac{1}{2})} \\ \underline{u}_{n(r)}^{(p+\frac{1}{2})} \end{bmatrix}_{(mn\times 1)} = \begin{bmatrix} C_2^{-1} & & & & & \\ & C_1^{-1} & & & & \\ & & C_2^{-1} & & \bigcirc & \\ & & & \ddots & & \\ & & \bigcirc & & C_1^{-1} & \\ & & & & & C_2^{-1} \end{bmatrix}_{(mn\times mn)} \begin{bmatrix} \hat{C}_2\underline{u}_{1(r)}^{(p)}+r\underline{u}_{1(r)}^{(p+1/4)} \\ \hat{C}_1\underline{u}_{2(r)}^{(p)}+r\underline{u}_{2(r)}^{(p+1/4)} \\ \vdots \\ \hat{C}_1\underline{u}_{n-1(r)}^{(p)}+r\underline{u}_{n-1(r)}^{(p+1/4)} \\ \hat{C}_2\underline{u}_{n(r)}^{(p)}+r\underline{u}_{n(r)}^{(p+1/4)} \end{bmatrix}_{(mn\times} \tag{7.2.26}$$

For computational purposes, we will then have,

$$\left.\begin{aligned} \underline{u}_{j(r)}^{(p+\frac{1}{2})} &= C_2^{-1}[\hat{C}_2\underline{u}_{j(r)}^{(p)}+r\underline{u}_{j(r)}^{(p+1/4)}] \ , && (7.2.27a) \\ \underline{u}_{j+1(r)}^{(p+\frac{1}{2})} &= C_1^{-1}[\hat{C}_1\underline{u}_{j+1(r)}^{(p)}+r\underline{u}_{j+1(r)}^{(p+1/4)}] \ , && (7.2.27b) \end{aligned}\right\} j=1,3,\ldots,n-2,$$

and
$$\underline{u}_{n(r)}^{(p+\frac{1}{2})} = C_2^{-1}[\hat{C}_2\underline{u}_{n(r)}^{(p)}+r\underline{u}_{n(r)}^{(p+1/4)}] \ . \tag{7.2.27c}$$

By denoting $r_2 = \frac{c}{4r_1}$ and $r_3 = \frac{r}{r_1}$ the above equations can be written component-wise as follows:

(a) for equations (7.2.27a) and (7.2.27c),

$$\left.\begin{aligned} u_{ij}^{(p+\frac{1}{2})} &= [r_1v_{ij}-a_1\bar{v}_{ij}]/\Delta \\ u_{i+1,j}^{(p+\frac{1}{2})} &= [-a_1v_{ij}+r_1\bar{v}_{ij}]/\Delta \end{aligned}\right\} \text{for } j=1,3,\ldots,n;\ i=1,3,\ldots,m-2, \tag{7.2.28a}$$

$$u_{mj}^{(p+\frac{1}{2})} = r_2u_{mj}^{(p)}+r_3u_{mj}^{(p+1/4)} \quad , \ j=1,3,\ldots,n,$$

where
$$v_{ij} = \frac{c}{4}\,u_{ij}^{(p)}+a_1u_{i+1,j}^{(p)}+ru_{ij}^{(p+1/4)}$$

and
$$\bar{v}_{ij} = a_1u_{ij}^{(p)} + \frac{c}{4}\,u_{i+1,j}^{(p)}+ru_{i+1,j}^{(p+1/4)} .$$

(b) for equation (7.2.27b)

$$u_{1j}^{(p+\frac{1}{2})} = r_2u_{1j}^{(p)}+r_3u_{1j}^{(p+1/4)} \quad , \ j=2,4,\ldots,n-1,$$

$$\left.\begin{aligned} u_{ij}^{(p+\frac{1}{2})} &= [r_1 v_{ij} - a_1 \overline{v}_{ij}]/\Delta \\ u_{i+1,j}^{(p+\frac{1}{2})} &= [-a_1 v_{ij} + r_1 \overline{v}_{ij}]/\Delta \end{aligned}\right\} j=2,4,\ldots,n-1;\ i=2,4,\ldots,m-1, \tag{7.2.28b}$$

where v_{ij} and $\overline{v}_{ij}$ are given as in (7.2.28a).

(iii) At the third intermediate level (the $(p+3/4)^{th}$ iterate)

If we *reorder* the mesh points *column-wise* parallel to the y-axis, we have,

$$\underline{u}_{(c)} = (\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_m)^T \text{ with } \underline{u}_i = (u_{i1}, u_{i2}, \ldots, u_{in})^T$$

for $i=1,2,\ldots,m$.

We also find that,

$$(G_3+G_4)\underline{u}_{(r)} = (\hat{G}_1+\hat{G}_2)\underline{u}_{(c)}$$

and

$$G_3\underline{u}_{(r)} = \hat{G}_1\underline{u}_{(c)},\ G_4\underline{u}_{(r)} = \hat{G}_2\underline{u}_{(c)}\ . \tag{7.2.29}$$

Hence the third equation of (7.2.11) is transformed to

$$(\hat{G}_1+rI)\underline{u}_{(c)}^{(p+3/4)} = \hat{G}_1\underline{u}_{(c)}^{(p)} + r\underline{u}_{(c)}^{(p+\frac{1}{2})}$$

or

$$\underline{u}_{(c)}^{(p+3/4)} = (\hat{G}_1+rI)^{-1}[\hat{G}_1\underline{u}_{(c)}^{(p)} + r\underline{u}_{(c)}^{(p+\frac{1}{2})}]\ . \tag{7.2.30}$$

Let the matrices P_1 and P_2 be exactly of the same forms as C_1 and C_2 of (7.2.15) and (7.2.16) but of order $(n\times n)$. We will then have

$$\hat{G}_1+rI = \left(\begin{array}{c|c|c|c|c|c} P_1 & & & & & \\ \hline & P_2 & & & & \\ \hline & & P_1 & & O & \\ \hline & & & \ddots & & \\ \hline & O & & & P_2 & \\ \hline & & & & & P_1 \end{array}\right)_{(mn\times mn)} ,$$

$$(\hat{G}_1+rI)^{-1} = \left(\begin{array}{c|c|c|ccc|c|c}
P_1^{-1} & & & & & & & \\ \hline
 & P_2^{-1} & & & & & & \\ \hline
 & & P_1^{-1} & & & & & \\ \hline
 & & & \ddots & & O & & \\
 & & & O & & & & \\ \hline
 & & & & & & P_2^{-1} & \\ \hline
 & & & & & & & P_1^{-1}
\end{array}\right)$$

and

$$\begin{pmatrix} \underline{u}_{1(c)}^{(p+3/4)} \\ \hline \underline{u}_{2(c)}^{(p+3/4)} \\ \hline \vdots \\ \hline \underline{u}_{m-1(c)}^{(p+3/4)} \\ \hline \underline{u}_{m(c)}^{(p+3/4)} \end{pmatrix} = \left(\begin{array}{c|c|c|ccc|c|c}
P_1^{-1} & & & & & & & \\ \hline
 & P_2^{-1} & & & & & & \\ \hline
 & & P_1^{-1} & & & O & & \\ \hline
 & & & \ddots & & & & \\ \hline
 & & O & & & & P_2^{-1} & \\ \hline
 & & & & & & & P_1^{-1}
\end{array}\right) \begin{pmatrix} \hat{P}_1\underline{u}_{1(c)}^{(p)}+r\underline{u}_{1(c)}^{(p+\frac{1}{2})} \\ \hline \hat{P}_2\underline{u}_{2(c)}^{(p)}+r\underline{u}_{2(c)}^{(p+\frac{1}{2})} \\ \hline \hat{P}_1\underline{u}_{3(c)}^{(p)}+r\underline{u}_{3(c)}^{(p+\frac{1}{2})} \\ \hline \vdots \\ \hline \hat{P}_2\underline{u}_{m-1(c)}^{(p)}+r\underline{u}_{(m-1)(c)}^{(p+\frac{1}{2})} \\ \hline \hat{P}_1\underline{u}_{m(c)}^{(p)}+r\underline{u}_{m(c)}^{(p+\frac{1}{2})} \end{pmatrix}$$

where $\hat{P}_1 \equiv P_1$ with r_1 replaced by $\frac{c}{4}$,

and $\hat{P}_2 \equiv P_2$ with r_1 replaced by $\frac{c}{4}$.

The following equations are therefore obtained for computation at the $(p+3/4)^{th}$ level:

$$\underline{u}_{i(c)}^{(p+3/4)} = P_1^{-1}[\hat{P}_1\underline{u}_{i(c)}^{(p)}+r\underline{u}_{i(c)}^{(p+\frac{1}{2})}] \text{ for } i=1,3,\ldots,m \qquad (7.2.31a)$$

and

$$\underline{u}_{i(c)}^{(p+3/4)} = P_2^{-1}[\hat{P}_2\underline{u}_{i(c)}^{(p)}+r\underline{u}_{i(c)}^{(p+\frac{1}{2})}] \text{ for } i=2,4,\ldots,m-1. \qquad (7.2.31b)$$

which component-wise yields

(a) for equation (7.2.31a)

$$\begin{aligned} u_{i1}^{(p+3/4)} &= r_2u_{i1}^{(p)}+r_3u_{i1}^{(p+\frac{1}{2})}, \quad i=1,3,\ldots,m, \\ &\left.\begin{aligned} u_{i,j}^{(p+3/4)} &= [r_1w_{ij}-a_1\tilde{w}_{ij}]/\Delta, \\ u_{i,j+1}^{(p+3/4)} &= [-a_1w_{ij}+r_1\tilde{w}_{ij}]/\Delta, \end{aligned}\right\} i=1,3,\ldots,m;\ j=2,4,\ldots,n-1, \end{aligned} \qquad (7.2.32a)$$

where

$$w_{ij} = \frac{c}{4}u_{ij}^{(p)}+a_1u_{i,j+1}^{(p)}+ru_{ij}^{(p+\frac{1}{2})}$$

and

$$\bar{w}_{ij} = a_1u_{ij}^{(p)}+\frac{c}{4}u_{i,j+1}^{(p)}+ru_{i,j+1}^{(p+\frac{1}{2})} \quad ,$$

(b) for equation (7.2.31b)

$$\left.\begin{aligned} u_{i,j}^{(p+3/4)} &= [r_1w_{ij}-a_1\bar{w}_{ij}]/\Delta \\ u_{i,j+1}^{(p+3/4)} &= [-a_1w_{ij}+r_1\bar{w}_{ij}]/\Delta \end{aligned}\right\}, \; i=2,4,\ldots,m-1;\; j=1,3,\ldots,n-2, \qquad (7.2.32b)$$

$$u_{i,n}^{(p+3/4)} = r_2u_{i,n}^{(p)}+r_3u_{i,n}^{(p+\frac{1}{2})} \quad ,$$

where w_{ij} and $\bar{w}_{ij}$ are given as in (7.2.32a).

(iv) At the fourth intermediate level (the (p+1)th iterate)

By virtue of (7.2.29), the last equation of (7.2.11) is transformed to

$$(\hat{G}_2+rI)\underline{u}_{(c)}^{(p+1)} = \hat{G}_2\underline{u}_{(c)}^{(p)}+r\underline{u}_{(c)}^{(p+3/4)}$$

or

$$\underline{u}_{(c)}^{(p+1)} = (\hat{G}_2+rI)^{-1}[\hat{G}_2\underline{u}_{(c)}^{(p)}+r\underline{u}_{(c)}^{(p+3/4)}]$$

which leads to the following formulae:

$$\underline{u}_{i(c)}^{(p+1)} = P_2^{-1}[\hat{P}_2\underline{u}_{i(c)}^{(p)}+r\underline{u}_{i(c)}^{(p+3/4)}] \;,\; i=1,3,\ldots,m, \qquad (7.2.33a)$$

and

$$\underline{u}_{i(c)}^{(p+1)} = P_1^{-1}[\hat{P}_1\underline{u}_{i(c)}^{(p)}+r\underline{u}_{i(c)}^{(p+3/4)}] \;,\; i=2,4,\ldots,m-1 \;. \qquad (7.2.33b)$$

For computational purposes, we have,

(a) for equation (7.2.33a),

$$\left.\begin{aligned} u_{ij}^{(p+1)} &= [r_1z_{ij}-a_1\bar{z}_{ij}]/\Delta \\ u_{i,j+1}^{(p)} &= [-a_1z_{ij}+r_1\bar{z}_{ij}]/\Delta \end{aligned}\right\}, \; i=1,3,\ldots,m;\; j=1,3,\ldots,n-2 \qquad (7.2.34a)$$

$$u_{i,n}^{(p+1)} = r_2u_n^{(p)}+r_3u_{i,n}^{(p+3/4)} \;,\; i=1,3,\ldots,m$$

and

(b) for equation (7.2.33b)

$$\left.\begin{aligned} u_{i,1}^{(p+1)} &= r_2u_{i,1}^{(p)}+r_3u_{i,1}^{(p+3/4)} \quad, \; i=2,4,\ldots,m-1, \\ u_{ij}^{(p+1)} &= [r_1z_{ij}-a_1\bar{z}_{ij}]/\Delta \\ u_{i,j+1}^{(p+1)} &= [-a_1z_{ij}+r_1\bar{z}_{ij}]/\Delta \end{aligned}\right\}, \; i=2,4,\ldots,m-1; \; j=2,4,\ldots,n-1, \qquad (7.2.34b)$$

where

$$z_{ij} = \frac{c}{4}\,u_{ij}^{(p)}+a_1u_{i,j+1}^{(p)}+ru_{ij}^{(p+3/4)}$$

and $$\bar{z}_{ij} = a_1u_{ij}^{(p)}+\frac{c}{4}u_{i,j+1}^{(p)}+ru_{i,j+1}^{(p+3/4)} \quad .$$

Hence, the AGE scheme corresponds to sweeping through the mesh parallel to the coordinate x and y axes involving at each stage the solution of 2×2 block systems. The iterative procedure is continued until convergence is reached, that is, when the requirement $|u_{ij}^{(p+1)}-u_{ij}^{(p)}|\leqslant\varepsilon$ is met where ε is the convergence criterion.

7.3 THE AGE METHOD TO SOLVE THREE-DIMENSIONAL PARABOLIC PROBLEMS

We will now develop the AGE method for three-dimensional problems in exactly the same manner for two-dimensional equations. Consider the following heat equation in three dimensions,

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} + h(x,y,z,t), \quad (x,y,z,t) \in R \times (0,T] \tag{7.3.1}$$

with the initial condition

$$U(x,y,z,0) = F(x,y,z), \quad (x,y,z,t) \in R \times \{0\} \tag{7.3.1a}$$

and the boundary conditions

$$U(x,y,z,t) = G(x,y,z,t), \quad (x,y,z,t) \in \partial R \times (0,T] \tag{7.3.1b}$$

where R is the cube $0<x,y,z<1$ and ∂R its boundary. Let i,j,k and N be the indices in the x,y,z and t-direction respectively with increments $\Delta x,\Delta y,\Delta z$ and Δt (for a cube, $0\leq i,j,k\leq (m+1)$, $N=0,1,\ldots$ and $\Delta x=\Delta y=\Delta z=\frac{1}{(m+1)}$). At the point $P(x_i,y_j,z_k,t_N)$ in the solution domain, the value of $U(x_i,y_j,z_k,t_N)$ is denoted by $U_{i,j,k}^{[N]}$. A weighted finite-difference approximation to (7.3.1) at the point $(x_i,y_j,z_k,t_{N+\frac{1}{2}})$ is given by (with $0\leq\theta\leq 1$),

$$\frac{(u_{i,j,k}^{[N+1]}-u_{i,j,k}^{[N]})}{\Delta t} = \frac{1}{(\Delta x)^2}\{\theta(\delta_x^2+\delta_y^2+\delta_z^2)u_{i,j,k}^{[N+1]}+(1-\theta)(\delta_x^2+\delta_y^2+\delta_z^2)u_{i,j,k}^{[N]}\}+ h_{i,j,k}^{[N+\frac{1}{2}]}, \quad i,j,k=1,2,\ldots,m, \tag{7.3.2}$$

which leads to the *seven-point formula*

$$\begin{aligned}&-\lambda\theta u_{i-1,j,k}^{[N+1]}+(1+6\lambda\theta)u_{i,j,k}^{[N+1]}-\lambda\theta u_{i+1,j,k}^{[N+1]}-\lambda\theta u_{i,j-1,k}^{[N+1]}-\lambda\theta u_{i,j,k-1}^{[N+1]}-\\ &\lambda\theta u_{i,j+1,k}^{[N+1]}-\lambda\theta u_{i,j,k+1}^{[N+1]} = \lambda(1-\theta)u_{i-1,j,k}^{[N]}+(1-6\lambda(1-\theta))u_{i,j,k}^{[N]}\\ &+\lambda(1-\theta)u_{i+1,j,k}^{[N]}+\lambda(1-\theta)u_{i,j-1,k}^{[N]}+\lambda(1-\theta)u_{i,j,k-1}^{[N]}+\lambda(1-\theta)u_{i,j+1,k}^{[N]}\\ &+\lambda(1-\theta)u_{i,j,k+1}^{[N]}+\Delta t\, h_{i,j,k}^{[N+\frac{1}{2}]} \quad .\end{aligned} \tag{7.3.3}$$

As before, the formula (7.3.2) corresponds to the explicit, the Crank-Nicolson and the fully implicit schemes when θ is equal to $0, \frac{1}{2}$ and 1 respectively. By considering our approximations as sweeps parallel to the xy-plane of the cube $R=\{(x,y,z: 0<x,y,z<1\}$, the equation (7.3.3) can be written in matrix form as,

$$\begin{aligned} A\underline{u}^{[N+1]} &= B\underline{u}^{[N]} + \underline{b} + \underline{g} \\ &= \underline{f}\ , \end{aligned} \qquad (7.3.4)$$

where $\underline{b}$ consists of the boundary values,

$$\underline{g} = (\underline{g}_1, \underline{g}_2, \ldots, \underline{g}_m)^T$$

with $\underline{g}_k = (g_{1,1,k}, g_{2,1,k}, \ldots, g_{m,1,k}, g_{1,2,k}, g_{2,2,k}, \ldots, g_{m,2,k}, g_{1,m,k},$

$$g_{2,m,k}, \ldots, g_{m,m,k})^T \text{ , for } k=1,2,\ldots,m$$

and $g_{i,j,k} = \Delta t h_{i,j,k}^{[N+\frac{1}{2}]}$. The vector $\underline{u}^{[N]}$ denotes the u-values at time level N - and the coefficient matrix A is given by,

$$A = \left[\begin{array}{c|c|c|c|c|c} A_1 & A_2 & & & & \\ \hline A_2 & A_1 & A_2 & & & \\ \hline & \ddots & \ddots & \ddots & O & \\ \hline & O & & A_2 & A_1 & A_2 \\ \hline & & & & A_2 & A_1 \end{array}\right]_{(m^3 \times m^3)} \qquad (7.3.5)$$

where the block matrices A_i of size $m^2 \times m^2$ take the forms

$$A_1 = \left[\begin{array}{c|c|c|c|c|c} H_1 & H_2 & & & & \\ \hline H_2 & H_1 & H_2 & & & \\ \hline & \ddots & \ddots & \ddots & O & \\ \hline & & O & H_2 & H_1 & H_2 \\ \hline & & & & H_2 & H_1 \end{array}\right]_{(m^2 \times m^2)} , \ A_2 = \text{diag}(a_1) \text{ of order } (m^2 \times m^2)$$

with

$$H_1 = \begin{pmatrix} c & a_1 & & & \bigcirc \\ a_1 & c & a_1 & & \\ & \ddots & \ddots & \ddots & \\ \bigcirc & & a_1 & c & a_1 \\ & & & a_1 & c \end{pmatrix}, \quad H_2=\text{diag}(a_1) \text{ of order } (m\times m)$$

and $\quad c = 1+6\lambda\theta, \quad a_1 = -\lambda\theta \;, \quad \lambda = \dfrac{\Delta t}{(\Delta x)^2}$.

Similarly, the block matrix B is defined by,

$$B = \left(\begin{array}{c|c|c|c|c} B_1 & B_2 & & & \\ \hline B_2 & B_1 & B_2 & & \bigcirc \\ \hline & \ddots & \ddots & \ddots & \\ \hline \bigcirc & & B_2 & B_1 & B_2 \\ \hline & & & B_2 & B_1 \end{array}\right)_{(m^3\times m^3)} \tag{7.3.6}$$

where,

$$B_1 = \left(\begin{array}{c|c|c|c|c} J_1 & J_2 & & & \\ \hline J_2 & J_1 & J_2 & & \bigcirc \\ \hline & \ddots & \ddots & \ddots & \\ \hline \bigcirc & & J_2 & J_1 & J_2 \\ \hline & & & J_2 & J_1 \end{array}\right)_{(m^2\times m^2)}, \quad B_2=\text{diag}(e_1) \text{ of order } (m^2\times m^2),$$

$$J_1 = \begin{pmatrix} d & e_1 & & & & \\ e_1 & d & e_1 & & \bigcirc & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & \bigcirc & & e_1 & d & e_1 \\ & & & & e_1 & d \end{pmatrix}_{(m\times m)} \text{ and } J_2=\text{diag}(e_1) \text{ of order } (m\times m)$$

with $d = 1-6\lambda(1-\theta)$ and $e_1=\lambda(1-\theta)$.

For our approximations which are taken as sweeps parallel to the xy-plane, the u values are evaluated at points lying on planes which are parallel to the xy-plane and perpendicular to the z-axis and on each of these planes, the points are ordered row-wise (parallel to the x-axis) as depicted in Figure 7.3.1 below.

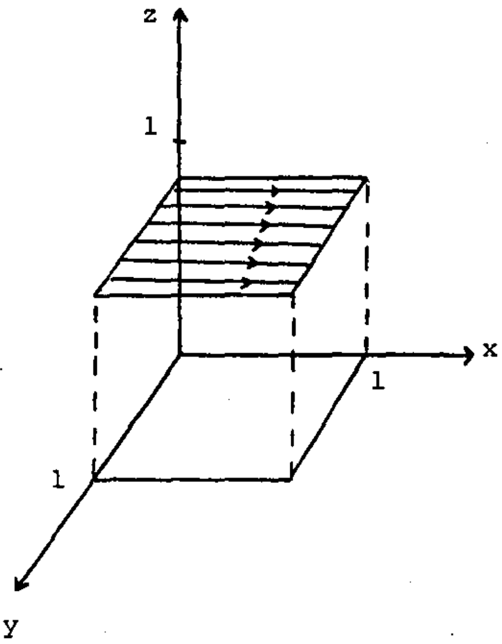


FIGURE 7.3.1

Equation (7.3.4) is therefore written as,

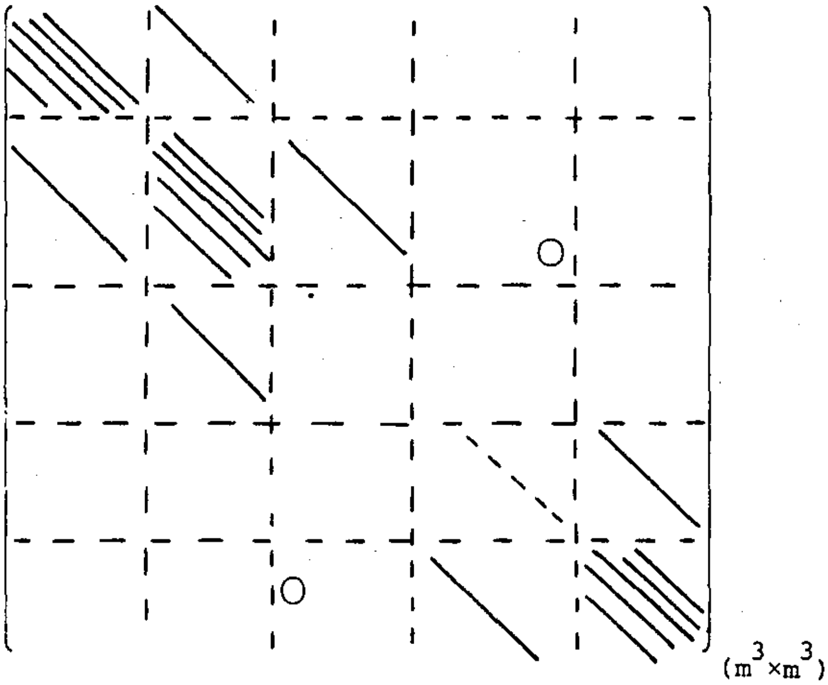
$$\begin{aligned} A\underline{u}^{[N+1]}_{[xy]} &= B\underline{u}^{[N]}_{[xy]} + \underline{b} + \underline{g} \\ &= \underline{f}\ . \end{aligned} \qquad (7.3.7)$$

By splitting A into the sum of its constituent symmetric and positive definite matrices G_1, G_2, G_3, G_4, G_5 and G_6, we have

$$A = G_1 + G_2 + G_3 + G_4 + G_5 + G_6 \qquad (7.3.8)$$

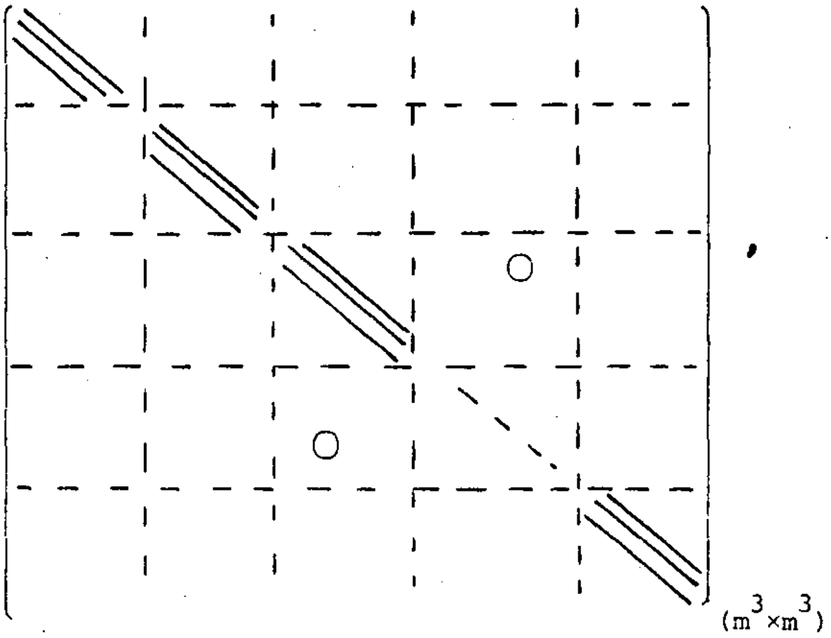
and,

A =



with

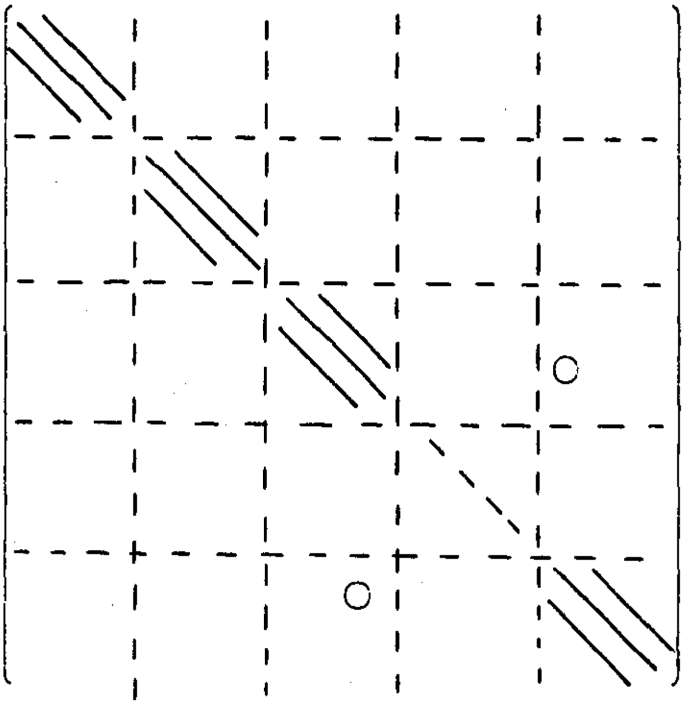
(G_1+G_2) =



,

$$\text{diag}(G_1+G_2) = \frac{1}{3}\,\text{diag}(A)$$

$(G_3+G_4) =$

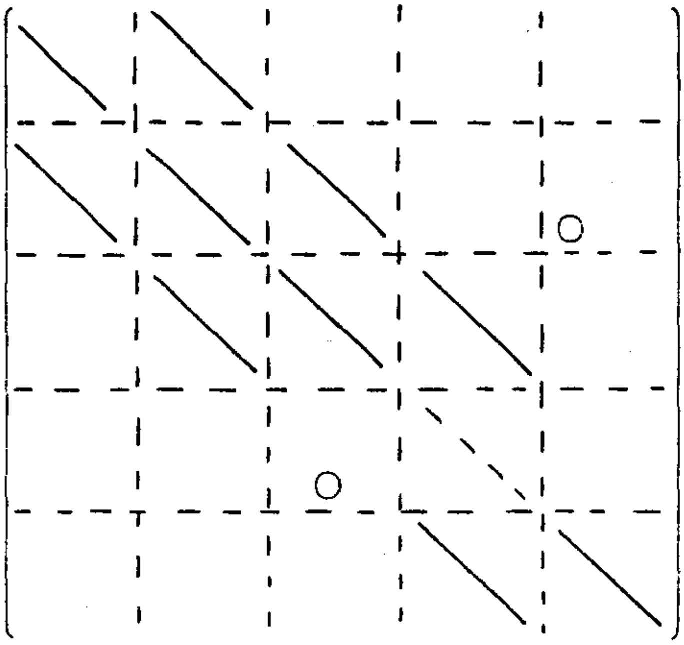


$(m^3 \times m^3)$

$$\text{diag}(G_3+G_4) = \frac{1}{3}\,\text{diag}(A)$$

and

$(G_5+G_6) =$



,

$(m^3 \times m^3)$

$$\text{diag}(G_5+G_6) = \frac{1}{3}\,\text{diag}(A).$$

As a natural extension to (7.2.11), the Douglas-Rachford formula for the AGE fractional scheme takes the form,

$$(G_1+rI)\underline{u}_{[xy]}^{(p+1/6)} = (rI-G_1-2G_2-2G_3-2G_4-2G_5-2G_6)\underline{u}_{[xy]}^{(p)} + 2\underline{f},$$

$$\left.\begin{aligned}
(G_2+rI)\underline{u}_{[xy]}^{(p+1/3)} &= G_2\underline{u}_{[xy]}^{(p)} + r\underline{u}_{[xy]}^{(p+1/6)} ,\\
(G_3+rI)\underline{u}_{[xy]}^{(p+1/2)} &= G_3\underline{u}_{[xy]}^{(p)} + r\underline{u}_{[xy]}^{(p+1/3)} ,\\
(G_4+rI)\underline{u}_{[xy]}^{(p+2/3)} &= G_4\underline{u}_{[xy]}^{(p)} + r\underline{u}_{[xy]}^{(p+1/2)} ,\\
(G_5+rI)\underline{u}_{[xy]}^{(p+5/6)} &= G_5\underline{u}_{[xy]}^{(p)} + r\underline{u}_{[xy]}^{(p+2/3)} ,\\
(G_6+rI)\underline{u}_{[xy]}^{(p+1)} &= G_6\underline{u}_{[xy]}^{(p)} + r\underline{u}_{[xy]}^{(p+5/6)} .
\end{aligned}\right\} \quad (7.3.9)$$

We now consider the above iterative formula at each of the six intermediate levels:

(i) At the first intermediate level (the (p+1/6)th iterate)

By virtue of (7.3.8), the first equation of (7.3.9) may be rewritten as

$$(G_1+rI)\underline{u}_{[xy]}^{(p+1/6)} = ((rI+G_1)-2A)\underline{u}_{[xy]}^{(p)} + 2\underline{f} . \quad (7.3.10)$$

By assuming that *m is odd*, we find that,

$$(rI+G_1) = \left(\begin{array}{c|c|c|ccc|c|c}
\bar{C}_1 & & & & & & & \\ \hline
 & \bar{\bar{C}}_2 & & & & & & \\ \hline
 & & \bar{C}_1 & & & & & \\ \hline
 & & & \ddots & & O & & \\
 & & & & \ddots & & & \\
 & & & O & & \ddots & & \\ \hline
 & & & & & & \bar{C}_2 & \\ \hline
 & & & & & & & \bar{C}_1
\end{array}\right)_{(m^3\times m^3)} \quad (7.3.10a)$$

where

$$\bar{c}_1 = \begin{pmatrix} c_1 & & & & & & \\ & c_2 & & & & & \\ & & c_1 & & & & \\ & & & c_2 & & O & \\ & & & & \ddots & & \\ & & & O & & c_2 & \\ & & & & & & c_1 \end{pmatrix}_{(m^2\times m^2)}, \tag{7.3.11}$$

$$\bar{c}_2 = \begin{pmatrix} c_2 & & & & & & \\ & c_1 & & & & & \\ & & c_2 & & & & \\ & & & c_1 & & O & \\ & & & & \ddots & & \\ & & & O & & c_1 & \\ & & & & & & c_2 \end{pmatrix}_{(m^2\times m^2)}, \tag{7.3.12}$$

$$c_1 = \begin{pmatrix} r_1 & & & & & & & \\ & r_1 & a_1 & & & & & \\ & a_1 & r_1 & & & & & \\ & & & r_1 & a_1 & & O & \\ & & & a_1 & r_1 & & & \\ & & & & & \ddots & & \\ & & & O & & & r_1 & a_1 \\ & & & & & & a_1 & r_1 \end{pmatrix}_{(m\times m)}$$

and

$$C_2 = \left(\begin{array}{cc|cc|c|cc|c} r_1 & a_1 & & & & & & \\ a_1 & r_1 & & & & & & \\ \hline & & r_1 & a_1 & & & & \\ & & a_1 & r_1 & & O & & \\ \hline & & & & \ddots & & & \\ \hline & & & & & r_1 & a_1 & \\ & & & O & & a_1 & r_1 & \\ \hline & & & & & & & r_1 \end{array}\right)_{(m\times m)} . \quad (7.3.13)$$

Therefore,

$$(rI+G_1)^{-1} = \left(\begin{array}{c|c|c|c|c|c} \bar{C}_1^{-1} & & & & & \\ \hline & \bar{C}_2^{-1} & & & & \\ \hline & & \bar{C}_1^{-1} & & O & \\ \hline & & & \ddots & & \\ \hline & & O & & \bar{C}_2^{-1} & \\ \hline & & & & & \bar{C}_1^{-1} \end{array}\right)_{(m^3\times m^3)} \quad (7.3.14)$$

where,

$$\bar{C}_1^{-1} = \left(\begin{array}{c|c|c|c|c|c} C_1^{-1} & & & & & \\ \hline & C_2^{-1} & & & & \\ \hline & & C_1^{-1} & & O & \\ \hline & & & \ddots & & \\ & & O & & & \\ \hline & & & & C_2^{-1} & \\ \hline & & & & & C_1^{-1} \end{array}\right)_{(m^2\times m^2)} \quad (7.3.15)$$

and similarly for C_2^{-1} with,

$$C_1^{-1} = \frac{1}{\Delta}\left(\begin{array}{c|cc|cc|c|cc} \frac{\Delta}{r_1} & & & & & & & \\ \hline & r_1 & -a_1 & & & & & \\ & -a_1 & r_1 & & & & & \\ \hline & & & r_1 & -a_1 & & & \\ & & & -a_1 & r_1 & & \multicolumn{2}{c}{O} \\ \hline & & & & & \ddots & & \\ \hline & & O & & & & r_1 & -a_1 \\ & & & & & & -a_1 & r_1 \end{array}\right)_{(m\times m)}$$

$$C_2^{-1} = \frac{1}{\Delta}\left(\begin{array}{cc|cc|c|cc|c} r_1 & -a_1 & & & & & & \\ -a_1 & r_1 & & & & & & \\ \hline & & r_1 & -a_1 & & & & \\ & & -a_1 & r_1 & & & & \\ \hline & & & & \ddots & \multicolumn{2}{c}{O} & \\ \hline & & & O & & r_1 & -a_1 & \\ & & & & & -a_1 & r_1 & \\ \hline & & & & & & & \frac{\Delta}{r_1} \end{array}\right)_{(m\times m)} \tag{7.3.16}$$

and $$r_1 = r+c/6, \quad \Delta = (r_1+a_1)(r_1-a_1) \ . \tag{7.3.17}$$

For simplicity we assume that the boundary value $\underline{b}=\underline{0}$.

Consequently, we have,

$$\underline{f} = B\underline{u}_{[xy]}^{[N]} + \underline{g}$$

and equation (7.3.10) becomes,

$$(G_1+rI)\underline{u}_{[xy]}^{(p+1/6)} = ((rI+G_1)-2A)\underline{u}_{[xy]}^{(p)}+2B\underline{u}_{[xy]}^{[N]}+2\underline{g}$$

or $$\underline{u}^{(p+1/6)}_{[xy]} = (G_1+rI)^{-1}[((rI+G_1)-2A)\underline{u}^{(p)}_{[xy]}+2B\underline{u}^{[N]}_{[xy]}+2\underline{g}] . \qquad (7.3.18)$$

If we write

$$D_1 = \bar{C}_1-2A_1, \quad D_2 = \bar{C}_2-2A_1, \quad E_1 = -2A_2, \; F_1 = 2B_1 \text{ and } F_2 = 2B_2$$

then $((rI+G_1)-2A)\underline{u}^{(p)}_{[xy]}+2B\underline{u}^{[N]}_{[xy]}+2\underline{g}$ becomes,

$$\left(\begin{array}{c|c|c|c|c|c}
D_1 & E_1 & & & & \\ \hline
E_1 & D_2 & E_1 & & & \\ \hline
 & E_1 & D_1 & E_1 & & \\ \hline
 & & \ddots & \ddots & \ddots \quad O & \\
 & & O \quad \ddots & \ddots & \ddots & \\ \hline
 & & & E_1 & D_2 & E_1 \\ \hline
 & & & & E_1 & D_1
\end{array}\right)
\left(\begin{array}{c}
\underline{u}^{(p)}_{1[xy]} \\ \hline
\underline{u}^{(p)}_{2[xy]} \\ \hline
\underline{u}^{(p)}_{3[xy]} \\ \hline
\vdots \\ \hline
\underline{u}^{(p)}_{m-1[xy]} \\ \hline
\underline{u}^{(p)}_{m[xy]}
\end{array}\right) +$$

$$\left(\begin{array}{c|c|c|c|c|c}
F_1 & F_2 & & & & \\ \hline
F_2 & F_1 & F_2 & & & \\ \hline
 & F_2 & F_1 & F_2 & & \\ \hline
 & & \ddots & \ddots & \ddots \quad O & \\
 & & O \quad \ddots & \ddots & \ddots & \\ \hline
 & & & F_2 & F_1 & F_2 \\ \hline
 & & & & F_2 & F_1
\end{array}\right)
\left(\begin{array}{c}
\underline{u}^{[N]}_{1[xy]} \\ \hline
\underline{u}^{[N]}_{2[xy]} \\ \hline
\underline{u}^{[N]}_{3[xy]} \\ \hline
\vdots \\ \hline
\underline{u}^{[N]}_{m-1[xy]} \\ \hline
\underline{u}^{[N]}_{m[xy]}
\end{array}\right) + 2
\left(\begin{array}{c}
\underline{g}_{1[xy]} \\ \hline
\underline{g}_{2[xy]} \\ \hline
\vdots \\ \hline
\underline{g}_{m[xy]}
\end{array}\right) . \qquad (7.3.19)$$

Hence, using (7.3.18) we obtain the following set of equations for the computation of the AGE scheme at the $(p+1/6)^{th}$ iterate:

$$\underline{u}^{(p+1/6)}_{1[xy]} = \bar{C}_1^{-1}(D_1\underline{u}^{(p)}_{1[xy]}+E_1\underline{u}^{(p)}_{2[xy]}+F_1\underline{u}^{[N]}_{1[xy]}+F_2\underline{u}^{[N]}_{2[xy]}+2\underline{g}_{1[xy]}) \qquad (7.3.20a)$$

$$\underline{u}^{(p+1/6)}_{k[xy]} = \bar{C}_2^{-1}(E_1(\underline{u}^{(p)}_{k-1[xy]}+\underline{u}^{(p)}_{k+1[xy]})+D_2\underline{u}^{(p)}_{k[xy]}+F_2(\underline{u}^{[N]}_{k-1[xy]}+$$
$$\underline{u}^{[N]}_{k+1[xy]})+F_1\underline{u}^{[N]}_{k[xy]}+2\underline{g}_{k[xy]}), \ k=2,4,\ldots,m-1 \qquad (7.3.20b)$$

$$\underline{u}^{(p+1/6)}_{k[xy]} = \bar{C}_1^{-1}(E_1(\underline{u}^{(p)}_{k-1[xy]}+\underline{u}^{(p)}_{k+1[xy]})+D_1\underline{u}^{(p)}_{k[xy]}+F_2(\underline{u}^{[N]}_{k-1[xy]}+$$
$$\underline{u}^{[N]}_{k+1[xy]})+F_1\underline{u}^{[N]}_{k[xy]}+2\underline{g}_{k[xy]}), \ k=3,5,\ldots,m-2 \qquad (7.3.20c)$$

$$\underline{u}^{(p+1/6)}_{m[xy]} = \bar{C}_1^{-1}(E_1\underline{u}^{(p)}_{m-1[xy]}+D_1\underline{u}^{(p)}_{m[xy]}+F_2\underline{u}^{[N]}_{m-1[xy]}+F_1\underline{u}^{[N]}_{m[xy]}+$$
$$2\underline{g}_{m[xy]}) \ . \qquad (7.3.20d)$$

Let $r_2=r_1-2c$. When the above equations are written component-wise, we have,

(a) for (7.3.20a)

(1) $$u^{(p+1/6)}_{1,j,1} = [r_2u^{(p)}_{1,j,1}-2a_1(u^{(p)}_{2,j,1}+u^{(p)}_{1,j,2}+u^{(p)}_{1,j+1,1})+2du^{[N]}_{1,j,1}+2e_1$$
$$(u^{[N]}_{2,j,1}+u^{[N]}_{1,j,2}+u^{[N]}_{1,j+1,1})+2g_{1,j,1}]/r_1, \text{ for } j=1,3,\ldots,m$$

$$\left.\begin{aligned} u^{(p+1/6)}_{i,j,1} &= (r_1a_{i,j,1}-a_1b_{i,j,1})/\Delta \\ u^{(p+1/6)}_{i+1,j,1} &= (-a_1a_{i,j,1}+r_1b_{i,j,1})/\Delta \end{aligned}\right\} j=1,3,\ldots,m;\ i=2,4,\ldots,m-1,$$

(7.3.21a)

with $$u^{(p)}_{1,m+1,1} = u^{(p)}_{1,m-1,1}, \quad u^{[N]}_{1,m+1,1} = u^{[N]}_{1,m-1,1} \ ;$$

$$u^{(p)}_{i,m+1,1} = u^{(p)}_{i,m-1,1}, \quad u^{[N]}_{i,m+1,1} = u^{[N]}_{i,m-1,1} \ ,$$

and
$$a_{i,j,1} = r_2u^{(p)}_{i,j,1}-2a_1(u^{(p)}_{i,j,2}+u^{(p)}_{i-1,j,1}+u^{(p)}_{i,j+1,1})-a_1u^{(p)}_{i+1,j,1}+$$
$$2e_1(u^{[N]}_{i,j,2}+u^{[N]}_{i-1,j,1}+u^{[N]}_{i+1,j,1}+u^{[N]}_{i,j+1,1})+2du^{[N]}_{i,j,1}+2g_{i,j,1},$$

$$b_{i,j,1} = r_2u^{(p)}_{i+1,j,1}-2a_1(u^{[N]}_{i+1,j+1,1}+u^{[N]}_{i+1,j,2})-a_1u^{(p)}_{i,j,1}+2e_1$$
$$(u^{[N]}_{i,j,1}+u^{[N]}_{i+1,j+1,1}+u^{[N]}_{i+1,j,2})+2du^{[N]}_{i+1,j,1}+2g_{i+1,j,1}.$$

(2) $$\left.\begin{aligned} u^{(p+1/6)}_{i,j,1} &= (r_1\bar{a}_{i,j,1}-a_1\bar{b}_{i,j,1})/\Delta \\ u^{(p+1/6)}_{i+1,j,1} &= (-a_1\bar{a}_{i,j,1}+r_1\bar{b}_{i,j,1})/\Delta \end{aligned}\right\} j=2,4,\ldots,m-1;\ i=1,3,\ldots,m-2,$$

$$\begin{aligned} u^{(p+1/6)}_{m,j,1} = {} & [r_2u^{(p)}_{m,j,1}-2a_1(u^{(p)}_{m,j-1,1}+u^{(p)}_{m-1,j,1}+u^{(p)}_{m,j+1,1}+u^{(p)}_{m,j,2}) \\ & +2e_1(u^{[N]}_{m,j-1,1}+u^{[N]}_{m-1,j,1}+u^{[N]}_{m,j+1,1}+u^{[N]}_{m,j,2})+2du^{[N]}_{m,j,1}+ \\ & 2g_{m,j,1}]/r_1,\ \text{for } j=2,4,\ldots,m-1\ , \end{aligned} \qquad (7.3.21b)$$

where,

$$\begin{aligned} \bar{a}_{i,j,1} = {} & r_2u^{(p)}_{i,j,1}-2a_1(u^{(p)}_{i,j-1,1}+u^{(p)}_{i,j+1,1}+u^{(p)}_{i,j,2})-a_1u_{i+1,j,1} \\ & +2e_1(u^{[N]}_{i,j-1,1}+u^{[N]}_{i+1,j,1}+u^{[N]}_{i,j+1,1}+u^{[N]}_{i,j,2})+2du^{[N]}_{i,j,1} \\ & +2g_{i,j,1} \end{aligned}$$

and $$\begin{aligned} \bar{b}_{i,j,1} = {} & r_2u^{(p)}_{i+1,j,1}-2a_1(u^{(p)}_{i+1,j-1,1}+u^{(p)}_{i+2,j,1}+u^{(p)}_{i+1,j+1,1}+ \\ & u^{(p)}_{i+1,j,2})-a_1u_{i,j,1}+2e_1(u^{[N]}_{i+1,j-1,1}+u^{[N]}_{i,j,1}+u^{[N]}_{i+2,j,1}+ \\ & u^{[N]}_{i+1,j+1,1}+u^{[N]}_{i+1,j,2})+2du^{[N]}_{i+1,j,1}+2g_{i+1,j,1}\ . \end{aligned}$$

(b) for (7.3.20b)

(1) $$\left.\begin{aligned} u^{(p+1/6)}_{i,j,k} &= (r_1c_{i,j,k}-a_1d_{i,j,k})/\Delta \\ u^{(p+1/6)}_{i+1,j,k} &= (-a_1c_{i,j,k}+r_1d_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} & k=2,4,\ldots,m-1;\ j=1,3,\ldots,m; \\ & i=1,3,\ldots,m-2; \end{aligned}$$

$$\begin{aligned} u^{(p+1/6)}_{m,j,k} = {} & [r_2u^{(p)}_{m,j,k}-2a_1(u^{(p)}_{m,j,k-1}+u^{(p)}_{m,j,k+1}+u^{(p)}_{m-1,j,k}+u^{(p)}_{m,j+1,k}) \\ & +2e_1(u^{[N]}_{m,j,k-1}+u^{[N]}_{m,j,k+1}+u^{[N]}_{m-1,j,k}+u^{[N]}_{m,j+1,k})+ \\ & 2du^{[N]}_{m,j,k}+2g_{m,j,k}]/r_1,\ \text{for } k=2,4,\ldots,m-1;\ j=1,3,\ldots,m, \end{aligned}$$

with (7.3.22a)

$$u^{(p)}_{m,m+1,k} = u^{(p)}_{m,m-1,k}\ ,\quad u^{[N]}_{m,m+1,k} = u^{[N]}_{m,m-1,k}\ ;$$

$$u^{(p)}_{i,m+1,k} = u^{(p)}_{i,m-1,k}\ ,\quad u^{[N]}_{i,m+1,k} = u^{[N]}_{i,m-1,k}\ ;$$

$$c_{i,j,k} = r_2u^{(p)}_{i,j,k}-2a_1(u^{(p)}_{i,j,k-1}+u^{(p)}_{i,j,k+1}+u^{(p)}_{i,j+1,k})-a_1u^{(p)}_{i+1,j,k}$$
$$+2e_1(u^{[N]}_{i,j,k-1}+u^{[N]}_{i,j,k+1}+u^{[N]}_{i+1,j,k}+u^{[N]}_{i,j+1,k})+2du^{[N]}_{i,j,k}$$
$$+2g_{i,j,k}$$

and

$$d_{i,j,k} = r_2u^{(p)}_{i+1,j,k}-2a_1(u^{(p)}_{i+1,j,k-1}+u^{(p)}_{i+1,j,k+1}+u^{(p)}_{i+2,j,k}+u^{(p)}_{i+1,j+1,k})$$
$$-a_1u^{(p)}_{i,j,k}+2e_1(u^{[N]}_{i+1,j,k-1}+u^{[N]}_{i+1,j,k+1}+u^{[N]}_{i,j,k}+u^{[N]}_{i+2,j,k}+$$
$$u^{[N]}_{i+1,j+1,k})+2du^{[N]}_{i+1,j,k}+2g_{i+1,j,k}\ .$$

(2) $$u^{(p+1/6)}_{1,j,k} = [r_2u^{(p)}_{1,j,k}-2a_1(u^{(p)}_{1,j,k-1}+u^{(p)}_{1,j,k+1}+u^{(p)}_{1,j-1,k}+u^{(p)}_{2,j,k}$$
$$+u^{(p)}_{1,j+1,k})+2e_1(u^{[N]}_{1,j,k-1}+u^{[N]}_{1,j,k+1}+u^{[N]}_{1,j-1,k}+u^{[N]}_{2,j,k}+$$
$$u^{[N]}_{1,j+1,k})+2du^{[N]}_{1,j,k}+2g_{1,j,k}]/r_1$$

$$\left.\begin{aligned} u^{(p+1/6)}_{i,j,k} &= (r_1\bar{c}_{i,j,k}-a_1\bar{d}_{i,j,k})/\Delta \\ u^{(p+1/6)}_{i+1,j,k} &= (-a_1\bar{c}_{i,j,k}+r_1\bar{d}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &k=2,4,\ldots,m-1;\ j=2,4,\ldots,m-1; \\ &i=2,4,\ldots,m-1, \end{aligned}$$

(7.3.22b)

where

$$\bar{c}_{i,j,k} = r_2u^{(p)}_{i,j,k}-2a_1(u^{(p)}_{i,j,k-1}+u^{(p)}_{i,j,k+1}+u^{(p)}_{i,j-1,k}+u^{(p)}_{i-1,j,k}+$$
$$u^{(p)}_{i,j+1,k})-a_1u^{(p)}_{i+1,j,k}+2e_1(u^{[N]}_{i,j,k-1}+u^{[N]}_{i,j,k+1}+u^{[N]}_{i,j-1,k}$$
$$+u^{[N]}_{i-1,j,k}+u^{[N]}_{i+1,j,k}+u^{[N]}_{i,j+1,k})+2du^{[N]}_{i,j,k}+2g_{i,j,k}$$

and

$$\bar{d}_{i,j,k} = r_2u^{(p)}_{i+1,j,k}-2a_1(u^{(p)}_{i+1,j,k-1}+u^{(p)}_{i+1,j,k+1}+u^{(p)}_{i+1,j-1,k}+$$
$$u^{(p)}_{i+1,j+1,k})-a_1u^{(p)}_{i,j,k}+2e_1(u^{[N]}_{i+1,j,k-1}+u^{[N]}_{i+1,j,k+1}+u^{[N]}_{i+1,j-1,k}$$
$$u^{[N]}_{i,j,k}+u^{[N]}_{i+1,j+1,k})+2du^{[N]}_{i+1,j,k}+2g_{i+1,j,k}\ .$$

(c) For (7.3.20c)

(1) $$u^{(p+1/6)}_{1,j,k} = [r_2u^{(p)}_{1,j,k}-2a_1(u^{(p)}_{1,j,k-1}+u^{(p)}_{1,j,k+1}+u^{(p)}_{2,j,k}+u^{(p)}_{1,j+1,k})$$
$$+2e_1(u^{[N]}_{1,j,k-1}+u^{[N]}_{1,j,k+1}+u^{[N]}_{2,j,k}+u^{[N]}_{1,j+1,k})+2du^{[N]}_{1,j,k}$$
$$+2g_{1,j,k}]/r_1\ ,$$

$$\left.\begin{aligned} u^{(p+1/6)}_{i,j,k} &= (r_1 e_{i,j,k} - a_1 f_{i,j,k})/\Delta \\ u^{(p+1/6)}_{i+1,j,k} &= (-a_1 e_{i,j,k} + r_1 f_{i,j,k})/\Delta \end{aligned}\right\} \begin{array}{l} k=3,5,\ldots,m-2;\ j=1,3,\ldots,m; \\ i=2,4,\ldots,m-1, \end{array}$$

(7.3.23a)

with

$$u^{(p)}_{1,m+1,k} = u^{(p)}_{1,m-1,k}, \quad u^{[N]}_{1,m+1,k} = u^{[N]}_{1,m-1,k};$$

$$u^{(p)}_{i,m+1,k} = u^{(p)}_{i,m-1,k}, \quad u^{[N]}_{i,m+1,k} = u^{[N]}_{i,m-1,k};$$

$$\begin{aligned} e_{i,j,k} = {} & r_2 u^{(p)}_{i,j,k} - 2a_1(u^{(p)}_{i,j,k-1} + u^{(p)}_{i,j,k+1} + u^{(p)}_{i-1,j,k} + u^{(p)}_{i,j+1,k}) - a_1 u^{(p)}_{i+1,j,k} \\ & + 2e_1(u^{[N]}_{i,j,k-1} + u^{[N]}_{i,j,k+1} + u^{[N]}_{i-1,j,k} + u^{[N]}_{i+1,j,k} + u^{[N]}_{i,j+1,k}) + 2du^{[N]}_{i,j,k} \\ & + 2g_{i,j,k} \end{aligned}$$

and

$$\begin{aligned} f_{i,j,k} = {} & r_2 u^{(p)}_{i+1,j,k} - 2a_1(u^{(p)}_{i+1,j,k-1} + u^{(p)}_{i+1,j,k+1} + u^{(p)}_{i+1,j+1,k}) - a_1 u^{(p)}_{i,j,k} \\ & + 2e_1(u^{[N]}_{i+1,j,k-1} + u^{[N]}_{i+1,j,k+1} + u^{[N]}_{i,j,k} + u^{[N]}_{i+1,j+1,k}) + 2du^{[N]}_{i+1,j,k} + \\ & 2g_{i+1,j,k}, \end{aligned}$$

(2)

$$\begin{aligned} u^{(p+1/6)}_{m,j,k} = {} & [r_2 u^{(p)}_{m,j,k} - 2a_1(u^{(p)}_{m,j,k-1} + u^{(p)}_{m,j,k+1} + u^{(p)}_{m,j-1,k} + u^{(p)}_{m-1,j,k} + \\ & u^{(p)}_{m,j+1,k}) + 2e_1(u^{[N]}_{m,j,k-1} + u^{[N]}_{m,j,k+1} + u^{[N]}_{m,j-1,k} + u^{[N]}_{m-1,j,k} + \\ & u^{[N]}_{m,j+1,k}) + 2du^{[N]}_{m,j,k} + 2g_{m,j,k}]/r_1, \end{aligned}$$

$$\left.\begin{aligned} u^{(p+1/6)}_{i,j,k} &= (r_1 \bar{e}_{i,j,k} - a_1 \bar{f}_{i,j,k})/\Delta \\ u^{(p+1/6)}_{i+1,j,k} &= (-a_1 \bar{e}_{i,j,k} + r_1 \bar{f}_{i,j,k})/\Delta \end{aligned}\right\} \begin{array}{l} k=3,5,\ldots,m-2;\ j=2,4,\ldots,m-1; \\ i=1,3,\ldots,m-2, \end{array}$$

where,

(7.3.23b)

$$\begin{aligned} \bar{e}_{i,j,k} = {} & r_2 u^{(p)}_{i,j,k} - 2a_1(u^{(p)}_{i,j,k-1} + u^{(p)}_{i,j,k+1} + u^{(p)}_{i,j-1,k} + u^{(p)}_{i,j+1,k}) - a_1 u^{(p)}_{i+1,j,k} \\ & + 2e_1(u^{[N]}_{i,j,k-1} + u^{[N]}_{i,j,k+1} + u^{[N]}_{i,j-1,k} + u^{[N]}_{i+1,j,k} + u^{[N]}_{i,j+1,k}) \\ & + 2du^{[N]}_{i,j,k} + 2g_{i,j,k} \end{aligned}$$

and
$$\begin{aligned}\bar{f}_{i,j,k} = {} & r_2u^{(p)}_{i+1,j,k}-2a_1(u^{(p)}_{i+1,j,k-1}+u^{(p)}_{i+1,j,k+1}+u^{(p)}_{i+1,j-1,k} \\ & +u^{(p)}_{i+2,j,k}+u^{(p)}_{i+1,j+1,k})-a_1u^{(p)}_{i,j,k}+2e_1(u^{[N]}_{i+1,j,k-1} \\ & +u^{[N]}_{i+1,j,k+1}+u^{[N]}_{i+1,j-1,k}+u^{[N]}_{i,j,k}+u^{[N]}_{i+2,j,k}+u^{[N]}_{i+1,j+1,k}) \\ & +2du^{[N]}_{i+1,j,k}+2g_{i+1,j,k} .\end{aligned}$$

(d) for (7.3.20d)

(1)
$$\begin{aligned}u^{(p+1/6)}_{1,j,m} = {} & [r_2u^{(p)}_{1,j,m}-2a_1(u^{(p)}_{1,j,m-1}+u^{(p)}_{2,j,m}+u^{(p)}_{1,j+1,m})+2e_1 \\ & (u^{[N]}_{1,j,m-1}+u^{[N]}_{2,j,m}+u^{[N]}_{1,j+1,m})+2du^{[N]}_{1,j,m}+2g_{1,j,m}]/r_1, \\ & \qquad j=1,3,\ldots,m,\end{aligned}$$

$$\left.\begin{aligned}u^{(p+1/6)}_{i,j,m} &= (r_1p_{i,j,m}-a_1q_{i,j,m})/\Delta \\ u^{(p+1/6)}_{i+1,j,m} &= (-a_1p_{i,j,m}+r_1q_{i,j,m})/\Delta\end{aligned}\right\},j=1,3,\ldots,m;\ i=2,4,\ldots,m-1, \tag{7.3.24a}$$

with
$$\begin{aligned}u^{(p)}_{1,m+1,m} &= u^{(p)}_{1,m-1,m}\ , \quad u^{[N]}_{1,m+1,m} = u^{[N]}_{1,m-1,m}\ ; \\ u^{(p)}_{i,m+1,m} &= u^{(p)}_{i,m-1,m}\ , \quad u^{[N]}_{1,m+1,m} = u^{[N]}_{1,m-1,m}\ ;\end{aligned}$$

$$\begin{aligned}p_{i,j,m} = {} & r_2u^{(p)}_{i,j,m}-2a_1(u^{(p)}_{i,j,m-1}+u^{(p)}_{i-1,j,m}+u^{(p)}_{i,j+1,m})-a_1u^{(p)}_{i+1,j,m} \\ & +2e_1(u^{[N]}_{i,j,m-1}+u^{[N]}_{i+1,j,m}+u^{[N]}_{i-1,j,m}+u^{[N]}_{i,j+1,m})+2du^{[N]}_{i,j,m}+2g_{i,j,m}\end{aligned}$$

and
$$\begin{aligned}q_{i,j,m} = {} & r_2u^{(p)}_{i+1,j,m}-2a_1(u^{(p)}_{i+1,j,m-1}+u^{(p)}_{i+1,j+1,m})+2e_1(u^{[N]}_{i+1,j,m-1}+ \\ & u^{[N]}_{i,j,m}+u^{[N]}_{i+1,j+1,m})+2du^{[N]}_{i+1,j,m}+2g_{i+1,j,m} .\end{aligned}$$

(2)
$$\begin{aligned}u^{(p+1/6)}_{m,j,m} = {} & [r_2u^{(p)}_{m,j,m}-2a_1(u^{(p)}_{m,j,m-1}+u^{(p)}_{m,j-1,m}+u^{(p)}_{m-1,j,m}+u^{(p)}_{m,j+1,m}) \\ & +2e_1(u^{[N]}_{m,j,m-1}+u^{[N]}_{m,j-1,m}+u^{[N]}_{m-1,j,m}+u^{[N]}_{m,j+1,m})+2du^{[N]}_{m,j,m} \\ & +2g_{m,j,m}]/r_1 ,\end{aligned}$$

$$\left.\begin{aligned} u^{(p+1/6)}_{i,j,m} &= (r_1\bar{p}_{i,j,m}-a_1\bar{q}_{i,j,m})/\Delta \\ u^{(p+1/6)}_{i+1,j,m} &= (-a_1\bar{p}_{i,j,m}+r_1\bar{q}_{i,j,m})/\Delta \end{aligned}\right\} \begin{aligned} &j=2,4,\ldots,m-1; \\ &i=1,3,\ldots,m-2, \end{aligned}$$

where (7.3.24b)

$$\begin{aligned} \bar{p}_{i,j,m} = {} & r_2u^{(p)}_{i,j,m}-2a_1(u^{(p)}_{i,j,m-1}+u^{(p)}_{i,j-1,m}+u^{(p)}_{i,j+1,m})-a_1u^{(p)}_{i+1,j,m} \\ & +2e_1(u^{[N]}_{i,j,m-1}+u^{[N]}_{i,j-1,m}+u^{[N]}_{i+1,j,m}+u^{[N]}_{i,j+1,m})+2du^{[N]}_{i,j,m} \\ & +2g_{i,j,m} \end{aligned}$$

and

$$\begin{aligned} \bar{q}_{i,j,m} = {} & r_2u^{(p)}_{i+1,j,m}-2a_1(u^{(p)}_{i+1,j,m-1}+u^{(p)}_{i+1,j-1,m}+u^{(p)}_{i+2,j,m}+u^{(p)}_{i+1,j+1,m}) \\ & -a_1u^{(p)}_{i,j,m}+2e_1(u^{[N]}_{i+1,j,m-1}+u^{[N]}_{i+1,j-1,m}+u^{[N]}_{i,j,m}+u^{[N]}_{i+2,j,m} \\ & +u^{[N]}_{i+1,j+1,m})+2du^{[N]}_{i+1,j,m}+2q_{i+1,j,m}\ . \end{aligned}$$

(ii) At the second intermediate level (the $(p+1/3)^{th}$ iterate)

From the second equation of (7.3.9) we have,

$$(G_2+rI)\underline{u}^{(p+1/3)}_{[xy]} = G_2u^{(p)}_{[xy]} + r\underline{u}^{(p+1/6)}_{[xy]}$$

which gives $$\underline{u}^{(p+1/3)}_{[xy]} = (G_2+rI)^{-1}\{G_2\underline{u}^{(p)}_{[xy]}+r\underline{u}^{(p+1/6)}_{[xy]}\}. \qquad (7.3.25)$$

Now,

$$(rI+G_2) = \begin{pmatrix} \bar{C}_2 & & & & & & \\ & \bar{C}_1 & & & & & \\ & & \bar{C}_2 & & & O & \\ & & & \ddots & & & \\ & & O & & \bar{C}_2 & & \\ & & & & & \bar{C}_1 & \\ & & & & & & \bar{C}_2 \end{pmatrix}_{(m^3\times m^3)} . \qquad (7.3.26)$$

By letting $\hat{\bar{C}}_2 \equiv \bar{C}_2$ with diagonal elements r_1 be replaced by c/6 and $\hat{\bar{C}}_1 \equiv \bar{C}_1$ with diagonal elements r_1 be replaced by c/6 we find from (7.3.25) that

$$\underline{u}_{k[xy]}^{(p+1/3)} = \begin{cases} \bar{C}_2^{-1}\ \{\hat{\bar{C}}_2 \underline{u}_{k[xy]}^{(p)} + r\underline{u}_{k[xy]}^{(p+1/6)}\} \text{ for } k=1,3,\ldots,m & (7.3.27a) \\ \bar{C}_1^{-1}\ \{\hat{\bar{C}}_1 \underline{u}_{k[xy]}^{(p)} + r\underline{u}_{k[xy]}^{(p+1/6)}\} \text{ for } k=2,4,\ldots,m-1 & (7.3.27b) \end{cases}$$

The computation of the AGE algorithm is carried out as follows:

(1) for (7.3.27a)

$$u_{m,j,k}^{(p+1/3)} = (\frac{c}{6}\,u_{m,j,k}^{(p)} + ru_{m,j,k}^{(p+1/6)})/r_1\ , \text{ for } k=1,3,\ldots,m;\ j=1,3,\ldots,m,$$

$$\left.\begin{aligned} u_{i,j,k}^{(p+1/3)} &= (r_1a_{i,j,k} - a_1\bar{a}_{i,j,k})/\Delta \\ u_{i+1,j,k}^{(p+1/3)} &= (-a_1a_{i,j,k} + r_1\bar{a}_{i,j,k})/\Delta \end{aligned}\right\}\ , k=1,3,\ldots,m;\ j=1,3,\ldots,m, \qquad (7.3.28a)$$

$$u_{1,j,k}^{(p+1/3)} = (\frac{c}{6}u_{1,j,k}^{(p)} + ru_{1,j,k}^{(p+1/6)})/r_1, \text{ for } k=1,3,\ldots,m;\ j=2,4,\ldots,m-1,$$

$$\left.\begin{aligned} u_{i,j,k}^{(p+1/3)} &= (r_1a_{i,j,k} - a_1\bar{a}_{i,j,k})/\Delta \\ u_{i+1,j,k}^{(p+1/3)} &= (-a_1a_{i,j,k} + r_1\bar{a}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &k=1,3,\ldots,m;\ j=2,4,\ldots,m-1, \\ &i=2,4,\ldots,m-1. \end{aligned}$$

and

(2) for (7.3.27b)

$$u_{1,j,k}^{(p+1/3)} = (\frac{c}{6}u_{1,j,k}^{(p)} + ru_{1,j,k}^{(p+1/6)})/r_1 \text{ for } k=2,4,\ldots,m-1;\ j=1,3,\ldots,m,$$

$$\left.\begin{aligned} u_{i,j,k}^{(p+1/3)} &= (r_1a_{i,j,k} - a_1\bar{a}_{i,j,k})/\Delta \\ u_{i+1,j,k}^{(p+1/3)} &= (-a_1a_{i,j,k} + r_1\bar{a}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &k=2,4,\ldots,m-1;\ j=2,4,\ldots,m-1, \\ &i=1,3,\ldots,m-2, \end{aligned} \qquad (7.3.28b)$$

$$u_{m,j,k}^{(p+1/3)} = (\frac{c}{6}\,u_{m,j,k}^{(p)} + ru_{m,j,k}^{(p+1/6)})/r_1 \text{ for } k=2,4,\ldots,m-1;\ j=2,4,\ldots,m-1,$$

$$\left.\begin{aligned} u_{i,j,k}^{(p+1/3)} &= (r_1a_{i,j,k} - a_1\bar{a}_{i,j,k})/\Delta \\ u_{i+1,j,k}^{(p+1/3)} &= (-a_1a_{i,j,k} + r_1\bar{a}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &k=2,4,\ldots,m-1;\ j=2,4,\ldots,m-1, \\ &i=1,3,\ldots,m-2, \end{aligned}$$

where $$a_{i,j,k} = \frac{c}{6} u_{i,j,k}^{(p)} + a_1 u_{i+1,j,k}^{(p)} + ru_{i,j,k}^{(p+1/6)}$$

and $$\bar{a}_{i,j,k} = a_1 u_{i,j,k}^{(p)} + \frac{c}{6} u_{i+1,j,k}^{(p)} + ru_{i+1,j,k}^{(p+1/6)} .$$

If we take our approximations as sweeps parallel to the yz-plane the u values are then evaluated at points lying on planes which are parallel to the yz-plane and on each of these planes, the points are *reordered row-wise* (parallel to the y-axis, as in Figure 7.3.2), such that,

$$(G_3+G_4)\underline{u}_{[xy]} = (G_1+G_2)\underline{u}_{[yz]} \tag{7.3.29}$$

where,

$$\underline{u}_{[yz]} = (\underline{u}_{1[yz]},\underline{u}_{2[yz]},\ldots,\underline{u}_{m[yz]})^T$$

with $$\underline{u}_{i[yz]} = (u_{i,1,1},u_{i,2,1},\ldots,u_{i,m,1},u_{i,1,2},u_{i,2,2},\ldots,$$
$$u_{i,m,2},\ldots,u_{i,m,m})^T , \text{ for } i=1,2,\ldots,m.$$

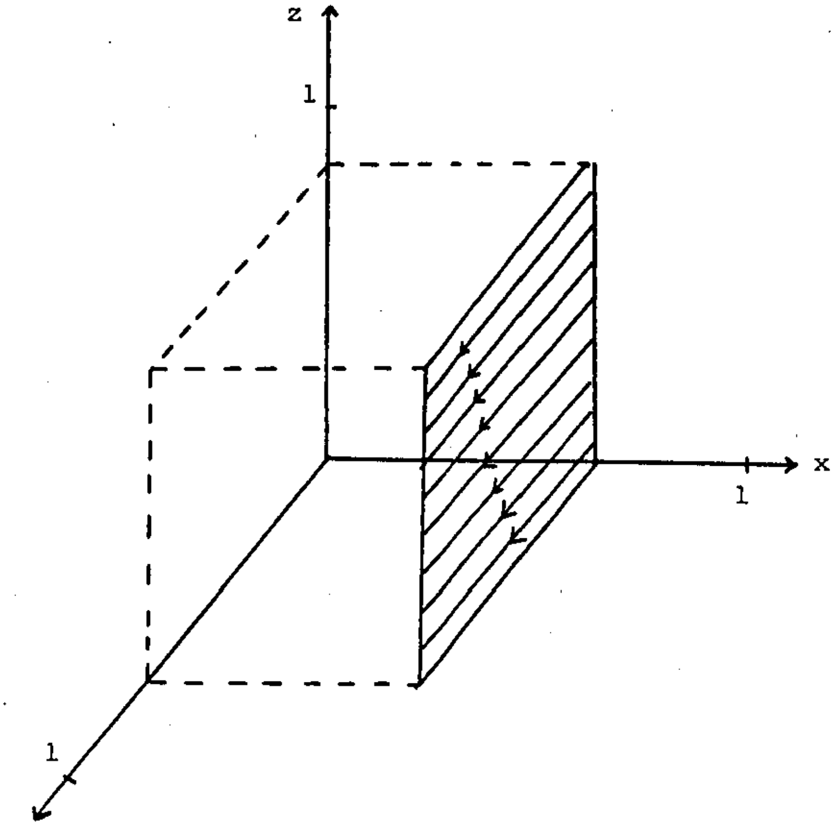


FIGURE 7.3.2

With the above reordering of points we are able to derive the computational formulae for the third and fourth intermediate levels as follows:

(iii) At the third intermediate level (the $(p+\frac{1}{2})^{th}$ *iterate)*

The third equation of (7.3.9) is transformed to

$$(G_1+rI)\underline{u}^{(p+\frac{1}{2})}_{[yz]} = G_1\underline{u}^{(p)}_{[yz]} + r\underline{u}^{(p+1/3)}_{[yz]}$$

or
$$\underline{u}^{(p+\frac{1}{2})}_{[yz]} = (G_1+rI)^{-1}[G_1\underline{u}^{(p)}_{[yz]} + r\underline{u}^{(p+1/3)}_{[yz]}] \qquad (7.3.30)$$

which leads to,

$$\underline{u}^{(p+\frac{1}{2})}_{i[yz]} = \begin{cases} \bar{C}_1^{-1}[\hat{\bar{C}}_1\underline{u}^{(p)}_{i[yz]} + r\underline{u}^{(p+1/3)}_{i[yz]}] \text{ for } i=1,3,\ldots,m & (7.3.31a) \\ \bar{C}_2^{-1}[\hat{\bar{C}}_2\underline{u}^{(p)}_{i[yz]} + r\underline{u}^{(p+1/3)}_{i[yz]}] \text{ for } i=2,4,\ldots,m-1. & (7.3.31b) \end{cases}$$

Hence, we find that,

(a) for (7.3.31a)

$$u^{(p+\frac{1}{2})}_{i,1,k} = (\frac{c}{6}u^{(p)}_{i,1,k} + ru^{(p+1/3)}_{i,1,k})/r_1, \text{ for } i=1,3,\ldots,m;\ k=1,3,\ldots,m$$

$$\left.\begin{aligned} u^{(p+\frac{1}{2})}_{i,j,k} &= (r_1b_{i,j,k} - a_1\bar{b}_{i,j,k})/\Delta \\ u^{(p+\frac{1}{2})}_{i,j+1,k} &= (-a_1b_{i,j,k} + r_1\bar{b}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &i=1,3,\ldots,m;\ k=1,3,\ldots,m; \\ &j=2,4,\ldots,m-1. \end{aligned} \qquad (7.3.32a)$$

$$u^{(p+\frac{1}{2})}_{i,m,k} = (\frac{c}{6}u^{(p)}_{i,m,k} + ru^{(p+1/3)}_{i,m,k})/r_1 \text{ for } i=1,3,\ldots,m;\ k=2,4,\ldots,m-1,$$

$$\left.\begin{aligned} u^{(p+\frac{1}{2})}_{i,j,k} &= (r_1b_{i,j,k} - a_1\bar{b}_{i,j,k})/\Delta \\ u^{(p+\frac{1}{2})}_{i,j+1,k} &= (-a_1b_{i,j,k} + r_1\bar{b}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &i=1,3,\ldots,m;\ k=2,4,\ldots,m-1; \\ &j=1,3,\ldots,m-2. \end{aligned}$$

and

(b) for (7.3.31b)

$$u^{(p+\frac{1}{2})}_{i,m,k} = (\frac{c}{6}u^{(p)}_{i,m,k} + ru^{(p+1/3)}_{i,m,k})/r_1, \text{ for } i=2,4,\ldots,m-1;\ k=1,3,\ldots,m,$$

$$\left.\begin{aligned} u^{(p+\frac{1}{2})}_{i,j,k} &= (r_1 b_{i,j,k}-a_1\bar{b}_{i,j,k})/\Delta \\ u^{(p+\frac{1}{2})}_{i,j+1,k} &= (-a_1 b_{i,j,k}+r_1\bar{b}_{i,j,k})/\Delta \end{aligned}\right\} \begin{array}{l} i=2,4,\ldots,m-1;\ k=1,3,\ldots,m; \\ j=1,3,\ldots,m-2, \end{array}$$

(7.3.32b)

$$u^{(p+\frac{1}{2})}_{i,1,k} = (\tfrac{c}{6}u^{(p)}_{i,1,k}+ru^{(p+1/3)}_{i,1,k})/r_1\ ,\ \text{for } i=2,4,\ldots,m-1;\ k=2,4,\ldots,m-1,$$

$$\left.\begin{aligned} u^{(p+\frac{1}{2})}_{i,j,k} &= (r_1 b_{i,j,k}-a_1\bar{b}_{i,j,k})/\Delta \\ u^{(p+\frac{1}{2})}_{i,j+1,k} &= (-a_1 b_{i,j,k}+r_1\bar{b}_{i,j,k})/\Delta \end{aligned}\right\} \begin{array}{l} i=2,4,\ldots,m-1;\ k=1,3,\ldots,m; \\ j=2,4,\ldots,m-1, \end{array}$$

where

$$b_{i,j,k} = \frac{c}{6}\,u^{(p)}_{i,j,k}+a_1u^{(p)}_{i,j+1,k}+ru^{(p+1/3)}_{i,j,k}$$

and

$$\bar{b}_{i,j,k} = a_1u^{(p)}_{i,j,k} + \frac{c}{6}u^{(p)}_{i,j+1,k}+ru^{(p+1/3)}_{i,j+1,k}\ .$$

(iv) At the fourth intermediate level (the $(p+2/3)^{th}$ *iterate)*

The fourth equation of (7.3.9) is transformed to

$$(G_2+rI)\underline{u}^{(p+2/3)}_{[yz]} = G_2\underline{u}^{(p)}_{[yz]} + r\underline{u}^{(p+\frac{1}{2})}_{[yz]}$$

or

$$\underline{u}^{(p+2/3)}_{[yz]} = (G_2+rI)^{-1}[G_2\underline{u}^{(p)}_{[yz]} + r\underline{u}^{(p+\frac{1}{2})}_{[yz]}] \tag{7.3.33}$$

which leads to

$$\underline{u}^{(p+2/3)}_{i[yz]} = \begin{cases} \bar{C}_2^{-1}[\hat{\bar{C}}_2\underline{u}^{(p)}_{i[yz]} + r\underline{u}^{(p+\frac{1}{2})}_{i[yz]}]\ ,\ \text{for } i=1,3,\ldots,m, & (7.3.34a) \\ \bar{C}_1^{-1}[\hat{\bar{C}}_1\underline{u}^{(p)}_{i[yz]} + r\underline{u}^{(p+\frac{1}{2})}_{i[yz]}]\ ,\ \text{for } i=2,4,\ldots,m-1. & (7.3.34b) \end{cases}$$

Therefore, we obtain

(a) for (7.3.34a)

$$u^{(p+2/3)}_{i,m,k} = (\frac{c}{6}\,u^{(p)}_{i,m,k} + ru^{(p+\frac{1}{2})}_{i,m,k})/r_1,\ \text{for } i=1,3,\ldots,m;\ k=1,3,\ldots,m,$$

$$\left.\begin{aligned} u^{(p+2/3)}_{i,j,k} &= (r_1 c_{i,j,k}-a_1\bar{c}_{i,j,k})/\Delta \\ u^{(p+2/3)}_{i,j+1,k} &= (-a_1 c_{i,j,k}+r_1\bar{c}_{i,j,k})/\Delta \end{aligned}\right\} \begin{array}{l} i=1,3,\ldots,m;\ k=1,3,\ldots,m; \\ j=1,3,\ldots,m-2, \end{array}$$

(7.3.35a)

$$u^{(p+2/3)}_{i,1,k} = (\tfrac{c}{6}u^{(p)}_{i,1,k}+ru^{(p+\frac{1}{2})}_{i,1,k})/r_1\ ,\ \text{for } i=1,3,\ldots,m;\ k=2,4,\ldots,m-1,$$

$$\left.\begin{aligned} u^{(p+2/3)}_{i,j,k} &= (r_1c_{i,j,k}-a_1\overline{c}_{i,j,k})/\Delta \\ u^{(p+2/3)}_{i,j+1,k} &= (-a_1c_{i,j,k}+r_1\overline{c}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &i=1,3,\ldots,m;\ k=2,4,\ldots,m-1; \\ &j=2,4,\ldots,m-1, \end{aligned}$$

and

(b) for (7.3.34b)

$$u^{(p+2/3)}_{i,1,k} = (\tfrac{c}{6}u^{(p)}_{i,1,k}+ru^{(p+\frac{1}{2})}_{i,1,k})/r_1 \text{ for } i=2,4,\ldots,m-1;\ k=1,3,\ldots,m,$$

$$\left.\begin{aligned} u^{(p+2/3)}_{i,j,k} &= (r_1c_{i,j,k}-a_1\overline{c}_{i,j,k})/\Delta \\ u^{(p+2/3)}_{i,j+1,k} &= (-a_1c_{i,j,k}+r_1\overline{c}_{i,j,k})/\Delta \end{aligned}\right\}, \begin{aligned} &i=2,4,\ldots,m-1;\ k=1,3,\ldots,m; \\ &j=2,4,\ldots,m-1,, \end{aligned}$$

(7.3.35b)

$$u^{(p+2/3)}_{i,m,k} = (\tfrac{c}{6}\ u^{(p)}_{i,m,k}+ru^{(p+\frac{1}{2})}_{i,m,k})/r_1 \text{ for } i=2,4,\ldots,m-1;\ k=2,4,\ldots,m-1,$$

$$\left.\begin{aligned} u^{(p+2/3)}_{i,j,k} &= (r_1c_{i,j,k}-a_1\overline{c}_{i,j,k})/\Delta \\ u^{(p+2/3)}_{i,j+1,k} &= (-a_1c_{i,j,k}+r_1\overline{c}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &i=2,4,\ldots,m-1;\ k=2,4,\ldots,m-1; \\ &j=1,3,\ldots,m-2, \end{aligned}$$

where $$c_{i,j,k} = \frac{c}{6}\ u^{(p)}_{i,j,k} + a_1u^{(p)}_{i,j+1,k}+ru^{(p+\frac{1}{2})}_{i,j,k}$$

and $$\overline{c}_{i,j,k} = a_1u^{(p)}_{i,j,k} + \frac{c}{6}u^{(p)}_{i,j+1,k}+ru^{(p+\frac{1}{2})}_{i,j+1,k}\ .$$

To determine the AGE equations at the fifth and sixth intermediate levels, it is necessary that we consider our approximations as sweeps parallel to the xz-plane and then evaluate the u values at points lying on each of these planes as illustrated in Figure 7.3.3.

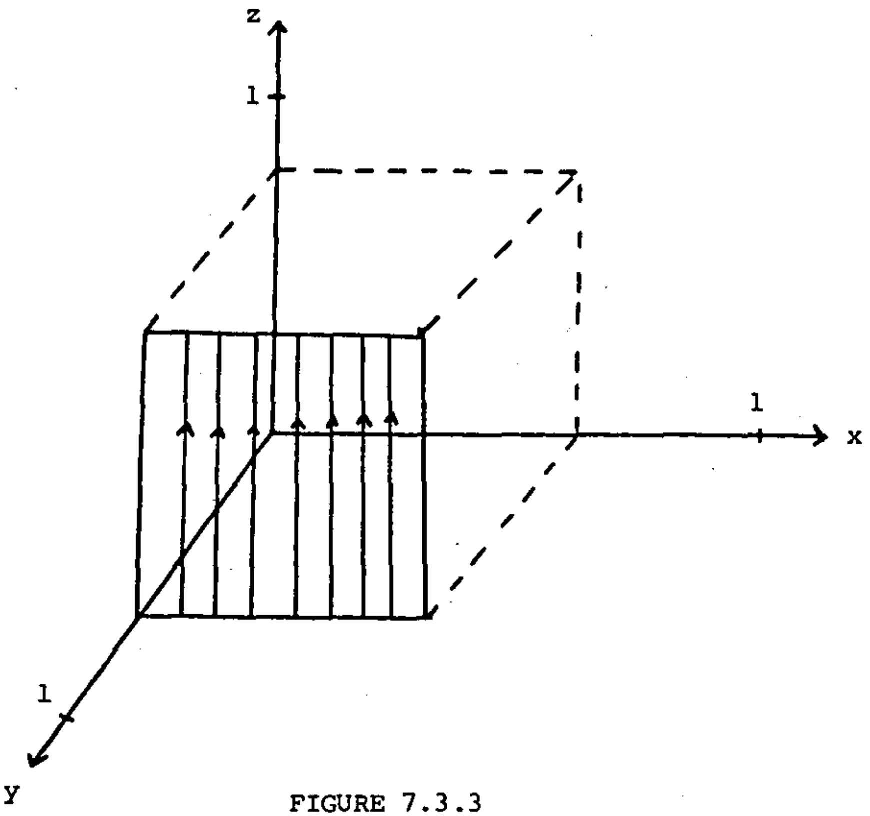


FIGURE 7.3.3

In this case, the points are *reordered column-wise* (parallel to the z-axis) such that,

$$(G_5+G_6)\underline{u}_{[xy]} = (G_1+G_2)\underline{u}_{[xz]} , \qquad (7.3.36)$$

where $\underline{u}_{[xz]} = (\underline{u}_{1[xz]},\underline{u}_{2[xz]},\ldots,\underline{u}_{m[xz]})^T$

with $\underline{u}_{j[xz]} = (u_{1,j,1},u_{1,j,2},\ldots,u_{1,j,m},u_{2,j,1},u_{2,j,2},\ldots,u_{2,j,m},$

$\ldots,u_{m,j,1},u_{m,j,2},\ldots,u_{m,j,m})^T$, for $j=1,2,\ldots,m$.

With the above reordering of points, the fifth and sixth equations are transformed to

$$(G_1+rI)\underline{u}_{[xz]}^{(p+5/6)} = G_1\underline{u}_{[xz]}^{(p)}+r\underline{u}_{[xz]}^{(p+2/3)} \qquad (7.3.37)$$

and

$$(G_2+rI)\underline{u}_{[xz]}^{(p+1)} = G_2\underline{u}_{[xz]}^{(p)}+r\underline{u}_{[xz]}^{(p+5/6)} \qquad (7.3.38)$$

respectively. We now derive the computational formulae at the fifth and sixth intermediate levels.

(v) At the fifth intermediate level (the $(p+5/6)^{th}$ iterate)

From the equation (7.3.37) we have,

$$\underline{u}^{(p+5/6)}_{[xz]} = (G_1+rI)^{-1}[G_1\underline{u}^{(p)}_{[xz]}+r\underline{u}^{(p+2/3)}_{[xy]}], \qquad (7.3.39)$$

which leads to,

$$\underline{u}^{(p+5/6)}_{j[xz]} = \begin{cases} \bar{C}_1^{-1}(\hat{C}_1\underline{u}^{(p)}_{j[xz]}+r\underline{u}^{(p+2/3)}_{j[xz]}), \text{ for } j=1,3,\ldots,m, & (7.3.40a) \\ \bar{C}_2^{-1}(\hat{C}_2\underline{u}^{(p)}_{j[xz]}+r\underline{u}^{(p+2/3)}_{j[xz]}), \text{ for } j=2,4,\ldots,m-1. & (7.3.40b) \end{cases}$$

Therefore, we obtain,

(a) for (7.3.40a)

$$\underline{u}^{(p+5/6)}_{i,j,1} = (\tfrac{c}{6}u^{(p)}_{i,j,1}+ru^{(p+2/3)}_{i,j,1})/r_1, \text{ for } j=1,3,\ldots,m;\ i=1,3,\ldots,m,$$

$$\left.\begin{aligned} u^{(p+5/6)}_{i,j,k} &= (r_1d_{i,j,k}-a_1\bar{d}_{i,j,k})/\Delta \\ u^{(p+5/6)}_{i,j,k+1} &= (-a_1d_{i,j,k}+r_1\bar{d}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &j=1,3,\ldots,m;\ i=1,3,\ldots,m \\ &k=2,4,\ldots,m-1, \qquad (7.3.41a) \end{aligned}$$

$$u^{(p+5/6)}_{i,j,m} = (\tfrac{c}{6}u^{(p)}_{i,j,m}+ru^{(p+2/3)}_{i,j,m})/r_1, \text{ for } j=1,3,\ldots,m;\ i=2,4,\ldots,m-1,$$

$$\left.\begin{aligned} u^{(p+5/6)}_{i,j,k} &= (r_1d_{i,j,k}-a_1\bar{d}_{i,j,k})/\Delta \\ u^{(p+5/6)}_{i,j,k+1} &= (-a_1d_{i,j,k}+r_1\bar{d}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &j=1,3,\ldots,m;\ i=2,4,\ldots,m-1, \\ &k=1,3,\ldots,m-2, \end{aligned}$$

and

(b) for (7.3.40b)

$$u^{(p+5/6)}_{i,j,m} = (\tfrac{c}{6}\ u^{(p)}_{i,j,m}+ru^{(p+2/3)}_{i,j,m})/r_1, \text{ for } j=2,4,\ldots,m-1;\ i=1,3,\ldots,m,$$

$$\left.\begin{aligned} u^{(p+5/6)}_{i,j,k} &= (r_1d_{i,j,k}-a_1\bar{d}_{i,j,k})/\Delta \\ u^{(p+5/6)}_{i,j,k+1} &= (-a_1d_{i,j,k}+r_1\bar{d}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &j=2,4,\ldots,m-1;\ i=1,3,\ldots,m; \\ &k=1,3,\ldots,m-2, \end{aligned} \qquad (7.3.41b)$$

$$u^{(p+5/6)}_{i,j,1} = (\tfrac{c}{6}u^{(p)}_{i,j,1}+ru^{(p+2/3)}_{i,j,1})/r_1, \text{ for } j=2,4,\ldots,m-1;\ i=2,4,\ldots,m-1,$$

$$\left.\begin{aligned} u^{(p+5/6)}_{i,j,k} &= (r_1d_{i,j,k}-a_1\bar{d}_{i,j,k})/\Delta \\ u^{(p+5/6)}_{i,j,k+1} &= (-a_1d_{i,j,k}+r_1\bar{d}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &\text{for } j=2,4,\ldots,m-1;\ i=2,4,\ldots,m-1; \\ &\quad k=2,4,\ldots,m-1, \end{aligned}$$

where

$$d_{i,j,k} = \frac{c}{6}u^{(p)}_{i,j,k}+a_1u^{(p)}_{i,j,k+1}+ru^{(p+2/3)}_{i,j,k}$$

and
$$\bar{d}_{i,j,k} = a_1u^{(p)}_{i,j,k}+\frac{c}{6}u^{(p)}_{i,j,k+1}+ru^{(p+2/3)}_{i,j,k+1} .$$

(vi) At the sixth intermediate level (the (p+1)th level)

From the equation (7.3.38), we have

$$\underline{u}^{(p+1)}_{[xz]} = (G_2+rI)^{-1}[G_2\underline{u}^{(p)}_{[xz]}+r\underline{u}^{(p+5/6)}_{[xz]}], \qquad (7.3.42)$$

which leads to

$$\underline{u}^{(p+1)}_{j[xz]} = \begin{cases} \bar{C}_2^{-1}(\hat{C}_2\underline{u}^{(p)}_{j[xz]}+r\underline{u}^{(p+5/6)}_{j[xz]}) \text{ for } j=1,3,\dots,m, & (7.3.43a) \\ \bar{C}_1^{-1}(\hat{C}_1\underline{u}^{(p)}_{j[xz]}+r\underline{u}^{(p+5/6)}_{j[xz]}) \text{ for } j=2,4,\dots,m-1. & (7.3.43b) \end{cases}$$

Therefore, we obtain,

(a) for (7.3.43a)

$$u^{(p+1)}_{i,j,m} = (\frac{c}{6}u^{(p)}_{i,j,m}+ru^{(p+5/6)}_{i,j,m})/r_1, \text{ for } j=1,3,\dots,m;\ i=1,3,\dots,m,$$

$$\left.\begin{aligned} u^{(p+1)}_{i,j,k} &= (r_1e_{i,j,k}-a_1\bar{e}_{i,j,k})/\Delta \\ u^{(p+1)}_{i,j,k+1} &= (-a_1e_{i,j,k}+r_1\bar{e}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &j=1,3,\dots,m;\ i=1,3,\dots,m; \\ &k=1,3,\dots,m-2, \end{aligned} \qquad (7.3.44a)$$

$$u^{(p+1)}_{i,j,1} = (\frac{c}{6}u^{(p)}_{i,j,1}+ru^{(p+5/6)}_{i,j,1})/r_1 \text{ , for } j=1,3,\dots,m;\ i=2,4,\dots,m-1,$$

$$\left.\begin{aligned} u^{(p+1)}_{i,j,k} &= (r_1e_{i,j,k}-a_1\bar{e}_{i,j,k})/\Delta \\ u^{(p+1)}_{i,j,k+1} &= (-a_1e_{i,j,k}+r_1\bar{e}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &j=1,3,\dots,m;\ i=2,4,\dots,m-1; \\ &k=2,4,\dots,m-1, \end{aligned}$$

and

(b) for (7.3.43b)

$$u^{(p+1)}_{i,j,1} = (\frac{c}{6}u^{(p)}_{i,j,1}+ru^{(p+5/6)}_{i,j,1})/r_1, \text{ for } j=2,4,\dots,m-1;\ i=1,3,\dots,m,$$

$$\left.\begin{aligned} u^{(p+1)}_{i,j,k} &= (r_1e_{i,j,k}-a_1\bar{e}_{i,j,k})/\Delta \\ u^{(p+1)}_{i,j,k+1} &= (-a_1e_{i,j,k}+r_1\bar{e}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &j=2,4,\dots,m-1;\ i=1,3,\dots,m; \\ &k=2,4,\dots,m-1, \end{aligned}$$

$$(7.3.44b)$$

$$u_{i,j,m}^{(p+1)} = (\frac{c}{6}u_{i,j,m}^{(p)}+ru_{i,j,m}^{(p+5/6)})/r_1 \text{ , for } j=2,4,\ldots,m-1;\ i=2,4,\ldots,m-1,$$

$$\left.\begin{aligned} u_{i,j,k}^{(p+1)} &= (r_1e_{i,j,k}-a_1\bar{e}_{i,j,k})/\Delta \\ u_{i,j,k+1}^{(p+1)} &= (-a_1e_{i,j,k}+r_1\bar{e}_{i,j,k})/\Delta \end{aligned}\right\} \begin{aligned} &j=2,4,\ldots,m-1;\ i=2,4,\ldots,m-1; \\ &k=1,3,\ldots,m-2, \end{aligned}$$

where

$$e_{i,j,k} = \frac{c}{6}u_{i,j,k}^{(p)}+a_1u_{i,j,k+1}^{(p)}+ru_{i,j,k}^{(p+5/6)}$$

and

$$\bar{e}_{i,j,k} = a_1u_{i,j,k}^{(p)}+\frac{c}{6}u_{i,j,k+1}^{(p)}+ru_{i,j,k+1}^{(p+5/6)} .$$

Thus, we see that the AGE algorithm corresponds to sweeping through the mesh lying on planes (where the appropriate reordering of the points are done) parallel to the xy, yz and xz-planes. This process involves tridiagonal systems which in turn entails at each stage the solution of (2×2) block systems. The iterative procedure is continued until convergence is reached, that is, when the requirement $|u_{i,j,k}^{(p+1)}-u_{i,j,k}^{(p)}|\leq\varepsilon$ is met, where ε is the convergence criterion.

7.4 NUMERICAL EXPERIMENTS

The AGE algorithm was tested on the following examples:

(i) Two-dimensional problems

(a) Problem 1

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} \ , \ 0 \leqslant x \leqslant 2, \ 0 \leqslant y \leqslant 1, \ t \geqslant 0 \ , \tag{7.4.1}$$

subject to the initial condition

$$U(x,y,0) = \sin(\pi y)\sin(2\pi x) \ , \tag{7.4.1a}$$

and the boundary conditions,

$$U(0,y,t) = U(2,y,t) = U(x,0,t) = U(x,1,t) = 0 \ . \tag{7.4.1b}$$

The exact solution is given by (Johnson and Riess (1982))

$$U(x,y,t) = e^{-5\pi^2 t}\sin(2\pi x)\sin(\pi y), \ 0 \leqslant x \leqslant 2, \ 0 \leqslant y \leqslant 1, \ t \geqslant 0. \tag{7.4.2}$$

(b) Problem 2

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + h(x,y,t) \ , \ 0 \leqslant x \leqslant 1, \ 0 \leqslant y \leqslant 1, \ t \geqslant 0, \tag{7.4.3}$$

with,

$$h(x,y,t) = \sin x \sin y \ e^{-t} - 4,$$

where the theoretical solution is given by (Gourlay and MacGuire (1971))

$$U(x,y,t) = \sin x \sin y \ e^{-t} + x^2 + y^2, \ 0 \leqslant x \leqslant 1, \ 0 \leqslant y \leqslant 1, \ t \geqslant 0. \tag{7.4.4}$$

The initial and boundary conditions are defined so as to agree with the exact solution.

The numerical results for both problems are displayed in Tables 7.4.1-7.4.7 and Tables 7.4.8-7.4.13 for different mesh sizes and values of x and y. For Problem 1, the convergence of the AGE schemes over the whole xy-plane at each time level was considered using a convergence requirement of $\varepsilon = 10^{-4}$. It is evident that the AGE method employing the Crank-Nicolson formula (AGE-CN) where the truncation error is $O((\Delta x)^2+(\Delta t)^2)$ is more accurate than the corresponding fully implicit formula (AGE-IMP) where the truncation error is $O((\Delta x)^2+\Delta t)$. This can be inferred from an examination of the absolute errors at the grid points as well as from the average of the errors given in Table 7.4.7. From the same table, we also find that both methods require only 4 iterations to complete the iterative procedure.

For Problem 2, the AGE solutions were compared with the corresponding results of the GE schemes obtained by Evans and Abdullah (1983). As above, the AGE-CN scheme is found to be more superior than AGE-IMP. It is generally observed from Tables 7.4.8-7.4.13 that the AGE formulae using a more stringent requirement of $\varepsilon=10^{-8}$ produce more accurate solutions than those with $\varepsilon=10^{-4}$. For AGE-CN in particular, the accuracies of the solutions improve by about one place of decimal. To a convergence requirement of $\varepsilon=10^{-4}$, the AGE-CN scheme can have comparable accuracies with that of the GE schemes such as (S)AGE and (D)AGE at some of the grid points. However, with $\varepsilon=10^{-8}$, the AGE-CN formula is obviously more accurate over the whole plane than any of the GE methods. Nevertheless, this is achieved at the expense of introducing an iterative procedure as we readily see from Table 7.4.13.

(ii) A three-dimensional problem

This problem involved the solution of

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} + h(x,y,z,t),\ 0\leqslant x,y,z\leqslant 1;\ t\geqslant 0\ , \quad (7.4.5)$$

$$(h(x,y,z,t) = (3\pi^2-1)\sin(\pi x)\sin(\pi y)\sin(\pi z)e^{-t})$$

subject to the initial condition,

$$U(x,y,z,0) = \sin(\pi x)\sin(\pi y)\sin(\pi z) \ , \tag{7.4.5a}$$

and the boundary conditions,

$$U(0,y,z,t) = U(1,y,z,t) = U(x,0,z,t) = U(x,1,z,t) = U(x,y,0,t) = U(x,y,1,t) = 0 \ . \tag{7.4.5b}$$

The exact solution is given by,

$$U(x,y,z,t) = \sin(\pi x)\sin(\pi y)\sin(\pi z)e^{-t} \ . \tag{7.4.6}$$

The absolute errors of the AGE solutions to the above problem for $\lambda=0.05$ and at some of the grid points on the cube are shown in Tables 7.4.14-7.4.17. As in the one and two-dimensional cases, because of the high order truncation errors of the Crank-Nicolson formula, the AGE-CN scheme displays more accurate results than AGE-IMP. No attempt was made to obtain the AGE solutions for larger values of t because of the prohibitively large computations involved which in turn can lead to larger rounding errors and consequently make convergence difficult.

y=0.02

$\lambda=0.5$, $\Delta x= \Delta y = 0.02$, $\Delta t=0.0002$, $t=0.0018$, $r=1.0$, $\varepsilon=10^{-4}$

x Method	0.02	0.3	0.58	0.86	1.14	1.42	1.7	1.98
AGE-IMP	5.21×10^{-6}	2.02×10^{-5}	1.59×10^{-5}	1.42×10^{-5}	2.13×10^{-5}	6.23×10^{-6}	2.36×10^{-5}	1.2×10^{-6}
AGE-CN	1.45×10^{-6}	9.88×10^{-6}	5.28×10^{-6}	7.9×10^{-6}	8.24×10^{-6}	4.82×10^{-6}	1.0×10^{-5}	1.28×10^{-6}
EXACT SOLUTION	0.0072008	0.0546416	-0.0276785	-0.0442688	0.0442688	0.0276785	-0.0546416	-0.0072008

TABLE 7.4.1: The absolute errors of the AGE solutions for 2D Problem 1 (7.4.1) at y=0.02, t=0.0018

y=0.5

$\lambda=0.5$, $\Delta x=\Delta y=0.02$, $\Delta t=0.0002$, $t=0.0018$, $r=1.0$, $\varepsilon=10^{-4}$

x Method	0.02	0.3	0.58	0.86	1.14	1.42	1.7	1.98
AGE-IMP	6.76×10^{-5}	5.33×10^{-4}	2.56×10^{-4}	4.37×10^{-4}	4.2×10^{-4}	2.79×10^{-4}	3.93×10^{-4}	6.95×10^{-5}
AGE-CN	2.18×10^{-5}	1.66×10^{-4}	8.26×10^{-5}	1.35×10^{-4}	1.33×10^{-4}	8.48×10^{-5}	1.65×10^{-4}	2.36×10^{-5}
EXACT SOLUTION	0.1146805	0.8702209	-0.4408067	-0.7050230	0.7050230	0.4408067	-0.8702209	-0.1146805

TABLE 7.4.2: The absolute errors of the AGE solutions for 2D Problem 1 (7.4.1) at y=0.5, t=0.0018

y=0.98

$\lambda=0.5$, $\Delta x=\Delta y=0.02$, $\Delta t=0.0002$, $t=0.0018$, $r=1.0$, $\varepsilon=10^{-4}$

x Method	0.02	0.3	0.58	0.86	1.14	1.42	1.7	1.98
AGE-IMP	4.05×10^{-6}	5.38×10^{-5}	1.94×10^{-5}	4.65×10^{-5}	3.69×10^{-5}	3.27×10^{-5}	4.91×10^{-5}	1.11×10^{-5}
AGE-CN	1.4×10^{-6}	1.18×10^{-5}	5.53×10^{-6}	9.75×10^{-6}	9.18×10^{-6}	6.31×10^{-6}	1.15×10^{-5}	1.83×10^{-6}
EXACT SOLUTION	0.0072008	0.0546416	-0.0276785	-0.0442688	0.0442688	0.0276785	-0.0546416	-0.0072008

TABLE 7.4.3: The absolute errors of the AGE solutions for 2D Problem 1 (7.4.1) at y=0.98, t=0.0018

y=0.02

$\lambda=1.0$, $\Delta x=\Delta y=0.02$, $\Delta t=0.0004$, $t=0.0036$, $r=1.0$, $\varepsilon=10^{-4}$

x Method	0.02	0.3	0.58	0.86	1.14	1.42	1.7	1.98
AGE-IMP	3.77×10^{-5}	1.25×10^{-4}	8.61×10^{-5}	1.57×10^{-4}	2.71×10^{-5}	1.68×10^{-4}	3.57×10^{-5}	5.64×10^{-5}
AGE-CN	3.92×10^{-6}	4.66×10^{-6}	1.13×10^{-5}	4.18×10^{-7}	1.14×10^{-5}	3.89×10^{-6}	1.0×10^{-5}	1.22×10^{-6}
EXACT SOLUTION	0.0065888	0.0499973	-0.0253259	-0.0405061	0.0405061	0.0253259	-0.0499973	-0.0065888

TABLE 7.4.4: The absolute errors of the AGE solutions for 2D Problem 1 (7.4.1) at y=0.02, t=0.0036

y=0.5

$\lambda=1.0,\ \Delta x=\Delta y=0.02, \Delta t=0.0004,\ t=0.0036,\ r=1.0,\ \varepsilon=10^{-4}$

x Method	0.02	0.3	0.58	0.86	1.14	1.42	1.7	1.98
AGE-IMP	1.92×10^{-4}	1.51×10^{-3}	6.8×10^{-4}	1.25×10^{-3}	1.15×10^{-3}	8.22×10^{-4}	1.46×10^{-3}	3.17×10^{-4}
AGE-CN	3.53×10^{-5}	2.76×10^{-4}	1.27×10^{-4}	2.28×10^{-4}	2.13×10^{-4}	1.48×10^{-4}	2.68×10^{-4}	5.26×10^{-5}
EXACT SOLUTION	0.1049331	0.7962559	-0.4033401	-0.6450991	0.6450991	0.4033401	-0.7962559	-0.1049331

TABLE 7.4.5: The absolute errors of the AGE solutions for 2D Problem 1 (7.4.1) at y=0.5, t=0.0036

y=0.98

$\lambda=1.0,\ \Delta x=\Delta y=0.02,\ \Delta t=0.0004,\ t=0.0036,\ r=1.0,\ \varepsilon=10^{-4}$

x Method	0.02	0.3	0.58	0.86	1.14	1.42	1.7	1.98
AGE-IMP	8.93×10^{-6}	3.72×10^{-4}	2.64×10^{-5}	3.62×10^{-4}	1.62×10^{-4}	3.01×10^{-4}	2.75×10^{-4}	1.09×10^{-4}
AGE-CN	1.39×10^{-6}	3.77×10^{-5}	8.2×10^{-6}	3.46×10^{-5}	2.12×10^{-5}	2.67×10^{-5}	3.12×10^{-5}	9.29×10^{-6}
EXACT SOLUTION	0.0065888	0.0499973	-0.0253259	-0.0405061	0.0405061	0.0253259	-0.0499973	-0.0065888

TABLE 7.4.6: The absolute errors of the AGE solutions for 2D Problem 1 (7.4.1) at y=0.98, t=0.0036

Method		Average of All Absolute Errors		Number of Iterations	
λ	t	AGE-IMP	AGE-CN	AGE-IMP	AGE-CN
0.5	0.0018	2.33×10^{-4}	7.26×10^{-5}	4	4
1.0	0.0036	6.98×10^{-4}	1.21×10^{-4}	4	4

TABLE 7.4.7

x=0.1

$\lambda=0.1$, $\Delta x=\Delta y=0.1$, $\Delta t=0.001$, $t=0.1$, $r=1.5$

y Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
GER		7.0×10^{-6}	5.0×10^{-6}	1.3×10^{-5}	1.1×10^{-5}	1.9×10^{-5}	1.8×10^{-5}	2.3×10^{-5}	2.2×10^{-5}	2.2×10^{-5}
(S)AGE		2.0×10^{-6}	4.0×10^{-6}	2.0×10^{-6}	7.0×10^{-6}	4.0×10^{-6}	9.0×10^{-6}	6.0×10^{-6}	9.0×10^{-6}	5.0×10^{-6}
(D)AGE		3.0×10^{-6}	3.0×10^{-6}	5.0×10^{-7}	5.0×10^{-6}	1.0×10^{-6}	6.0×10^{-6}	2.0×10^{-6}	5.0×10^{-6}	4.0×10^{-7}
AGE-IMP	$(\varepsilon=10^{-4})$	2.0×10^{-5}	3.9×10^{-5}	5.7×10^{-5}	7.3×10^{-5}	8.4×10^{-5}	9.0×10^{-5}	8.9×10^{-5}	7.6×10^{-5}	4.8×10^{-5}
	$(\varepsilon=10^{-8})$	1.96×10^{-6}	3.86×10^{-6}	5.59×10^{-6}	7.07×10^{-6}	8.15×10^{-6}	8.66×10^{-6}	8.39×10^{-6}	7.06×10^{-6}	4.38×10^{-6}
AGE-CN	$(\varepsilon=10^{-4})$	1.75×10^{-5}	3.4×10^{-5}	4.98×10^{-5}	6.29×10^{-5}	7.24×10^{-5}	7.69×10^{-5}	7.42×10^{-5}	6.21×10^{-5}	3.78×10^{-5}
	$(\varepsilon=10^{-8})$	1.23×10^{-6}	2.42×10^{-6}	3.5×10^{-6}	4.43×10^{-6}	5.1×10^{-6}	5.42×10^{-6}	5.25×10^{-6}	4.42×10^{-6}	2.74×10^{-6}
EXACT SOLUTION		0.029018	0.067946	0.126695	0.205177	0.303308	0.421006	0.558194	0.714801	0.890760

TABLE 7.4.8: The absolute errors of the numerical solutions to 2D Problem 2 (7.4.3)

x=0.5

$\lambda=0.1$, $\Delta x=\Delta y=0.1$, $\Delta t=0.001$, $t=0.1$, $r=1.5$

Method \ y	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
GER	1.9×10^{-5}	6.0×10^{-6}	1.3×10^{-5}	8.0×10^{-6}	1.1×10^{-5}	8.0×10^{-6}	1.2×10^{-5}	1.0×10^{-6}	1.7×10^{-5}
(S)AGE	1.6×10^{-5}	1.2×10^{-5}	4.0×10^{-6}	1.9×10^{-5}	3.0×10^{-6}	2.3×10^{-5}	6.0×10^{-6}	1.8×10^{-5}	1.0×10^{-6}
(D)AGE	1.6×10^{-5}	1.1×10^{-5}	6.0×10^{-6}	1.7×10^{-5}	4.0×10^{-7}	2.0×10^{-7}	2.0×10^{-6}	1.5×10^{-5}	5.0×10^{-6}
AGE-IMP ($\varepsilon=10^{-4}$)	8.41×10^{-5}	1.66×10^{-4}	2.41×10^{-4}	3.07×10^{-4}	3.58×10^{-4}	3.87×10^{-4}	3.84×10^{-4}	3.34×10^{-4}	2.15×10^{-4}
($\varepsilon=10^{-8}$)	8.15×10^{-6}	1.6×10^{-5}	2.33×10^{-5}	2.95×10^{-5}	3.43×10^{-5}	3.67×10^{-5}	3.59×10^{-5}	3.05×10^{-5}	1.92×10^{-5}
AGE-CN ($\varepsilon=10^{-4}$)	7.24×10^{-5}	1.42×10^{-4}	2.07×10^{-4}	2.62×10^{-4}	3.04×10^{-4}	3.25×10^{-4}	3.16×10^{-4}	2.68×10^{-4}	1.65×10^{-4}
($\varepsilon=10^{-8}$)	5.1×10^{-6}	1.0×10^{-5}	1.46×10^{-5}	1.85×10^{-5}	2.14×10^{-5}	2.3×10^{-5}	2.24×10^{-5}	1.91×10^{-5}	1.2×10^{-5}
EXACT SOLUTION	0.303308	0.376183	0.468197	0.578931	0.707976	0.854943	1.019463	1.201191	1.399809

TABLE 7.4.9: The absolute errors of the numerical solutions to 2D Problem 2 (7.4.3)

x=0.9

$\lambda=0.1$, $\Delta x=\Delta y=0.1$, $\Delta t=0.001$, $t=0.1$, $r=1.5$

y Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
GER		2.22×10^{-5}	1.0×10^{-5}	1.8×10^{-5}	1.0×10^{-5}	1.7×10^{-5}	7.0×10^{-5}	1.8×10^{-5}	1.0×10^{-6}	1.1×10^{-5}
(S)AGE		1.9×10^{-5}	1.3×10^{-5}	1.1×10^{-5}	1.9×10^{-5}	6.0×10^{-6}	2.0×10^{-5}	3.0×10^{-6}	1.7×10^{-5}	6.0×10^{-6}
(D)AGE		2.0×10^{-5}	1.3×10^{-5}	1.2×10^{-5}	1.7×10^{-5}	8.0×10^{-6}	1.8×10^{-5}	6.0×10^{-6}	1.4×10^{-5}	9.0×10^{-6}
AGE-IMP	$(\varepsilon=10^{-4})$	4.84×10^{-5}	9.59×10^{-5}	1.4×10^{-4}	1.81×10^{-4}	2.15×10^{-4}	2.4×10^{-4}	2.48×10^{-4}	2.29×10^{-4}	1.63×10^{-4}
	$(\varepsilon=10^{-8})$	4.38×10^{-6}	8.64×10^{-6}	1.27×10^{-5}	1.62×10^{-5}	1.92×10^{-5}	2.11×10^{-5}	2.15×10^{-5}	1.95×10^{-5}	1.35×10^{-5}
AGE-CN	$(\varepsilon=10^{-4})$	3.78×10^{-5}	7.45×10^{-5}	1.09×10^{-4}	1.4×10^{-4}	1.65×10^{-4}	1.81×10^{-4}	1.83×10^{-4}	1.63×10^{-4}	1.09×10^{-4}
	$(\varepsilon=10^{-8})$	2.74×10^{-6}	5.4×10^{-6}	7.92×10^{-6}	1.02×10^{-5}	1.2×10^{-5}	1.32×10^{-5}	1.35×10^{-5}	1.22×10^{-5}	8.41×10^{-6}
EXACT SOLUTION		0.890760	0.990813	1.109460	1.246013	1.399809	1.570209	1.756611	1.958450	2.175209

TABLE 7.4.10: The absolute errors of the numerical solutions to 2D Problem 2 (7.4.3)

x=0.5

$\lambda=1.0$, $\Delta x=\Delta y=0.1$, $\Delta t=0.01$, $t=0.5$, $r=1.5$

y Method		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
(S)AGE		1.49×10^{-4}	6.5×10^{-5}	1.8×10^{-5}	9.7×10^{-5}	1.69×10^{-4}	9.7×10^{-5}	1.31×10^{-4}	2.57×10^{-4}	1.11×10^{-4}
(D)AGE		1.16×10^{-4}	4.4×10^{-5}	7.3×10^{-5}	7.6×10^{-5}	2.6×10^{-5}	1.16×10^{-4}	1.1×10^{-5}	1.28×10^{-4}	3.2×10^{-5}
AGE-IMP	$(\varepsilon=10^{-4})$	2.8×10^{-5}	1.1×10^{-4}	8.86×10^{-5}	1.57×10^{-4}	1.14×10^{-4}	1.55×10^{-4}	1.04×10^{-4}	6.52×10^{-5}	7.46×10^{-5}
	$(\varepsilon=10^{-8})$	3.03×10^{-5}	5.92×10^{-5}	8.52×10^{-5}	1.07×10^{-4}	1.22×10^{-4}	1.28×10^{-4}	1.23×10^{-4}	1.02×10^{-4}	6.3×10^{-5}
AGE-CN	$(\varepsilon=10^{-4})$	1.56×10^{-5}	3.37×10^{-5}	4.33×10^{-5}	5.66×10^{-5}	5.88×10^{-5}	6.24×10^{-5}	5.45×10^{-5}	4.21×10^{-5}	2.61×10^{-5}
	$(\varepsilon=10^{-8})$	4.26×10^{-6}	8.33×10^{-6}	1.2×10^{-5}	1.5×10^{-5}	1.71×10^{-5}	1.8×10^{-5}	1.72×10^{-5}	1.44×10^{-5}	8.86×10^{-6}
EXACT SOLUTION		0.289030	0.347770	0.425933	0.523238	0.639410	0.774190	0.927330	1.098597	1.287781

TABLE 7.4.11: The absolute errors of the numerical solutions to 2D Problem 2 (7.4.3)

x=0.5

$\lambda=1.0$, $\Delta x=\Delta y=0.1$, $\Delta t=0.01$, $t=1.2$, $r=1.5$

Method \ y		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
(D)AGE		8.7×10^{-5}	2.3×10^{-5}	1.14×10^{-4}	5.5×10^{-5}	1.31×10^{-4}	7.7×10^{-5}	1.53×10^{-4}	9.9×10^{-5}	1.2×10^{-4}
AGE-IMP	$(\varepsilon=10^{-4})$	1.4×10^{-5}	5.49×10^{-5}	4.42×10^{-5}	7.82×10^{-5}	5.7×10^{-5}	7.7×10^{-5}	5.16×10^{-5}	3.25×10^{-5}	3.71×10^{-5}
	$(\varepsilon=10^{-8})$	1.5×10^{-5}	2.94×10^{-5}	4.23×10^{-5}	5.3×10^{-5}	6.04×10^{-5}	6.35×10^{-5}	6.08×10^{-5}	5.08×10^{-5}	3.13×10^{-5}
AGE-CN	$(\varepsilon=10^{-4})$	2.52×10^{-5}	5.28×10^{-5}	7.01×10^{-5}	9.05×10^{-5}	9.63×10^{-5}	1.02×10^{-4}	8.97×10^{-5}	6.94×10^{-5}	4.21×10^{-5}
	$(\varepsilon=10^{-8})$	2.11×10^{-6}	4.14×10^{-6}	5.95×10^{-6}	7.45×10^{-6}	8.49×10^{-6}	8.93×10^{-6}	8.56×10^{-6}	7.14×10^{-6}	4.4×10^{-6}
EXACT SOLUTION		0.274416	0.318688	0.382673	0.466232	0.569229	0.691534	0.833025	0.993586	1.173113

TABLE 7.4.12: The absolute errors of the numerical solutions to 2D Problem 2 (7.4.3)

Method		Average of All Absolute Errors				Number of Iterations			
		AGE-IMP		AGE-CN		AGE-IMP		AGE-CN	
λ	t	$(\varepsilon=10^{-4})$	$(\varepsilon=10^{-8})$	$(\varepsilon=10^{-4})$	$(\varepsilon=10^{-8})$	$(\varepsilon=10^{-4})$	$(\varepsilon=10^{-8})$	$(\varepsilon=10^{-4})$	$(\varepsilon=10^{-8})$
0.1	0.1	2.13×10^{-4}	1.98×10^{-5}	1.74×10^{-4}	1.24×10^{-5}	2	9	3	10
1.0	0.5	7.28×10^{-5}	6.85×10^{-5}	3.17×10^{-5}	9.64×10^{-6}	4	22	4	11
1.0	1.2	3.63×10^{-5}	3.4×10^{-5}	5.13×10^{-5}	4.79×10^{-6}	4	21	3	11

TABLE 7.4.13

y=0.025, z=0.025

$\lambda=0.05$, $\Delta x=\Delta y=\Delta z=0.025$, $\Delta t=3.125\times 10^{-5}$, $t=3.125\times 10^{-5}$, $\varepsilon=10^{-4}$, $r=2.5$

x Method	0.025	0.5	0.975	Number of Iterations
AGE-IMP	1.81×10^{-6}	3.89×10^{-5}	6.99×10^{-7}	9
AGE-CN	7.06×10^{-8}	1.41×10^{-5}	5.57×10^{-8}	9
EXACT SOLUTION	0.0004830	0.0061556	0.0004830	-

TABLE 7.4.14: The absolute errors of the AGE solutions to 3D Problem (7.4.5)

y=0.975, z=0.025

$\lambda=0.05$, $\Delta x=\Delta y=\Delta z=0.025$, $\Delta t=3.125\times 10^{-5}$, $t=6.25\times 10^{-5}$, $\varepsilon=10^{-4}$, $r=2.5$

x Method	0.025	0.5	0.975	Number of Iterations
AGE-IMP	7.0×10^{-6}	2.93×10^{-4}	7.02×10^{-5}	8
AGE-CN	1.67×10^{-6}	2.16×10^{-4}	7.46×10^{-5}	9
EXACT SOLUTION	0.00048295	0.00615546	0.00048295	-

TABLE 7.4.15: The absolute errors of the AGE solutions to 3D Problem (7.4.5)

y=0.975, z=0.5

λ=0.05, $\Delta x=\Delta y=\Delta z$=0.025, $\Delta t=3.125\times10^{-5}$, $t=1.25\times10^{-4}$, $\varepsilon=10^{-4}$, r=2.5

Method \ x	0.025	0.5	0.975	Number of Iterations
AGE-IMP	9.58×10^{-4}	2.59×10^{-2}	8.57×10^{-4}	8
AGE-CN	2.81×10^{-4}	2.16×10^{-2}	5.16×10^{-4}	8
EXACT SOLUTION	0.00615507	0.0784494	0.00615507	-

TABLE 7.4.16: The absolute errors of the AGE solutions to 3D Problem (7.4.5)

y=0.5, z=0.5

λ=0.05, $\Delta x=\Delta y=\Delta z$=0.025, $\Delta t=3.125\times10^{-5}$, $t=1.5625\times10^{-4}$, $\varepsilon=10^{-4}$, r=2.5

Method \ x	0.025	0.5	0.975	Number of Iterations
AGE-IMP	9.83×10^{-3}	1.88×10^{-1}	1.55×10^{-2}	8
AGE-CN	3.95×10^{-3}	8.77×10^{-2}	3.89×10^{-3}	8
EXACT SOLUTION	0.07844683	0.99984376	0.07844683	-

TABLE 7.4.17: The absolute errors of the AGE solutions to 3D Problem (7.4.5)

SUMMARY AND CONCLUSION

Throughout the course of this thesis, new *explicit* methods were proposed as alternative finite-difference schemes to solve a given hyperbolic or parabolic equation and emphasis on the convergence, stability, consistency, accuracy and efficiency of these schemes was stressed. The GE methods employed to solve first and second-order hyperbolic equations can be very competitive in terms of accuracy and efficiency. Unfortunately, we are limited in our choice of larger time steps because of restrictions on stability. The same observation also applies to parabolic problems with cylindrical and spherical symmetries, where the author encountered theoretical difficulties in establishing conditions of stability using Brauer's theorem. One possible reason is that the upper bounds given by Brauer's theorem were not very close to the largest eigenvalue of the amplification matrix under investigation. This leaves open the scope for further research on the possibility of finding stable, asymmetric semi-implicit approximations which when appropriately coupled at the grid points lead to stable explicit equations thus qualifying as a difference scheme in the GE class of methods. An obvious way to achieve this is to look into the possibility of finding the appropriate discretisation of the derivatives appearing in the differential equation.

The flexibility of the AGE schemes is demonstrated by our freedom to use the appropriate well-known implicit difference formula to approximate the given differential equation which is guaranteed of stability. We also have a choice of the basic iterative formula - the PR or DR variant. The PR variant is generally favoured because of its $O([\Delta x]^2+[\Delta t]^2)$ accuracy and given a reasonable convergence criterion, for example $\varepsilon=10^{-4}$, it requires a low iteration count.

The DR variant, on the other hand, is always used for two and higher dimensional problems because unlike the PR variant, it is unconditionally stable. However, it is only $O([\Delta x]^2+\Delta t)$ accurate. It, therefore, becomes essential that further studies be conducted to develop a more accurate as well as stable variant which ensures the simple splitting of the coefficient matrix of the system of linear equations thus enabling the application of the AGE algorithm.

Both the GE and AGE iterative class of methods have the merits that they are explicit and accurate. Furthermore, in the case of the AGE schemes, they are always stable. Hence, with improved stability and convergence characteristics, the two schemes can be used effectively on parallel computers. For higher dimensional problems, the ADI method is frequently used up to the present time. This employs the Thomas algorithm to solve the resulting system of equations in each of the appropriate coordinate directions. For such implicit methods, often we are not able to exploit to the full the implicit parallelism within the algorithm. Thus, although the AGE class of methods appear to require a slightly more computational work and incur more rounding errors than the Thomas algorithm on a sequential machine, they however, serve as an efficient technique for achieving parallelism.

REFERENCES

ABDULLAH, A.R.: *The study of some numerical methods for solving parabolic partial differential equations*, Ph.D. Thesis, Loughborough University of Technology, 1983.

ALBASINY, E.L.: *On the numerical solution of a cylindrical heat-conduction problem*, Q.J.Mech.Appl.Math. 13, pp.374-384 (1960).

AMES, W.F.: *Non-linear partial differential equations in engineering*, New York: Academic Press, 1965.

AMES, W.F.: *Numerical methods for partial differential equations (2nd Edition)*, New York: Academic Press, 1977.

CALDWELL, J. and SMITH, P.: *Solution of Burger's equation with a large Reynolds number*, Appl.Math. Modelling 6, pp.381-385 (1982).

CARRIER, G.F.: *On the non-linear vibration problem of the elastic string*, Quart.Appl.Math. 3, pp.157-165 (1945).

CASULLI, V., CHENG, R.T. and BULGARELLI, U.: *Eulerian-Lagrangian solution of convection dominated diffusion problems*, Numerical Methods for Non-linear Problems 2, pp.962-971, (ed. Taylor, C. Hinton, E. and Owen, D.R.J.), Swansea: Pineridge Press, 1984.

COLE, J.D.: *On a quasilinear parabolic equation occurring in aerodynamics*, A.Appl.Maths. 9, pp.225-236 (1951).

CRANK, J. and NICOLSON, P.: *A practical method for numerical evaluation of solutions of partial differential equations of the heat conduction type*, Proc.Camb.Phil.Soc. 32, pp.50-67 (1947).

DANAEE, A.: *A study of hopscotch methods for solving parabolic partial differential equations*, Ph.D. Thesis, Loughborough University of Technology, 1980.

DOUGLAS, J.: *A survey of numerical methods for parabolic differential equations*, Advances in Computers 2, pp.1-54, New York: Academic Press, 1961.

DOUGLAS, J.: *On the numerical integration of* $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial u}{\partial t}$ *by implicit methods*, J.Soc.Indust.Appl.Math. 3, pp.42-65 (1955).

DOUGLAS, J.: *Alternating direction methods for three space variables*, Numerische Mathematik 4, pp.41-63 (1962).

DOUGLAS, J. and GUNN, J.E.: *A general formulation of alternating direction methods*, Part I. Numerische Mathematik 6, pp.428-453 (1964).

DOUGLAS, J. and JONES, B.F.: *On predictor-corrector methods for non-linear parabolic differential equations*, J.Soc.Indust.Appl. Math. 11, pp. 195-204 (1963).

DOUGLAS, J. and RACHFORD, H.H.: *On the numerical solution of heat conduction problems in two or three space variables*, Trans. Amer.Math.Soc. 82, pp.421-439 (1956).

EISEN, D.: *Stability and convergence of finite difference schemes with singular coefficients*, SIAM J.Numer.Anal. 3, pp.545-552 (1966).

EISEN, D.: *Consistency conditions for difference schemes with singular coefficients*, Maths.Comp. 22, pp.347-351 (1968).

EVANS, D.J.: *Group explicit iterative methods for solving large linear systems*, Intern.J.Computer Math. 17, pp.81-108 (1985).

EVANS, D.J. and ABDULLAH, A.R.B.: *A new explicit method for the solution of* $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$, Intern.J.Computer Math. 14, pp.325-353 (1983).

EVANS, D.J. and ABDULLAH, A.R.B.: *Group explicit methods for parabolic equations*, Intern.J. Computer Math. 14, pp.73-105 (1983).

EVANS, D.J. and ABDULLAH, A.R.B.: *The group explicit method for the solution of Burger's equation*, Computing 32, pp.239-253 (1984).

FOX, L.: *Numerical solution of ordinary and partial differential equations*, Oxford: Pergamon Press, 1962.

FOX, P.: *The solution of hyperbolic partial differential equations by difference methods*, Mathematical Methods for Digital Computers, Vol. 2, pp.165-188 (ed. Ralston, A. and Wilf, H.S.), John Wiley and Sons, 1967.

FORSYTHE, G.E. and WASOW, W.R.: *Finite difference methods for partial differential equations*, New York: John Wiley, 1960.

FRIEDMAN, A.: *Partial differential equations of parabolic type*, Englewood Cliffs: Prentice-Hall Inc. 1964.

GANE, C.R.: *Computational techniques for the numerical solution of parabolic and elliptic partial differential equations*, Ph.D. Thesis, Loughborough University of Technology, 1974.

GLADWELL, I. and WAIT, R.: *A survey of numerical methods for partial differential equations*, Oxford: Clarendon Press 1979.

GOULT, R.J., HOSKINS, R.F., MILNER, J.A. and PRATT, M.J.: *Computational methods in linear algebra*, London: Stanley Thornes (Publishers) Ltd. 1974.

GOURLAY, A.R.: *Hopscotch: A fast second-order partial differential equation solver*, J.Inst.Math.Appl. 6, pp. 375-390 (1970).

GOURLAY, A.R.: *Splitting methods for time-dependent partial differential equations*, The State of the Art in Numerical Analysis, pp. 757-796 (ed. Jacobs, D.A.H.), London: Academic Press, 1977.

GOURLAY, A.R. and GANE, C.R.: *Block hopscotch procedures for second order parabolic differential equations*, J.Inst.Math.Applics. 19, pp.205-216 (1977).

GOURLAY, A.R. and MCGUIRE, G.R.: *General hopscotch algorithm for the numerical solution of partial differential equations*, J.Inst. Maths.Applics. 7, pp. 216-227 (1971).

GOURLAY, A.R. and MCKEE, S.: *The construction of hopscotch methods for parabolic and elliptic equations in two space dimensions with a mixed derivative*, J.Comp.Appl.Maths. 3, pp. 201-206, (1977).

GOURLAY, A.R. and MORRIS, J.Ll.: *A multi-step formulation of the optimised Lax-Wendroff method for non-linear hyperbolic systems in two space variables*, Math.Comp. 22, pp.715-720 (1968).

HAMDAN, A.R.: *Private Communication*, 1984.

HENRICI, P.: *Discrete variable methods in ordinary differential equations*, New York: John Wiley and Sons 1962.

HILDEBRAND, F.B.: *Introduction to numerical analysis*, New York: McGraw-Hill, 1956.

ISAACSON, E. and KELLER, H.B.: *Analysis of numerical methods*, New York: John Wiley and Sons, 1966.

JAIN, M.K.: *Numerical solution of differential equations (2nd Edition)*, Wiley Eastern Ltd., 1984.

JOHNSON, L.W. and REISS, R.D.: *Numerical analysis (2nd Edition)*, Reading: Addison-Wesley Publishing Company 1982.

KOPAL, Z.: *Numerical anslysis*, New York: John Wiley and Sons, 1961.

LAPIDUS, L. and PINDER, G.F.: *Numerical solution of partial differential equations in science and engineering*, New York: John Wiley and Sons, 1982.

LEES, M.: *Approximate solution of parabolic equations*, J.Soc.Indust. Appl.Math. 7, pp.167-183 (1959).

LEES, M.: *Alternating direction method for hyperbolic differential equations*, J.Soc.Indust.Appl.Math. 10, pp.610-616 (1962).

LEES, M.: *A linear three level difference scheme for quasi-linear parabolic equations*, Maths.Comp. 20, pp.516-522 (1966).

LERAT, A. and PEYRET, R.: *Proprietes dispersives et dissipatives d'une classe de schemes aux differences pour les systemes hyperboliques nonlineares*, Rech. Aerosp. 2, pp.61-79 (1975).

LOWAN, A.N.: *The operator approach to problems of stability and convergence of solutions of difference equations and the convergence of various iteration processes*, Scripta Mathematica Studies (1957).

MACCORMACK, R.W.: *Numerical solution of the interaction of a shock wave with a laminar boundary layer*, Proceedings of the Second International Conference on Numerical Methods in Fluid Dynamics in Lecture Notes in Physics (ed. HOLT, M.) Springer-Verlag, 1971.

MADSEN, N.K. and SINCOVEC, R.F.: *General software for partial differential equations in numerical methods for differential system*, pp.229-242 (ed. Lapidus, L. and Schiesser, W.E.), New York: Academic Press, 1976.

MARCHUK, G.I.: *Methods of numerical mathematics*, New York: Springer-Verlag, 1975.

MCGUIRE, G.R. and MORRIS, J. Ll.: *A Class of implicit second order accurate dissipative schemes for solving systems of conservation laws*, J.Comp.Phys. 14, pp.126-147 (1974).

MITCHELL, A.R.: *Computational methods in partial differential equations*, London: John Wiley and Sons 1969.

MITCHELL, A.R. and FAIRWEATHER, G.: *Improved forms of the alternating direction methods of Douglas, Peaceman and Rachford for solving parabolic and elliptic equations*, Numer.Maths. 6, pp.285-292 (1964).

MITCHELL, A.R. and GRIFFITHS, D.F.: *The finite difference methods in partial differential equations*, Chichester: John Wiley and Sons, 1980.

MITCHELL, A.R. and PEARCE, R.P.: *Explicit difference methods for solving the cylindrical heat conduction equation*, Maths.Comp. 17, pp.426-432 (1963).

MORRIS, J. Ll. and GREIG, I.S.: *A hopscotch method for the Kortweg-de-Vries Equation*, J.Comp.Phys., 20, pp.64-80 (1976).

NOBLE, B.: *Applied linear algebra*, Englewood Cliffs: Prentice-Hall Inc. 1969.

O'BRIEN, G.G., HYMAN, M.A. and KAPLAN, S.: *Numerical solution of partial differential equations*, J.Math.Phys. 29, pp.223-251 (1951).

PEACEMAN, D.W.: *Fundamentals of numerical reservoir simulation*, New York: Elsevier Scientific Publishing Company, 1977.

PEACEMAN, D.W. and RACHFORD, H.H.: *The numerical solution of parabolic and elliptic differential equations*, J.Soc. Indust.Appl. Maths. 3, pp.28-41 (1955).

PEACEMAN, D.W. and RACHFORD, H.H.: *Numerical calculation of multi-dimensional miscible displacement*, Soc. Petroleum Engineering Journal 24, pp. 327-338 (1962).

POLAK, S.J.: *An increased accuracy scheme for diffusion equations in cylindrical coordinates*, J.Inst.Maths.Applics. 14, pp.197-201 (1974).

RAMOS, J.I.: *Numerical solution of reactive-diffusive systems, Part 1: Explicit methods*, Intern.Jour.Comp.Maths. 18, pp.43-66 (1985), *Part 2: Methods of lines and implicit algorithms*, Intern.Jour.Comp.Maths. 18, pp. 141-162 (1985), *Part 3: Time linearisation and operator-splitting techniques*, Intern.Jour. Comp.Maths. 18, pp. 289-310 (1985).

REKTORYS, K.: *Survey of applicable mathematics*, London: Iliffe Books Ltd. 1969.

RICHTMYER, R.D. and MORTON, K.W.: *Difference methods for initial value problems (2nd Edition)*, New York: Interscience Publishers, 1967.

ROBERTS, K.V. and WEISS, N.O.: *Convective difference schemes*, Maths. Comp. 20, pp. 272-299 (1966).

RUBIN, E.L. and BURNSTEIN, S.Z.: *Difference methods for the inviscid and viscous equations of a compressible gas*, J.Comp.Phys. 2, pp. 178-196 (1967).

SANUGI, B.: *Private Communication*, 1983.

SAULEV, V.K.: *Integration of equations of parabolic type by the method of nets*, Oxford: Pergamon Press, 1964.

SHANKAR, V. and ANDERSON, D.: *Numerical solutions of supersonic corner flow*, J.Comp.Phys. 17, pp.160-180 (1975).

SMITH, G.D.: *Numerical solution of partial differential equations: finite difference methods (2nd Edition)*, Oxford: Clarendon Press, 1979.

STEPHENSON, G.: *An introduction to partial differential equations for science students (2nd Edition)*, London: Longman, 1978.

STEPHENSON, G.: *Mathematical methods for science students (2nd Edition)*, London: Longman, 1982.

TURKEL, E.: *Phase error and stability of second order methods for hyperbolic problems I*, J.Comp.Phys. 15, pp. 226-250 (1974).

VARGA, R.S.: *Matrix iterative analysis*, London: Prentice-Hall International, 1962.

VEMURI, V. and KARPUS, W.J.: *Digital computer treatment of partial differential equations*, Englewood Cliffs: Prentice-Hall, Inc. 1981.

VON ROSENBERG, D.U.: *Methods for the numerical solution of partial differential equations*, New York: American Elsevier Publishing Company, Inc. 1969.

WACHSPRESS, E.L.: *Iterative solution of elliptic systems*, Englewood Cliffs: Prentice-Hall, Inc. 1966.

WARMING, R.F., KUTLER, P. and LOMAX, H.: *Second and third order non-centred schemes for non-linear hyperbolic equations*, AIAA J. 11, pp. 189-195 (1973).

YANENKO, N.N.: *The method of fractional steps*, New York: Springer-Verlag, 1971.

YATES, R.C.: *Differential equations (1st Edition)*, New York: McGraw-Hill Book Company, Inc. 1952.

YOUSIF, W.S.: *New block iterative methods for the numerical solution of boundary value problems*, Ph.D. Thesis, Loughborough University of Technology, 1984.

APPENDIX

SOME SELECTED COMPUTER PROGRAMS

```
c     GE methods for parabolic problems with cylindrical symmetry.
c     This problem is taken from Mitchell and Pearce and even number
c     of intervals are used. The NAG subroutine generates Bessel values.
c
      library 'nagf'
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/p,q/bl2/u,ut/bl3/a1,a2
      dimension p(15),q(15),u(15,210),ut(15,210)
      open (unit=5, file='dcopeven4c',form='formatted')
      do 1000 lr=1,2
      ifail=1
      alpha=1.0
      beta=2.405
      read(5,*) nc,tl,sl,h,fk,mp1,np1,inj
c     note that rl  denotes lambda or mesh ratio
      rl=fk/(h**2)
      a3=1.0-2.0*rl
      a2=(alpha+1.0)*rl
      a1=2.0*a2
      n=np1-1
      ns=np1-2
      nd=np1-4
      m=mp1-1
      do 10 i=2,m
      fi=i
      p(i)=(1.0-alpha/(2.0*(fi-1.0)))*rl
   10 q(i)=(1.0+alpha/(2.0*(fi-1.0)))*rl
      write(1,120) alpha,tl,sl,h,fk,rl
c     define initial and boundary values
      do 1 i=1,mp1
      fi=i
      x=beta*(fi-1.0)*h
      bessel=s17aef(x,ifail)
      u(i,1)=bessel
      do 1 j=1,np1
      fj=j
      x1=-(beta**2)*(fj-1.0)*fk
      u(mp1,j)=0.0
    1 ut(i,j)=bessel*exp(x1)
c     solve the given differential equation by grouping points as
c     a result of coupling the relevant difference equations
      write(1,100)
      write(1,105)
c     fully explicit scheme
      if (nc .eq. 1) then
          write(1,104)
          do 32 j=1,n
          u(2,j+1)=p(2)*u(1,j)+q(2)*u(3,j)+a3*u(2,j)
          u(1,j+1)=(u(1,j)+a1*u(2,j+1))/(1.0+a1)
          do 32 i=2,m
   32     u(i,j+1)=(1.0-2.0*rl)*u(i,j)+p(i)*u(i-1,j)+q(i)*u(i+1,j)
          call output(u,1,mp1,1,np1,inj)
          write(1,140)
          call abserr(mp1,np1)
          call averag(u,aver,1,m,1,np1,1,inj)
c     group explicit with right ungrouped point (ger) scheme
      else if (nc .eq. 2) then
          write(1,106)
          do 31 j=1,n
          u(1,j+1)=u(1,j)+a1*(u(2,j)-u(1,j))
```

```
   31     call ger(m,mp1,j)
          call output(u,1,mp1,1,np1,inj)
          write(1,140)
          call abserr(mp1,np1)
          call averag(u,aver,1,m,1,np1,1,inj)
c     group explicit with left ungrouped point (gel) scheme
      else if (nc .eq. 3) then
          write(1,107)
          do 40 j=1,n
          u(1,j+1)=((1.0+p(2))*(1.0-a2)*u(1,j)+(2.0+p(2)-q(2))
     1    *a2*u(2,j)+a2*q(2)*u(3,j))/(1.0+p(2)+a2)
   40     call gel(m,j)
          call output(u,1,mp1,1,np1,inj)
          write(1,140)
          call abserr(mp1,np1)
          call averag(u,aver,1,m,1,np1,1,inj)
c     single alternating group explicit (sage) scheme
      else if (nc .eq. 4) then
          write(1,108)
          call esage(m,mp1,ns)
          call output(u,1,mp1,1,np1,inj)
          write(1,140)
          call abserr(mp1,np1)
          call averag(u,aver,1,m,1,np1,1,inj)
c     double alternating group explicit (dage) scheme
      else if (nc .eq. 5) then
          write(1,109)
          call edage(m,mp1,nd)
          call output(u,1,mp1,1,np1,inj)
          write(1,140)
          call abserr(mp1,np1)
          call averag(u,aver,1,m,1,np1,1,inj)
      end if
c     formats
  100 format(/,'2nd order parabolic diffusion equation with even number
     1of intervals',/)
  104 format(/,'using (a) fully explicit scheme',/)
  105 format(/,'soln of the coupled difference equations is',/)
  106 format(/,'using (b) ger scheme',/)
  107 format(/,'using (c) gel scheme',/)
  108 format(/,'using (d) sage scheme',/)
  109 format(/,'using (e) dage scheme',/)
  120 format('alpha=',d20.10/
     1        'max value in the t-direction=',d20.10/
     1        'max value in the x-direction=',d20.10/
     1        'increment h along the x-axis=',d20.10/
     1        ''increment k along the t-axis=',d20.10/
     1        'lambda=',d20.10)
  130 format(/,'theoretical soln is given by',/)
  140 format(/,'the absolute error at each mesh point is',/)
  141 format(/,'the percentage error at each mesh point is',/)
 1001 continue
 1000 continue
      write(1,130)
      call output(ut,1,mp1,1,np1,inj)
      call exit
      end
c
c
c
```

```
c
c
      subroutine abserr(l,n)
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl2/a,b
      dimension a(15,210),b(15,210)
      do 1 i=1,l
      do 1 j=1,n
    1 a(i,j)=abs(a(i,j)-b(i,j))
      return
      end
c
c
c
c
c
      subroutine percer(l,n)
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl2/a,b
      dimension a(15,210),b(15,210)
      do 1 i=1,l
      do 1 j=1,n
    1 a(i,j)=a(i,j)*100/abs(b(i,j))
      return
      end
c
c
c
c
c
      subroutine esage(m,mp1,ns)
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/p,q/bl2/u,ut/bl3/a1,a2
      dimension p(15),q(15),u(15,210),ut(15,210)
      do 1000 j=1,ns,2
      u(1,j+1)=u(1,j)+a1*(u(2,j)-u(1,j))
      call ger(m,mp1,j)
      u(1,j+2)=((1.0+p(2))*(1.0-a2)*u(1,j+1)+(2.0+p(2)-q(2))
     1*a2*u(2,j+1)+a2*q(2)*u(3,j+1))/(1.0+p(2)+a2)
 1000 call gel(m,j+1)
      return
      end
c
c
c
c
c
      subroutine edage(m,mp1,nd)
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/p,q/bl2/u,ut/bl3/a1,a2
      dimension p(15),q(15),u(15,210),ut(15,210)
      do 1000 j=1,nd,4
      u(1,j+1)=u(1,j)+a1*(u(2,j)-u(1,j))
      call ger(m,mp1,j)
      do 999 k=1,2
      u(1,j+k+1)=((1.0+p(2))*(1.0-a2)*u(1,j+k)+(2.0+p(2)-q(2))
     1*a2*u(2,j+k)+a2*q(2)*u(3,j+k))/(1.0+p(2)+a2)
  999 call gel(m,j+k)
      u(1,j+4)=u(1,j+3)+a1*(u(2,j+3)-u(1,j+3))
 1000 call ger(m,mp1,j+3)
```

```
      return
      end
c
c
c
c
c
      subroutine output(a,i1,i2,j1,j2,inj)
      implicit real*8(a-h,o-z),integer*2(i-n)
      dimension a(15,210)
      do 1000 j=j1,j2,inj
      write(1,1) j
 1000 write(1,2) (a(i,j),i=i1,i2)
    1 format('j=',i3)
    2 format(6d20.10/)
      return
      end
c
c
c
c
c
      subroutine averag(a,aver,i1,i2,j1,j2,icount,inj)
      implicit real*8(a-h,o-z),integer*2(i-n)
      dimension a(15,210)
      if (icount .eq. 0) then
          i3=i2
      else if (icount .eq. 1) then
          i3=i2
      end if
      do 1000 j=j1,j2,inj
      aver=0.
      write(1,1) j
      write(1,2) (a(i,j),i=i1,i2)
      do 1001 i=i1,i2
 1001 aver=aver+a(i,j)
      aver=aver/float(i3)
      write(1,3) aver
 1000 continue
    1 format('j=',i3)
    2 format(6d20.10/)
    3 format(/,'average of all errors=',d20.10)
      return
      end
c
c
c
c
c
      subroutine ger(m,mp1,j)
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/p,q/bl2/u,ut/bl3/a1,a2
      dimension p(15),q(15),u(15,210),ut(15,210)
      do 4 i=3,mp1,2
      if (i .eq. mp1) then
              u(m,j+1)=(p(m)*u(m-1,j)+(1.0-p(m))*u(m,j)
     1        +q(m)*u(m+1,j+1))/(1.0+q(m))
          else
              u(i-1,j+1)=((1.0+p(i))*p(i-1)*u(i-2,j)+(1.0+p(i))
     1        *(1.0-p(i-1))*u(i-1,j)+q(i-1)*(1.0-q(i))*u(i,j)
```

```
     1          +q(i)*q(i-1)*u(i+1,j))/(1.0+p(i)+q(i-1))
                u(i,j+1)=(p(i)*p(i-1)*u(i-2,j)+p(i)*(1.0-p(i-1))
     1          *u(i-1,j)+(1.0+q(i-1))*(1.0-q(i))*u(i,j)+(1.0+q(i-1))
     1          *q(i)*u(i+1,j))/(1.0+p(i)+q(i-1))
      end if
    4 continue
      return
      end
c
c
c
c
c
      subroutine gel(m,j)
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/p,q/bl2/u,ut/bl3/a1,a2
      dimension p(15),q(15),u(15,210),ut(15,210)
      do 7 i=2,m,2
      if (i .eq. 2) then
                u(2,j+1)=(p(2)*(1.0-a2)*u(1,j)+((1.0+a2)*(1.0-q(2))
     1          +p(2)*a2)*u(2,j)+(1.0+a2)*q(2)*u(3,j))/(1.0+p(2)+a2)
      else
                u(i-1,j+1)=((1.0+p(i))*p(i-1)*u(i-2,j)+(1.0+p(i))
     1          *(1.0-p(i-1))*u(i-1,j)+q(i-1)*(1.0-q(i))*u(i,j)
     1          +q(i)*q(i-1)*u(i+1,j))/(1.0+p(i)+q(i-1))
                u(i,j+1)=(p(i)*p(i-1)*u(i-2,j)+p(i)*(1.0-p(i-1))
     1          *u(i-1,j)+(1.0+q(i-1))*(1.0-q(i))*u(i,j)+(1.0+q(i-1))
     1          *q(i)*u(i+1,j))/(1.0+p(i)+q(i-1))
      end if
    7 continue
      return
      end
```

```
c     The Ramos non-linear reaction-diffusion problem is solved by
c     means of the AGE algorithm which employs the fourth time
c     linearisation technique. Odd number of internal points are used.
c
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/u,u1,u2,x
      dimension u(901),u1(901),u2(901),x(901)
      open (unit=5, file='dt4tl',form='formatted')
      v=1.0/sqrt(2.0)
      do 1000 lr=1,2
      read(5,*) tl,sl1,sl2,h,fk,r,m,n,ic1,ic2,omega,eps
      nt1=n-1
      data kmax/30/
      rl=fk/(h**2)
c     input data for matrix of coefficients
c
      d=(2.0-omega)*r
c     define initial values
c
      do 1 i=1,m
      fi=i
      u(i)=1.0/(1.0+exp(v*((fi)*h-50.0)))
    1 x(i)=u(i)
      write(1,3)
      if (omega .eq. 0.0) then
         write(1,8)
      else
         write(1,9)
      end if
      write(1,4) tl,sl1,sl2,h,fk,rl,r,m,n,eps
      do 2 jt=1,nt1
      fjt=jt
      t1=fjt*fk
      jt2=jt+1
      do 28 k=1,kmax
      ivct=1
c     define u(k+0.5)=u1(i)
c      equations for
c              u(k+0.5)=(g1+ri)<-1>((ri-g2)u(k)+f)
c
      a1=2.0*rl+5.0*(2.0-fk*x(1)*(2.0-3.0*x(1)))/6.0
      b1=-rl+(1.0-0.5*fk*x(2)*(2.0-3.0*x(2)))/6.0
      c1=-rl+(1.0+0.5*fk)/6.0
      s1=r-a1/2.0
      r1=r+a1/2.0
      f1=(5.0*(2.0-fk*x(1)*(2.0-3.0*x(1)))/6.0-2.0*rl)
     1*x(1)+((1.0-0.5*fk*x(2)*(2.0-3.0*x(2)))/6.0+rl)
     2*x(2)+fk*(10.0*(x(1)**2)*(1.0-x(1))+(x(2)**2)
     3*(1.0-x(2)))/6.0+2.0*rl
      u1(1)=(s1*u(1)-b1*u(2)+f1)/r1
      do 60 i=2,m-3,2
      a1=2.0*rl+5.0*(2.0-fk*x(i)*(2.0-3.0*x(i)))/6.0
      a2=2.0*rl+5.0*(2.0-fk*x(i+1)*(2.0-3.0*x(i+1)))/6.0
      b1=-rl+(1.0-0.5*fk*x(i+1)*(2.0-3.0*x(i+1)))/6.0
      b2=-rl+(1.0-0.5*fk*x(i+2)*(2.0-3.0*x(i+2)))/6.0
      c1=-rl+(1.0-0.5*fk*x(i-1)*(2.0-3.0*x(i-1)))/6.0
      c2=-rl+(1.0-0.5*fk*x(i)*(2.0-3.0*x(i)))/6.0
      r1=r+a1/2.0
      r2=r+a2/2.0
```

```
      s1=r-a1/2.0
      s2=r-a2/2.0
      p1=1.0-0.5*fk*x(i-1)*(2.0-3.0*x(i-1))
      p2=2.0-fk*x(i)*(2.0-3.0*x(i))
      p3=1.0-0.5*fk*x(i+1)*(2.0-3.0*x(i+1))
      p4=(x(i-1)**2)*(1.0-x(i-1))
      p5=(x(i)**2)*(1.0-x(i))
      p6=(x(i+1)**2)*(1.0-x(i+1))
      p10=1.0-0.5*fk*x(i)*(2.0-3.0*x(i))
      p20=2.0-fk*x(i+1)*(2.0-3.0*x(i+1))
      p30=1.0-0.5*fk*x(i+2)*(2.0-3.0*x(i+2))
      p60=(x(i+2)**2)*(1.0-x(i+2))
      f1=(p1/6.0+rl)*x(i-1)+(5.0*p2/6.0-2.0*rl)*x(i)+(p3/6.0+rl)*
     1x(i+1)+fk*(10.0*p5+p4+p6)/6.0
      f2=(p10/6.0+rl)*x(i)+(5.0*p20/6.0-2.0*rl)*x(i+1)+(p30/6.0+rl)*
     1x(i+2)+fk*(10.0*p6+p5+p60)/6.0
      alpha=r1*r2-b1*c2
      ss1=-c1*u(i-1)+s1*u(i)+f1
      ss2=s2*u(i+1)-b2*u(i+2)+f2
      u1(i)=(r2*ss1-b1*ss2)/alpha
   60 u1(i+1)=(-c2*ss1+r1*ss2)/alpha
      a1=2.0*rl+5.0*(2.0-fk*x(m-1)*(2.0-3.0*x(m-1)))/6.0
      a2=2.0*rl+5.0*(2.0-fk*x(m)*(2.0-3.0*x(m)))/6.0
      b1=(1.0-0.5*fk*x(m)*(2.0-3.0*x(m)))/6.0-rl
      c1=(1.0-0.5*fk*x(m-2)*(2.0-3.0*x(m-2)))/6.0-rl
      c2=(1.0-0.5*fk*x(m-1)*(2.0-3.0*x(m-1)))/6.0-rl
      r1=r+a1/2.0
      r2=r+a2/2.0
      s1=r-a1/2.0
      s2=r-a2/2.0
      p1=1.0-0.5*fk*x(m-2)*(2.0-3.0*x(m-2))
      p2=2.0-fk*x(m-1)*(2.0-3.0*x(m-1))
      p3=1.0-0.5*fk*x(m)*(2.0-3.0*x(m))
      p4=(x(m-2)**2)*(1.0-x(m-2))
      p5=(x(m-1)**2)*(1.0-x(m-1))
      p6=(x(m)**2)*(1.0-x(m))
      p10=1.0-0.5*fk*x(m-1)*(2.0-3.0*x(m-1))
      p20=2.0-fk*x(m)*(2.0-3.0*x(m))
      f1=(p1/6.0+rl)*x(m-2)+(5.0*p2/6.0-2.0*rl)*x(m-1)+(p3/6.0+rl)
     1*x(m)+fk*(10.0*p5+p4+p6)/6.0
      f2=(p10/6.0+rl)*x(m-1)+(5.0*p20/6.0-2.0*rl)*x(m)
     1+fk*(10.0*p6+p5)/6.0
      alpha=r1*r2-b1*c2
      ss1=-c1*u(m-2)+s1*u(m-1)+f1
      ss2=s2*u(m)+f2
      u1(m-1)=(r2*ss1-b1*ss2)/alpha
      u1(m)=(-c2*ss1+r1*ss2)/alpha
c
c     define u(k+1)=u2(i)
c     equations for
c       u(k+1)=(g2+ri)<-1>((g2-(1-w)ri)u(k)+(2-w)ru(k+0.5))
      do 61 i=1,m-2,2
      a1=2.0*rl+5.0*(2.0-fk*x(i)*(2.0-3.0*x(i)))/6.0
      a2=2.0*rl+5.0*(2.0-fk*x(i+1)*(2.0-3.0*x(i+1)))/6.0
      b1=-rl+(1.0-0.5*fk*x(i+1)*(2.0-3.0*x(i+1)))/6.0
      c2=-rl+(1.0-0.5*fk*x(i)*(2.0-3.0*x(i)))/6.0
      r1=r+a1/2.0
      r2=r+a2/2.0
      q1=a1/2.0-(1.0-omega)*r
      q2=a2/2.0-(1.0-omega)*r
```

```
      alpha=r1*r2-b1*c2
      ss1=q1*u(i)+b1*u(i+1)+d*u1(i)
      ss2=c2*u(i)+q2*u(i+1)+d*u1(i+1)
      u2(i)=(r2*ss1-b1*ss2)/alpha
   61 u2(i+1)=(-c2*ss1+r1*ss2)/alpha
      a1=2.0*r1+5.0*(2.0-fk*x(m)*(2.0-3.0*x(m)))/6.0
      r1=r+a1/2.0
      q1=a1/2.0-(1.0-omega)*r
      u2(m)=(q1*u(m)+d*u1(m))/r1
c     generate solns on each time level. set ivct=1 for successful convergence
c     and 0 otherwise. begin iterative process.
      do 27 i=1,m
      if(abs(u2(i)-u(i))-eps) 27,27,29
   29 ivct=0
   27 continue
      do 32 i=1,m
   32 u(i)=u2(i)
      if (ivct .ne. 1) go to 28
      do 80 i=1,m
   80 x(i)=u2(i)
c     theoretical solution
      do 34 i=1,m
      fi=i
   34 u1(i)=1.0/(1.0+exp(v*((fi*h-50.0)-v*t1)))
      do 320 jw=ic1,nt1,ic1
      if (jt .eq. jw) then
      write(1,11)jt2,t1
      write(1,12)k
      write(1,14) (u(i),i=2,m-1,ic2)
      write(1,5)
      write(1,14) (u1(i),i=2,m-1,ic2)
      write(1,96)
      call abserr(1,m)
      write(1,14) (u2(i),i=2,m-1,ic2)
      call averag(1,m)
      sum=0.0
      do 420 i=1,m
  420 sum=sum+(u(i)**2)*(1.0-u(i))
      vtrap=h*sum
      vper=abs(vtrap-v)*100.0/v
      write(1,144) v,vtrap,vper
      go to 320
      else
      go to 320
      end if
  320 continue
      go to 2
   28 continue
      write(1,37) kmax,jt2
      go to 1000
    2 continue
 1000 continue
   90 call exit
c     formats
c
    3 format(//,'Non-linear reaction-diffusion equation',/
     1          'with odd number of internal mesh points',/)
    4 format('maximum value in the t-direction=',d20.10/
     2       'minimum value in the x-direction=',d20.10/
     2       'maximum value in the x-direction=',d20.10/
```

```
     3      'increment h along the x-axis=',d20.10/
     5      'increment k along the y-axis=',d20.10/
     6      'lambda=',d20.10/
     7      'parameter r=',d20.10/
     8      'number of points used in the x-direction=',i5/
     1      'number of time levels=',i5/
     2      'accuracy:convergence criterion eps=',d20.10)
    5 format(/,'theoretical soln at selected points is given by',/)
    7 format(6d20.10)
    8 format('using the peaceman rachford variant',/)
    9 format('using the douglas rachford variant',/)
   11 format(/,'age iterative soln at time level n=',i3/
     1        'time t=',d20.10/
     1        'employing ramos implicit approximation',/
     2        'with 4th time linearisation - 4tl',/)
   12 format('method converges with k=',i4,' iterations',/)
   13 format(/,'the absolute error at each mesh point is',/)
   14 format(6d20.10)
   15 format(/,'the percentage error at each mesh point is',/)
   37 format('method fails to converge in',i4,'iterations',/
     1      'at time level n=',i3)
   96 format(/,'the absolute errors at selected points are',/)
   97 format(/,'the percentage errors at selected points are',/)
  144 format(/,'the steady state wave speed v=',d20.10/
     1        'the computed wave speed vtrap=',d20.10/
     2        'the percentage error of the wave speed is=',d20.10)
      end
c
c
c
c
c
      subroutine abserr(ic2,m)
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/u,u1,u2,x
      dimension u(901),u1(901),u2(901),x(901)
      do 1 i=1,m,ic2
    1 u2(i)=abs(u1(i)-u(i))
      return
      end
c
c
c
c
c
      subroutine averag(ic2,m)
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/u,u1,u2,x
      dimension u(901),u1(901),u2(901),x(901)
      aver=0.
      m1=((m-1)/ic2)+1
      do 1001 i=1,m,ic2
 1001 aver=aver+u2(i)
      aver=aver/float(m1)
      write(1,3) aver
    1 format('j=',i3)
    2 format(6d20.10/)
    3 format(/,'average of all errors=',d20.10)
      return
      end
```

```
c
c
c
c
c
      subroutine percer(ic2,m)
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/u,u1,u2,x
      dimension u(901),u1(901),u2(901),x(901)
      do 1 i=1,m,ic2
    1 u2(i)=u2(i)*100/abs(u1(i))
      return
      end
```

```
c
c
c
c
c
      subroutine percer(ic2,m)
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/u,u1,u2,x
      dimension u(901),u1(901),u2(901),x(901)
      do 1 i=1,m,ic2
    1 u2(i)=u2(i)*100/abs(u1(i))
      return
      end
```

```
c     This two-dimensional heat problem is taken from Gourlay and MacGuire
c     and is solved using the AGE algorithm. The size of the coefficient
c     matrix is assumed to be odd.
c
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/u,u1,u2,v
      dimension u(110,110),u1(110,110),u2(110,110),v(110,110)
      open (unit=5, file='dage2d2', form='formatted')
      do 1000 lr=1,9
      read(5,*) tl,dt,sx,sy.h,fk,r,m,n,nt,ic1,ic2,icount
      nt1=nt-1
      data eps,kmax/0.00000001,40/
      rl=dt/(h**2)
c     input data for matrix of coefficients
c
c     FULLY IMPLICIT METHOD
      if (icount .eq. 1) then
         write(1,300)
         th=1.0
c     CRANK NICOLSON METHOD
      else if (icount .eq. 2) then
         write(1,301)
         th=0.5
      end if
      th1=rl*th
      th2=rl*(1.0-th)
      th3=th1+th2
      c=1.0+4.0*th1
      a1=-th1
      d=1.0-4.0*th2
      e1=th2
      r1=r+c/4.0
      r2=r1*c/4.0-a1**2
      r3=a1*r
      r4=c/(4.0*r1)
      r5=r/r1
      del=r1**2-a1**2
      p1=r1-2.0*c
      p2=-2.0*a1
      p3=2.0*d
      p4=2.0*e1
c     define initial values u(i,j)
c
      do 1 j=1,n
      fj=j
      y=fj*h
      do 1 i=1,m
      fi=i
      x=fi*h
      u(i,j)=(sin(x)*sin(y))+(x**2)+(y**2)
    1 v(i,j)=u(i,j)
      write(1,3)
      write(1,9)
      write(1,4) tl,dt,sx,h,sy,fk,rl,r.m,n,nt,eps
      do 2 jt=1,nt1
      fjt=jt
      f1=fjt*dt
      f2=(fjt-0.5)*dt
      f3=(fjt-1.0)*dt
      jt2=jt+1
```

```
      do 28 k=1,kmax
      ivct=1
c     approximation parallel to x-axis for level (n+0.5)
c     define u(k+0.25)=u1(i,j)
c     1. equations for
c             U1(K+0.25)=C1<-1>(D1*U1(K)+E1*U2(K)+F1*U1[N]+F2*U2[N])
c
      u1(1,1)=(p1*u(1,1)+p2*(u(2,1)+u(1,2))+p3*v(1,1)
     2+2.0*(dt*(((sin(h))**2)*exp(-f2)-4.0)+(2.0*(h**2))*th3)
     2+p4*(v(2,1)
     1+v(1,2)))/r1
      do 60 i=2,m-3,2
      fi=i
      x=fi*h
      x1=(fi+1.0)*h
      u1(i,1)=(r1*(p1*u(i,1)+p2*(u(i-1,1)+u(i,2))+p3*v(i,1)
     1+2.0*(dt*(sin(x)*sin(h)*exp(-f2)-4.0)+(x**2)*th3)
     1+p4*
     1(v(i-1,1)+v(i,2)+v(i+1,1))-a1*u(i+1,1))-a1*(p1*u(i+1,1)
     2+p2*(u(i+1,2)+u(i+2,1))+p3*v(i+1,1)
     1+2.0*(dt*(sin(x1)*sin(h)*exp(-f2)-4.0)+(x1**2)*th3)
     1+p4*(v(i,1)+v(i+1,2)
     3+v(i+2,1))-a1*u(i,1)))/del
   60 u1(i+1,1)=(-a1*(p1*u(i,1)+p2*(u(i-1,1)+u(i,2))+p3*v(i,1)
     1+2.0*(dt*(sin(x)*sin(h)*exp(-f2)-4.0)+(x**2)*th3)+p4
     1*(v(i-1,1)+v(i,2)+v(i+1,1))-a1*u(i+1,1))+r1*(p1*u(i+1,1)
     2+p2*(u(i+1,2)+u(i+2,1))+p3*v(i+1,1)
     1+2.0*(dt*(sin(x1)*sin(h)*exp(-f2)-4.0)+(x1**2)*th3)
     1+p4*(v(i,1)+v(i+1,2)
     3+v(i-2,1))-a1*u(i,1)))/del
      xm1=float(m-1)*h
      xm=float(m)*h
      u1(m-1,1)=(r1*(p1*u(m-1,1)+p2*(u(m-2,1)+u(m-1,2))+p3*v(m-1,1)
     1+2.0*(dt*(sin(xm1)*sin(h)*exp(-f2)-4.0)+(xm1**2)*th3)
     1+p4*(v(m-2,1)+v(m-1,2)+v(m,1))-a1*u(m,1))-a1*(p1*u(m,1)+p2*u(m,2)
     1+p3*v(m,1)+2.0*(dt*(sin(xm)*sin(h)*exp(-f2)-4.0)
     2+((xm**2)+(1.0+h**2))*th3+sin(1.0)*sin(h)*(th2*exp(-f3)
     3+th1*exp(-f1)))
     2+p4*(v(m-1,1)+v(m,2))-a1*u(m-1,1)))/del
      u1(m,1)=(-a1*(p1*u(m-1,1)+p2*(u(m-2,1)+u(m-1,2))
     1+p3*v(m-1,1)+2.0*(dt*(sin(xm1)*sin(h)*exp(-f2)-4.0)
     2+(xm1**2)*th3)
     1+p4*(v(m-2,1)+v(m-1,2)+v(m,1))-a1*u(m,1))+r1*(p1*u(m,1)+p2*u(m,2)
     1+p3*v(m,1)+2.0*(dt*(sin(xm)*sin(h)*exp(-f2)-4.0)
     2+((xm**2)+(1.0+h**2))*th3+sin(1.0)*sin(h)*(th2*exp(-f3)
     3+th1*exp(-f1)))
     2+p4*(v(m-1,1)+v(m,2))-a1*u(m-1,1)))/del
c     2. equations for
c           UJ(K+0.25)=C2<-1>(E1*(UJ-1(K)+UJ+1(K))+D2*UJ(K)
c           +F2(UJ-1[N]+UJ+1[N])+F1*UJ[N]) J=2,4,...,N-1
      do 61 j=2,n-1,2
      fj=j
      y=fj*h
      u1(1,j)=(r1*(p1*u(1,j)+p2*(u(1,j-1)+u(1,j+1))+p3*v(1,j)
     1+2.0*(dt*(sin(h)*sin(y)*exp(-f2)-4.0)+th3*(y**2))+p4*
     1(v(1,j-1)+v(1,j+1)+v(2,j))-a1*u(2,j))-a1*(p1*u(2,j)+p2*(u(2,j-1)
     2+u(2,j+1)+u(3,j))+p3*v(2,j)
     2+2.0*dt*(sin(2.0*h)*sin(y)*exp(-f2)-4.0)
     3+p4*(v(1,j)+v(2,j-1)+v(2,j+1)+v(3,j))
     3-a1*u(1,j)))/del
```

```
   61 u1(2,j)=(-a1*(p1*u(1,j)+p2*(u(1,j-1)+u(1,j+1))+p3*v(1,j)
     1+2.0*(dt*(sin(h)*sin(y)*exp(-f2)-4.0)+th3*(y**2))+p4*
     2(v(1,j-1)+v(1,j+1)+v(2,j))-a1*u(2,j))+r1*(p1*u(2,j)+p2*(u(2,j-1)
     3+u(2,j+1)+u(3,j))+p3*v(2,j)
     4+2.0*dt*(sin(2.0*h)*sin(y)*exp(-f2)-4.0)
     5+p4*(v(1,j)+v(2,j-1)+v(2,j+1)+v(3,j))
     6-a1*u(1,j)))/del
      do 62 j=2,n-1,2
      fj=j
      y=fj*h
      do 62 i=3,m-2,2
      fi=i
      x=fi*h
      x1=(fi+1.0)*h
      u1(i,j)=(r1*(p1*u(i,j)+p2*(u(i-1,j)+u(i,j-1)+u(i,j+1))+p3*v(i,j)
     1+2.0*dt*(sin(x)*sin(y)*exp(-f2)-4.0)
     1+p4*(v(i-1,j)+v(i,j-1)+v(i,j+1)+v(i+1,j))-a1*u(i+1,j))-a1*(p1
     2*u(i+1,j)+p2*(u(i+1,j-1)+u(i+1,j+1)+u(i+2,j))+p3*v(i+1,j)
     2+2.0*dt*(sin(x1)*sin(y)*exp(-f2)-4.0)+p4
     3*(v(i,j)+v(i+1,j-1)+v(i+1,j+1)+v(i+2,j))-a1*u(i,j)))/del
   62 u1(i+1,j)=(-a1*(p1*u(i,j)+p2*(u(i-1,j)+u(i,j-1)+u(i,j+1))
     1+2.0*dt*(sin(x)*sin(y)*exp(-f2)-4.0)
     1+p3*v(i,j)+p4*(v(i-1,j)+v(i,j-1)+v(i,j+1)+v(i+1,j))-a1
     2*u(i+1,j))+r1*(p1*u(i+1,j)+p2*(u(i+1,j-1)+u(i+1,j+1)
     3+u(i+2,j))+p3*v(i+1,j)
     1+2.0*dt*(sin(x1)*sin(y)*exp(-f2)-4.0)
     2+p4*(v(i,j)+v(i+1,j-1)+v(i+1,j+1)
     4+v(i+2,j))-a1*u(i,j)))/del
      do 63 j=2,n-1,2
      fj=j
      y=fj*h
   63 u1(m,j)=(p1*u(m,j)+p2*(u(m-1,j)+u(m,j-1)+u(m,j+1))+p3*v(m,j)
     1+2.0*(dt*(sin(float(m)*h)*sin(y)*exp(-f2)-4.0)
     1+sin(1.0)*sin(y)*(th2*exp(-f3)+th1*exp(-f1))
     2+th3*(1.0+y**2))
     1+p4*(v(m-1,j)+v(m,j-1)+v(m,j+1)))/r1
c     3. for equations
c               UJ(K+0.25)=C1<-1>(E1*(UJ-1(K)+UJ+1(K))+D1UJ(K)
c                          +F2*(UJ-1[N]+UJ+1[N])+F1*UJ[N]) J=3,5,...,N-2
      do 64 j=3,n-2,2
      fj=j
      y=fj*h
      u1(1,j)=(p1*u(1,j)+p2*(u(1,j-1)+u(1,j+1)+u(2,j))+p3*v(1,j)
     1+2.0*(dt*(sin(h)*sin(y)*exp(-f2)-4.0)+(y**2)*th3)+p4
     1*(v(1,j-1)+v(1,j+1)+v(2,j)))/r1
      u1(m-1,j)=(r1*(p1*u(m-1,j)+p2*(u(m-2,j)+u(m-1,j-1)+u(m-1,j+1))
     1+2.0*dt*(sin(float(m-1)*h)*sin(y)*exp(-f2)-4.0)
     1+p3*v(m-1,j)+p4*(v(m-2,j)+v(m-1,j-1)+v(m-1,j+1)+v(m,j))-a1
     2*u(m,j))-a1*(p1*u(m,j)+p2*(u(m,j-1)+u(m,j+1))+p3*v(m,j)
     1+2.0*(dt*(sin(float(m)*h)*sin(y)*exp(-f2)-4.0)
     2+sin(1.0)*sin(y)*(th2*exp(-f3)+th1*exp(-f1))
     3+th3*(1.0+y**2))+p4
     3*(v(m,j-1)+v(m,j+1)+v(m-1,j))-a1*u(m-1,j)))/del
   64 u1(m,j)=(-a1*(p1*u(m-1,j)+p2*(u(m-2,j)+u(m-1,j-1)+u(m-1,j+1))
     1+2.0*dt*(sin(float(m-1)*h)*sin(y)*exp(-f2)-4.0)
     1+p3*v(m-1,j)+p4*(v(m-2,j)+v(m-1,j-1)+v(m-1,j+1)+v(m,j))-a1
     2*u(m,j))+r1*(p1*u(m,j)+p2*(u(m,j-1)+u(m,j+1))+p3*v(m,j)
     1+2.0*(dt*(sin(float(m)*h)*sin(y)*exp(-f2)-4.0)
     2+sin(1.0)*sin(y)*(th2*exp(-f3)+th1*exp(-f1))
     3+th3*(1.0+y**2))+p4
```

```
     3*(v(m,j-1)+v(m,j+1)+v(m-1,j))-a1*u(m-1,j)))/del
      do 65 j=3,n-2,2
      fj=j
      y=fj*h
      do 65 i=2,m-3,2
      fi=i
      x=fi*h
      x1=(fi+1.0)*h
      u1(i,j)=(r1*(p1*u(i,j)+p2*(u(i-1,j)+u(i,j-1)+u(i,j+1))+p3*v(i,j)
     2+2.0*dt*(sin(x)*sin(y)*exp(-f2)-4.0)
     1+p4*(v(i-1,j)+v(i,j-1)+v(i,j+1)+v(i+1,j))-a1*u(i+1,j))-a1*(p1
     2*u(i+1,j)+p2*(u(i+1,j-1)+u(i+1,j+1)+u(i+2,j))+p3*v(i+1,j)
     2+2.0*dt*(sin(x1)*sin(y)*exp(-f2)-4.0)
     3+p4*(v(i,j)+v(i+1,j-1)+v(i+1,j+1)+v(i+2,j))-a1*u(i,j)))/del
   65 u1(i+1,j)=(-a1*(p1*u(i,j)+p2*(u(i-1,j)+u(i,j-1)+u(i,j+1))+p3
     1*v(i,j)+2.0*dt*(sin(x)*sin(y)*exp(-f2)-4.0)
     2+p4*(v(i-1,j)+v(i,j-1)+v(i,j+1)+v(i+1,j))-a1*u(i+1,j))
     2+r1*(p1*u(i+1,j)+p2*(u(i+1,j-1)+u(i+1,j+1)+u(i+2,j))+p3*v(i+1,j)
     2+2.0*dt*(sin(x1)*sin(y)*exp(-f2)-4.0)
     3+p4*(v(i,j)+v(i+1,j-1)+v(i+1,j+1)+v(i+2,j))-a1*u(i,j)))/del
c     4. for equations
c               UN(K+0.25)=C1<-1>(E1*UN-1(K)+D1*UN(K)+F2*UN-1[N]
c                          +F1*UN[N])
      yn=float(n)*h
      u1(1,n)=(p1*u(1,n)+p2*(u(1,n-1)+u(2,n))+p3*v(1,n)
     1+2.0*(th3*((yn**2)+((h**2)+1.0))
     2+sin(h)*sin(1.0)*(th2*exp(-f3)+th1*exp(-f1))
     3+dt*(sin(h)*sin(yn)*exp(-f2)-4.0))+p4*(v(1,n-1)
     1+v(2,n)))/r1
      do 66 i=2,m-3,2
      fi=i
      x=fi*h
      x1=(fi+1.0)*h
      u1(i,n)=(r1*(p1*u(i,n)+p2*(u(i-1,n)+u(i,n-1))+p3*v(i,n)
     1+2.0*(sin(x)*sin(1.0)*(th2*exp(-f3)+th1*exp(-f1))
     2+(1.0+x**2)*th3
     1+dt*(sin(x)*sin(yn)*exp(-f2)-4.0))+p4
     1*(v(i-1,n)+v(i,n-1)+v(i+1,n))-a1*u(i+1,n))-a1*(p1*u(i+1,n)+p2
     2*(u(i+1,n-1)+u(i+2,n))+p3*v(i+1,n)
     2+2.0*(sin(x1)*sin(1.0)*(th2*exp(-f3)+th1*exp(-f1))
     2+(1.0+x1**2)*th3
     2+dt*(sin(x1)*sin(yn)*exp(-f2)-4.0))
     2+p4*(v(i,n)+v(i+1,n-1)
     3+v(i+2,n))-a1xu(i,n)))/del
   66 u1(i+1,n)=(-a1*(p1*u(i,n)+p2*(u(i-1,n)+u(i,n-1))+p3*v(i,n)
     1+2.0*(sin(x)*sin(1.0)*(th2*exp(-f3)+th1*exp(-f1))
     2+(1.0+x**2)*th3
     2+dt*(sin(x)*sin(yn)*exp(-f2)-4.0))+p4
     1*(v(i-1,n)+v(i,n-1)+v(i+1,n))-a1*u(i+1,n))+r1*(p1*u(i+1,n)+p2
     2*(u(i+1,n-1)+u(i+2,n))+p3*v(i+1,n)
     2+2.0*(sin(x1)*sin(1.0)*(th2*exp(-f3)+th1*exp(-f1))
     2+(1.0+x1**2)*th3
     1+dt*(sin(x1)*sin(yn)*exp(-f2)-4.0))
     2+p4*(v(i,n)+v(i+1,n-1)
     3+v(i+2,n))-a1*u(i,n)))/del
      xm1=float(m-1)*h
      xm=float(m)*h
      u1(m-1,n)=(r1*(p1*u(m-1,n)+p2*(u(m-2,n)+u(m-1,n-1))+p3*v(m-1,n)

     1+2.0*(sin(xm1)*sin(1.0)*(th2*exp(-f3)+th1*exp(-f1))
```

```
     2+(1.0+xm1**2)*th3
     2+dt*(sin(xm1)*sin(yn)*exp(-f2)-4.0))
     1+p4*(v(m-2,n)+v(m-1,n-1)+v(m,n))-a1*u(m,n))-a1*(p1*u(m,n)+p2
     2*u(m,n-1)+p3*v(m,n)
     2+2.0*(dt*(sin(xm)*sin(yn)*exp(-f2)-4.0)
     3+(th2*exp(-f3)+th1*exp(-f1))*sin(1.0)
     4*(sin(xm)+sin(yn))+(2.0+(xm**2)+(yn**2))*th3)
     2+p4*(v(m-1,n)+v(m,n-1))-a1*u(m-1,n)))/del
      u1(m,n)=(-a1*(p1*u(m-1,n)+p2*(u(m-2,n)+u(m-1,n-1))+p3*v(m-1,n)
     1+2.0*(sin(xm1)*sin(1.0)*(th2*exp(-f3)+th1*exp(-f1))
     2+(1.0+xm1**2)*th3
     2+dt*(sin(xm1)*sin(yn)*exp(-f2)-4.0))
     1+p4*(v(m-2,n)+v(m-1,n-1)+v(m,n))-a1*u(m,n))+r1*(p1*u(m,n)+p2
     2*u(m,n-1)+p3*v(m,n)
     2+2.0*(dt*(sin(xm)*sin(yn)*exp(-f2)-4.0)
     3+(th2*exp(-f3)+th1*exp(-f1))*sin(1.0)
     4*(sin(xm)+sin(yn))+(2.0+(xm**2)+(yn**2))*th3)
     2+p4*(v(m-1,n)+v(m,n-1))-a1*u(m-1,n)))/del
c     define u(k+0.5)=u2(i,j)
      do 36 j=1,n,2
      do 36 i=1,m-2,2
      u2(i,j)=(r2*u(i,j)+r3*u(i+1,j)+r*(r1*u1(i,j)-a1*u1(i+1,j)))/del
   36 u2(i+1,j)=(r3*u(i,j)+r2*u(i+1,j)+r*(-a1*u1(i,j)+r1*u1(i+1,j)))/del
      do 50 j=1,n,2
   50 u2(m,j)=r4*u(m,j)+r5*u1(m,j)
      do 16 j=2,n-1,2
   16 u2(1,j)=r4*u(1,j)+r5*u1(1,j)
      do 17 j=2,n-1,2
      do 17 i=2,m-1,2
      u2(i,j)=(r2*u(i,j)+r3*u(i+1,j)+r*(r1*u1(i,j)-a1*u1(i+1,j)))/del
   17 u2(i+1,j)=(r3*u(i,j)+r2*u(i+1,j)+r*(-a1*u1(i,j)+r1*u1(i+1,j)))/del
c     approximation parallel to y-axis for level (n+1)
c     define u(k+0.75)=u1(i,j) by overwriting
      do 18 i=1,m,2
   18 u1(i,1)=r4*u(i,1)+r5*u2(i,1)
      do 19 i=2,m-1,2
   19 u1(i,n)=r4*u(i,n)+r5*u2(i,n)
      do 20 i=1,m,2
      do 20 j=2,n-1,2
      u1(i,j)=(r2*u(i,j)+r3*u(i,j+1)+r*(r1*u2(i,j)-a1*u2(i,j+1)))/del
   20 u1(i,j+1)=(r3*u(i,j)+r2*u(i,j+1)+r*(-a1*u2(i,j)+r1*u2(i,j+1)))/del
      do 21 i=2,m-1,2
      do 21 j=1,n-2,2
      u1(i,j)=(r2*u(i,j)+r3*u(i,j+1)+r*(r1*u2(i,j)-a1*u2(i,j+1)))/del
   21 u1(i,j+1)=(r3*u(i,j)+r2*u(i,j+1)+r*(-a1*u2(i,j)+r1*u2(i,j+1)))/del
c     define u(k+1)=u2(i,j) by overwriting
      do 22 i=1,m,2
      do 22 j=1,n-2,2
      u2(i,j)=(r2*u(i,j)+r3*u(i,j+1)+r*(r1*u1(i,j)-a1*u1(i,j+1)))/del
   22 u2(i,j+1)=(r3*u(i,j)+r2*u(i,j+1)+r*(-a1*u1(i,j)+r1*u1(i,j+1)))/del
      do 23 i=1,m,2
   23 u2(i,n)=r4*u(i,n)+r5*u1(i,n)
      do 24 i=2,m-1,2
   24 u2(i,1)=r4*u(i,1)+r5*u1(i,1)
      do 25 i=2,m-1,2
      do 25 j=2,n-1,2
      u2(i,j)=(r2*u(i,j)+r3*u(i,j+1)+r*(r1*u1(i,j)-a1*u1(i,j+1)))/del
   25 u2(i,j+1)=(r3*u(i,j)+r2*u(i,j+1)+r*(-a1*u1(i,j)+r1*u1(i,j+1)))/del
c     generate solns on each time level. set ivct=1 for successful convergence
c     and 0 otherwise. begin iterative process.
```

```
      do 26 j=1,n
      do 27 i=1,m
      if(abs(u2(i,j)-u(i,j))-eps) 27,29,29
   29 ivct=0
   27 continue
   26 continue
      do 32 j=1,n
      do 32 i=1,m
   32 u(i,j)=u2(i,j)
      if (ivct .ne. 1) go to 28
      do 80 j=1,n
      do 80 i=1,m
   80 v(i,j)=u2(i,j)
      if(jt2 .eq. nt) then
      write(1,11)jt2,f1
      write(1,12)k
      do 33 i=1,m,ic2
      fi=i
      x=fi*h
      write(1,6)i,x
   33 write(1,14) (u(i,j),j=1,n,ic1)
c     theoretical solution
      do 34 j=1,n
      fj=j
      y=fj*h
      do 34 i=1,m
      fi=i
      x=fi*h
   34 u1(i,j)=sin(x)*sin(y)*exp(-f1)+(x**2)+(y**2)
      write(1,5)
      do 35 i=1,m,ic2
      fi=i
      x=fi*h
      write(1,6)i,x
   35 write(1,14) (u1(i,j),j=1,n,ic1)
      write(1,96)
      call abserr(1,1,m,n)
      do 99 i=1,m,ic2
      fi=i
      x=fi*h
      write(1,6)i,x
   99 write(1,14) (u2(i,j),j=1,n,ic1)
      call averag(1,1,m,n)
      call percer(1,1,m,n)
      write(1,97)
      do 95 i=1,m,ic2
      fi=i
      x=fi*h
      write(1,6)i,x
   95 write(1,14) (u2(i,j),j=1,n,ic1)
      call averag(1,1,m,n)
      go to 2
      else
      go to 2
      end if
   28 continue
      write(1,37) kmax,jt2
      go to 1000
    2 continue
 1000 continue
```

```
   90 call exit
c     formats
c
    3 format(//,'Second order 2 dimensional parabolic equation',/
     1          'with a source term g(x,y,t)',/)
    4 format('maximum value in the t-direction=',d20.10/
     1       'increment dt along the t-axis=',d20.10/
     2       'maximum value in the x-direction=',d20.10/
     3       'increment h along the x-axis=',d20.10/
     4       'maximum value in the y-direction=',d20.10/
     5       'increment k along the y-axis=',d20.10/
     6       'lambda=',d20.10/
     7       'parameter r=',d20.10/
     8       'number of points used in the x-direction=',i5/
     9       'number of points used in the y-direction=',i5/
     1       'number of time levels=',i5/
     2       'ACCURACY: convergence criterion eps=',d20.10)
    5 format(/,'Theoretical soln at selected points is given by',/)
    6 format('i=',i3/
     1       'x=',d20.10/)
    7 format(6d20.10)
    8 format('using the PEACEMAN RACHFORD variant',/)
    9 format('using the DOUGLAS RACHFORD variant',/)
   11 format(/,'AGE iterative soln at time level n=',i3/
     1         'time t=',d20.10)
   12 format('method converges with k=',i4,' iterations',/)
   13 format(/,'the absolute error at each mesh point is',/)
   14 format(6d20.10)
   15 format(/,'the percentage error at each mesh point is',/)
   37 format('method fails to converge in',i4,'iterations',/
     1       'at time level n=',i3)
   96 format(/,'the absolute errors at selected points are',/)
   97 format(/,'the percentage errors at selected points are',/)
  300 format(/,'FULLY IMPLICIT PARABOLIC',/)
  301 format(/,'CRANK-NICOLSON PARABOLIC',/)
      end
c
c
c
c
c
      subroutine abserr(ic1,ic2,m,n)
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/u,u1,u2,v
      dimension u(110,110),u1(110,110),u2(110,110),v(110,110)
      do 1 j=1,n,ic1
      do 1 i=1,m,ic2
    1 u2(i,j)=abs(u1(i,j)-u(i,j))
      return
      end
c
c
c
c
c
      subroutine averag(ic1,ic2,m,n)
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/u,u1,u2,v
      dimension u(110,110),u1(110,110),u2(110,110),v(110,110)
      aver=0.
```

```
      m1=((m-1)/ic2)+1
      n1=((n-1)/ic1)+1
      do 1001 j=1,n,ic1
      do 1001 i=1,m,ic2
 1001 aver=aver+u2(i,j)
      aver=aver/float(m1*n1)
      write(1,3) aver
    1 format('j=',i3)
    2 format(6d20.10/)
    3 format(/,'average of all errors=',d20.10)
      return
      end
c
c
c
c
c
      subroutine percer(ic1,ic2,m,n)
      implicit real*8(a-h,o-z),integer*2(i-n)
      common /bl1/u,u1,u2,v
      dimension u(110,110),u1(110,110),u2(110,110),v(110,110)
      do 1 j=1,n,ic1
      do 1 i=1,m,ic2
    1 u2(i,j)=u2(i,j)*100/abs(u1(i,j))
      return
      end
```