THE NUMERICAL SOLUTION OF BOUNDARY VALUE PROBLEMS IN PARTIAL DIFFERENTIAL EQUATIONS

Patrick Keast

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

1967

THE NUMERICAL SOLUTION OF
BOUNDARY VALUE PROBLEMS
IN
PARTIAL DIFFERENTIAL EQUATIONS.

A thesis presented by Patrick Keast, B.Sc.,
to
the University of St. Andrews,
in application for the degree of
Doctor of Philosophy.

DECLARATION.

I declare that the following thesis is a record of research work carried out by me, that the thesis is my own composition, and that it has not previously been presented in application for a higher degree.

PREFACE.

In October 1960, I matriculated at the University of Edinburgh, and read for a degree in Mathematical Science. In July 1964, I graduated with First Class Honours in Mathematical Science. In October 1964, I was admitted, under Ordinance 16, as a part-time Research Student in St. Salvator's College Mathematics Department, in the University of St. Andrews, under the supervision of Dr. A. R. Mitchell.

CERTIFICATE.

I certify that Patrick Keast has carried out part-time research work under my direction, and has fulfilled the requirements of Ordinance 16, and is thus qualified to submit the accompanying thesis in application for the degree of Doctor of Philosophy.

Research Supervisor.

CHAPTER 0.

INTRODUCTION.

0.1 Preliminary Definitions.

Let B be the Banach Space of functions, u, depending on the variables $\underset{\sim}{x} = (x_0, x_1, \ldots, x_p)$, i.e. B is a normed linear space of functions. Let A be a linear differential operator involving derivatives with respect to the variables x_i, $i = 0, 1, \ldots, p$. Then A is a linear mapping

$$A : \quad B \rightarrow B,$$

and by a differential equation for u, we mean an equation of the form

$$A(u) = F(\underset{\sim}{x}), \quad F \in B \tag{1}$$

A solution of this equation is any function $u \in B$, which is transformed into F by A.

The variables $\underset{\sim}{x} = (x_1, x_2, \ldots, x_p)$ are called space-like variables, and the variable x_0 is called a time-like variable. The reason for this distinction is that, in physical situations, the operator A is not given alone, but together with certain restrictions on u. For example, the exact form of u, and of some of its derivatives, may be given when one variable, say x_0, is zero. This is called the initial state of the system, and a solution of equation (1) is then required for $x_0 > 0$. It is, therefore, natural to think of x_0 as a time variable, and so x_0 will usually be denoted by t. The remaining variables will represent lengths, displacements, etc., in the problems considered in this thesis.

The order of the operator A is taken to mean the highest space derivative occurring in (1); and the order in time is the order of the highest derivative with respect to time appearing in (1).

For equation (1) to have a unique solution, further conditions must be placed on u. A pure initial value problem consists of equation (1), together with conditions specifying the values of u, and all its time derivatives, of order less than the order of A in time, on the hyperplane t = 0. i.e. :

$$\left.\begin{aligned} u(\underset{\sim}{x},0) &= f_0(\underset{\sim}{x}) \\ \frac{\partial u}{\partial t}(\underset{\sim}{x},0) &= f_1(\underset{\sim}{x}) \\ &\vdots \\ \frac{\partial^{s-1} u(\underset{\sim}{x},0)}{\partial\ t^{s-1}} &= f_{s-1}(\underset{\sim}{x}) \end{aligned}\right\} \quad \forall\, \underset{\sim}{x}, \tag{1a}$$

where A is of order s in time, and f_i $(i = 0,1,\ldots, s-1)$, are known functions of the variables $\underset{\sim}{x}$. This problem is called the Cauchy problem for u. For some equations of the form (1), the conditions (1a) are sufficient to determine a unique solution. This thesis, however, will be concerned with initial-boundary value problems. A set of initial conditions of the form (1a), is given, on a closed region D in $\underset{\sim}{x}$ space, on the hyperplane t = 0. If ∂D represents the boundary of the region D, and $\partial D \times [t \geq 0]$ represents the open boundary of the region $D \times [t \geq 0]$ in $(\underset{\sim}{x},t)$ space, then the boundary conditions are given on $\partial D \times [t \geq 0]$, and the solution of equation (1) is sought in the region $D \times [t \geq 0]$. The boundary conditions considered, will be of the form

$$N_i(u) = \phi_i(\underset{\sim}{x},t) \tag{1b}$$

on $\partial D \times [t \geq 0]$, where N_i $(i = 1,2,\ldots,q)$, for some q to be specified, are linear differential operators of order less than the order of A, both in the time and space derivatives, and ϕ_i $(i = 1,2,\ldots,q)$, are given functions of $\underset{\sim}{x}$ and t, which will usually be assumed to be bounded for all $t \geq 0$.

In this thesis, three different types of equation, (defined in the next section), are considered. The operator A will be considered <u>together</u> with the boundary operators N_i, $i = 1,\ldots,q$, giving an operator $(A, N_1,\ldots,N_q) : B \to B$, under which equation (1), subject to (1a) and (1b), has a unique solution. Numerical approximations to this composite operator will be examined, and their <u>stability properties</u> (defined later) will be shown to depend on the form of the boundary operators. The solutions of the differential systems will also be examined, and it will be shown that their asymptotic behaviour (as $t \to \infty$) is affected by the form of the boundary operators.

0.2 Classification Of Differential Equations.

We shall consider only operators of order two in $\underset{\sim}{x}$, and of order one or two in time. The general linear operator of order two in the variables $\underset{\sim}{x}$ may be written in the form :

$$A = \sum_{i,j=0}^{p} a_{i,j}(\hat{\underset{\sim}{x}}) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=0}^{p} b_i(\hat{\underset{\sim}{x}}) \frac{\partial}{\partial x_i} + c(\hat{\underset{\sim}{x}}) \qquad (2)$$

where the coefficients $a_{i,j}$, b_i, and c, $(i,j = 0,1,\ldots,p)$ are in B. We associate with A the matrix Q, where

$$Q = [a_{i,j}],$$

and Q is assumed to be symmetric for all $\hat{\underset{\sim}{x}}$. Three cases will be distinguished here :

(a) Q is positive definite for all $\hat{\underset{\sim}{x}}$.

(b) At least one eigenvalue of Q is zero.

(c) One eigenvalue of Q is negative, and the rest are positive.

By means of a transformation of variables, (e.g. Garabedian [18], p. 71), the second order terms in A may be put into the respective canonical forms :

(a) $$\sum_{i=0}^{p} e_i(\hat{\underset{\sim}{y}}) \frac{\partial^2}{\partial y_i^2} .$$

(b) $$\sum_{i=1}^{p} e_i(\hat{\underset{\sim}{y}}) \frac{\partial^2}{\partial y_i^2} .$$

(c) $$\sum_{i=1}^{p} e_i(\hat{\underset{\sim}{y}}) \frac{\partial^2}{\partial y_i^2} - e_0(\hat{\underset{\sim}{y}}) \frac{\partial^2}{\partial y_0^2} .$$

where $e_i \in B$, $i = 0,1,\ldots,p$, and $\hat{\underset{\sim}{y}} = (y_0,y_1,\ldots,y_p)$ are the transformed variables. Equation (1) will be assumed to be in one of these canonical forms, and the variables $\underset{\sim}{x}$ will be used.

The three problems discussed in this thesis are now described as follows :

(a) The problem in this class is <u>Poisson's equation</u> in two or more variables, (and the associated <u>Laplace's equation</u>). When this equation occurs in physical problems, it usually describes non-transient phenomena, i.e. phenomena independent of time. Thus, no variable is time-like, and x_i, $i = 1,2,\ldots,p$, are space like variables. A properly posed problem, in the sense that a solution (possibly non-unique) exists, can be shown to require boundary conditions on the boundary ∂D of a closed region D, in $\underset{\sim}{x}$ space. The solution is then to be found in D. Since the order of A is two, it can be shown that only one boundary operator N, of degree one in the variables x_i, $i = 1,\ldots,p$, is necessary. The equation which is considered is, therefore,

$$\sum_{i=1}^{p} \frac{\partial^2 u}{\partial x_i^2} = g(\underset{\sim}{x}) \tag{3}$$

in a closed region D in $\underset{\sim}{x}$-space, bounded by a surface ∂D, on which

$$N(u) \equiv \frac{\partial u}{\partial \nu} - H(\underset{\sim}{x})u = \phi(\underset{\sim}{x}), \tag{3a}$$

where ν is the inward normal to D, and H and ϕ are functions of position on ∂D. Clearly, this is not the most general N which could be considered, (tangential derivatives could be involved), but it is the one which occurs in many physical problems, and which we shall discuss.

(b) In problems of this type, the variable x_0 occurs in derivatives of first order only. It is usually time-like, and is denoted by t. The most common physical problem of this class involves the <u>equation of heat conduction</u>, which gives rise to an initial- boundary value problem.

The time derivative is of order one, and so only u need be specified initially, on D. The space derivatives are of second order, and so boundary conditions of the type (3a) are given on $\partial D \times [t \geq 0]$, where ∂D is the boundary of D. The solution has then to be found in $D \times [t \geq 0]$. The equation to be considered is thus of the form :

$$\frac{\partial u}{\partial t} = \sum_{i=1}^{p} \frac{\partial^2 u}{\partial x_i^2} \tag{4}$$

in the region $D \times [t \geq 0]$, subject to the initial conditions

$$u(\underset{\sim}{x}, 0) = f(\underset{\sim}{x}), \qquad \underset{\sim}{x} \in D, \tag{4a}$$

and the boundary conditions

$$\frac{\partial u}{\partial \nu} - H(\underset{\sim}{x})u = \phi(\underset{\sim}{x}, t), \quad \underset{\sim}{x} \in \partial D, \quad t \geq 0. \tag{4b}$$

(c) An example of a problem of this class is given by the wave equation, in which the variable x_0 represents time, and is denoted by t. Since the order in time is two, two conditions are required initially, and the boundary operator may involve the first derivative of u with respect to time. In physical situations, this is usually an initial-boundary value problem, and so it is considered in the form :

$$\frac{\partial^2 u}{\partial t^2} = \sum_{i=1}^{p} \frac{\partial^2 u}{\partial x_i^2}, \tag{5}$$

in the region $D \times [t \geq 0]$, subject to the initial conditions

$$\left.\begin{aligned} u(\underset{\sim}{x}, 0) &= f(\underset{\sim}{x}) \\ \frac{\partial u}{\partial t}(\underset{\sim}{x}, 0) &= g(\underset{\sim}{x}) \end{aligned}\right\} \quad \underset{\sim}{x} \in D, \tag{5a}$$

and the boundary conditions

$$\frac{\partial u}{\partial \nu} - H(x)u - K(x)\frac{\partial u}{\partial t} = \phi(x, t), \quad x \in \partial D, \ t \geq 0, \tag{5b}$$

where H, K are functions of position on ∂D.

It will be shown that the boundary operators N play a significant part in the study of a differential system, and of numerical approximations to it.

0.3 Discrete Approximations to Partial Differential Equations.

In this section time dependent problems are discussed. The operator A is assumed to be linear, of order K in time, and the coefficients in A are assumed to be constants. The problems (b) and (c), above, are examples with K = 1, and K = 2, respectively. A linear partial differential operator of order K in time gives an equation which may be reduced to a system of K linear equations of first order in time, involving in turn the unknown function and its first K-1 derivatives with respect to time. For example, the equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2},$$

which is of type (c), may be reduced to a system which is first order in time, by putting

$$\frac{\partial u}{\partial t} = v,$$

so that

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial v}{\partial t} ;$$

i.e.

$$\frac{\partial}{\partial t}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{\partial^2}{\partial x^2} & 0 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} .$$

Putting $\underset{\sim}{u} = (u, v)$, this equation may be written in the form

$$\frac{\partial \underset{\sim}{u}}{\partial t} = A\underset{\sim}{u} ,$$

where A is the linear differential operator

$$\begin{bmatrix} 0 & 1 \\ \frac{\partial^2}{\partial x^2} & 0 \end{bmatrix}$$

We therefore consider equations of the form :

$$\frac{d\underset{\sim}{u}}{dt} = A\underset{\sim}{u} \tag{6}$$

where

$$\underset{\sim}{u} = \left(u, \frac{\partial u}{\partial t}, \frac{\partial^2 u}{\partial t^2}, \ldots\ldots, \frac{\partial^{K-1} u}{\partial t^{K-1}}\right),$$

and A is a <u>linear matrix operator</u>, with elements which are space derivatives and constants. The initial vector, $\underset{\sim}{u}(\underset{\sim}{x},0)$, is given in accordance with the remarks of the previous section. If equation (6) refers to an initial-boundary value problem for $\underset{\sim}{u}$, then there are boundary conditions given on $\partial D \times [t \geq 0]$, where ∂D is the boundary of some closed region D, on which $\underset{\sim}{u}(\underset{\sim}{x},0)$ is given.

A <u>discrete or difference approximation</u> to equation (6) is obtained by regarding t not as a continuous variable, but as a <u>discrete variable</u>, taking the values $0, \Delta t, 2\Delta t, \ldots, n\Delta t$, where n is an integer, and Δt is called the <u>time increment</u>. The derivative $\frac{d\underset{\sim}{u}}{dt}$ can be approximated, then, by, for example,

$$\frac{d\underset{\sim}{u}(n\Delta t)}{dt} \doteq \frac{\underset{\sim}{u}((n+1)\Delta t) - \underset{\sim}{u}(n\Delta t)}{\Delta t} .$$

In addition, the values of the space variables $(x_1, x_2, \ldots, x_p)$ are considered to be discrete, and the increment in the variable x_i, $i = 1, \ldots, p$, is denoted by Δx_i. The space derivatives are then approximated by linear combinations of the values of the function $\underset{\sim}{u}$ at the <u>mesh points</u>, i.e. at the points which are used in the discretisation. For example

$$\frac{\partial^2 \underset{\sim}{u}(x_i)}{\partial x_i^2} \doteq \frac{\underset{\sim}{u}(x_i+\Delta x_i) - 2\underset{\sim}{u}(x_i) + \underset{\sim}{u}(x_i-\Delta x_i)}{(\Delta x_i)^2}$$

As an application of this technique, consider the system of equations

$$\frac{d}{dt}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{\partial^2}{\partial x^2} & 0 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix}$$

obtained from the wave equation $u_{tt} = u_{xx}$. Then, using the approximations to $\frac{d}{dt}$, and $\frac{\partial^2}{\partial x^2}$, suggested above, we obtain the difference equation

$$\begin{bmatrix} u_i^{n+1} \\ v_i^{n+1} \end{bmatrix} - \begin{bmatrix} u_i^{n} \\ v_i^{n} \end{bmatrix} = \left(\frac{\Delta t}{\Delta x}\right)^2 \begin{bmatrix} 0 & (\Delta x)^2 \\ -2 & 0 \end{bmatrix}\begin{bmatrix} u_i^{n} \\ v_i^{n} \end{bmatrix}$$

$$+ \left(\frac{\Delta t}{\Delta x}\right)^2 \begin{bmatrix} 0 \\ v_{i+1}^n + v_{i-1}^n \end{bmatrix},$$

where (u_i^m, v_i^m) represents $\underset{\sim}{u}(m\Delta t, x_i)$, for $m = n, n+1$, etc. If we replace the expression

$$v_{i+1}^n - 2v_i^n + v_{i-1}^n,$$

by

$$\delta^2 v_i^n,$$

where δ is a difference operator, we obtain the equation

$$\underset{\sim}{u}_i^{n+1} = \begin{bmatrix} 1 & \Delta t \\ \left(\frac{\Delta t}{\Delta x}\right)^2 \delta^2 & 1 \end{bmatrix} \underset{\sim}{u}_i^{n}, \qquad (7)$$

where $\underset{\sim}{u}_i^n$ is the approximation to $\underset{\sim}{u}(n\Delta t, x_i)$. For a more general equation, we can still replace the derivatives occurring, by discrete approximations as above, and letting $\underset{\sim}{u}^n$ be the approximation to $\underset{\sim}{u}(n\Delta t, x)$ we obtain the equation

$$H_1 \underset{\sim}{u}^{n+1} = H_2 \underset{\sim}{u}^n, \qquad (8)$$

where

$$H_j = H_j(\Delta t, \Delta x_1, \ldots, \Delta x_p),\ j = 1,2,$$

are matrices of order K, involving difference operators such as δ. Equation (8) is assumed to be solvable for $\underset{\sim}{u}^{n+1}$, so that, given $\underset{\sim}{u}^0$, we can find $\underset{\sim}{u}^n$, with n = 1,2,..., by a step-by-step process, such as (8). This assumption implies that H_1 is invertible, and so (8) may be written in the form

$$\underset{\sim}{u}^{n+1} = C\underset{\sim}{u}^n, \tag{9}$$

where

$$C = H_1^{-1}H_2 .$$

If $H_1 = I$, the identity matrix, equation (8) is said to represent an <u>explicit method</u>. Otherwise (8) is an <u>implicit method</u>. (Occasionally, the term <u>fully implicit</u> is used to describe the case when $H_2 = I$).

It is natural in a definition of convergence of a difference approximation to the differential equation, to suppose that numerical calculations are carried out using equation (8) with increasingly finer meshes, i.e. $\Delta t, \Delta x_i$ $(i = 1,2,\ldots,p) \to 0$. It is assumed that these increments do not tend to zero arbitrarily, but that the increments Δx_i, for each i, are functions of Δt; i.e.

$$\Delta x_i = \alpha_i(\Delta t),\ i = 1,2,\ldots,p,$$

where

$$\alpha_i(\Delta t) \to 0, \text{ as } \Delta t \to 0.$$

Then equation (9) may be written in the form :

$$\underset{\sim}{u}^{n+1} = C(\Delta t)\, \underset{\sim}{u}^n \tag{10}$$

Let $\underset{\sim}{u}$ be the <u>exact</u> solution of the <u>differential equation</u>, (6). Let $\underset{\sim}{v}$ be the <u>exact</u> solution of the <u>difference equation</u> (8); i.e. the solution which would be obtained if infinite accuracy could be achieved in the numerical calculations. Finally, let $\tilde{\underset{\sim}{v}}$ be the <u>numerical</u> solution of the <u>difference equation</u> (8); i.e. the solution actually computed. Then, following O'Brien, Hyman, and Kaplan, [37], the definitions below are made :

(1) The quantity $\underset{\sim}{u} - \underset{\sim}{v}$ is the Truncation Error, (T.E.).

(2) The difference scheme (8) is said to Converge to the differential equation (6), as $\Delta t \to 0$, if the T.E. tends to zero; i.e. if

$$||\underset{\sim}{u} - \underset{\sim}{v}|| \to 0, \text{ as } \Delta t \to 0,$$

where $||.||$ is the norm in B.

(3) The quantity $\underset{\sim}{v} - \underset{\sim}{\hat{v}}$ is the Numerical Error, (N.E.).

(4) The difference method (8) is said to be stable if the N.E. tends to zero as $\Delta t \to 0$, or if the N.E. is bounded everywhere in the region of integration; i.e. if

$$||\underset{\sim}{v} - \underset{\sim}{\hat{v}}|| \text{ is bounded as } \Delta t \to 0,$$

where $||.||$ is the norm in the finite dimensional vector space.

These definitions indicate the principal phenomena which must be examined in a discussion of difference methods, viz. stability, and convergence. In order to obtain a practical method of measuring the quantities defined in (1) and (3) above, we follow the analysis of Lax and Richtmeyer, [33,41].

Equation (10) may be written in the form

$$\underset{\sim}{u}^{n+1} = C^{n+1}(\Delta t)\, \underset{\sim}{u}^{0},$$

since C is a matrix with constant elements. The solution of equation (6), if it exists, may also be written in this way : i.e.

$$\underset{\sim}{u}(t) = F(t)\underset{\sim}{u}(0),$$

for some continuous operator F. Equation (6) is then said to be properly posed if the solution $\underset{\sim}{u}(t)$ depends continuously on the initial data; i.e. if the operator F is uniformly bounded for $0 \leq t \leq T$, for some fixed T.

In chapter 2, we discuss the position when boundedness of F is required for all t.

Convergence of the difference equation to the differential equation requires that, for some sequence Δt_i, $i = 1,2,\ldots$, which tends to zero as $i \to \infty$, the norms of the operators F and C^n tend to one another; i.e.

$$||C^{n_i}(\Delta t_i) - F(t)|| \to 0,$$

where $n_i \Delta t_i \to t$, as $\Delta t_i \to 0$.

The difference method is stable if

$$\exists M, \exists \tau, \forall \Delta t, \quad (0 \leq \Delta t \leq \tau,\ 0 \leq n\Delta t \leq T) \Rightarrow ||C^n(\Delta t)|| \leq M.$$

Finally, if (9) is an approximation to (6), then

$$\frac{\underset{\sim}{u}^{n+1} - \underset{\sim}{u}^n}{\Delta t} = \frac{(C - I)}{\Delta t}\, \underset{\sim}{u}^n ,$$

must be an approximation to $\frac{d\underset{\sim}{u}}{dt}$. Thus, if

$$\lim_{\Delta t \to 0} \left|\left| \frac{C(\Delta t)-I}{\Delta t} - A \right|\right| = 0, \quad 0 \leq t \leq T,$$

then the method is said to be consistent. The truncation error is obtained in a manner which verifies consistency, as follows : The quantity

$$\left\{ \frac{C(\Delta t)-I}{\Delta t} - A \right\} \underset{\sim}{u}^n$$

is calculated, (using Taylor expansions about the point $(n\Delta t, \underset{\sim}{x})$, to find $\left\{\frac{C(\Delta t)-I}{\Delta t}\right\} \underset{\sim}{u}^n$), and this is the T.E. The order of the T.E. is the order, or power, of Δt and Δx_i, $i = 1,2,\ldots,p$, in the principal part of the T.E., (i.e. in the lowest term of the T.E.). Consistency, therefore, requires a T.E. of order at least one, in both Δt and Δx_i for each i.

For a pure initial value problem, Lax and

Richtmyer, [33,41] have proved that if the problem is properly posed, and the method of approximation is consistent, then the method is stable if, and only if, it is also convergent. It will be suggested, below, that this equivalence theorem may also hold for initial-boundary value problems, if "properly posed" is interpreted in a certain manner.

0.4 The Analysis Of Stability Properties Of Difference Schemes.

The method of von Neumann [10,37], for testing the stability of a difference scheme is as follows : let

$$\underset{\sim}{\varepsilon}^n = \underset{\sim}{y}^n - \underset{\sim}{\hat{y}}^n ,$$

be the numerical error. Then

$$\underset{\sim}{\varepsilon}^{n+1} = C(\Delta t)\underset{\sim}{\varepsilon}^n \qquad (11)$$

A Fourier decomposition of each error $\underset{\sim}{\varepsilon}^n$ at the point $\underset{\sim}{x} = (x_1, x_2, \ldots, x_p)$, is made; i.e.

$$\underset{\sim}{\varepsilon}^n = a_n e^{i\underset{\sim}{l}\cdot\underset{\sim}{\theta}} , \quad n = 0, 1, 2, \ldots,$$

where

$$\underset{\sim}{l} = (l_1, l_2, \ldots, l_p)$$

is a vector with positive or negative integer components, $\underset{\sim}{\theta}$ is a p-component vector, and a_n, $(n \geq 0)$ are constant vectors. Substituting $\underset{\sim}{\varepsilon}^n$ into equation (11), we obtain the equation

$$a_{n+1} = G(\underset{\sim}{l}, \underset{\sim}{\theta})a_n ; \qquad (12)$$

i.e.

$$a_{n+1} = G^{n+1}(\underset{\sim}{l}, \underset{\sim}{\theta})a_0 . \qquad (13)$$

The quantity G is a matrix of order K, and is called the <u>amplification matrix</u>. From (13), the stability condition is

$$\exists M > 0, \exists \tau > 0, \forall n > 0, \quad (0 \leq \Delta t \leq \tau, n\Delta t = t \leq T) \Rightarrow \|G^n\| \leq M \qquad (14),$$

for <u>all</u> $\underset{\sim}{\theta}$ and $\underset{\sim}{l}$, and for some T. The symbol $\|\cdot\|$ denotes some matrix norm. This gives a definition of stability for an initial value problem, in terms of the norms of a set of matrix operators. The analysis of these norms, and sets of necessary and sufficient conditions for stability, have been discussed at length by several authors, in particular, Kreiss, [27,28,29,30,31,32], and Buchanan [5,6,8]. Strictly, these results are not applicable to initial-boundary value problems, but the following remarks suggest a method for extending the work of Kreiss and Buchanan to cover the case of boundary value problems.

If it is assumed that each error vector ε^n can be expanded in a Fourier series, i.e.

$$\varepsilon^n = a_n \cos(l.(\theta+\phi)),$$

where a_n, $n \geq 0$, are functions of n only; $l = (l_1,\ldots,l_p)$ is a vector with integer components; θ is a p-component vector; and ϕ is a parameter, then an amplification matrix $G(l,\theta,\phi)$ is obtained as before. The parameter ϕ, is eliminated using the boundary conditions, and the condition (14) is then imposed for all θ, and l. This suggestion is not backed by a theoretical argument, but appears to work in practice. However, it will not be followed in this thesis, since a direct approach to the question of stability is made (c.f. Todd [46]).

Let the set of values of u^n at the points on the time level $t = n\Delta t$ be ordered in some way. There are a finite number of points at this level, since the hyperplane $t = n\Delta t$ intersects the cylinder $D \times [t \geq 0]$ in a closed region. Then the corresponding errors ε^n are ordered in the same way, and may be written as one vector E^n of all the errors at this time level. Equation (11) may then be replaced by an equation of the form

$$E^{n+1} = R(\Delta t)E^n$$

i.e.

$$E^{n+1} = R^{n+1}(\Delta t)\, E^0, \tag{15}$$

since R is a matrix with elements independent of t. The important points to note here are that :

(1) R is affected by the form of the boundary conditions. For example, for boundary conditions of the form (4b), i.e.

$$\frac{\partial u}{\partial \nu} - H(\underset{\sim}{x})u = \phi(\underset{\sim}{x}, t), \quad \underset{\sim}{x} \in \partial D, \ t > 0,$$

for an equation which is first order in time, and second order in the space derivatives, the matrix R will involve $H(\underset{\sim}{x})$.

(2) The order of R becomes infinitely large as Δt, and, therefore, Δx_i, for all i, tend to zero.

The definition of stability suggested by equation (15) is thus :

$$\exists K > 0, \exists \epsilon > 0, \quad (\Delta t < \epsilon, \ n\Delta t = t \leq T) \Rightarrow ||R^n(\Delta t)|| \leq K .$$

This is the basis of the direct approach, first suggested by Todd [46]. In [12], Douglas also suggests this approach, but his results are not complete, as the following discussion shows.

The condition for stability is that the norms of R^n should stay bounded, as $n \to \infty$, where $n\Delta t = T$, and $\Delta t \to 0$. [A more restrictive definition is one that requires boundedness of R^n as $n \to \infty$, $\Delta t \to 0$, and for all T, however large, c.f. Gary [19]. This point is taken up again in chapter 1.] As $\Delta t \to 0$, the order of R tends to ∞. It can be shown (e.g. [20], Godunov and Ryabenki) that a necessary condition for stability is that the maximum modulus eigenvalue of R, i.e. the spectral radius $\rho(R)$, should be less than, or equal to, one. This is not, however, a sufficient condition, unless R is symmetric, as the following example shows :

Let the matrix R be given by

$$R = \begin{bmatrix} 1-r & 1 & & & & \\ 0 & 1-r & 1 & & & \\ & 0 & 1-r & 1 & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & & 0 & 1-r \end{bmatrix}$$

and let it be of order N, where N will be allowed to tend to infinity. Clearly

$$\rho(R) \leq 1, \quad \text{if}$$

$$0 \leq r \leq 2.$$

Thus, for a fixed N, $R^n \to [0]$, as $n \to \infty$, (e.g. Varga [47]). But, if $N \to \infty$ at the same rate as n, then it can be shown that the norms of R^n become infinitely large, unless $0 \leq r \leq 1$. This particular example has been considered by Godunov and Ryabenki [20], where it is shown that the spectrum of each matrix R^n (i.e. the set of eigenvalues) is inside, or on the unit circle if $0 \leq r \leq 2$, $\forall n$, but that the spectrum of the family $\{R^n\}$ of N^{th} order matrices, as N and $n \to \infty$ is inside, or on, the unit circle if, and only if, $0 \leq r \leq 1$. Douglas, however, states that $\rho(R) \leq 1$ is a sufficient condition (provided only that any eigenvalues of modulus 1 are not multiple eigenvalues), in the sense that $R^n \to 0$, where R is a matrix of fixed order. This is a misleading definition of stability, since it does not allow Δx_1 and Δt to tend to zero at some fixed point, i.e. $\Delta t \to 0$, with $n\Delta t = T$, for some fixed T. If R is symmetric, however, it can be shown (e.g. see the following section) that $\rho(R) \leq 1$ is a sufficient condition for stability.

0.5 Norms Of Matrices.

A brief note on the norm which will be used, is added here. The norm which is employed is the third matrix norm :

$$||R|| = \sqrt{\mu}$$

where

$$\mu = \rho(R^*R),$$

R^* being the transposed conjugate of R, and ρ being the spectral radius, defined above. The spectral radius is not itself a norm, but (e.g. [14])

$$\rho(R) \leqslant ||R||,$$

where $||.||$ is any norm. Thus, since $\rho(R^n) = \rho^n(R)$, then if $\rho(R) > 1$, $||R^n||$ is unbounded, as $n \to \infty$. If, however, R is symmetric, then

$$\rho(R) = \sqrt{\mu} = \sqrt{(\rho(R^*R))},$$

so that

$$\rho(R) \leqslant 1,$$

is a necessary and sufficient condition for boundedness of the norms of R^n.

The inequalities

$$\rho^n(R) \leqslant ||R^n|| \leqslant ||R||^n,$$

follow at once from the above remarks, and from the definition of the norm of a product. Thus, an obvious sufficient, but not necessary, condition for stability is $||R|| \leqslant 1$. In fact, this may sometimes be weakened to $||R|| \leqslant 1+O(\Delta t)$, since then

$$||R||^n \leqslant 1+O(n\Delta t) \leqslant 1+O(T),$$

which is bounded, if T is fixed. This point will be raised again in chapter 1.

0.6 Boundary Value Problems.

In this section we return to second order boundary value problems, of the type (a), i.e. problems which involve no time-like variable, and in which the value of the unknown function is given on the boundary of a closed region. In the same manner as in section 0.3, the operator, such as the one in equation (3), is replaced by a discrete approximation. Ordering the values of the function u, as in section 0.3, we obtain a set of linear equations, which can be written in the form :

$$R\underset{\sim}{u} = \underset{\sim}{k}, \tag{16}$$

where $\underset{\sim}{k}$ is a known vector, and R may be singular. As above, R involves the coefficients in the boundary conditions, which are of the form (3a). The solution of (16) is not usually carried out directly, but by some <u>iterative process</u>, which is a rule whereby successive approximations $\underset{\sim}{u}^n$, $n = 0, 1, \ldots,$ to $\underset{\sim}{u}$ may be calculated, where $\underset{\sim}{u}^0$ is some arbitrary vector. For example, one such iterative process is

$$\underset{\sim}{u}^{n+1} = S\underset{\sim}{u}^n + \underset{\sim}{k}, \tag{17}$$

where $I-S = R$. The error of the <u>iterative process</u> is

$$\underset{\sim}{\varepsilon}^{n+1} = \underset{\sim}{u}^{n+1} - \underset{\sim}{u}^n,$$

and this satisfies the equation

$$\underset{\sim}{\varepsilon}^{n+1} = S\underset{\sim}{\varepsilon}^n. \tag{18}$$

The problem is not one of convergence of a difference method to a differential equation, for it is assumed that the approximation (16) is a good (i.e. consistent at least) approximation to the differential equation. Instead, the problem is whether or not the iterates $\underset{\sim}{u}^n$ tend to the vector $\underset{\sim}{u}$, i.e. whether or not $\underset{\sim}{\varepsilon}^n \to \underset{\sim}{0}$, as $n \to \infty$. From (18), since S has a fixed order, the convergence condition is

$$\rho(S) < 1.$$

It may be noted that the matrix R is singular if, and only if, S has a real eigenvalue equal to ± 1. In the problems considered in this thesis it will be shown that

$$\rho(S) < 1,$$

is a necessary, and sufficient condition for convergence of the iterative processes used.

0.7 Conclusion.

From the above remarks, it will be seen that the difference systems always take the form of an approximation to the differential equation and the boundary conditions combined; i.e. the boundary conditions are regarded as part of the differential operator. This will be the theme followed throughout this thesis. First, the effect of boundary conditions on the stability of some difference approximations to a simple equation of type (b) will be considered. Then the effect of the boundary conditions on the differential operator will be discussed. The same programme will be carried out for equations of the type (a) and (c). It will be shown that the boundary conditions decide whether the problem is properly posed, or not, in the sense of section 0.3, i.e. whether the solutions of the differential equations are bounded or not. It will also be shown that the boundary conditions for which the difference methods examined are unstable, correspond exactly to the boundary conditions for which the problems are not properly posed.

CHAPTER 1.

THE EFFECT OF BOUNDARY CONDITIONS ON THE STABILITY OF A DIFFERENCE APPROXIMATION TO A PARABOLIC DIFFERENTIAL EQUATION.

1.1 Introduction.

In this chapter we consider a parabolic equation with boundary conditions involving linear combinations of the function and its first derivative. A class of approximating difference methods is examined, and the stability of these methods is shown to depend on the form of the boundary conditions.

1.2 The Differential Equation.

Consider the equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{1}$$

in the region

$$R = [0 \leqslant x \leqslant 1] X [t \geqslant 0]$$

for the function $u(x,t)$, subject to the initial condition

$$u(x,0) = f(x) \qquad 0 \leqslant x \leqslant 1 \tag{1a}$$

and the boundary conditions

$$\begin{aligned} \frac{\partial u}{\partial x} - pu &= \phi_0(t) \qquad x=0, \quad t \geqslant 0 \\ \frac{\partial u}{\partial x} + qu &= \phi_1(t) \qquad x=1, \quad t \geqslant 0 \end{aligned} \tag{1b}$$

where p, q are constants.

It is assumed that there are no discontinuities in the initial or boundary conditions, that ϕ_0 and ϕ_1 are bounded as $t \to \infty$, and that there are no discontinuities at the corners of R. That is

$$\begin{aligned} D f(0) - pf(0) &= \phi_0(0) \\ D f(1) + qf(1) &= \phi_1(1) \end{aligned} \qquad D \equiv \frac{d}{dx}, \tag{2}$$

The question of discontinuities in the initial data, or at the corners of R, will be mentioned, briefly, later. The conditions (1b), for various values of p and q give all possible linear boundary conditions involving both u and its space derivative. When the values of u alone are specified on the boundaries, the problem is simple and has been examined by many authors. The mixed problem, with u specified on one boundary, and a linear combination of u and $\frac{\partial u}{\partial x}$ on the other, is considered later.

Several authors [e.g. 7, 16, 21, 34, 42] have considered the problem (1) subject to (1a), (1b) for the particular case of non-negative p, q. This restriction is not imposed here.

1.3 The Numerical Approximation.

The region R is covered by a rectangular net whose mesh points are $(j\Delta x, n\Delta t)$, where $j = 0,1,2,\ldots,N$, $N\Delta x = 1$, Δx is the space increment, $n \geq 0$, and Δt is the time increment. The time increment is assumed to be a function of Δx; for this problem

$$\frac{\Delta t}{(\Delta x)^2} = r$$

where the number r is fixed and is called the mesh ratio.

The class of finite difference methods used involves the six points marked in figure [1], and is given by

$$v_j^{n+1} - v_j^n = r[\theta\delta^2 v_j^{n+1} + (1-\theta)\delta^2 v_j^n] \tag{3}$$

where v_j^n is the difference approximation to $u(j\Delta x, n\Delta t)$, θ is a parameter in the range $0 \leq \theta \leq 1$, and δ is the usual central difference operator given by

$$\delta^2 v_j^m = v_{j+1}^m - 2v_j^m + v_{j-1}^m \qquad m = n, n+1.$$

The approximation (3) is applied for values of $n \geq 0$, with the initial values v_j^0 given by $f(j\Delta x)$, and for each n, (3) is applied for values of $j = 1,2,\ldots\ldots\ldots,N-1$.

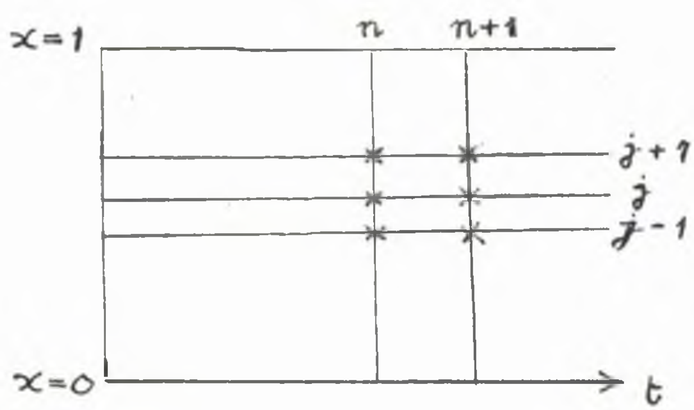

Figure [1]

When $j = 0$ or N the boundary conditions (1b) must be used. A particular boundary difference replacement is considered first, and others are considered later. When $j = 0$ or N in (3) values of v at points outside R are introduced. These are eliminated using the boundary conditions (1b) with the derivative approximated by

$$\frac{\partial v_j^m}{\partial x} = \frac{1}{2\Delta x}\left(v_{j+1}^m - v_{j-1}^m\right), \quad j = 0, N, \quad m = n, n+1 \qquad (3a)$$

If $\theta = 0$ or $\theta = 1$ the boundary conditions are applied only at one time level, n or $n+1$ respectively. The above boundary replacement is used because the error introduced by it is $O(\Delta x^2)$, which is of the same order as the error introduced by equation (3).

The set of difference equations may now be written in the form

$$A\underset{\sim}{\omega}^{n+1} = B\underset{\sim}{\omega}^{n} + \underset{\sim}{k}^{n} \qquad (4)$$

where $\underset{\sim}{k}^n$ is an $(N+1)$ component column vector involving $\phi_0(t)$ and $\phi_1(t)$ at $t = n\Delta t$ and $t = (n+1)\Delta t$, $\underset{\sim}{\omega}^m$ $(m = n, n+1)$ is a column vector given by

$$\underset{\sim}{\omega}^m = \{v_0^m, v_1^m, \ldots\ldots, v_N^m\}^T \qquad m = n, n+1,$$

and A, B are matrices of order $(N+1)$ given by

$$\left.\begin{aligned} A &= I + r\theta U \\ B &= I - r(1-\theta)U \end{aligned}\right\} \qquad (4a)$$

with

$$U = \begin{bmatrix} 2(1+\frac{p}{N}) & -2 & 0 & \cdots\cdots\cdots & 0 \\ -1 & 2 & -1 & 0 \cdots\cdots & 0 \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot & \cdot \\ & & & -1 & 2 & -1 \\ & & & & -2 & 2(1+\frac{q}{N}) \end{bmatrix} \qquad (5)$$

Let $\underset{\sim}{\omega}^{n+1}$ be the exact (i.e. theoretical) solution of equation (4). Let $\hat{\underset{\sim}{\omega}}^{n+1}$ be the numerical solution obtained by computation. Then the numerical error

$$\underset{\sim}{\epsilon}^{n+1} = \underset{\sim}{\omega}^{n+1} - \hat{\underset{\sim}{\omega}}^{n+1}$$

satisfies the homogeneous equation

$$A\underset{\sim}{\epsilon}^{n+1} = B\underset{\sim}{\epsilon}^{n}$$

or, if A is non-singular (which we assume to be the case),

$$\underset{\sim}{\epsilon}^{n+1} = A^{-1}B\underset{\sim}{\epsilon}^{n} \qquad (6)$$

1.4 Stability Analysis.

The method of von Neumann for deciding stability questions may be used only for pure initial value problems, or for problems with periodic boundary conditions. This method cannot therefore be applied to the above problem. [But c.f. the remarks in chapter 0, section 3]. The stability criterion used is the one applied by, for example, Godunov and Ryabenki [20] (see chapter 0).

Equation (6) may be rewritten in the form

$$\underset{\sim}{\epsilon}^n = (A^{-1}B)^n \underset{\sim}{\epsilon}^0 \qquad n \geq 0.$$

Stability follows if and only if, for some suitable norm $||\ .\ ||$, we have $||\ \underset{\sim}{\epsilon}^n\ ||$ bounded. That is, since

$$||\ \underset{\sim}{\epsilon}^n\ || \leq ||(A^{-1}B)^n||.||\ \underset{\sim}{\epsilon}^0\ ||,$$

if and only if a constant K can be found such that

$$||(A^{-1}B)^n|| \leq K \tag{7}$$

for all n. The condition (7) is required to hold in the limit as Δx, $\Delta t \to 0$, with $n\Delta t = t$, for all $t \leq T$, for some fixed T. Recalling the definition of stability given in chapter 1, therefore, equation (6) is stable in a given norm if and only if, for some fixed T, and $r = \Delta t/(\Delta x)^2$ fixed,

$$(\exists K>0)\ (\forall \epsilon>0)\ \ (\Delta t < \epsilon,\ n\Delta t = t \leq T) \Rightarrow ||(A^{-1}B)^n|| \leq K.$$

Clearly, as $\Delta x \to 0$, the order of $A^{-1}B$ (which is N+1) also tends to ∞.

For any matrix C, the spectral radius of C, $\rho(C)$, (the maximum modulus eigenvalue of C), is related to the norm of C, $||C||$, by the inequality

$$\rho^n(C) \leq ||\ C^n|| \leq ||\ C\ ||^n \qquad \forall\ n > 0. \tag{8}$$

Thus it follows that, as stated in chapter 0,

(1). The condition $||\ A^{-1}B\ || \leq 1$ is sufficient to guarantee stability of equation (6), but not necessary. This is the condition usually applied on the grounds of simplicity.

(2) A necessary condition for stability is $\rho(A^{-1}B) \leq 1$, but this is not sufficient.

If, however, C is a symmetric matrix, the <u>equalities</u> in (8) hold. Further, if C is similar to a symmetric matrix $\tilde{C}$, i.e.

$$C = P^{-1}\tilde{C}P,$$

where P is some non-singular matrix, then

$$C^n = P^{-1}\tilde{C}^n P$$
$$\tilde{C}^n = P\,C^n P^{-1}$$

Thus

$$\|C^n\| \leq \|P^{-1}\|.\|\tilde{C}^n\|.\|P\|$$

and

$$\|\tilde{C}^n\| \leq \|P\|.\|C^n\|.\|P^{-1}\|.$$

Thus if $\|P\|$ and $\|P^{-1}\|$ are bounded as $n \to \infty$, then $\|C^n\|$ is bounded as $n \to \infty$, if and only if $\|\tilde{C}^n\|$ is bounded. But, since $\rho(C) = \rho(\tilde{C})$ and $\tilde{C}$ is symmetric, $\|\tilde{C}^n\|$ is bounded if and only if $\rho(C)$ is bounded.

For the matrix U in (5), the matrices $\tilde{U}$ and $\widetilde{A^{-1}B}$ given by

$$\tilde{U} = D^{-1}U\,D \quad \text{and} \quad \widetilde{A^{-1}B} = D^{-1}(A^{-1}B)D,$$

where

$$D = \begin{vmatrix} \sqrt{2} & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & \cdot & & & & \\ & & & & \cdot & & & \\ & & & & & \cdot & & \\ & & & & & & 1 & \\ & & & & & & & \sqrt{2} \end{vmatrix}$$

are symmetric, and the norms $\|D\|$ and $\|D^{-1}\|$ are both bounded for all n.

The necessary and sufficient condition for (6) to be stable is then

$$[\rho(A^{-1}B)]^n \leqslant K$$

or

$$\rho(A^{-1}B) \leqslant 1 \tag{9}$$

Many authors [e.g. 4,33,41] weaken (9), when $A^{-1}B$ is similar to a symmetric matrix, as follows. Since what is required is that $\rho^n(A^{-1}B)$ be bounded for all n, with $n\Delta t \leqslant T$, as $\Delta t \to 0$, then

$$\rho(A^{-1}B) \leqslant 1 + O(\Delta t) \tag{9a}$$

will be a necessary and sufficient condition for (9) to hold, as shown in chapter 6. Gary [19] discusses this point in a more general context, and gives an alternative definition of stability which differs from the formal definition above only in the removal of the condition $n\Delta t \leqslant T$. His definition is in effect a definition of stability which is uniform in time, i.e., stability which is still maintained as $t \to \infty$. Such a definition is necessary for equations which must be considered for large time e.g. dissipative systems. Kreiss [32] and Gary both derive a stability condition which, when applied to the above problem, becomes the necessary and sufficient condition

$$\rho(A^{-1}B) < 1.$$

It will be shown that for the problem under consideration the condition

$$\rho(A^{-1}B) \leqslant 1$$

is also sufficient.

1.5 Eigenvalues of $A^{-1}B$.

Let λ_j $(j = 0,1,2,\ldots\ldots,N)$ be the eigenvalues of U. These are real since U is similar to a symmetric matrix. Then μ_j $(j = 0, 1,2,\ldots\ldots,N)$, the eigenvalues of $A^{-1}B$, are given by

$$\mu_j = (1 + r\theta\lambda_j)^{-1}(1 - r(1-\theta)\lambda_j). \qquad (10)$$

Instability will occur if and only if

$$\rho(A^{-1}B) = \max_j |\mu_j| > 1 .$$

Since A is assumed to be non-singular, it follows that $1 + r\theta\lambda_j \neq 0$, $(j = 0,1,\ldots\ldots,N)$. Then $\rho > 1$ if and only if one of the following conditions holds :

(a) $\exists j \quad -1/r\theta < \lambda_j < 0.$

(b) $\forall\theta,\ 0 \leqslant \theta \leqslant 1/2,\ \exists j, \quad \lambda_j < -1/r$, or $\lambda_j > \dfrac{1}{r(1/2-\theta)}$.

(c) $\theta = 1/2, \ \exists j \quad \lambda_j < -2/r.$

(d) $\forall\theta$, $1/2 < \theta \leqslant 1, \quad \exists j \quad -\dfrac{1}{r(\theta-1/2)} < \lambda_j < -1/r\theta.$

In addition the spectral radius of $A^{-1}B$ is 1 when no one of (a),(b),(c),(d) is satisfied, and either

(e) $\exists j \quad \lambda_j = 0 \quad \forall\theta$, $0 \leqslant \theta \leqslant 1.$

Or

(f) $\exists j \quad \lambda_j = \dfrac{1}{r(1/2-\theta)}, \quad \forall\theta , \quad 0 \leqslant \theta \leqslant 1.$

These conditions may be interpreted as follows: (a) and (e) are conditions on U. If U has any negative eigenvalues the scheme is unstable for any r, according to (a). If one eigenvalue of U is zero and the rest are such that (a),(b),(c),(d) are not satisfied, then $\rho(A^{-1}B) = 1$; thus since $\widetilde{A^{-1}B}$ is symmetric and similar to $A^{-1}B$, $||(A^{-1}B)^n||$ is bounded, and the scheme is stable. The latter instance is of some importance and will be discussed further, later. The second condition, (b), is a condition on r when λ_j is given for each j. Thus it is necessary that

$$r \leqslant 1/[\lambda_j(1/2-\theta)]\forall j \qquad (b)'$$

for stability. The difference scheme is then conditionally stable for $0 \leq \theta \leq 1/2$ (i.e. stable under a condition on r), provided none of the conditions (a), (c),(d) are satisfied. The conditions (c) and (d) and the first condition in (b) are ignored for the reasons given below.

1.6 The Eigenvalues of U.

By applying the theorem of Gerschgorin to the matrix U, the eigenvalues λ_j $(j = 0,1,\ldots,N)$ are shown to lie in the union of the intervals

$$\begin{gathered} 0 \leqslant \lambda_j \leqslant 4 \\ \frac{2p}{N} \leqslant \lambda_j \leqslant 4 + \frac{2p}{N} \\ \frac{2q}{N} \leqslant \lambda_j \leqslant 4 + \frac{2q}{N} \end{gathered} \qquad (11)$$

i.e. in the interval

$$\min(0,\tfrac{2p}{N},\tfrac{2q}{N}) \leqslant \lambda_j \leqslant \max(4,4+\tfrac{2p}{N},4+\tfrac{2q}{N}).$$

Thus, if p and q are both non-negative, the eigenvalues are non-negative. If, when one of p, q is negative, there is a negative λ_j, it is $O(1/N)$ in magnitude, at most. This observation is the reason for dropping the cases (c) and (d), and part of (b), since, for reasonable choices of N and r, these cases cannot arise.

When a negative eigenvalue occurs it is at most $O(1/N)$ in magnitude. Then

$$\rho(A^{-1}B) = 1 + O(1/N)$$

Thus

$$\rho^n(A^{-1}B) = [1 + O(1/N)]^n$$

Now, if the formal definition of convergence is applied i.e. if the restriction $n\Delta t = t \leqslant T$ is imposed, then, since $\Delta t = r/N^2$, where r is fixed, as $N \to \infty$, and Δx, Δt, $\to 0$,

$$\begin{aligned} \rho^n &= [1 + O(1/N)]^{t/\Delta t} \\ &= [1 + O(1/N)]^{N^2 t/r} \\ &\doteq e^{Nt/r} \to \infty \text{ as } N \to \infty. \end{aligned}$$

The condition

$$\rho = 1 + O(1/N)$$

is in fact

$$\rho \quad = 1 + O(\sqrt{(\Delta t)})$$

and is not sufficient to satisfy even the inequality (9a), which is the less strict condition for stability. If, as was suggested in section 4, the the restriction $n\Delta t \leqslant$ a fixed T is removed, then the occurence of a negative eigenvalue will ensure instability, even if the above estimate of $\rho = 1+O(1/N)$ is too rough.

For $0 \leqslant \theta < 1/2$, there is the condition (b)' to consider yet. For the problem with u specified on the boundaries, it is commonly supposed that the von Neumann stability criterion may be applied. This yields the stability condition on r

$$r \quad \leqslant \frac{1}{2(1-2\theta)} \tag{12}$$

for the above difference scheme. Now, if there are any eigenvalues of U greater than 4, the condition (b)' is only slightly stricter than the condition (12), which would be expected to apply; this is because any eigenvalue greater than 4 is, at most, $4 + O(1/N)$, and the decrease in the upper limit of r is negligible. Thus the occurrence of eigenvalues greater than 4 is ignored here as far as stability considerations are concerned. The conditions on p and q are sought which ensure that all eigenvalues are non-negative.

The characteristic equation of U is

$$f(\lambda) \equiv \quad | U - \lambda I | = 0$$

This is expanded to give a polynomial of degree N+1 in λ, viz. :

$$f(\lambda) \equiv [(\lambda-2)^2-4+ 4pq/N^2]T_{N-1}(\lambda) + 2(p+q)/N\ [T_N(\lambda) - T_{N-2}(\lambda)] = 0 \tag{13}$$

where the polynomials $T_m(\lambda)$, $(m \geqslant 1)$, are the characteristic equations of matrices of order m, and are given by

$$T_m(\lambda) = \begin{vmatrix} (2-\lambda) & -1 & 0 & & & \\ -1 & (2-\lambda) & -1 & 0 & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & -1 & (2-\lambda) & -1 \\ & & & & -1 & (2-\lambda) \end{vmatrix} \qquad (14)$$

The polynomials $T_m(\lambda)$ have been discussed by Rutherford [43,44], and the following properties can either be found in these papers, or can be deduced from the results there:

1. $T_m(\lambda)$ satisfies the recurrence relation

$$T_m(\lambda) + T_{m-2}(\lambda) = T_{m-1}(\lambda).$$

2. If $\lambda = 2(1-\cos\phi)$ then

$$T_m(\lambda) = \frac{\sin(m+1)\phi}{\sin\phi}.$$

3. The zeros of $T_m(\lambda)$ are given by

$$\beta_k = 4\sin^2\frac{\pi k}{2(m+1)}, \quad k = 1,2,\ldots.,m.$$

4.

$$T_m(\lambda) = \sum_{r=0}^{[m/2]} C_r^{m-r}(2-\lambda)^{m-2r}(-1)^r$$

where

$$[m/2] = \begin{cases} m/2, & m \text{ even,} \\ (m-1)/2, & m \text{ odd.} \end{cases}$$

5.

$$T_m(\lambda) = \sum_{r=0}^{m} C_r^{2m-r+1}(-\lambda)^{m-r}.$$

6. $T_m(0) = m+1, \qquad T_m(4) = (-1)^m(m+1).$

7. The function $T_m(\lambda) - T_{m-2}(\lambda)$ is denoted by $Q_m(\lambda)$, and from 2, it follows that

$$Q_m(\lambda) = 2\cos m\phi .$$

Thus the zeros of $Q_m(\lambda)$ are given by

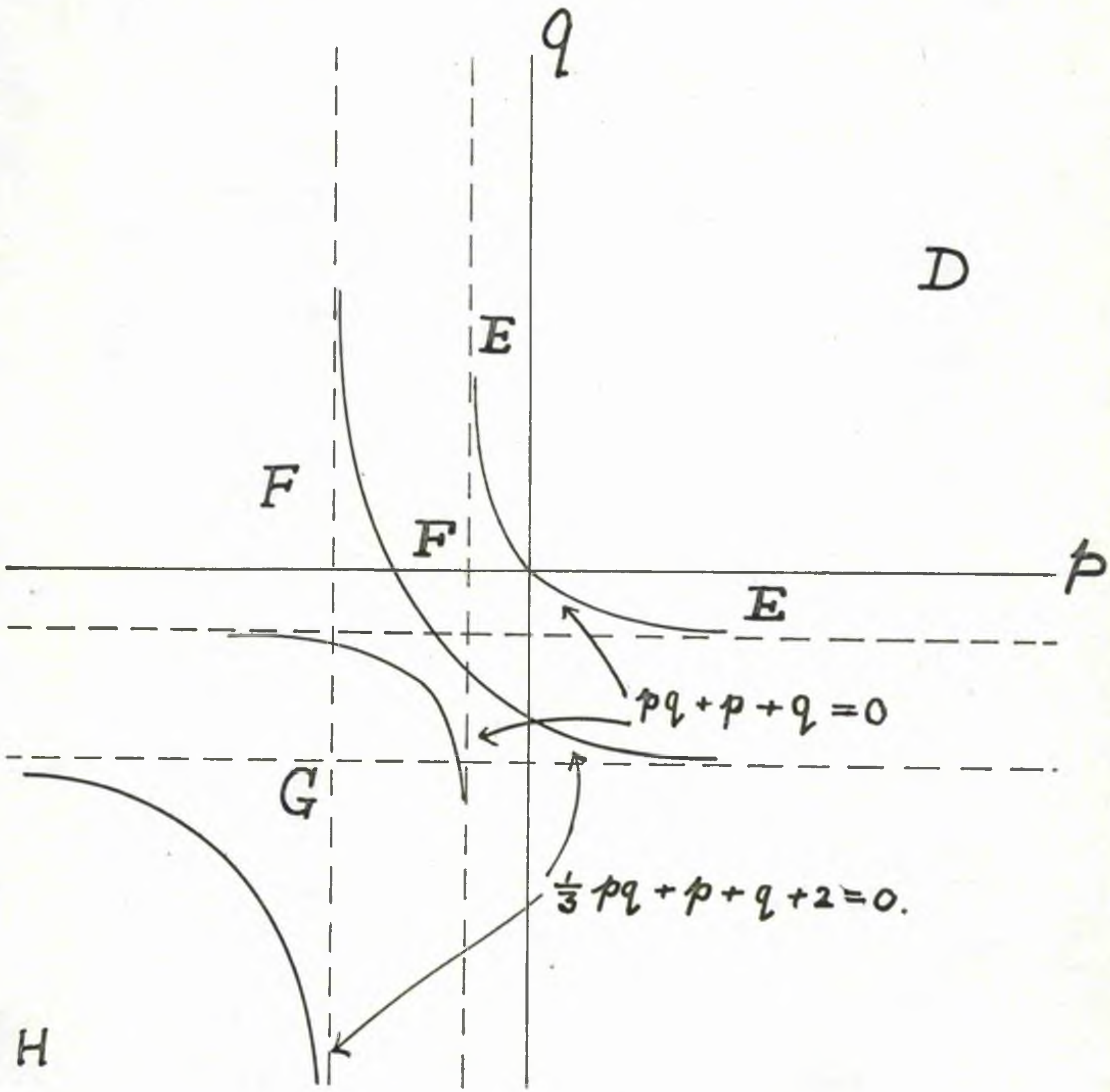

Figure 2

$$y_k = 4\sin^2\frac{(2k-1)\pi}{4(m+1)} \quad , \ k = 1,2,\ldots,m.$$

The polynomial equation $f(\lambda) = 0$ is now written in the form

$$S(\lambda)\, T_{N-1}(\lambda) + 2(p+q)/N\, Q_N(\lambda) = 0 \qquad (15)$$

where

$$S(\lambda) = (\lambda-2)^2-4 + 4pq/N^2 \ .$$

We now refer to figure 2. The (p,q) plane has been partitioned into sections as follows:

D is $p \geqslant 0$, $q \geqslant 0$, $p+q > 0$.

E (both parts)is the region in which the functions $pq+p+q$, and $p+q$ are both positive or zero.

F is the region between, but not including, the two branches of $pq+p+q = 0$G is the region between the lower branch of $pq+p+q=0$ and the lower branch of $1/3pq +p +q = 0$.

The reasons for this partition of the plane will become clearer later.

In the region D, all points p,q give positive roots for $f(\lambda) = 0$. This is obvious, since Gersohgorin's theorem shows that the roots are non-negative, and putting $\lambda = 0$ in (15) shows that this is a root if and only if (using 6) $pq + p + q = 0$; i.e. for (p,q) inside D, there are no zero roots.

In the regions E, $pq \leqslant 0$; thus, since the roots of $S(\lambda) = 0$ are

$$4pq/N^2 \quad \text{and} \quad 4-4pq/N^2$$

then $S(\lambda)$ stays negative in $(0,4)$ for values of p,q taken from E. In addition when (p,q) is inside E, $p+q > 0$ except at $p=q=0$. The case $p=q=0$ is an easy case to dispose of, since the polynomial equation becomes

$$[(\lambda-2)^2- 4]T_{N-1}(\lambda) = 0$$

with roots at $\lambda = 0$, and $\lambda = 4$, and $\lambda_k = \sin \frac{\pi k}{2N}$, $k = 1,2,\ldots,N-1$. If the roots of $T_{N-1}(\lambda)$ and $Q_N(\lambda)$ are called β_k ($k = 1,2,\ldots,N-1$) and γ_k ($k = 1, 2,\ldots,N$) respectively, then 3 and 4 show that

$0 < \gamma_1 < \beta_1 < \text{------} < \beta_{N-1} < \gamma_N < 4$, and that consequently

$$\mathrm{sign}[T_{N-1}(\gamma_k)] = (-1)^{k+1}$$

and

$$\mathrm{sign}[Q_N(\beta_k)] = (-1)^k.$$

Thus, since, for (p,q) inside E, $S(\lambda)$ and $(p+q)$ are fixed in sign, $f(\lambda)$ changes sign in each of the intervals (γ_k, γ_{k+1}) for $k = 1,2,\ldots,N$. This argument locates $(N-1)$ roots of $f(\lambda) = 0$, and only two remain to be considered. As $\lambda \to \infty$, using 5 we see that $f(\lambda)$ has the sign of $(-1)^{N-1}$. At $\lambda = 4$, the sign of $f(4)$ is (using 6) the same as the sign of $(-1)^N(pq+p+q)$; thus in E the sign of $f(4)$ is $(-1)^N$. Therefore $f(\lambda)$ has a root $\lambda > 4$. Except on the curve $pq+p+q = 0$, the sign of $f(0)$, using 6, is the sign of $pq+p+q$, i.e. positive, and the sign of $f(\gamma_1)$ is the sign of $S(\gamma_1)T_{N-1}(\gamma_1)$, i.e., negative. Thus there is a root λ in the interval $(0,\gamma_1)$. When (p,q) is on the curve $pq+p+q = 0$ this last root is $\lambda = 0$. There are therefore no negative roots for (p,q) inside the upper branch of $pq+p+q = 0$, and one zero root when (p,q) is on this upper branch.

In the region F, since $f(0)$ has the sign of $pq+p+q$, i.e. negative, and since $f(-\infty)$ is positive (as 5 readily shows) then $f(\lambda)$ has a sign change between $\lambda = -\infty$ and $\lambda = 0$. By arguments similar to the ones above, it can be shown that there is only one negative root when (p,q) is in the region F, and not a larger odd number. The knowledge, however, that there is one, is sufficient for the present problem.

For large N it can be shown, using properties 1 - 7, that the sign of the derivative of $f(\lambda)$ at $\lambda = 0$ is the sign of

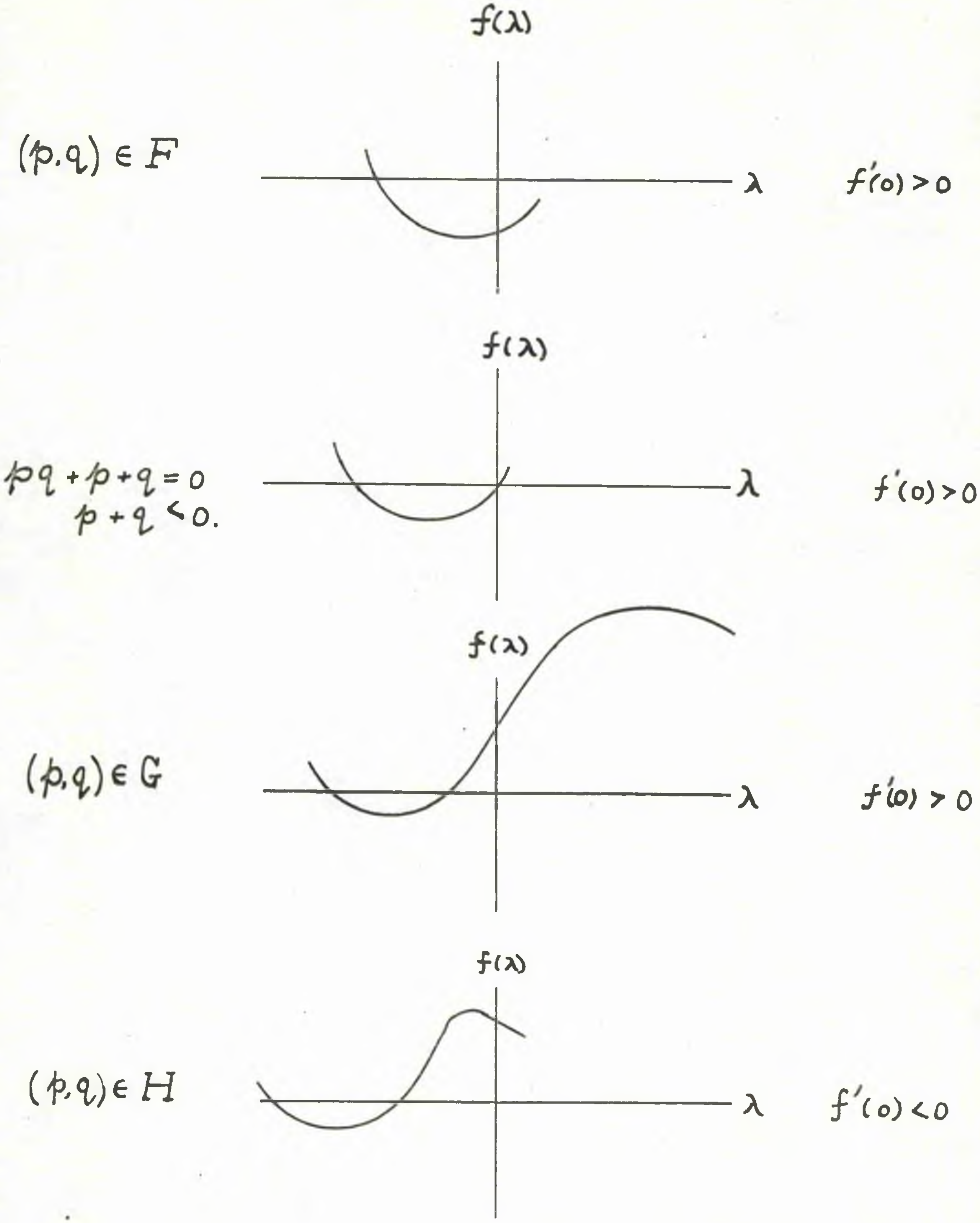

FIGURE 3.

$-(p + q + 1/3pq + 2)$. Thus, when (p,q) is on the lower branch of $pq+p+q = 0$ (where $f(\lambda)$ has a zero root), and when (p,q) is in the region G, the gradient of $f(\lambda)$ at $\lambda = 0$ is positive. As figure 3 shows, this means that there is a turning point of f in $\lambda < 0$, when (p,q) is in G, and hence that there are at least two roots in $\lambda < 0$, since $f(0)$ is positive in G. When (p,q) is in H (figure 2) it seems reasonable to assume that $f(\lambda) = 0$ has two roots in $\lambda < 0$, as in G. This fact (plausible from figure 3) is estlablished, later, by numerical examination of the polynomial $f(\lambda)$.

The conclusions of this section may be summarised thus:

(a) $\forall(p,q)$ in $pq+p+q \geqslant 0$, $p+q \geqslant 0$, $f(\lambda)$ has no negative roots.

(b) $\forall(p,q)$ in $pq+p+q = 0$, $p+q \geqslant 0$, $f(\lambda)$ has one zero root, and the other roots are positive.

(c) $\forall(p,q)$ in $pq+p+q < 0$, $f(\lambda)$ has one root which is negative, while the others are positive.

(d) $\forall(p,q)$ in $pq+p+q = 0$, $p+q < 0$, $f(\lambda)$ has one root $\lambda < 0$, and one zero root.

(e) $\forall(p,q)$ in $pq+p+q > 0$, $p+q < 0$, $f(\lambda)$ has two negative roots.

If p,q are such that any one of (c),(d), (e) is satisfied, then the difference method is unstable.

1.7 Alternative Boundary Replacement.

As an alternative to the replacement for $\frac{\partial u}{\partial x}$ given by (3a), there is the following difference replacement used by Lowan [35] and others. When $j = 1$ or $N-1$, in (3), the values of v_0^m and v_N^m $(m = n, n+1)$, are obtained by using the boundary conditions, with the derivatives in the difference form:

$$\frac{\partial u_j^m}{\partial x} = \frac{1}{\Delta x}\left(v_{j+1}^m - v_j^m\right), \quad j = 0, N-1, \quad m = n, n+1 \qquad (16)$$

The matrices A and B in (4a) are now of order N-1 and U is given by

$$U = \begin{vmatrix} \frac{1+2p/N}{1+p/N} & -1 & 0 & & & \\ -1 & 2 & -1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & -1 & 2 & -1 \\ & & & & -1 & \frac{1+2q/N}{1+q/N} \end{vmatrix} \qquad (17)$$

The matrix U is now symmetric. The main objection to this difference replacement of the derivatives is that it is accurate only to $O(\Delta x)$ and is therefore of a lower accuracy than the difference replacement of the differential equation. The characteristic equation for U is given by

$$g(\lambda) \equiv [pq/N^2 + (p+q)/N]T_{N-1}(\lambda) - [(p+q)/N + \lambda]T_{N-2}(\lambda) = 0 \qquad (18)$$

It is assumed that $|p|, |q| << N$. The equation (18) is seen to have a root $\lambda = 0$ only when $pq+p+q = 0$.

Using the properties of $T_m(\lambda)$ given in section 6, it is seen that $g(0) = 1/N(pq+p+q)$ and $g(-\infty) > 0$ for large N. Thus, when (p, q) is in the region F, $f(0) < 0$, and $f(-\infty) > 0$. There is therefore a negative eigenvalue. The Gerschgorin theorem applied to the matrix U in (18) shows that the eigenvalues lie in the interval

$$\min\left(\frac{p/N}{1+p/N},\frac{q/N}{1+q/N},0\right) \leq \lambda \leq \max\left(1+\frac{1+2p/N}{1+p/N},1+\frac{1+2q/N}{1+q/N},4\right)$$

Thus if $p,q \geq 0$, it follows that $\lambda_j \geq 0$, for $j = 1,2,\ldots,N-1$. Since g has a zero root only on $pq+p+q = 0$, it follows that $\lambda_j > 0$ for $j = 1,2,\ldots,N-1$ except when $p=q=0$, when one root is zero and the others are positive.

By 3 in section 6, the roots of $T_{N-1}(\lambda)$ are

$$\beta_k = 4\sin\frac{\pi k}{2N} \qquad k = 1,2,\ldots,N-1,$$

and the roots of $T_{N-2}(\lambda)$ are

$$\beta'_k = 4\sin\frac{\pi k}{2(N-1)} \quad k = 1,2,\ldots,N-2.$$

Thus these are spaced in the interval $(0,4)$ so that

$$\beta'_{k-1} < \beta_k < \beta'_k$$

since

$$\frac{k-1}{N-1} < \frac{k}{N} < \frac{k}{N-1}$$

when $k < N$.

Thus

$$0 < \beta_1 < \beta'_1 < \beta_2 < \cdots < \beta'_{N-2} < \beta_{N-1} < 4 \qquad (19).$$

If $p+q \geq 0$, then the sign of $(p+q)/N + \lambda$ is positive for λ in $(0,4)$. Thus because of (19), $g(\lambda)$ changes sign in each interval (β_k,β_{k+1}). This argument locates N-2 of the N-1 roots. When (p,q) is in E, $g(0) > 0$ and $g(-\infty) > 0$ so that the root which is still unlocated cannot be negative. Hence in D and E $g(\lambda)$ has no negative roots, and has a zero root only on the boundary of E, $pq+p+q = 0$.

In fact it can be shown that the number of negative roots which $g(\lambda)$ has in F and G (it has already been shown that there is a negative root in F), is the same as for the polynomial $f(\lambda)$. This may be proven algebraically but the proof is tedious; it is demonstrated numerically, later.

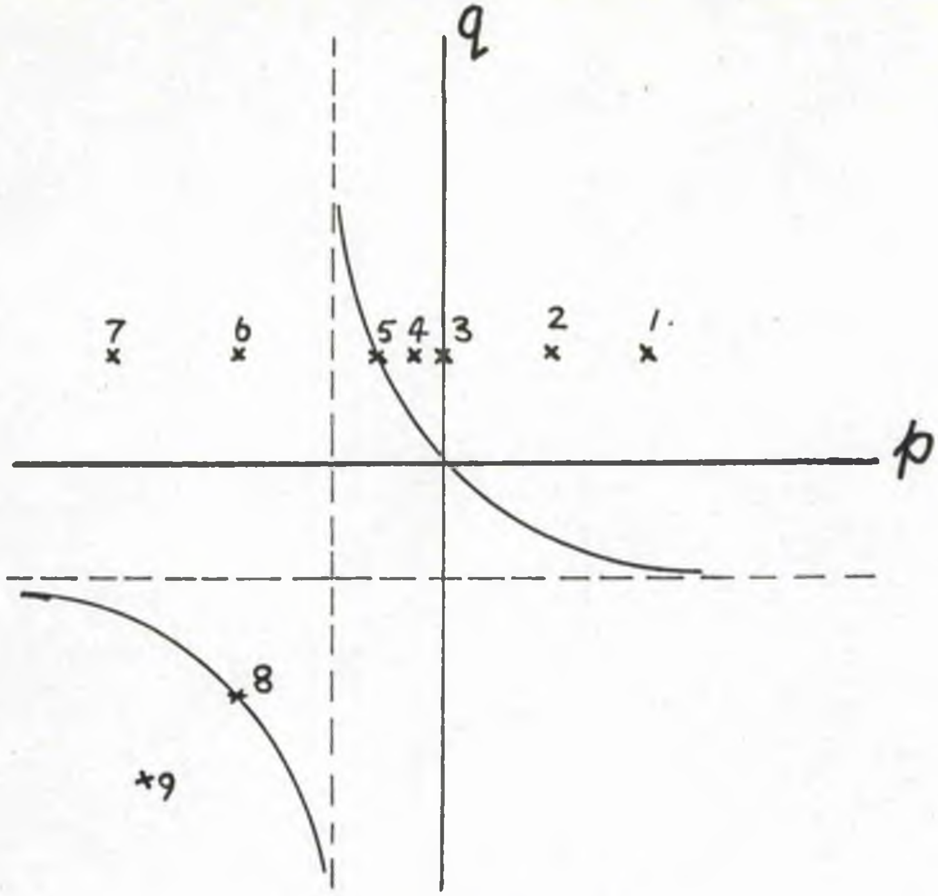

FIGURE 4

Pt.	(p, q)	$f(\lambda)$	$g(\lambda)$	$h(\lambda)$
1	1, 1	NONE	NONE	NONE
2	$\frac{1}{2}$, 1	NONE	NONE	NONE
3	0, 1	NONE	NONE	NONE
4	−1/4, 1	NONE	NONE	NONE
5	$-\frac{1}{2}$, 1	zero root	zero root	zero root
6	−1, 1	root = −·003	root = −·003	root = −·003
7	−2, 1	root = −·010	root = −·010	root = −·010
8	−2, 2	zero root root = −·015	zero root root = −·015	zero root root = −·015
9	−3, −3	root = −·016 root = −·027	root = −·016 root = −·027	root = −·020 root = −·030

Table 1

A third replacement of $\frac{\partial u}{\partial x}$ which, in common with the one in (16) eliminates v_0^m and v_N^m ($m = n, n+1$), but which is accurate to $O(\Delta x)^2$, was used by Batten [2]. This is given by

$$\frac{\partial v_0^m}{\partial x} = \frac{1}{2\Delta x}(-v_2^m + 4v_1^m - 3v_0^m)$$

and

$$\frac{\partial v_N^m}{\partial x} = \frac{1}{2\Delta x}(v_{N-2}^m - 4v_{N-1}^m + 3v_N^m)$$

for $m = n, n+1$. With this replacement, the characteristic polynomial for the matrix U which arises is

$$\begin{aligned} h(\lambda) = 4pq/N^2 T_{N-1}(\lambda) + 2(p+q)/N[T_{N-3}(\lambda) + 3T_{N-1}(\lambda) - 4T_{N-2}(\lambda)] \\ +(\lambda^2 + 4\lambda)T_{N-3}(\lambda) - 8\lambda T_{N-2}(\lambda) = 0 \end{aligned} \tag{21}$$

Equation (21) has a zero root again only on the curve $pq+p+q = 0$. The examination of $h(\lambda)$ is carried out numerically at the end of this section, where it is shown that negative roots arise for exactly the same values of (p, q) as for $f(\lambda)$ and $g(\lambda)$.

The numerical tests made on f, g, h were as follows: the polynomials were evaluated for $\lambda = 0(-\cdot001)-0\cdot1$, for the case $N=20$, and for various values of p and q. These are marked in figure 4. The sign changes of the polynomials were noted, together with the rate at which they were increasing at $\lambda = -0\cdot1$. The negative eigenvalues, if any, were in the interval $(-0\cdot1, 0)$, as is to be expected, since if there are any negative roots they are of magnitude $O(1/N)$, or $O(0\cdot5)$. A selection of the results is shownnin table 1. For each (p, q) the position of any negative roots is noted for f, g, and h.

From the position of the points (p, q) chosen, and from the results, it is evident that the conclusions (a) - (e) in section 6 are justified.

1.8 Mixed Boundary Conditions.

When the boundary conditions are given in the form

$$\begin{aligned} \frac{\partial u}{\partial x} - pu &= \phi_0(t), \quad x = 0, \quad t \geq 0 \\ u &= \phi_1(t), \quad x = 1, \quad t \geq 0 \end{aligned} \qquad (22)$$

the preceding analysis is not valid. For a problem with boundary conditions (22) the matrices occurring in the difference solution using the first boundary replacement, are of order N. With the other two replacements the matrices are still of order N-1. It is easy to show that the polynomial equations resulting from the boundary conditions (22) are those obtained by letting $q \to \infty$ in the equations $f(\lambda) = 0$, $g(\lambda) = 0$, $h(\lambda) = 0$, after first dividing by q. The polynomial equations are then:

$$\begin{aligned} f(\lambda) &\equiv 2p/NT_{N-1}(\lambda) + T_N(\lambda) - T_{N-2}(\lambda) = 0 \\ g(\lambda) &\equiv (p/N + 1)T_{N-1}(\lambda) - T_{N-2}(\lambda) = 0 \\ h(\lambda) &\equiv 2p/NT_{N-1}(\lambda) + T_{N-3}(\lambda) + 3T_{N-1}(\lambda) - 4T_{N-2}(\lambda) = 0 \end{aligned} \qquad (23)$$

The value of each of these polynomials at $\lambda = 0$ is

$$\begin{aligned} f(0) &= 2(p+1) \\ g(0) &= p+1 \\ h(0) &= 2(p+1) \end{aligned}$$

Thus, since the three polynomials are positive at $\lambda = -\infty$, there will a negative root if $p < -1$. If $p = -1$ there is a zero root. By means of simple arguments of the type used in the previous two sections, it is easily seen that there are no negative roots if $p \geq -1$. For the similar problem when $p = +\infty$ we require $q \geq -1$.

1.9 Numerical Results.

To demonstrate the results of the previous sections problem (1) was taken with the boundary conditions:

$$\frac{\partial u}{\partial x} - pu = \pi e^{-\pi^2 t}, \quad x = 0,\ t \geq 0,$$

$$\frac{\partial u}{\partial x} + qu = -\pi e^{-\pi^2 t}, \quad x = 1,\ t \geq 0,$$

and with the initial conditions

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

The solution of this in R is

$$u(x,t) = e^{-\pi^2 t}\sin\pi x, \ \forall p,q.$$

The difference method used was (3) with $\theta = \frac{1}{2}$ (i.e. the Crank-Nicolson method) and with the first boundary replacement. The values $N = 20$ and $r = 1$ were taken. The method was run for 400 time steps (i.e. to $t = 1$) and the errors were calculated at every 20^{th} time step. These errors are shown in table 2, for the values of (p,q) in table 1. All errors are quoted at $x = \frac{1}{2}$, where the solution is largest, and the value of the theoretical solution of the differential equation at this point is also given. K is the number of time steps.

When (p,q) lies on the upper branch of the hyperbola $pq+p+q = 0$, we have an instance of the phenomenon first noted by Parker and Crank [38]. For example, when $(p,q) = (-\frac{1}{2},1)$, the theoretical solution decays to zero, while the error tends to a non-zero limit. This was called a persistent error by Parker and Crank, and was at first ascribed to the presence of discontinuities in the initial conditions, or between the initial and boundary conditions. However, the preceding analysis shows that such errors are the result of the matrix $A^{-1}B$ having spectral radius unity.

TABLE 2.

K	Theoretical Solution.	p = 1·0 q = 1·0	0·5 1·0	0·0 1·0	−0·25 1·0
20	0·610 498	0·000 311	0·000 302	0·000 293	0·000 287
40	0·372 707	0·000 050	0·000 089	0·000 135	0·000 161
60	0·227 537	0·000 329	0·000 397	0·000 484	0·000 535
80	0·138 911	0·000 483	0·000 575	0·000 696	0·000 771
100	0·084 804	0·000 556	0·000 664	0·000 813	0·000 910
120	0·051 773	0·000 579	0·000 699	0·000 872	0·000 987
140	0·031 607	0·000 573	0·000 702	0·000 894	0·001 026
160	0·019 296	0·000 552	0·000 687	0·000 895	0·001 042
180	0·011 780	0·000 522	0·000 662	0·000 883	0·001 044
200	0·007 191	0·000 489	0·000 631	0·000 863	0·001 037
220	0·004 390	0·000 455	0·000 598	0·000 839	0·001 025
240	0·002 680	0·000 421	0·000 565	0·000 813	0·001 010
260	0·001 636	0·000 389	0·000 532	0·000 787	0·000 993
280	0·000 999	0·000 358	0·000 500	0·000 760	0·000 975
300	0·000 609	0·000 329	0·000 469	0·000 733	0·000 957
320	0·000 372	0·000 303	0·000 441	0·000 707	0·000 939
340	0·000 227	0·000 278	0·000 413	0·000 682	0·000 921
360	0·000 138	0·000 256	0·000 388	0·000 657	0·000 903
380	0·000 084	0·000 235	0·000 364	0·000 634	0·000 886
400	0·000 051	0·000 216	0·000 341	0·000 611	0·000 868

TABLE 2 (contd.)

K	p = -0·5 q = 1·0	-1·0 1·0	-2·0 1·0	-2·0 -2·0	-3·0 -3·0
20	0·000 282	0·000 270	0·000 270	0·000 173	0·000 099
40	0·000 191	0·000 258	0·000 445	0·000 841	0·001 450
60	0·000 594	0·000 738	0·001 196	0·002 076	0·003 945
80	0·000 859	0·001 087	0·001 913	0·003 413	0·007 704
100	0·001 026	0·001 341	0·002 637	0·004 957	0·013 676
120	0·001 129	0·001 534	0·003 424	0·006 859	0·023 505
140	0·001 193	0·001 691	0·004 322	0·009 295	0·039 936
160	0·001 233	0·001 827	0·005 380	0·012 483	0·067 572
180	0·001 258	0·001 951	0·006 651	0·016 694	0·114 164
200	0·001 273	0·002 070	0·008 192	0·022 284	0·192 776
220	0·001 282	0·002 189	0·010 071	0·029 721	0·325 457
240	0·001 288	0·002 308	0·012 368	0·039 623	0·549 418
260	0·001 291	0·002 431	0·015 181	0·052 816	0·927 472
280	0·001 294	0·002 559	0·018 628	0·070 395	1·565 649
300	0·001 295	0·002 692	0·022 854	0·093 821	2·642 938
320	0·001 296	0·002 831	0·028 037	0·125 040	4·461 485
340	0·001 296	0·002 977	0·034 393	0·166 648	7·531 319
360	0·001 297	0·003 130	0·042 190	0·222 099	12·7
380	0·001 297	0·003 290	0·051 754	0·296 000	21·4
400	0·001 297	0·003 459	0·063 485	0·394 492	36·2

It appears, from numerical calculations and from the results in the previous sections, that initial discontinuities do not have a lasting effect on the calculations. They will obviously be important for a short time, but as time increases in a problem such as the one considered here, the effect of the discontinuities will disappear. This has been shown by Pearson [40] and by Albasiny [1], and is referred to by Walsh([50], page 115).

It is clear, from Tables 1 and 2, that the numerical error grows, and becomes unbounded, only in the cases for which U has a negative eigenvalue.

1.10 An Alternative Difference Approximation.

It may appear that the results obtained in the previous sections arise from the difference scheme (3). Accordingly, another difference method is examined, viz. the Du-Fort Frankel scheme. This is an explicit method, using the five points shown in figure 5,

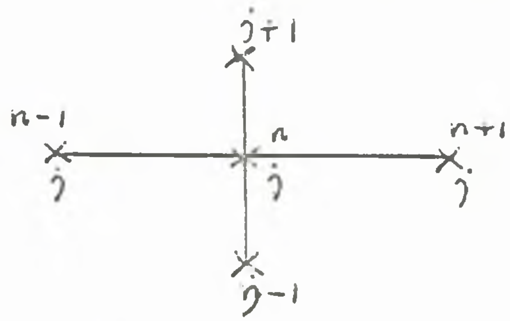

Figure 5

and given by

$$v_j^{n+1} - v_j^n = 2r(v_{j+1}^n - v_j^{n+1} - v_j^{n-1} + v_{j-1}^n) \tag{24}$$

When the first boundary replacement is used, (24) may be written in the matrix form

$$\underset{\sim}{v}^{n+1} = A\underset{\sim}{v}^n + \frac{2r}{1+2r}\underset{\sim}{v}^{n-1} + \underset{\sim}{k}^n \tag{25}$$

where $\underset{\sim}{k}^n$ is a vector involving the boundary conditions, and where

$$A = 2I - U \tag{26}$$

the matrix U being given by (5). The associated error equation is

$$\underset{\sim}{\epsilon}^{n+1} = A\underset{\sim}{\epsilon}^n + a\underset{\sim}{\epsilon}^{n-1} \tag{27}$$

where $a = \frac{2r}{1+2r}$, and may be written in the form:

$$\begin{bmatrix} \underset{\sim}{\epsilon}^{n+1} \\ \underset{\sim}{\epsilon}^{n} \end{bmatrix} = \begin{bmatrix} A & aI \\ I & 0 \end{bmatrix} \begin{bmatrix} \underset{\sim}{\epsilon}^{n} \\ \underset{\sim}{\epsilon}^{n-1} \end{bmatrix} \tag{28}$$

The condition for stability of (25) is thus that the norms of

$$\begin{bmatrix} A & aI \\ I & 0 \end{bmatrix}^n$$

be bounded for all n. If D is the diagonal matrix defined in section 4 then the matrix

$$\begin{bmatrix} \sqrt{a}D & 0 \\ 0 & D \end{bmatrix}^{-1} \begin{bmatrix} A & aI \\ I & 0 \end{bmatrix} \begin{bmatrix} \sqrt{a}D & 0 \\ 0 & D \end{bmatrix}$$

is symmetric and similar to the matrix

$$G = \begin{bmatrix} A & aI \\ I & 0 \end{bmatrix}$$

Thus (25) is stable if and only if $\rho(G) \leq 1$. Elementary row operations on the determinant

$$\left| G - \mu \begin{vmatrix} I & 0 \\ 0 & I \end{vmatrix} \right| = \begin{vmatrix} A - \mu I & aI \\ I & -\mu I \end{vmatrix} = 0$$

give

$$\begin{vmatrix} A - \left(\frac{\mu^2 - a}{\mu}\right) I & aI \\ 0 & -\mu I \end{vmatrix} = 0$$

Thus the eigenvalues μ_j $(j = 0,1,\ldots\ldots,2N+1)$, are the roots of

$$\frac{\mu_j^2 - a}{\mu_j} = \theta_j$$

where θ_j are the eigenvalues of A; or, since $A = 2I - U$,

$$\frac{\mu_j^2 - a}{\mu_j} = 2 - \lambda_j\,, \quad j = 0,1,\ldots,N.$$

where λ_j are the eigenvalues of U. It can be shown that for any r there is an eigenvalue of G greater than one in modulus if and only if U has a negative eigenvalue. Thus the Du-Fort Frankel scheme is stable or unstable for the same class of problems as the scheme (3).

1.11 Conclusion.

The main conclusion of this chapter is not that there are values of p and q for which the difference schemes (3) and (25) are unstable, but that both difference schemes and all three boundary replacements (and in fact other difference schemes not mentioned here) are stable or unstable for the same problems. There will be instability unless the two conditions

$$pq + p + q \geq 0, \text{ and } p + q \geq 0$$

are both satisfied. This motivates the investigations of the next chapter.

CHAPTER 2.

THE ASYMPTOTIC BEHAVIOUR OF THE SOLUTIONS OF SOME BOUNDARY VALUE PROBLEMS.

2.1 Introduction.

In the preceding chapter, the effect of boundary conditions on the stability of numerical solutions of a simple parabolic equation was discussed. It was shown that the stability properties of the difference systems were independent of the difference method used, and of the method of approximating the boundary conditions. A stable class of boundary conditions was defined independently of the fineness of the difference approximations, and this suggested that the instability which was found in the numerical calculations could be traced to the differential system. In this chapter, therefore, the theoretical solutions of some parabolic boundary value problems are obtained, and their asymptotic behaviour (i.e. as $t \to \infty$) is examined.

2.2 The One Dimensional Equation Of Heat.

Consider the differential equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \tag{1}$$

in the region

$$R = [0 \leqslant x \leqslant 1] \times [t \geqslant 0]$$

for the function $u(x,t)$ subject to the initial condition

$$u(x,0) = f(x), \quad 0 \leqslant x \leqslant 1 \tag{1a}$$

and the boundary conditions

$$\begin{aligned} \frac{\partial u}{\partial x} - pu &= \phi_0(t), \quad x = 0,\ t \geqslant 0, \\ \frac{\partial u}{\partial x} + qu &= \phi_1(t), \quad x = 1,\ t \geqslant 0. \end{aligned} \tag{1b}$$

The same continuity assumptions on f, ϕ_0, ϕ_1 are made as were made in chapter 1. The problem consisting of (1), (1a), and (1b) is linear, and its solution may therefore be obtained as the sum of the solutions of the three problems:

(a) (1),(1a),(1b) with $\phi_0 = \phi_1 = 0$.

(b) (1),(1a),(1b) with $f = \phi_1 = 0$.

(c) (1),(1a),(1b) with $f = \phi_0 = 0$.

Problem (a) is solved by a Sturm-Liouville method, while (b) and (c) are solved by means of the Laplace Transform.

2.3 The Sturm-Liouville Problem.

If $\phi_0 = \phi_1 = 0$ then equation (1) subject to (1a) and (1b) may be solved by separation of variables. Let

$$u(x,t) = X(x)T(t)$$

Then, from (1)

$$T' = -\lambda T \tag{2}$$

and

$$X'' = -\lambda X \tag{3}$$

where

$$\left.\begin{aligned} X'(0) - pX(0) &= 0 \\ X'(1) + qX(1) &= 0 \end{aligned}\right\} \tag{4}$$

The solution of equation (2) is

$$T(t) = e^{-\lambda t}.$$

Equations (3) and (4) give a Sturm-Liouville problem for $X(x)$. There are non-trivial solutions only when λ is an eigenvalue and the associated $X(x)$ is an eigenfunction. It is well known (see e.g. [11]) that the eigenvalues form a denumerable set of real numbers

$$\lambda_0 < \lambda_1 < \lambda_2 < \cdots\cdots$$

which have no limit points. The eigenfunction corresponding to the eigenvalue λ_n is X_n. The set $\{ X_n \mid n = 0, 1, \ldots \}$, when normalised, forms an orthonormal set of functions on the interval $[0,1]$. If f'' is continuous in $[0,1]$ it may be expressed in the form:

$$f(x) = \sum_{n=0}^{\infty} a_n X_n(x) \tag{5}$$

where

$$a_n = \int_0^1 X_n(x) f(x)\,dx \tag{6}$$

and the series (5) is uniformly convergent in $[0,1]$.

Equation (1) for the case (a) thus has the solution

$$u(x,t) = \sum_{n=0}^{\infty} a_n e^{-\lambda_n t} X_n(x), \tag{7}$$

where $\{a_n\}$ are given by (6). This series for $u(x,t)$ is uniformly and absolutely convergent for all t if and only if $\lambda_n \geq 0, \forall n$;i.e. if and only if $\lambda_0 \geq 0$. If $\lambda_0 < 0$, then $u(x,t)$ becomes unbounded as $t \to \infty$. Thus, whenever the Sturm-Liouville problem, (3) and (4), has a negative eigenvalue, the equation (1) has a solution growing exponentially in time. If all the eigenvalues are positive, $u(x,t) \to 0$ as $t \to \infty$; and if $\lambda_0 = 0$, then $u(x,t)$ does not tend to zero but reaches a "steady-state" value

$$u_{\infty}(x) = a_0 X_0(x).$$

The solution of equation (3) is

$$X(x) = A\cos\sqrt{\lambda}x + B\sin\sqrt{\lambda}x,$$

where, from the boundary conditions,

$$\left.\begin{aligned} \sqrt{\lambda}B - pA &= 0 \\ \sqrt{\lambda}(-A\sin\sqrt{\lambda} + B\cos\sqrt{\lambda}) + q(A\cos\sqrt{\lambda} + B\sin\sqrt{\lambda}) &= 0 \end{aligned}\right\} \tag{8}$$

The eigenvalues of the Sturm-Liouville problem are thus the roots of the equation

$$(p+q)\sqrt{\lambda}\cos\sqrt{\lambda} = (\lambda - pq)\sin\sqrt{\lambda}. \tag{9}$$

Since the roots are real, putting $\lambda = \theta^2$ gives the equation

$$(p+q)\theta\cos\theta = (\theta^2 - pq)\sin\theta \tag{10}$$

for θ, which is real or pure imaginary. Equation (9) has negative roots if and only if, putting $\theta = i\phi$, there are real roots ϕ of the equation

$$(p+q)\phi\cosh\phi = -(\phi^2 + pq)\sinh\phi. \tag{11}$$

This equation is discussed in section 5.

2.4 The Boundary Value Problem.

The problems (b) and (c) are similar, and the following solution of (b) applies to (c) with only slight modification. The initial condition is now

$$u(x,0) = 0 \, , \quad 0 \leq x \leq 1,$$

and the boundary conditions are

$$\frac{\partial u}{\partial x} - pu = \phi_0(t), \quad x = 0, \quad t \geq 0,$$

$$\frac{\partial u}{\partial x} + qu = 0, \quad x = 1, \quad t \geq 0.$$

The Laplace transform of u is denoted by $\hat{u}$, and is given by:

$$\hat{u}(x,s) = \int_0^\infty e^{-st} u(x,t)dt.$$

The transform $\hat{u}$ satisfies the differential equation

$$\frac{d^2\hat{u}}{d\,x^2} - s\hat{u} = 0 \tag{12}$$

subject to the boundary conditions

$$\left.\begin{aligned} \frac{d\hat{u}}{dx} - p\hat{u} &= \hat{\phi}_0(s) \\ \frac{d\hat{u}}{dx} + q\hat{u} &= 0 \end{aligned}\right\} \tag{13}$$

where $\hat{\phi}_0$ is the Laplace transform of ϕ_0.

In order to apply the Laplace transform method, it must be assumed that $u(x,t)$ is bounded for x in $[0,1]$, and for all $t \geq 0$. The solution of (12) subject to (13) is

$$\hat{u}(x,s) = -\hat{\phi}_0(s) \frac{\sqrt{s} \cosh\sqrt{s}(1-x) + p \sinh\sqrt{s}(1-x)}{(s+pq)\sinh\sqrt{s} + (p+q)\sqrt{s}\cosh\sqrt{s}} \tag{15}$$

The solution $u(x,t)$ is now obtained by inverting the Laplace transform; i.e.

$$u(x,t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{st}\, \hat{u}(x,s)\, ds \tag{15}$$

where $c > 0$ is a constant.

The integral in equation (15) is calculated by the method of residues, using a suitable contour enclosing the poles of the integrand. A series is obtained for $u(x,t)$, which is seen to be bounded and uniformly convergent for x in $[0,1]$ and $t \geq 0$, if and only if all the poles of (14) lie in the region

$$R(s) \leq 0.$$

The equation for these poles is

$$(s + pq)\sinh\sqrt{s} = -(p+q)\sqrt{s}\cosh\sqrt{s}.$$

Let $\lambda = -s$ and the equation becomes

$$(\lambda - pq)\sin\sqrt{\lambda} = (p+q)\sqrt{\lambda}\cos\sqrt{\lambda}, \tag{16}$$

with the condition for boundedness replaced by $\lambda \geq 0$, since the roots of (16) are real, as shown in the previous section, (equation (9)).

The condition for the solution of equation (1) subject to (1a) and (1b) to be bounded for all positive t, is therefore the condition, or conditions on p and q which ensure that equation (11) in ϕ has no real roots.

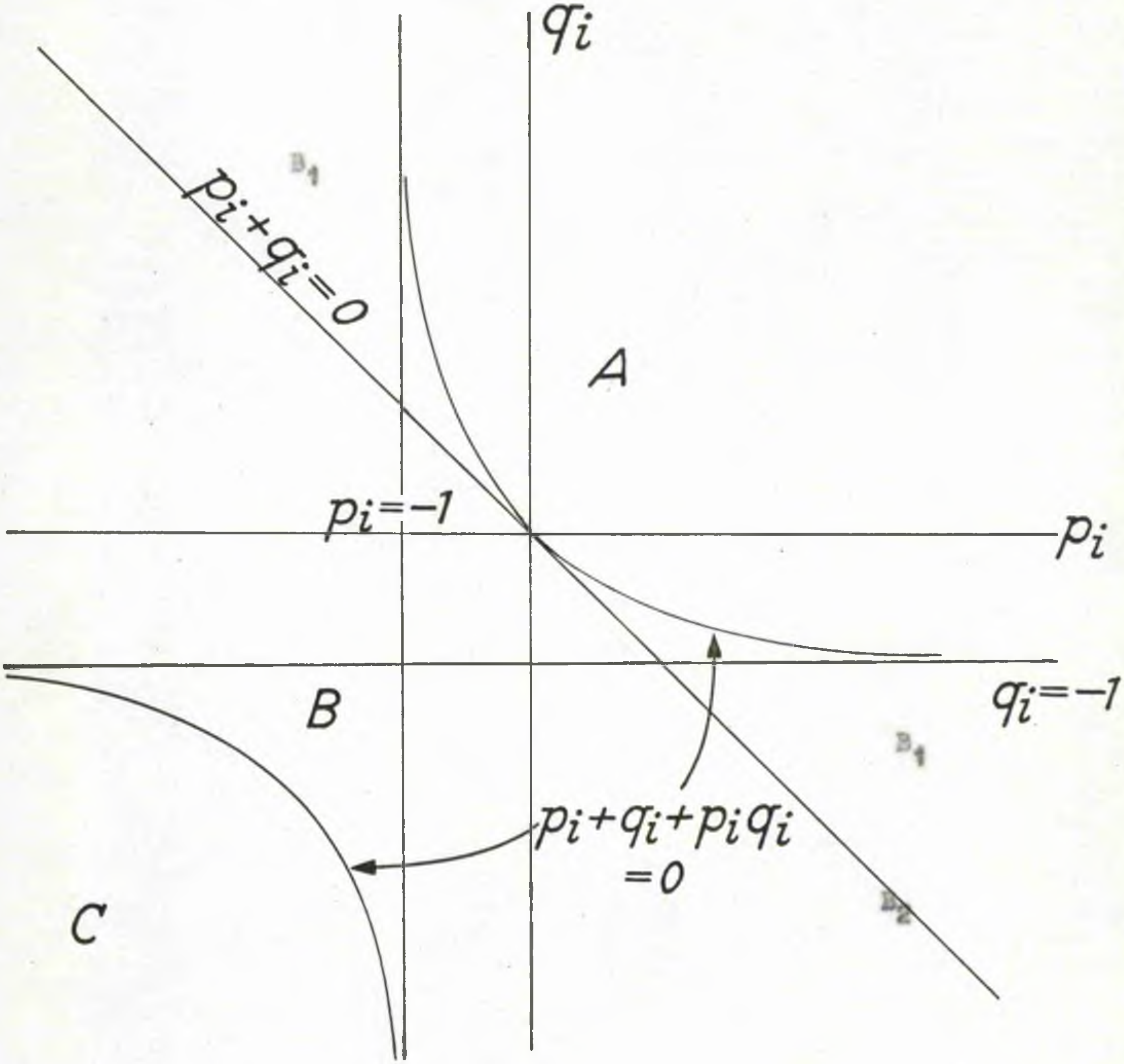

FIGURE 1a.

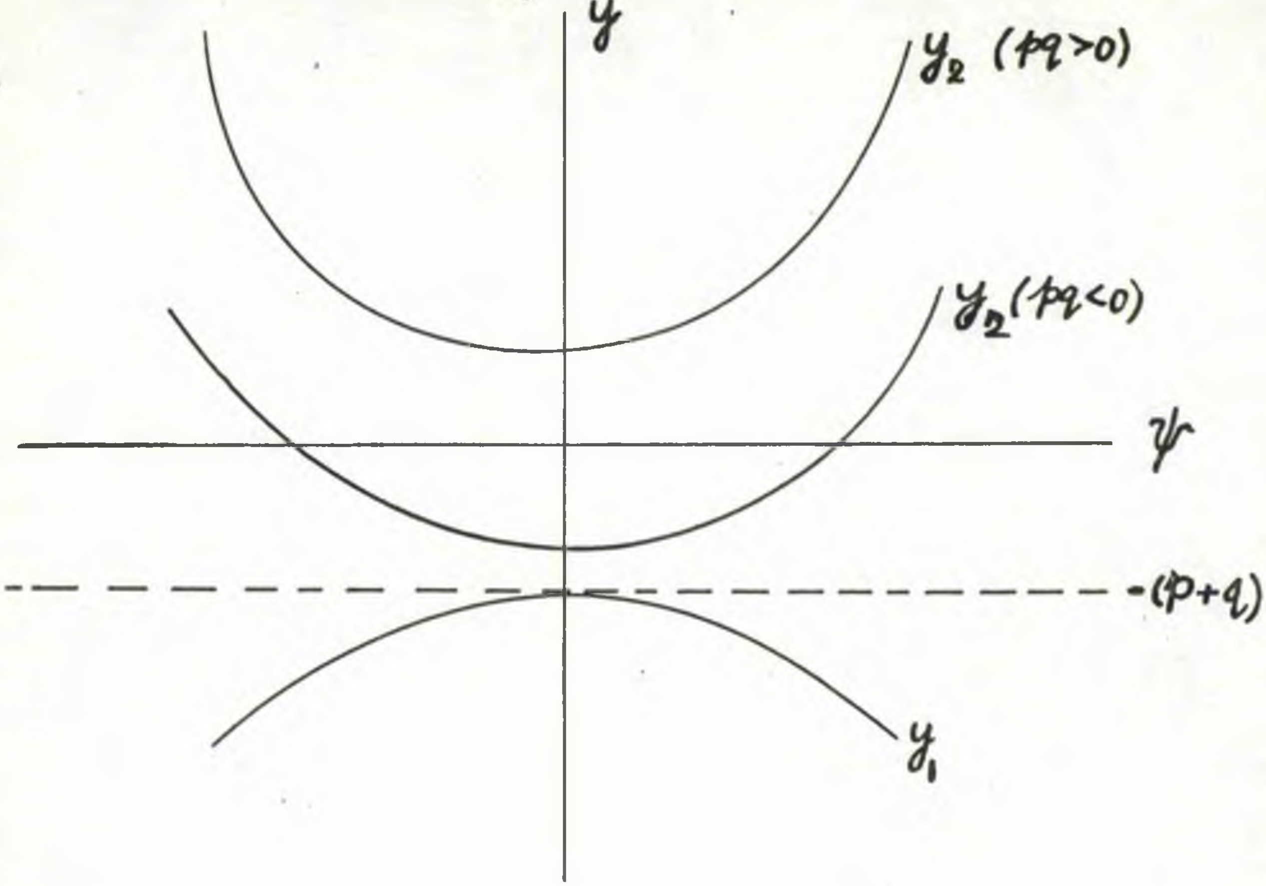

FIGURE 2

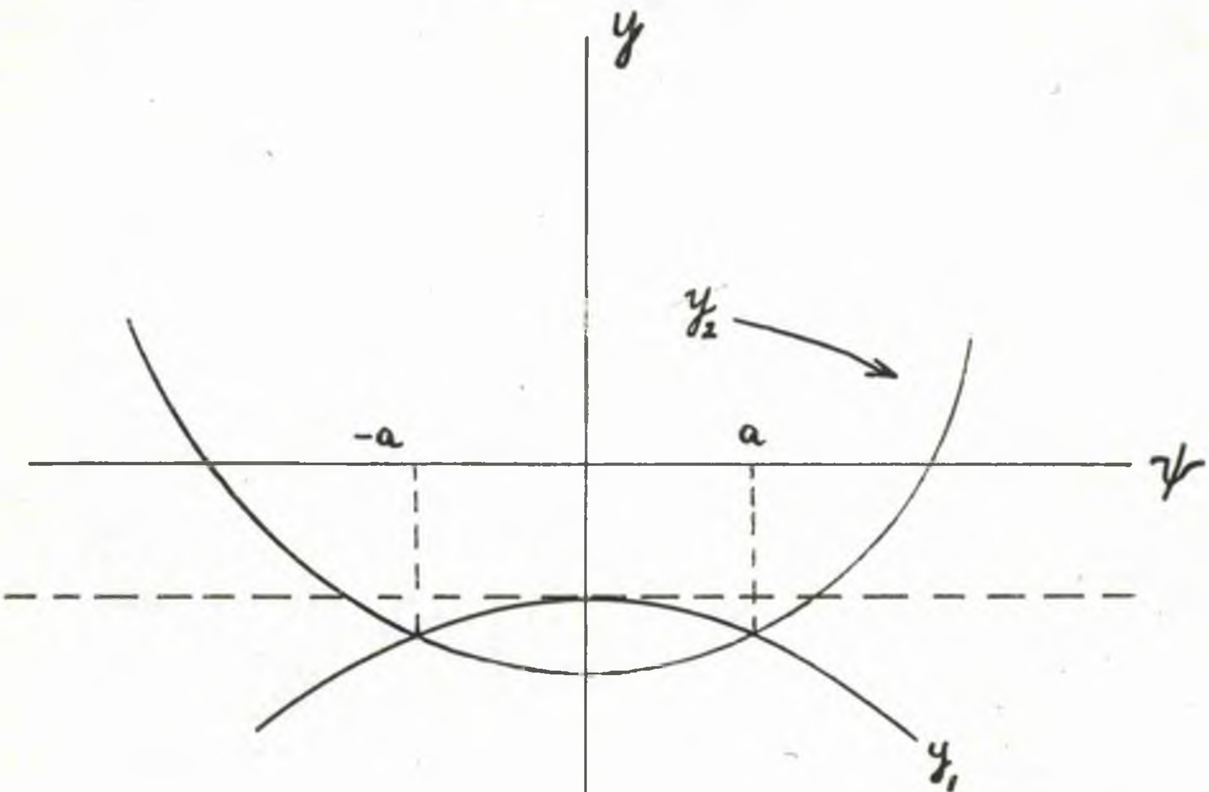

FIGURE 3

2.5 The Eigenvalues Of The Sturm-Liouville Problem.

Consider the two functions given by

$$y_1 = -(p+q)\phi\coth\phi$$

$$y_2 = \phi^2 + pq$$

Then the real roots of (11) are the intersections, if any, of y_1 and y_2. The (p,q) plane is partitioned into the regions:

$$A : \; pq + p + q \geqslant 0, \;\; p + q \geqslant 0.$$

$$B_1 : \; pq + p + q < 0, \;\; p + q > 0.$$

$$B_2 : \; p + q = 0, \;\; pq \neq 0,$$

$$B : \; pq + p + q < 0, \;\; p + q < 0.$$

$$C : \; pq + p + q \geqslant 0, \;\; p + q < 0.$$

These regions are shown in figure 1.

A : In region A, y_1 is as shown in figure 2, with a maximum at $\phi = 0$, given by $y_1 = -(p+q)$. The function y_2 has a minimum at $\phi = 0$, given by $y_2 = pq$. In A, pq may be positive or negative, but always $pq \geqslant -(p+q)$. (Both cases are shown in figure 2). Thus y_1 and y_2 do not intersect, except possibly when $\phi = 0$. This happens on the curve $pq + p + q = 0$. There are therefore no negative eigenvalues when (p,q) is in A, and one zero eigenvalue when (p,q) is on the hyperbola $pq + p + q = 0$.

B_1 : In the region B_1, y_1 and y_2 have the same form as in A (see figure 3), but now the minimum of y_2 is less than the maximum of y_1, since $pq < -(p+q)$, and so there are two roots $\phi = \pm a$, for some a. Thus there is one negative eigenvalue $\lambda = -a^2$.

B_2 : When $p+q = 0$, equation (11) takes the form

$$(\phi^2 + pq)\sinh\phi = 0.$$

Since $pq < 0$ in B_2, there are two roots $\phi = \pm\sqrt{(pq)}$, and hence one negative eigenvalue $\lambda = pq$, which is negative in B_2. The root $\phi = 0$ is neglected

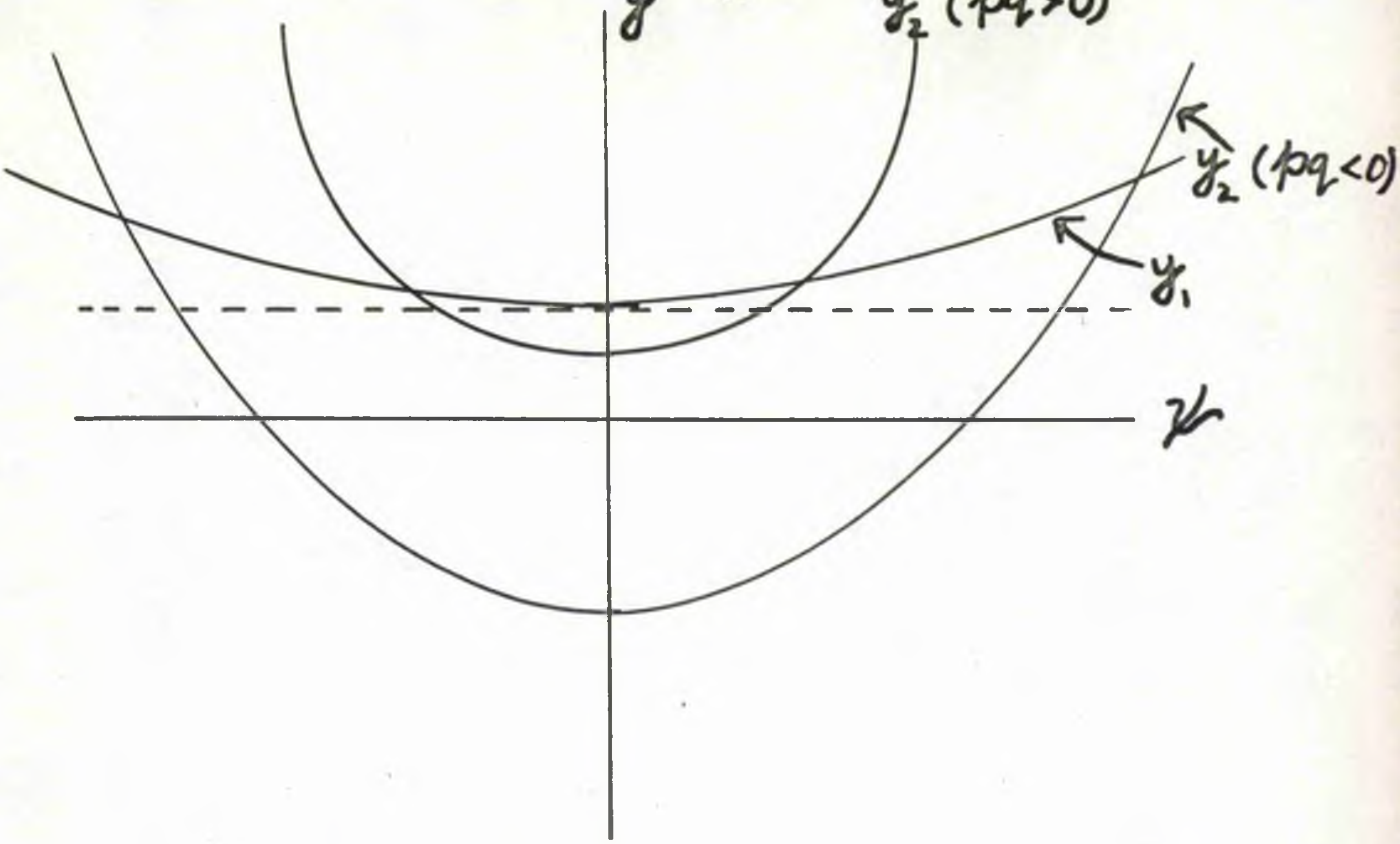

FIGURE 4

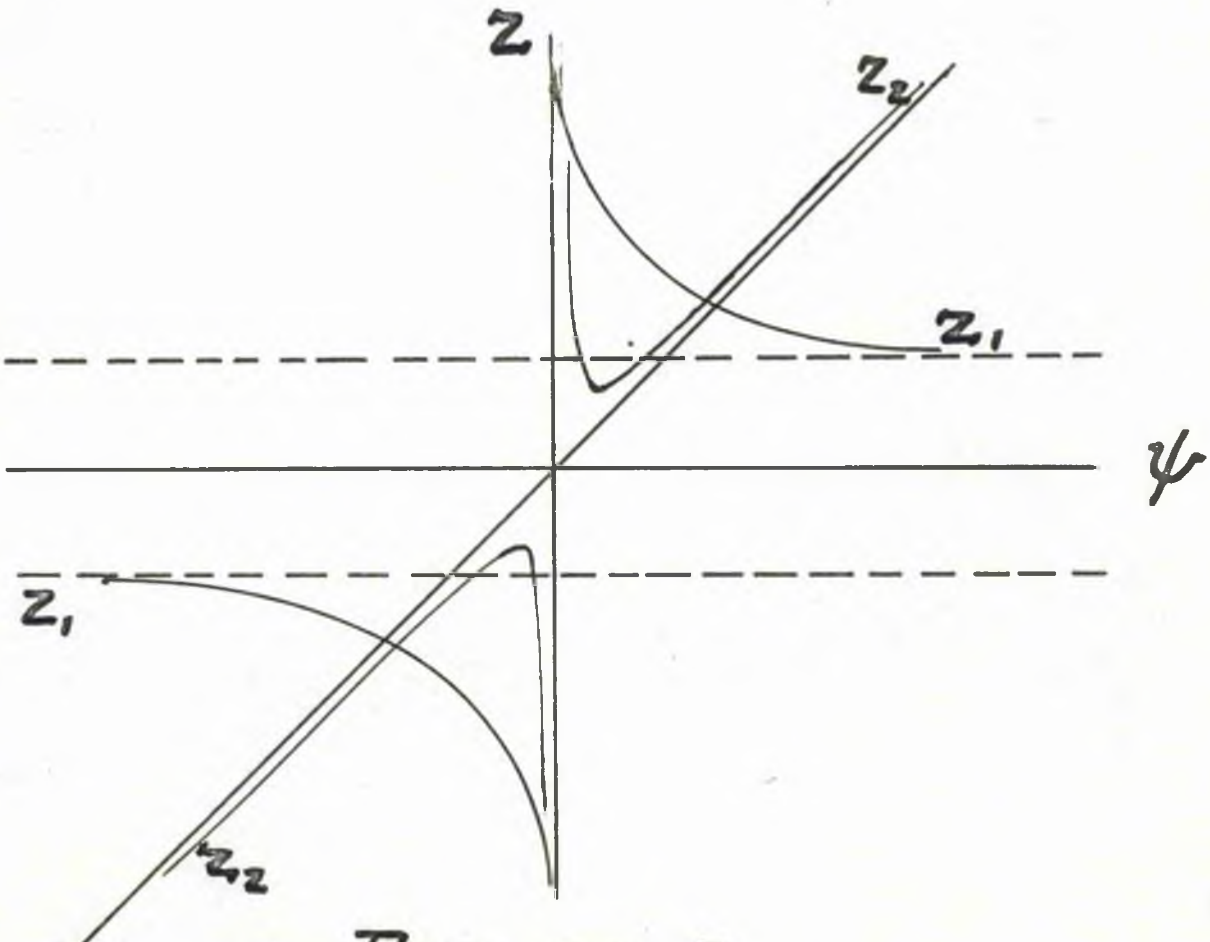

FIGURE 5

since $\lambda = 0$ is an eigenvalue only when there is an associated eigenfunction X satisfying the equation

$$X'' = 0$$

and the boundary conditions (4). This is possible only when

$$X(x) = 1 + px, \text{ and } pq + p + q = 0,$$

which is never the case in B_2.

B : In this region, y_1 is concave downwards, with a minimum at $\phi = 0$, where $y_1 = -(p+q)$. The forms of y_2 for $pq > 0$ and $pq < 0$ are shown in figure 4, where it is seen that y_2 has a minimum value $+pq$ where $\phi = 0$. Since $pq < -(p+q)$ in B, the curves y_1 and y_2 intersect. As $\phi \to \pm\infty$, y_1 behaves as $-(p+q)$, and y_2 as ϕ^2. Thus $y_1 < y_2$ for large ϕ, and from a consideration of the derivatives of y_1 and y_2, it can be shown that y_1 intersects y_2 only in the two points $\phi = \pm a$. Thus, when (p,q) is in B, there is one negative eigenvalue $\lambda = -a^2$.

C : When (p,q) is a point in C, it is better to consider the functions

$$z_1 = \coth\phi,$$

$$z_2 = \frac{-pq}{(p+q)\phi} - \frac{\phi}{p+q}.$$

There are real roots ϕ of equation (11) if and only if the graphs of z_1 and z_2 intersect. Clearly, $|z_1| > 1$ for all ϕ. The function z_2 is a hyperbola with asymptotes

$$z = \frac{-\phi}{p+q}, \quad \text{and} \quad \phi = 0.$$

Since in C, $p+q < 0$, then $z_2 \to +\infty$, as $\psi \to +\infty$ and since $pq > 0$, then $z_2 \to +\infty$, as $\phi \to 0^+$. Thus, the form of z_2 is as shown in figure 5. The minimum of z_2 is attained at $\phi = +\sqrt{(pq)}$, and here

$$z_2 = \frac{-2\sqrt{(pq)}}{p+q}.$$

Similarly, the maximum of z_2 on the lower branch is $2\sqrt{(pq)}/(p+q)$. In C, $pq + p + q > 0$, and $p+q < 0$, so that p and q are both negative. Thus, the minimum of z_2 satisfies

$$\min(z_2) = \frac{\sqrt{[(-p)(-q)]}}{-(p+q)/2} \leq \frac{-(p+q)/2}{-(p+q)/2} = 1,$$

by the arithmetic-geometric mean inequality. Thus, z_2 falls below z_1 for $\phi > 0$, and rises above z_1 for $\phi < 0$. This results in one negative eigenvalue.

As $\phi \to 0^+$, it can be shown that z_2 rises above z_1 once more. For:

$$z_2 - z_1 = \frac{(pq+p+q)\phi + O(\phi^3)}{-(p+q)\phi \sinh\phi}$$

$$= \frac{(pq+p+q)}{-(p+q)\sinh\phi} + O(\phi)$$

$$> 0, \text{ if } pq+p+q > 0, \text{ and } p+q < 0,$$

for ϕ sufficiently small. Thus, if $pq + p + q > 0$ and $p + q < 0$, there is a second negative eigenvalue.

To sum up, therefore,

(i) $pq + p + q > 0,\ p + q > 0$ :: no negative eigenvalue;

(ii) $pq + p + q = 0,\ p + q \geq 0$ :: one zero eigenvalue;

(iii) $pq + p + q < 0,\ p + q \gtrless 0$ :: one negative eigenvalue;

(iv) $pq + p + q = 0,\ p + q < 0$ :: one negative, and one zero eigenvalue;

(v) $pq + p + q > 0,\ p + q < 0$ :: two negative eigenvalues.

The conditions on p and q for the solutions of (1), subject to (1a), (1b), to be bounded, are thus the same as were required for the matrix U of chapter 1, to be positive semi-definite; and the conditions on p and q for the solutions of (1) to decay exponentially, are the conditions for U to be positive definite. In the following section an explanation is given, as to why the two sets of conditions are identical.

2.6 Asymptotic Instability.

If finite differences are used to approximate the space derivative in (1), together with the boundary conditions (1b), the totality of equations at a given time t, can be written as a first-order system viz. :

$$\frac{d\underset{\sim}{u}}{dt} = -U\,\underset{\sim}{u} + \underset{\sim}{l}^n \tag{17}$$

where $\underset{\sim}{u}$ is the vector of values $u(x,t)$ at the nodes $x_i = \frac{i}{N}$, $(i = 0,1,\ldots,N;\ Nh = 1)$, $\underset{\sim}{l}^n$ is a vector involving the boundary conditions at $t = nk$, and U is a matrix of order N+1. It is well known (e.g.[3]) that the system (17) of first order ordinary differential equations, is asymptotically stable (in the Liapunov sense) if and only if $\underset{\sim}{l}^n$ is bounded as $n \to \infty$, and the matrix U is positive semi-definite; and that the zero vector, $\underset{\sim}{0}$, is a stable solution (i.e. all solutions tend to zero as t tends to ∞) if and only if $\underset{\sim}{l}^n \to \underset{\sim}{0}$, as $n \to \infty$, and U is positive definite. Thus, whatever difference replacement is used for $\frac{\partial^2}{\partial x^2}$, the resulting matrix U will be positive semi-definite only when the solutions of the differential equation are bounded. If some implicit formula is used now, to replace the time derivative in (17), it is possible to have a difference method which is stable for all problems for which the differential equation has only bounded solutions; and, clearly, such a method must be unstable if the differential equation has an unbounded solution.

The origin of the persistent errors discussed in chapter 1, now becomes clear. If $\underset{\sim}{\epsilon}$ is the difference between the theoretical and numerical solutions of (17), then it satisfies the equation

$$\frac{d\underset{\sim}{\epsilon}}{dt} = -U\,\underset{\sim}{\epsilon} \tag{18}$$

Then, as $t \to \infty$, $\underset{\sim}{0}$ is an asymptotically stable solution of (18) if, and only if, U is positive definite. If U is singular and positive semi-definite, then $\underset{\sim}{\epsilon}$ stays bounded as $t \to \infty$, but does not tend to zero. It tends, instead, to a solution of the set of equations

$$U\underset{\sim}{\epsilon} = \underset{\sim}{0} \qquad (19)$$

The existence of a non-zero solution of (19) follows from the singularity of U, and corresponds to saying that there exists a non-zero solution of

$$\frac{d^2u}{d\,x^2} = 0,$$

subject to

$$u'(0)-pu(0) = 0$$

$$u'(1)+qu(1) = 0.$$

This is possible only when there are constants A, B, such that $Ax + B$ satisfies the boundary conditions. That is, if, and only if,

$$pq + p + q = 0,$$

in which case

$$u(x) = A(px + 1),$$

where A is arbitrary. If the initial error committed, in solving (17) is $O(1/N^2)$, (as is the case for the difference methods considered in chapter 2) then the "steady-state" value of $\underset{\sim}{\epsilon}$ will be $O(1/N^2)(px_j + 1)$, where $x_j = \frac{j}{N}$, $j = 0, 1, \ldots, N$.

Table 1 shows an example of such an error. The problem considered in chapter 1, was solved with $p = -0{\cdot}5$, and $q = 1{\cdot}0$, (i.e. $pq + p + q = 0$), and, in the table, the errors are shown at $t = 1$, by which time they had settled down to a steady value. The values of the error $\epsilon(j)$ at x_j are given for $j = 0, 1, \ldots, N$, and the same values are given, divided by $\epsilon(0)$, which is $O(1/N^2)$. The second vector, $\epsilon/\epsilon(0)$, is a close approximation to $-0 \cdot 5x + 1$, given in the last column.

j	error $\epsilon(j)$	$\frac{\text{error}}{\epsilon(0)}$	$-0 \cdot 5x + 1.$
0	0·001 730 46	1·000 000	1·000
1	0·001 687 16	0·974 978	0·975
2	0·001 643 86	0·949 956	0·950
3	0·001 600 55	0·924 927	0·925
4	0·001 557 25	0·899 905	0·900
5	0·001 513 95	0·874 883	0·875
6	0·001 470 65	0·849 861	0·850
7	0·001 427 35	0·824 838	0·825
8	0·001 384 06	0·799 822	0·800
9	0·001 340 77	0·774 806	0·775
10	0·001 297 49	0·749 795	0·750
11	0·001 254 22	0·724 790	0·725
12	0·001 210 95	0·699 785	0·700
13	0·001 167 69	0·674 786	0·675
14	0·001 124 44	0·649 792	0·650
15	0·001 080 20	0·624 805	0·625
16	0·001 037 96	0·599 817	0·600
17	0·000 994 73	0·574 836	0·575
18	0·000 951 50	0·549 854	0·550
19	0·000 908 28	0·524 875	0·525
20	0·000 865 06	0·499 902	0·500

TABLE 1.

2.7 Uniqueness Of The Solution.

The uniqueness of the solution of equation (1), subject to (1a) and (1b), has been established by other authors, when p and q are both non-negative. Friedman [15], appears to establish the uniqueness for all p and q, by an argument involving certain à priori bounds for the solutions. Since the solutions may be unbounded for certain values of p and q, his proof seems to be slightly unsatisfactory. By a slight modification of the usual method of proof, however, uniqueness may be proved when the conditions $pq+p+q \geqslant 0$, and $p+q \geqslant 0$, are satisfied.

The solution of equation (1), subject to (1a) and (1b), is unique if the equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \qquad \text{in } R, \tag{1)'$$

subject to

$$u(x,0) = 0, \qquad 0 \leqslant x \leqslant 1, \tag{1a)'$$

and

$$\left.\begin{aligned} \frac{\partial u}{\partial x} - pu &= 0, \quad x = 0, \quad t \geqslant 0, \\ \frac{\partial u}{\partial x} + qu &= 0, \quad x = 1, \quad t \geqslant 0, \end{aligned}\right\} \tag{1b)'$$

has no solution, other than the trivial solution $u(x,t) = 0$.

Let $\bar{R}$ be the region

$$[0 \leqslant x \leqslant 1] \times [0 \leqslant t \leqslant T],$$

and let $\partial\bar{R}$ be the boundary of $\bar{R}$. Then let

$$u(x,t) = a(x)\, v(x,t),$$

where $a(x)$ is the eigenfunction corresponding to the lowest eigenvalue, λ_0, of the problem

$$y''(x) = -\lambda y(x), \qquad 0 \leqslant x \leqslant 1, \tag{*}$$

where

$$y'(0) - py(0) = 0,$$
$$y'(1) + qy(1) = 0. \qquad (*)$$

Then $a(x) > 0$, in $(0,1)$, since the lowest eigenfunction has no zeros in the fundamental interval; and $\lambda_0 \geq 0$, provided

$$pq+p+q \geq 0, \text{ and } p+q \geq 0.$$

Then $v(x,t)$ satisfies the equation

$$\frac{\partial(av)}{\partial t} = \frac{\partial^2(av)}{\partial x^2}, \quad \text{in } \overline{R}, \qquad (20)$$

subject to

$$v(x,0) = 0, \quad 0 \leq x \leq 1,$$

and

$$\frac{\partial v}{\partial x} = 0, \quad x = 0, \quad t > 0,$$

$$\frac{\partial v}{\partial x} = 0, \quad x = 1, \quad t > 0.$$

From the divergence theorem the following identity may be obtained :

$$\int_{\overline{R}} \left[\frac{\partial}{\partial x}\left(a^2\frac{\partial v^2}{\partial x}\right) - \frac{\partial}{\partial t}(a^2v^2) \right] dxdt + \int_{\partial\overline{R}} \left[a^2v^2dx + a^2\frac{\partial v^2}{\partial x}dt \right] = 0.$$

Using equation (20) for v, and the boundary conditions, this becomes :

$$\int_{\overline{R}} [2a^2v_x^2 - 2aa''v^2]dxdt + \int_0^1 [a^2v^2]_{t=T}\, dx = 0.$$

Since $aa'' = -\lambda_0 a^2$, this is

$$\int_{\overline{R}} [2a^2v_x^2 + 2\lambda_0 a^2v^2]dxdt + \int_0^1 [a^2v^2]_{t=T}\, dt = 0.$$

Then all terms in this equation are positive, and their sum is zero, so that each term must be zero; i.e.

$$v(x,t) = 0, \text{ in } \overline{R}$$

Hence, $u(x,t) = 0$, in $\overline{R}$, and the uniqueness theorem is proved under the given conditions on p and q.

An alternative proof has been constructed by E.T. Copson [9]. The identity used is

$$\int_{\overline{R}} [v(w_{xx} - w_t) - w(v_{xx} + v_t)]dxdt \;+\; \int_{\partial\overline{R}} [vw_x - wv_x]dt + vwdx \;=\; 0.$$

Let $w = u^r$, where $r = 2m/n$, m and n being integers with no common factors, and where $u_{xx} = u_t$, and u satisfies the initial and boundary conditions (1a)' and (1b)'. Then, putting $v = \sin(kx + \epsilon)$, it follows that

$$\int_{\overline{R}} [r(r-1)u^{r-2}u_x^2 + k^2u^r]\sin(kx+\epsilon)dxdt \;+\; \int_0^1 u(x,T)^r\sin(kx+\epsilon)dx = 0, \quad (21)$$

provided

$$k\cos\epsilon - pr\sin\epsilon = 0,$$

$$k\cos(k+\epsilon) + qr\sin(k+\epsilon) = 0.$$

That is, k^2 is an eigenvalue, and $\sin(kx + \epsilon)$ is the associated eigenfunction for the problem (✱),above. If k^2 is the lowest eigenvalue, then $\sin(kx + \epsilon)$ has no zeros in $(0,1)$, and $k^2 \geq 0$, if, and only if,

$$p + q + rpq \geq 0, \text{ and } p + q \geq 0, \quad \text{(figure 6)}.$$

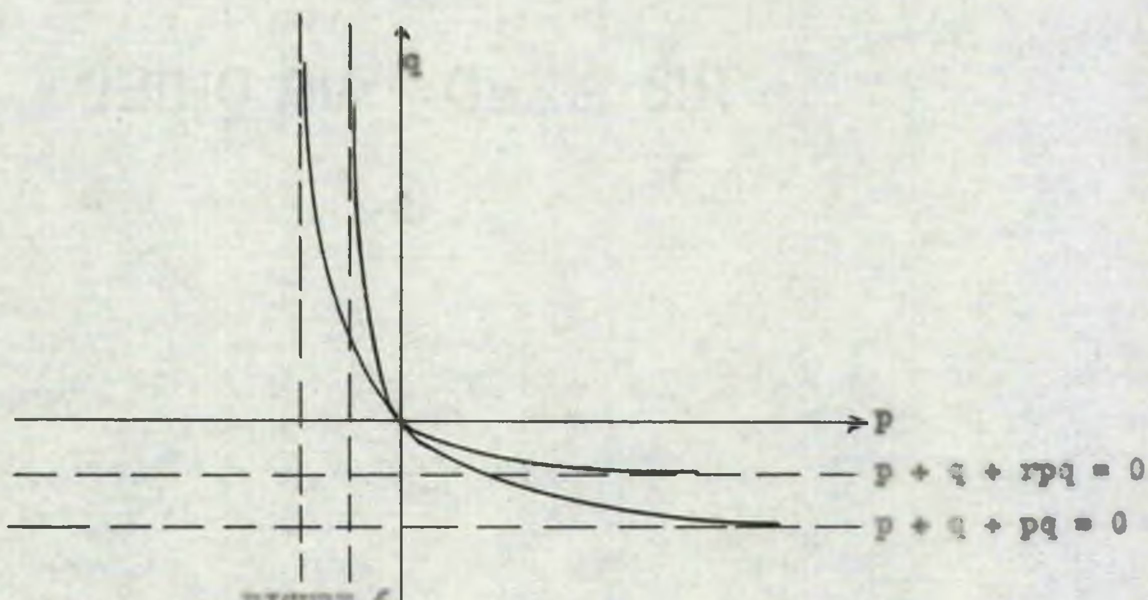

FIGURE 6.

If $r > 1$, this region (as shown in figure 6) is contained within the region

$$p + q + pq \geq 0, \; p + q \geq 0.$$

For the choice of r made above, each term in equation (21) is positive, and, as before,

$$u(x, t) = 0, \text{ in } \bar{R}.$$

This argument holds for all $r > 1$, and, in the limit, as $r \to 1$. Thus, uniqueness has been proved when $pq + p + q \geq 0$, and $p + q \geq 0$.

2.8 The Equation Of Heat In Higher Dimensions.

Let the operator ∇ in s space variables, be defined by :

$$\nabla = \left(\frac{\partial}{\partial x_1}, \ldots\ldots\ldots\ldots, \frac{\partial}{\partial x_s}\right),$$

and let $\underset{\sim}{x} = (x_1, x_2, x_3, \ldots\ldots, x_s)$. Let R be a closed region in s-space, such that $0 \leq x_i \leq 1$, $i = 1,2,\ldots,s$, and $\bar{R}$ is the region $R \times [t \geq 0]$. Then, consider the equation

$$\frac{\partial u}{\partial t} = \nabla^2 u, \quad \text{in } \bar{R}, \tag{22}$$

subject to

$$u(\underset{\sim}{x},0) = f(\underset{\sim}{x}), \quad \underset{\sim}{x} \in R, \tag{22a}$$

and

$$\begin{aligned} \frac{\partial u}{\partial x_i} - p_i u &= \phi_0^i(t), \quad x_i = 0, \quad t \geq 0, \\ \frac{\partial u}{\partial x_i} + q_i u &= \phi_1^i(t), \quad x_i = 1, \quad t \geq 0, \end{aligned} \tag{22b}$$

for $i = 1,2,\ldots,s$, where p_i and q_i are constants.

The solution of the initial-boundary value problem, when ϕ_0^i and ϕ_1^i are zero, will be considered. The case when ϕ_0^i and ϕ_1^i are non-zero is solved by means of Laplace Transforms, and produces results similar to the ones below. Let

$$u(\underset{\sim}{x},t) = T(t) \prod_{i=1}^{s} X_i(x_i)$$

Then the functions X_i $(i = 1,2,\ldots,s)$, satisfy the Sturm-Liouville equations

$$X_i''(x_i) = \lambda^{(i)} X_i(x_i)$$

with

$$X_i'(0) - p_i X_i(0) = 0,$$
$$X_i'(1) + q_i X_i(1) = 0;$$

and T satisfies

$$T'(t) = -(\lambda^{(1)} + \ldots\ldots + \lambda^{(s)})T(t).$$

The eigenvalues $\lambda^{(i)}$ are the roots of the equations :

$$\sqrt{\lambda^{(i)}}(p_i+q_i)\cos\sqrt{\lambda^{(i)}} = (\lambda^{(i)} - p_i q_i)\sin\sqrt{\lambda^{(i)}} \tag{23}$$

for $i = 1,2,\ldots,s$. Thus, as in section 3, the solutions of equation (22) are bounded if, and only if, for all roots of equation (23), we have

$$\lambda^{(1)} + \lambda^{(2)} + \ldots\ldots + \lambda^{(s)} \geqslant 0. \tag{24}$$

An obvious set of sufficient conditions for (24) to hold is

$$p_i q_i + p_i + q_i \geqslant 0, \text{ and } p_i + q_i \geqslant 0, \; i = 1,2,\ldots,s.$$

These conditions are not all necessary, however, since a negative $\lambda^{(i)}$ may be counterbalanced by other positive eigenvalues, so that (24) is still satisfied. In a later chapter, the numerical solutions of equation (22), subject to (22a) and (22b) will be shown to be stable for exactly the same conditions on p_i and q_i, $i = 1,2,\ldots,s$, as are given above.

A uniqueness proof for the solution of (22) may be constructed on the same lines as the proof in section 7.

It is commonly assumed that solutions of the equation (22), subject to (22a),(22b), with time independent boundary conditions, tend asymptotically, $(t \to \infty)$, to solutions of Laplace's equation

$$\nabla^2 u = 0, \quad \text{in } R,$$

subject to the boundary conditions (22b), with ϕ_0^i and ϕ_1^i functions of $\underline{x}$ only, where $\underline{x} \in \partial R$. The above remarks show, however, that this is true only when the solutions u stay bounded; i.e. when, and only when, the conditions

(24) are satisfied. The solutions of Laplace's equation with boundary coefficients which do not satisfy (24) are not, therefore, asymptotic solutions of the heat conduction equation. A particularly interesting case of Laplace's equation, arises when the problem is singular (in a sense to be defined). This is discussed in the following section.

2.9 The Singular Laplacian Operator.

We now consider the equation :

$$\nabla^2 u \quad = \quad F(\underset{\sim}{x}) + G(\underset{\sim}{x})u \tag{26}$$

in a closed region D, subject to the boundary conditions

$$\frac{\partial u}{\partial \nu} - H(\underset{\sim}{x})u = \phi(\underset{\sim}{x}) \tag{26a}$$

on ∂D, the boundary of D, where ν is the inward normal to D, and $H(\underset{\sim}{x})$ is a function of position on ∂D. The problem (26),(26a), is said to be singular if there is a solution $v(\underset{\sim}{x})$ in D, other than $v = 0$, of the equation

$$\begin{aligned} \nabla^2 v \quad &= \quad Gv, \quad \text{in D, with} \\ \frac{\partial v}{\partial \nu} - Hv \quad &= \quad 0, \text{ on } \partial D. \end{aligned} \tag{27}$$

When D is the region R, described previously, and H is defined as in (22b), then an example of a singular problem occurs when

$$p_i q_i + p_i + q_i \quad = \quad 0, \quad i = 1,2,\ldots,s.$$

The following results hold for a general, closed region D.

If the Laplacian operator, together with the boundary conditions, is regarded as a linear differential operator, L, it is necessary that the functions F and ϕ be in the range space of L. If L is non-singular, it maps the whole function space into itself, and F and ϕ are then in the range space of L. But, if L is singular, and $v(\underset{\sim}{x})$ is a solution of (27), (i.e. is in the null space of L), then F and ϕ must satisfy the condition :

$$\int_D v(\underset{\sim}{x})\, F(\underset{\sim}{x})\, dS \quad + \quad \int_{\partial D} v(\underset{\sim}{x})\, \phi(\underset{\sim}{x})\, ds \quad = \quad 0, \tag{28}$$

where dS is an element of D, and ds is an element of ∂D.

To prove this, consider the identity :

$$\int_D [\nabla w_1 . \nabla w_2 + w_2 \nabla^2 w_1] \, dS + \int_{\partial D} w_2 \frac{\partial w_1}{\partial \nu} \, ds = 0 \qquad (29)$$

which is obtained from the divergence theorem, (e.g. [18], p. 229).

In (29) put $w_1 = u$, a solution of (26), subject to (26a), and let $w_2 = v$, a solution of (27). [Equation (27) always has a solution, perhaps identically zero in D]. Then (29) becomes

$$\int_D [\nabla u . \nabla v + vF + vuG] \, dS + \int_{\partial D} [vHu + v\phi] \, ds = 0.$$

But, since $\frac{\partial v}{\partial \nu} = Hv$, on ∂D, this becomes

$$\int_D [\nabla u . \nabla v + vF + vuG] \, dS + \int_{\partial D} [v\phi + u\frac{\partial v}{\partial \nu}] \, ds = 0. \qquad (30)$$

Now, putting $w_1 = v$, and $w_2 = u$, in (29), we obtain

$$\int_D [\nabla u . \nabla v + u\nabla^2 v] \, dS + \int_{\partial D} u \frac{\partial v}{\partial \nu} \, ds = 0. \qquad (31)$$

Subtracting equation (31) from equation (30), and using the fact that $\nabla^2 v = Gv$, the result

$$\int_D vF(\underline{x}) \, dS + \int_{\partial D} v\phi(\underline{x}) \, ds = 0,$$

follows, which was to be proved.

The identity (29) may be used to prove uniqueness of the solution of equation (26), under restrictions on H. The solution of (26), subject to (26a), is unique, if equation (27) has no solution, except $v(\underline{x}) = 0$. Putting $w_1 = w_2 = v$, in equation (26), leads to

$$\int_D (\nabla v)^2 \, dS + \int_{\partial D} v \frac{\partial v}{\partial \nu} \, ds = 0;$$

i.e.

$$\int_D (\nabla v)^2\, dS \quad + \quad \int_\partial H v^2\, ds \quad = \quad 0. \tag{32}$$

Thus, if $H > 0$, then both terms in equation (32) are non-negative, and hence $v(\underset{\sim}{x}) = 0$. If $H = 0$, then, whenever $G = 0$, we have $\nabla v = 0$, and the solution of (26) is unique only up to a constant.

We now assume that $G = 0$, and that the region D is the region R, described previously, and the boundary conditions are of the form (22b), where ϕ_0^i and ϕ_1^i are time independent. Uniqueness may be proved under more general conditions on p_i and q_i, $(i = 1,2,\ldots,s)$. Let $a(\underset{\sim}{x})$ be the eigenfunction corresponding to the lowest eigenvalue, λ_0, of the equation

$$\nabla^2 y \quad = \quad -\lambda y, \qquad \text{in } R,$$

subject to

$$\frac{\partial y}{\partial \nu} - Hy \quad = \quad 0,$$

on ∂R, where H is defined as in (22b). Let

$$v(\underset{\sim}{x}) \quad = \quad a(\underset{\sim}{x})\, w(\underset{\sim}{x})$$

where $v(\underset{\sim}{x})$ is a solution of equation (27), if one exists, which is not identically zero. The following identity may be established from the divergence theorem :

$$\int_D \nabla.[a^2 w \nabla w\,]\, dS \quad + \quad \int_{\partial R} a^2 w \frac{\partial w}{\partial \nu}\, ds \quad = \quad 0 \tag{33}$$

Now, $w(\underset{\sim}{x})$ satisfies the equation

$$\nabla^2(aw) \quad = \quad 0, \qquad \text{in } R,$$

subject to

$$\frac{\partial w}{\partial \nu} \quad = \quad 0, \qquad \text{on } \partial R.$$

In addition

$$\begin{aligned}\nabla\cdot[\,a^2 w\nabla w\,] &= a^2(\nabla w)^2 + w\nabla\cdot[a^2\nabla w] \\ &= a^2(\nabla w)^2 + aw[\nabla^2(aw) - w\nabla^2 a] \\ &= a^2(\nabla w)^2 + a^2 w^2\lambda_o\end{aligned}$$

Thus (33) becomes

$$\int_R [a^2(\nabla w)^2 + a^2 w^2\lambda_o]\, dS = 0. \tag{34}$$

If $\lambda_o > 0$, (since $a(\underset{\sim}{x})$ has no zeros in R), then

$$a^2(\nabla w)^2 + a^2 w^2\lambda_o = 0;$$

i.e. $w = 0$ in R. This proves uniqueness of the solution.

If $\lambda_o = 0$, then (34) implies only that $\nabla w = \underset{\sim}{0}$; i.e. w is a constant. This is equivalent to saying that the solution of (26) is arbitrary to within a multiple of a solution of equation (27), if the operator L in (26),(26a) is singular. Thus, when the condition (24) is satisfied, with strict inequality, the solution has been proven unique.

2.10 Conclusion.

The instability observed in chapter 1, has, therefore, been shown to be caused by the existence of exponentially increasing solutions of the heat equation. It has also been shown that the class of problems for which the difference schemes in chapter 1 were unconditionally stable, is the same class as that for which the solutions of the differential equation were bounded. Unbounded solutions have been shown to arise in higher dimensions, and it remains to discuss the behaviour of difference methods of solution of the heat equation, and of Laplace's equation in two, or more, space variables.

CHAPTER 3.

THE THIRD BOUNDARY VALUE PROBLEM FOR PARABOLIC AND ELLIPTIC EQUATIONS.

3.1 Introduction.

The term "third boundary value problem" is applied to equations of the form considered in chapter 2, subject to boundary conditions involving linear combinations of the unknown function, and its first space derivatives. It has been shown, in chapter 2, that such a problem may have unbounded or bounded solutions, depending on the form of the boundary conditions. In this chapter, the effect of the boundary conditions on the stability of a numerical method of solution of the heat equation in two space variables, is examined. The results obtained are then applied to the numerical solution of Laplace's equation.

3.2 The Finite Difference Scheme.

The equation considered is

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \tag{1}$$

in the region

$$\bar{R} = [0 \leq x, y \leq 1] \times [t \geq 0]$$

subject to the initial condition

$$u(x,y,0) = f(x,y), \quad (x,y) \in \bar{R} \tag{1a}$$

and the boundary conditions

$$\begin{aligned}
\frac{\partial u}{\partial x} - p_1 u &= \phi_0(t), \quad x = 0,\ t \geq 0, \\
&\qquad\qquad\qquad\qquad 0 \leq y \leq 1, \\
\frac{\partial u}{\partial x} + q_1 u &= \phi_1(t), \quad x = 1,\ t \geq 0, \\
\frac{\partial u}{\partial y} - p_2 u &= \psi_0(t), \quad y = 0,\ t \geq 0, \\
&\qquad\qquad\qquad\qquad 0 \leq x \leq 1, \\
\frac{\partial u}{\partial y} + q_2 u &= \psi_1(t), \quad y = 1,\ t \geq 0,
\end{aligned} \tag{1b}$$

where p_i and q_i, $(i = 1,2)$, are constants. The same continuity conditions are assumed at the corners of the region, as were stated in chapter 2. Several authors, (e.g. 8,16,21,34), have considered the third boundary value problem, for equation (1), for the particular case when p_i, q_i, $i = 1,2$, are non-negative constants, or, e.g. [47], non-negative functions of position on the boundary of $\bar{R}$.

The region $\bar{R}$ is covered by a rectilinear net, with mesh points (ih, jh, nk), where : $n \geq 0$; $i,j = 0,1,2,\ldots,N$; $Nh = 1$; and (h,k) are the mesh spacings, in the space and time directions, respectively. The difference method used to approximate (1), is the

Peaceman-Rachford method [39], which takes the form :

$$
\begin{aligned}
(1 - \tfrac{1}{2}r\delta_x^2)\, v_{i,j}^{n+\frac{1}{2}} &= (1 + \tfrac{1}{2}r\delta_y^2)\, v_{i,j}^{n} \\
(1 - \tfrac{1}{2}r\delta_y^2)\, v_{i,j}^{n+1} &= (1 + \tfrac{1}{2}r\delta_x^2)\, v_{i,j}^{n+\frac{1}{2}}
\end{aligned}
\tag{2}
$$

where δ_x, and δ_y, are the usual central difference operators in the x and y directions, respectively; and $v_{i,j}^m$, ($m = n, n+1/2, n+1$), is the numerical approximation to $u(ih, jh, mk)$.

Equation (2) is applied for $i, j = 1, 2, \ldots, N-1$. When i, or j, is 0 or N, the values of v outside $\bar{R}$, which occur in (2), are eliminated, using the appropriate finite difference approximations to the boundary conditions (1b). The derivatives on the boundary are approximated by means of the $O(h^2)$ approximation used in chapter 1. For example

$$
\frac{\partial u}{\partial x}(ih, jh, nk) = \frac{v_{i+1,j}^n - v_{i-1,j}^n}{2h}
$$

when $i = 0, N$; for $j = 0, 1, \ldots, N$; and for all n. The other derivatives are approximated in a similar way. When $n = 0$, in equation (2), $v_{i,j}^0$ is obtained from the initial condition (1a), for $i, j = 0, 1, \ldots, N$.

The values of $v_{i,j}^m$ on each plane $t = mk$, are ordered, to give a vector with $(N+1)^2$ components :

$$
\underset{\sim}{v}^m = (v_{0,0}^m, v_{0,1}^m, \ldots, v_{0,N}^m, v_{1,0}^m, \ldots, v_{1,N}^m, \ldots, v_{N,0}^m, \ldots, v_{N,N}^m)^T
$$

Then the set of equations (2) may be written in the form :

$$
\begin{aligned}
(I + \tfrac{1}{2}rU_1)\, \underset{\sim}{v}^{n+\frac{1}{2}} &= (I - \tfrac{1}{2}rU_2)\, \underset{\sim}{v}^n + \underset{\sim}{l}_1^n \\
(I + \tfrac{1}{2}rU_2)\, \underset{\sim}{v}^{n+1} &= (I - \tfrac{1}{2}rU_1)\, \underset{\sim}{v}^{n+\frac{1}{2}} + \underset{\sim}{l}_2^n ,
\end{aligned}
\tag{3}
$$

where $\underset{\sim}{l}_1^n$ and $\underset{\sim}{l}_2^n$ are vectors involving the boundary functions ϕ_1 and ψ_1

$(i = 1,2)$; I is the unit matrix of order $(N+1)^2$; and the matrices U_1, and U_2, are given by

$$U_1 = V_1 \times I_{N+1}, \text{ and } U_2 = I_{N+1} \otimes V_2 \,. \tag{4}$$

The matrix I_{N+1} in (4), is the unit matrix of order $(N+1)$, and the symbol $\otimes$ represents the tensor product, as defined in Halmos [22]. The matrices V_i, $(i = 1,2)$, are also of order $(N+1)$, and are given by

$$V_i = \begin{bmatrix} 2(1+p_i/N) & -2 & 0 & \cdots\cdots & & \\ -1 & 2 & -1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & -1 & 2 & -1 \\ & & & & -2 & 2(1+q_i/N) \end{bmatrix} \tag{5}$$

3.3 The Matrices Of The Peaceman-Rachford Method.

An analysis of the properties of the matrices U_1 and U_2, defined by equations (4), and (5), will simplify the stability analysis of the following section.

The matrices U_1 and U_2 commute, since

$$U_1 U_2 = (V_1 \otimes I_{N+1})(I_{N+1} \otimes V_2)$$

$$= V_1 \otimes V_2,$$

and, similarly,

$$U_2 U_1 = V_1 \otimes V_2,$$

where multiplication of tensor products is defined as in Halmos [22, p.96].

The matrices V_1 and V_2 are of the same form as the matrix U, in chapter 1, and are, therefore, similar to symmetric matrices, $\tilde{V}_1$ and $\tilde{V}_2$, where

$$\tilde{V}_i = D_{N+1}^{-1} V_i D_{N+1}, \; i = 1,2,$$

the matrix D_{N+1} being of order N+1, and given by

$$D_{N+1} = \begin{bmatrix} \sqrt{2} & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & 1 & \\ & & & & & & \sqrt{2} \end{bmatrix}$$

Thus, V_i, $i = 1,2$, has a complete set of eigenvectors, denoted by x_j^i, $(i = 1,2;\; j = 0,1,\ldots,N)$, with associated real eigenvalues λ_j, μ_j. The eigenvectors of both U_1 and U_2, are

$$x_i^1 \otimes x_j^2, \quad i,j = 0,1,\ldots,N,$$

since

$$U_1\,(x_i^1 \otimes x_j^2) = V_1\, x_i^1 \otimes I_{N+1}\, x_j^2$$
$$= \lambda_j\, x_i^1 \otimes x_j^2$$

and

$$U_2\,(x_i^1 \otimes x_j^2) = I_{N+1}\, x_i^1 \otimes V_2\, x_j^2$$
$$= \mu_j\, x_i^1 \otimes x_j^2 .$$

The corresponding eigenvalues of U_1 and U_2 are, respectively, λ_1 and μ_j, each being repeated N+1 times.

If the matrix D, of order $(N+1)^2$ is defined by

$$D = D_{N+1} \otimes D_{N+1} ,$$

then

$$D^{-1} U_1 D = (D_{N+1}^{-1} \otimes D_{N+1}^{-1})(V_1 \otimes I_{N+1})(D_{N+1} \otimes D_{N+1})$$
$$= (D_{N+1}^{-1} V_1 D_{N+1}) \otimes I_{N+1}$$
$$= \tilde{V}_1 \otimes I_{N+1}$$
$$= \tilde{U}_1 ,$$

where $\tilde{U}_1$ is symmetric. Similarly

$$D^{-1} U_2 D = I_{N+1} \otimes (D_{N+1}^{-1} V_2 D_{N+1})$$
$$= I_{N+1} \otimes \tilde{V}_2$$
$$= \tilde{U}_2 ,$$

where $\tilde{U}_2$ is symmetric. Thus, U_1 and U_2 are similar to symmetric matrices, commute, and have a complete set of eigenvectors, and real eigenvalues.

3.4 The Stability Of The Difference Method.

Denoting the numerical error of the difference method (3) by $\underset{\sim}{\varepsilon}^m$, $m = n, n+\frac{1}{2}, n+1$, where

$$\underset{\sim}{\varepsilon}^m = \underset{\sim}{v}^m - \underset{\sim}{\bar{v}}^m,$$

$\underset{\sim}{v}^m$ and $\underset{\sim}{\bar{v}}^m$ being the theoretical and numerical solutions, respectively, of equation (3), then the equation for the error is

$$(I + \tfrac{1}{2}rU_1)\,\underset{\sim}{\varepsilon}^{n+\frac{1}{2}} = (I - \tfrac{1}{2}rU_2)\,\underset{\sim}{\varepsilon}^n$$

$$(I + \tfrac{1}{2}rU_2)\,\underset{\sim}{\varepsilon}^{n+1} = (I - \tfrac{1}{2}rU_1)\,\underset{\sim}{\varepsilon}^{n+\frac{1}{2}}$$

Elimination of the intermediate step, calculating $\underset{\sim}{\varepsilon}^{n+\frac{1}{2}}$, gives the equation

$$\underset{\sim}{\varepsilon}^{n+1} = (I+\tfrac{1}{2}rU_2)^{-1}(I-\tfrac{1}{2}rU_1)(I+\tfrac{1}{2}rU_1)^{-1}(I-\tfrac{1}{2}rU_2)\,\underset{\sim}{\varepsilon}^n$$

This is written in the form

$$\underset{\sim}{\varepsilon}^{n+1} = G\,\underset{\sim}{\varepsilon}^n$$

or

$$\underset{\sim}{\varepsilon}^{n+1} = G^{n+1}\underset{\sim}{\varepsilon}^0$$

where

$$G = (I+\tfrac{1}{2}rU_2)^{-1}(I-\tfrac{1}{2}rU_1)(I+\tfrac{1}{2}rU_1)^{-1}(I-\tfrac{1}{2}rU_2).$$

The equations(3) are stable if, and only if, the norm of the matrix G^n is bounded for all $n > 0$. But, since U_1 and U_2 areboth similar to a symmetric matrix, then G is also similar to a symmetric matrix $\tilde{G}$, where

$$\tilde{G} = D^{-1}G\,D,$$

since

$$D^{-1}G\,D = [D^{-1}(I+\tfrac{1}{2}rU_2)D]^{-1}[D^{-1}(I-\tfrac{1}{2}rU_1)D][D^{-1}(I+\tfrac{1}{2}rU_1)D]^{-1}[D^{-1}(I-\tfrac{1}{2}rU_2)D]$$

$$= (I+\tfrac{1}{2}r\tilde{U}_2)^{-1}(I-\tfrac{1}{2}r\tilde{U}_1)(I+\tfrac{1}{2}r\tilde{U}_1)^{-1}(I-\tfrac{1}{2}r\tilde{U}_2)$$

which is symmetric, since $\tilde{U}_1$ and $\tilde{U}_2$ are symmetric, and commute. Thus, the norms of the matrices G^n, for all n, may be taken as the spectral radius, $\rho(G^n)$, of G^n. But $\rho(G^n) = \rho^n(G)$, and so the necessary and sufficient condition for stability is

$$\exists K > 0, \ \forall n > 0, \ \rho^n(G) \leq K.$$

Following the same line of argument as was used in chapter 1, this is replaced by

$$\rho(G) \leq 1.$$

If the eigenvalues of G are denoted by $\mu_{i,j}$ $(i,j = 0,1,2,\ldots\ldots,N)$, then

$$\mu_{i,j} = \frac{(1-\frac{1}{2}r\mu_j)(1-\frac{1}{2}r\lambda_i)}{(1+\frac{1}{2}r\mu_j)(1+\frac{1}{2}r\lambda_i)} \quad . \tag{6}$$

The eigenvalues of V_i, $i = 1,2$, are the roots of the characteristic polynomial

$$[(\theta-2)^2 - 4 + 4p_iq_i/N^2 \,]\, T_{N-1}(\theta) \; + \; 2(p_i+q_i)/N \,[T_N(\theta) - T_{N-2}(\theta)] = 0, \tag{7}$$

where the polynomials $T_m(\theta)$, $m = N,N-1,N-2$, were defined in chapter 1. It was shown in chapter 1, that, for a reasonable value of r, it may be assumed that

$$1 + \tfrac{1}{2}r\mu_j > 0, \text{ and } \ 1 + \tfrac{1}{2}r\lambda_i > 0, \ i,j = 0,1,\ldots,N.$$

Thus the condition

$$\max_{i,j} \mu_{i,j} \leq 1,$$

will be satisfied if, and only if, the following conditions are satisfied:

$$\lambda_i + \mu_j \geq 0, \text{ and}$$

$$\lambda_i\mu_j > -4/r^2$$

for $i,j = 0,1,\ldots\ldots,N$.

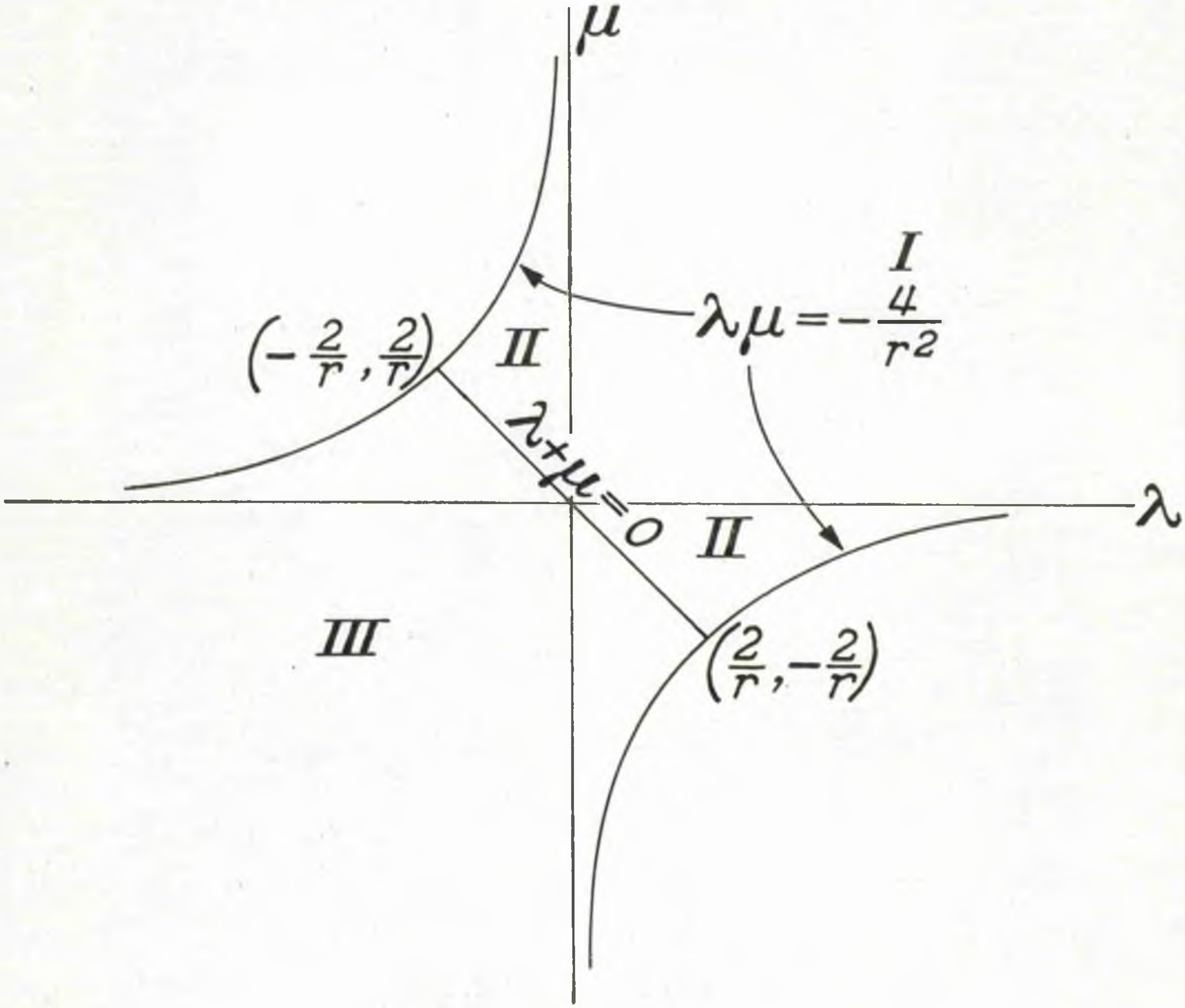

FIGURE 1.

The region is shown in figure 1, in which these conditions are satisfied.

The conditions on p_i and q_i, $i = 1,2$, for the equation (3) to be stable, now follow from an examination of equation (7). As was proved in chapter 1,

(i) If $p_i q_i + p_i + q_i > 0$, and $p_i + q_i > 0$, there are no negative roots.

(ii) If $p_i q_i + p_i + q_i = 0$, and $p_i + q_i \geq 0$, there is one zero root.

(iii) If $p_i q_i + p_i + q_i < 0$, there is one negative root.

(iv) If $p_i q_i + p_i + q_i = 0$, and $p_i + q_i < 0$, there is one zero root, and one negative root.

(v) If $p_i q_i + p_i + q_i > 0$, and $p_i + q_i < 0$, there are two negative roots.

Thus, if condition (i) is satisfied for both $i = 1$, and $i = 2$, then all values of λ_i and μ_j are in the region $\lambda + \mu \geq 0$, and equation (3) is stable. But conditions (8) may still be satisfied if one of λ_i or μ_j is negative. This corresponds to the fact noted in section 2.8, when it was shown that an eigenvalue of one of the Sturm-Liouville problems given after separation of variables, could be negative, while the solution of the differential equation remained bounded as $t \to \infty$.

	(a)	(b)	(c)	(d)	(e)
N	20	10	10	10	10
p_1	-3	2	1	0	$-\frac{1}{2}$
q_1	3	-2	1	0	1
p_2	-3	2	-1	0	$-\frac{1}{2}$
q_2	3	-2	1	0	1
Δ_1	-9	-4	+3	0	0
Δ_2	-9	-4	-1	0	0
δ_1	0	0	+2	0	$+\frac{1}{2}$
δ_2	0	0	0	0	$+\frac{1}{2}$
Case :	(iii)	(iii)	(iii) and (i)	(ii)	(ii)

TABLE 1.

K	(a)	K	(b)	(c)	(d)	(e)
1	•000 9	1	•006 088	•007 2	•006 5	•004 5
5	•002 3	2	•009 385	•006 4	•005 2	•004 8
10	•005 7	5	•025 160	•006 6	•005 0	•007 2
20	•018 3	10	•063 636	•006 1	•004 7	•007 4
40	•120 4	15	•145 111	•005 7	•004 6	•007 5
70	1•791 4	20	•323 066	•005 4	•004 6	•007 6
100	26•3	25	•715 069	•005 0	•004 6	•007 6
unstable			unstable	stable	← persistent error →	

TABLE 2.

3.5 Numerical Results.

The preceding results are now illustrated by a numerical example. The problem considered, consists of equation (1), in $\bar{R}$, subject to the initial condition

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

and the boundary conditions

$$\left.\begin{aligned} \frac{\partial u}{\partial x} - p_1 u &= \pi e^{-2\pi^2 t}\sin\pi y, \quad x = 0, \\ \frac{\partial u}{\partial x} + q_1 u &= -\pi e^{-2\pi^2 t}\sin\pi y, \quad x = 1, \end{aligned}\right\} \quad 0 \leq y \leq 1, \ t \geq 0,$$

$$\left.\begin{aligned} \frac{\partial u}{\partial y} - p_2 u &= \pi e^{-2\pi^2 t}\sin\pi x, \quad y = 0, \\ \frac{\partial u}{\partial y} + q_2 u &= -\pi e^{-2\pi^2 t}\sin\pi x, \quad y = 1, \end{aligned}\right\} \quad 0 \leq x \leq 1, \ t \geq 0.$$

The theoretical solution of the above problem is

$$u(x,y,t) \quad = \quad e^{-2\pi^2 t}\sin\pi x \sin\pi y, \quad \forall\ p_i, q_i, \ (i = 1,2).$$

Table 1 details the problems solved by the Peaceman-Rachford method. The values of N, p_1, p_2, q_1, q_2, are given, and the quantities $\Delta_i = p_i q_i + p_i + q_i$, and $\delta_i = p_i + q_i$, $(i = 1,2)$, are calculated. The number at the foot of each column, refers to the five cases (i),(ii),(iii),(iv),(v), of section 4.

The numerical errors are shown in table 2, after k time steps. The problems (a) and (b), are unstable, since there are eigenvalues λ_i and μ_j which are negative. In calculation (c), one of the eigenvalues μ_j is negative, but $\lambda_i + \mu_j > 0$, for all i,j, and the difference method is stable. The cases (d) and (e), are instances of a persistent error $(\Delta_i = 0,\ i = 1,2)$.

3.6 The Numerical Solution In s Space Variables.

Let the region $\bar{R}$ be

$$[0 \leq x_i \leq 1] \times [t \geq 0], \quad i = 1,2,\ldots,s,$$

and let the equation be

$$\frac{\partial u}{\partial t} = \nabla^2 u, \quad \text{in } \bar{R}, \qquad (8)$$

where

$$\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_s}\right).$$

The boundary conditions are now of the form

$$\left.\begin{aligned} \frac{\partial u}{\partial x_i} - p_i u &= \phi_0^i(t), \quad x_i = 0, \\ \frac{\partial u}{\partial x_i} + q_i u &= \phi_1^i(t), \quad x_i = 1, \end{aligned}\right\} t \geq 0,$$

for $i = 1,2,\ldots,s$; and the initial condition is

$$u(x_1,x_2,\ldots,x_s) = f(x_1,x_2,\ldots,x_s).$$

If a mesh is placed on the region $\bar{R}$, as before, and the difference approximation to (8) is obtained, then the analogue of (6) is

$$\mu_{i_1,i_2,\ldots,i_s} = \frac{(1-\frac{1}{2}r\lambda_{i}^{(1)})(\ldots\ldots)(1-\frac{1}{2}r\lambda_{j_s}^{(s)})}{(1+\frac{1}{2}r\lambda_{i}^{(1)})(\ldots\ldots)(1+\frac{1}{2}r\lambda_{j_s}^{(s)})}$$

where $\lambda_j^{(i)}$, $i = 1,2,\ldots,s$, $j = 0,1,\ldots,N$, are the eigenvalues of the matrices V_i given by equation (5).

A simple sufficient condition for stability is

$$\lambda_j^{(i)} \geq 0,$$

for each i and j, but, as for the case of two variables, it is possible for some $\lambda^{(i)}$ to be negative, while the difference equation remains stable, (c.f. the remarks in chapter 2).

3.7 Laplace's Equation.

The instability observed in the numerical solution of the heat conduction equation, was caused by the occurrence of a matrix with spectral radius $1 + O(\Delta t)$, where Δt was the time increment. The errors in the difference solution were shown to be dependent on the powers of this matrix, and, therefore, to grow with the number of time steps. In chapter 1 it was pointed out that a spectral radius of this order would give bounded errors in a closed region, $0 \leqslant t \leqslant T$, but that the errors would become unbounded as $t \to \infty$. Thus, problems of numerical instability in the solution of the third boundary value problem for the heat equation, which arise from the boundary conditions, are important only for large values of the time. However, this type of instability assumes a new importance in the solution of Laplace's equation.

There are two difference methods in common use, for the solution of Laplace's equation. Both of these involve obtaining a discrete approximation to the Laplacian operator, thus giving a system of linear equations to be solved. This system of equations is then solved by means of either

(1) An iterative method (see e.g. [36]), of which the two forms most commonly used are

(a) Alternating Direction Implicit Methods (A.D.I.)

and (b) Successive Over-relaxation (S.O.R)

or

(2) A direct method giving the inverse of the coefficient matrix of the system.

It will be shown that the asymptotic instability already observed, will prevent the use of the iterative methods (1), and make necessary the use of the computationally longer method 2. The singular problem (as defined in chapter 2) will also be considered numerically.

3.8 The Difference Equations.

Consider Laplace's equation in two independent variables

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{9}$$

in the region

$$R = \{0 \leq x,y \leq 1\}$$

subject to the boundary conditions

$$\left.\begin{aligned} \frac{\partial u}{\partial x} - p_1 u &= f_0(y), \quad x = 0, \\ \frac{\partial u}{\partial x} + q_1 u &= f_1(y), \quad x = 1, \end{aligned}\right\} \; 0 \leq y \leq 1,$$
$$\left.\begin{aligned} \frac{\partial u}{\partial y} - p_2 u &= g_0(x), \quad y = 0, \\ \frac{\partial u}{\partial y} + q_2 u &= g_1(x), \quad y = 1. \end{aligned}\right\} \; 0 \leq x \leq 1, \tag{9a}$$

The region R is covered by a rectilinear net, with mesh points $x_i = ih$, $y_j = jh$, where $i,j = 0,1,\ldots,N$, and $Nh = 1$. The Laplacian operator is replaced by the usual five point discretisation, given by

$$\delta_x^2 w_{i,j} + \delta_y^2 w_{i,j} = 0 \tag{10}$$

where $w_{i,j}$ is the numerical approximation to the function u at the point (x_i,y_j); and δ_x and δ_y are the usual central difference operators. Equation (10) holds for $i,j = 1,2,\ldots,N-1$, and when i or j is 0 or N, the boundary conditions (9a) are applied. The derivatives on the boundary are approximated in the (by now) usual manner. Letting $\underset{\sim}{w}$ be the vector given by

$$\underset{\sim}{w} = (w_{0,0}, w_{0,1}, \ldots, w_{0,N}, w_{1,0}, \ldots, w_{1,N}, \ldots, w_{N,0}, \ldots, w_{N,N})^T ,$$

then the system of equations (10) may be written in the form

$$A \underset{\sim}{w} = \underset{\sim}{b}, \tag{11}$$

where $\underset{\sim}{l}$ is a vector with $(N+1)^2$ components, involving the boundary functions, and where

$$A = U_1 + U_2,$$

the matrices U_1 and U_2 being given as in section 3, of this chapter.

The system (11) is a set of $(N+1)^2$ equations in $(N+1)^2$ unknowns, and for large N, it is not always feasible to invert A directly. Consequently, iterative methods were developed, which were quick and computationally economical methods for obtaining the solution of (11). Successive approximations $\underset{\sim}{w}^n$ to the solution of (11), are calculated according to some method. The Peaceman-Rachford A.D.I. method is given by

$$\left.\begin{aligned} \underset{\sim}{w}^{n+\frac{1}{2}} &= [I+rU_1]^{-1} [(I-rU_2)\, \underset{\sim}{w}^n + \underset{\sim}{l}\,] \\ \underset{\sim}{w}^{n+1} &= [I+rU_2]^{-1} [(I-rU_1)\, \underset{\sim}{w}^{n+\frac{1}{2}} + \underset{\sim}{l}\,] \end{aligned}\right\} \qquad (12)$$

where I is the unit matrix of order$(N+1)^2$. The number r is a parameter, and is usually varied from one iteration to the next, in order to accelerate convergence of the approximation $\underset{\sim}{w}^n$ to the solution of equation (11), (e.g.[47,48,49]). The measure of convergence of $\underset{\sim}{w}^n$ to $\underset{\sim}{w}$, is given by $\underset{\sim}{\epsilon}^n$, where

$$\underset{\sim}{\epsilon}^n = \underset{\sim}{w}^n - \underset{\sim}{w}\,.$$

Then $\underset{\sim}{w}^n \to \underset{\sim}{w}$, as $n \to \infty$, if, and only if, $\underset{\sim}{\epsilon}^n \to \underset{\sim}{0}$, i.e. $\|\underset{\sim}{\epsilon}^n\| \to 0$. But $\underset{\sim}{\epsilon}^n$ satisfies the equations

$$\underset{\sim}{\epsilon}^{n+\frac{1}{2}} = [I+rU_1]^{-1}[I-rU_2]\, \underset{\sim}{\epsilon}^n$$

$$\underset{\sim}{\epsilon}^{n+1} = [I+rU_2]^{-2}[I-rU_1]\, \underset{\sim}{\epsilon}^{n+\frac{1}{2}}$$

which, on elimination of $\underset{\sim}{\epsilon}^{n+\frac{1}{2}}$, give

$$\underset{\sim}{e}^{n+1} = [I+rU_2]^{-1}[I-rU_1][I+rU_1]^{-1}[I-rU_2]\,\underset{\sim}{e}^{n} \qquad (13)$$

But, from the relevant parts of sections 3 and 4, equation (13) gives $\underset{\sim}{e} \to \underset{\sim}{0}$, as n tends to infinity, if, and only if, the matrix A is positive definite.

The method of S.O.R., for the solution of (11), may be examined in the same way as above. The matrix A may be written in the form

$$A = \begin{bmatrix} B_0 & C_0 & & & & \\ D_1 & B_1 & C_1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & D_{N-1} & B_{N-1} & C_{N-1} \\ & & & & D_N & B_N \end{bmatrix}$$

where the matrices B_i, C_i, D_i, are of order N+1. The error equation may then be shown to be :

$$\underset{\sim}{e}^{n+1} = \begin{bmatrix} B_0 & & & \\ D_1 & B_1 & & \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ & & D_N & B_N \end{bmatrix}^{-1} \left\{ (1-r)I + r\begin{bmatrix} B_0 & & & \\ D_1 & B_1 & & \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ & & D_N & B_N \end{bmatrix} \begin{bmatrix} 0 & -C_0 & \\ & \cdot & \cdot \\ & & -C_{N-1} \\ & & 0 \end{bmatrix} \right\} \underset{\sim}{e}^{n} \qquad (14)$$

where r is a parameter, which may be varied.

It has been shown in Varga [47], that equation (14), like (13), gives $\underset{\sim}{e}^{n} \to \underset{\sim}{0}$, as $n \to \infty$, if, and only if, the matrix A is positive definite.

Thus, unless A is positive definite, or positive semi-definite, as the next section will show, the above iterative methods will not converge. Both the iterative methods, discussed above, are in fact solving heat conduction equations, giving the solution of Laplace's equation as the steady state solution of a heat equation. But, if the corresponding heat equation has solutions which grow exponentially in time, then the "steady state" solution is unbounded. This may be expressed otherwise, by saying that not all solutions of Laplace's equation may be obtained as steady state solutions of a heat conduction equation.

5.9 The Singular Case.

If A is singular, then equations (11) may have no solution, or many solutions. The question of whether there is a solution, is considered first. There will be a solution if, and only if, the vector $\underset{\sim}{l}$ lies in the range space of A; i.e. if, and only if,

$$\text{the rank of } A \;=\; \text{the rank of } \left[\begin{array}{c|c} & l_{0,0} \\ & l_{0,1} \\ A & \cdot \\ & \cdot \\ & \cdot \\ & l_{N,N} \end{array}\right]$$

It has been proved that, if A is singular, then the rank is $(N+1)^2-1$. It can be shown (e.g. Halmos [22]) that $\underset{\sim}{l}$ is in the range space of A, if, and only if, $\underset{\sim}{l}$ is in the orthogonal complement of the null space of A^T; i.e. $\underset{\sim}{x}^T\underset{\sim}{l} = 0$, $\forall \underset{\sim}{x}$, such that $A^T\underset{\sim}{x} = \underset{\sim}{0}$. But, since the rank of A^T is $(N+1)^2-1$, all such vectors $\underset{\sim}{x}$ are multiples of one vector. For example, if U_1 and U_2 are both singular, i.e.

$$p_i q_i + p_i + q_i \;=\; 0, \; i = 1,2,$$

then

$$\underset{\sim}{x} \;=\; \underset{\sim}{x}_1 \times \underset{\sim}{x}_2$$

where

$$x_i \;=\; \{\tfrac{1}{2},\; 1+p_i/N, \ldots\ldots,\; 1+p_i(N-1)/N,\; 1+\tfrac{1}{2}p_i\}^T, \quad i = 1,2.$$

The condition on $\underset{\sim}{l}$ for a solution to exist is thus

$$(\underset{\sim}{x}_1 \otimes \underset{\sim}{x}_2)^T\, \underset{\sim}{l} = 0 \tag{15}$$

The vector $\underset{\sim}{l}$ is given by

$$\mathbf{l} = -2/N\ \{f_0^0, f_0^1, \ldots, f_0^N, 0, 0, \ldots, 0, -f_1^0, -f_1^1, \ldots\ldots, -f_1^N\}^T +$$

$$-2/N\ \{g_0^0, 0, 0, \ldots, 0, -g_1^0, g_0^1, 0, \ldots, 0, -g_1^1, \ldots\ldots, g_0^N, 0, \ldots, 0, -g_1^N\}^T,$$

where

$$\left.\begin{array}{lcl} f_i^j & = & f_i(y_j) \\ g_i^j & = & g_i(x_j) \end{array}\right\}, \quad i = 0,1;\ j = 0,1,\ldots,N.$$

The condition (15) then becomes a trapezoidal approximation to the integral condition :

$$\int_0^1 (1+p_2y)\, f_0(y)\, dy - \int_0^1 (1+p_2y)(1+p_1)f_1(y)\, dy +$$
$$\int_0^1 (1+p_1x)\, g_0(x)\, dx - \int_0^1 (1+p_1x)(1+p_2)g_1(x)\, dx \qquad = \quad 0\ ;$$

i.e.

$$\int_{\partial R} (1+p_1x)(1+p_2y)\, F(x,y)\, ds \qquad = \qquad 0 \qquad\qquad (16),$$

where ∂R is the boundary of R; and F is the function which reduces to f_i, g_i on the different parts of ∂R. The relation (16) is, of course, a particular case of the relation (28), of section 2.9.

If condition (15) is satisfied, the equations (11) have a solution. To prove convergence of, for example, the A.D.I. method, it is necessary to show, not that $\underset{\sim}{\epsilon}^n \to \underset{\sim}{0}$, as $n \to \infty$, but that $\underset{\sim}{\epsilon}^n \to$ a vector in the null space of A. Provided A is positive semi-definite, this has been proved by Douglas and Pearcy [13]. A similar proof will show convergence of the S.O.R. method, when A is positive semi-definite.

The fact that $\underset{\sim}{\epsilon}^n$ does not tend to zero, as n tends to infinity, for either of the two iterative methods considered, <u>when A is positive semi-definite and singular,</u> is a reflection of the fact that equation (9), subject to (9a), does not have a unique solution. Solving (11) by an iterative method which requires the choice of a starting vector, would ensure a unique solution, if all calculations were carried out with infinite precision. The choice of starting values corresponds to being given initial data in a heat conduction equation. Errors committed at each iteration, however, will persist in the form of a vector in the null space of A.

3.10 The Unstable Case.

If the conditions (i), or (ii), on p_i and q_i, $i = 1,2$, given in section 3.4, are not satisfied, then A is not positive semi-definite, and the iterative methods of section 3.8 cannot be used. A direct method of inverting A must be employed. One, due to S. Schechter, [45], is used in the next section, to give numerical results in this unstable case, when A has one, or two, negative eigenvalues. This method is now described.

The matrix A may be written in the form

$$\begin{bmatrix} B_0 & 2I & & & & \\ I & B_1 & I & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & I & B_{N-1} & I \\ & & & & I & B_N \end{bmatrix}$$

where each submatrix is of order N+1. By a slight modification of Schechter's method, a sequence of matrices $\{H_n\}$, $n = 0,1,2,\ldots$, is formed, where

$$H_0 = 2I,$$

$$H_1 = B_0,$$

$$H_n = H_{n-1}B_{n-1} - H_{n-2}, \quad n = 2,3,\ldots,N+1.$$

The vector $\underline{l}$ is partitioned into N+1 vectors, $\underline{g}_i$, $i = 0,1,\ldots,N$, with N+1 components. A sequence of vectors $\{\underline{y}_n\}$, $n = 0,1,\ldots,N$, where

$$\underline{y}_0 = \underline{g}_0 \text{, and } \underline{y}_n = H_{n-1}\underline{g}_n - \underline{y}_{n-1},$$

is obtained; then, if the solution $\underline{w}$ is partitioned into N+1 vectors, $\underline{w}_0$, $\underline{w}_1,\ldots,\underline{w}_N$, we have :

$$\underset{\sim}{w}_N = H_{N+1}^{-1} \underset{\sim}{x}_N ,$$

$$\underset{\sim}{w}_{N-1} = \underset{\sim}{g}_N - B_N \underset{\sim}{w}_N ,$$

$$\underset{\sim}{w}_{n-1} = \underset{\sim}{g}_n - B_n \underset{\sim}{w}_n - \underset{\sim}{w}_{n+1} , \quad n = 1,2,\ldots,N-1.$$

This enables one to calculate the solution $\underset{\sim}{w}$, by inverting only one matrix H_{N+1}, of order $N+1$.

In [45], Schechter does not consider the case when H_{N+1} is singular. It is now shown that, in a rectangular region, H_N is singular if, and only if, A is singular.

It is easy to show that

$$B_0 = -V_2 - 2(1+p_1/N)I,$$

where V_2 is given by equation (5) of section (2); that

$$B_n = -V_2 - 2I, \quad n = 1,2,\ldots,N-1; \text{ and}$$

$$B_N = -\tfrac{1}{2}V_2 - (1+q_1/N)I.$$

Let the eigenvalues of H_n be $f_n(\mu_i)$, where $n = 0,1,\ldots,N+1$, and μ_i, $i = 0,1,\ldots,N$, are the eigenvalues of V_2; and where f_n is a polynomial, since each H_n is a polynomial in V_2 and I. Since $H_0 = 2I$, and H_1 is $-V_2 - 2(1+p_1/N)I$, then

$$f_0(\mu_i) = 2, \text{ and}$$

$$f_1(\mu_i) = -\mu_i - 2(1+p_1/N).$$

Using the recurrence relation for H_n, we obtain :

$$f_n(\mu_i) + f_{n-2}(\mu_i) = f_{n-1}(\mu_i)\,(-2-\mu_i), \quad n = 2,3,\ldots,N. \tag{17}$$

The eigenvalues μ_i, $i = 0,1,\ldots,N$, are the roots of the equation

$$[(\mu-2)^2 - 4 + 4p_2q_2/N^2\,]\,T_{N-1}(\mu) + 2(p_2+q_2)/N\,[T_N(\mu) - T_{N-2}(\mu)] = 0.$$

Putting $\mu = 2(1-\cos\phi)$, as in section 1.6, and using the properties of the polynomials T_n, we obtain :

$$p_2q_2/N^2 \sin N\phi \; + \; (p_2+q_2)/N \cos N\phi\sin\phi \; - \; \sin N\phi\sin^2\phi \; = \; 0.$$

The solution of the recurrence relation (17), for $f_n(\mu_i)$ is

$$f_n(\mu_i) \;=\; K\cos n(\theta+\pi) \;+\; L\sin n(\theta+\pi), \tag{18}$$

where, from (17), with $\mu = 4 \sin^2\phi/2$, we obtain

$$2 \cos(\theta+\pi) \;=\; -4 \sin^2\phi/2 \;-\; 2;$$

i.e.

$$\sin^2\theta/2 + \sin^2\phi/2 \;=\; 0. \tag{19}$$

For (18) to be the solution of (17), it must satisfy the initial conditions for f_0 and f_1. Thus

$$f_0(\mu) \;=\; 2 \;=\; K,$$

$$f_1(\mu) \;=\; -\mu - 2(1+p_1/N) = 2\cos(\theta+\pi) + L\sin(\theta+\pi);$$

i.e.

$$L \;=\; \frac{-2p_1}{N \sin\theta} .$$

Therefore

$$f_n(\mu_i) \;=\; 2(\cos n(\theta+\pi) + \frac{p_1}{N \sin\theta} \sin n(\theta+\pi)), \tag{20}$$

for $n = 2,3,\ldots,N$. Using the recurrence relation for H_n , with $n = N+1$, the result

$$f_{N+1}(\mu_i) \;=\; -f_{N-1}(\mu_i) + f_N(\mu_i)\,(-\tfrac{1}{2}\mu_i - 1 - q_1/N),$$

follows.

Taking f_{N-1} and f_N from equation (20), then H_{N+1} is singular if, and only if,

$$f_{N+1}(\mu_i) \;=\; 0, \text{ for some } i.$$

This gives the condition : if, and only if, θ satisfies

$$p_1q_1/N^2 \sin N(\theta+\pi) - (p_1+q_1)/N \cos N(\theta+\pi)\sin(\theta+\pi) - \sin N(\theta+\pi)\sin^2(\theta+\pi) = 0;$$

i.e.

$$p_1q_1/N^2 \sin N\theta + (p_1+q_1)/N \cos N\theta \sin\theta - \sin N\theta \sin^2\theta = 0.$$

But the roots of this equation give the eigenvalues λ_j of V_1, $j = 0,1,\ldots,N$, where

$$\lambda_j = 4\sin^2\theta/2.$$

Thus, using (19), H_{N+1} is singular if, and only if, there are eigenvalues λ_j, μ_i, of V_1 and V_2, respectively, such that

$$\lambda_j + \mu_i = 0.$$

This is the condition for A to be singular.

The method of Schechter may thus be used for all problems which cannot be solved iteratively, except when A is singular, and also has a negative eigenvalue, for example when

$$\left.\begin{aligned} p_iq_i + p_i + q_i &= 0, \text{ and} \\ p_i + q_i &< 0 \end{aligned}\right\} \quad i = 1,2.$$

For such a problem, a direct method may be used only if some further condition is applied to $\underset{\sim}{w}$, to make the system of equations non-singular.

It may be added as a note, that, for the case of several space variables, the same conclusions hold.

	N	p_1	q_1	p_2	q_2	Δ_1	Δ_2	δ_1	δ_2	case :	
(1)	10	1	1	1	1	3	3	2	2	(i)	stable
(2)	10	1	2	$-3/4$	1	5	$-\frac{1}{2}$	3	1/4	(iii),(i)	stable
(3)	10	-1	2	-1	2	-1	-1	1	1	(iii)	unstable

TABLE 3.

K	A.D.I.	S.O.R.	Direct.	A.D.I.	S.O.R.	Direct.	A.D.I.	S.O.R.	Direct.
10	•49499	•06244	$•158400 \times 10^{-9}$	•81861	•24207	$•827196 \times 10^{-10}$	3•108	2•215	$•248473 \times 10^{-10}$
20	•21124	•00350		•53682	•01527		4•642	7•892	
30	•09078	•00031		•34320	•00070		6•752	29•081	
40	•03927	•00004		•21900	•00018		9•810	107•	
50	•01678	-		•13973	•00000		14•252	-	
60	•00721	-		•08915	-		20•7051	-	
70	•00310	-		•05687	-		30•079	-	
80	•00133	-		•03629	-		-	-	
90	•00057	-		•02315	-		-	-	
100	•00000	-		•01477	-		-	-	
110	-	-		•00942	-		-	-	
120	-	-		•00601	-		-	-	
	(1)			(2)			(3)		

TABLE 4.

3.11 Numerical Results.

Some simple problems were solved, to demonstrate the preceding results: viz : that some boundary value Laplacian problems may be solved by A.D.I, or S.O.R, or by a direct method; but a class of problems exists which cannot be solved except by direct methods. The Peaceman-Rachford method, and the S.O.R. method, with fixed parameters, r, and Schechter's method, were each used to solve three problems.

The differential equation (9) was solved, subject to the boundary conditions (9a), where

$$\left.\begin{aligned} f_0(y) &= 1 + p_1 y \\ f_1(y) &= 1 + q_1(1-y) \end{aligned}\right\} \text{ and } \left.\begin{aligned} g_0(x) &= -1 - p_2 x \\ g_1(x) &= -1 + q_2(x-1) \end{aligned}\right\}$$

The solution of the differential equation is

$$u(x,y) = x - y, \text{ for all } p_i, q_i, \; i = 1,2.$$

The values of p_i, q_i are shown in table 3, for the problems solved, together with the values of the quantities Δ_i and δ_i $(i = 1,2)$, defined in section 5. The first two problems are solved equally well by all three methods, but only Schechter's method gives a solution of the last problem. In table 4, the maximum errors are shown after k iterations, for the Peaceman-Rachford method, and the S.O.R. method, and the maximum error is given for Schechter's method.

3.12 Conclusion.

It is thus clear that the type of instability discussed above, although not very important in heat conduction problems, in which the time is not allowed to become large, is of importance in the solution of Laplace's equation by iterative methods. It may be that the heat conduction equation is not always the best transient (i.e. time-dependent) equation to choose, in order to obtain the solution of Laplace's equation, as an asymptotic solution. Most iterative methods in use at the present time are of this form, but the question remains as to what other kinds of iterative methods might exist for solving second-order elliptic problems.

CHAPTER 4.

THE NUMERICAL SOLUTION OF THE WAVE EQUATION UNDER DERIVATIVE BOUNDARY CONDITIONS.

4.1 Introduction.

The wave equation in one space variable is considered, in an open rectangular region, with initial conditions given, at $t = 0$, and with boundary conditions which consist of linear combinations of the function and its first space and time derivatives, given on the two sides of the region. A class of difference-differential approximations to this system is examined, and conditions on the coefficients in the boundary conditions are obtained, to ensure that the solutions of the difference-differential system remain bounded for all $t > 0$. The solutions of the differential system are also examined, and their asymptotic behaviour, (as $t \to \infty$), is discussed.

4.2 The Differential And Difference-Differential Systems.

Consider the equation

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

in the region

$$R = [0 \leqslant x \leqslant 1] \times [t \geqslant 0],$$

subject to the initial conditions

$$\left.\begin{aligned} u(x,0) &= f_1(x) \\ u_t(x,0) &= f_2(x) \end{aligned}\right\} \quad 0 \leqslant x \leqslant 1, \quad t = 0, \qquad (1a)$$

and the boundary conditions

$$\left.\begin{aligned} u_x - p_1 u - p_2 u_t &= \phi_0(t), \quad x = 0, \\ u_x + q_1 u + q_2 u_t &= \phi_1(t), \quad x = 1, \end{aligned}\right\} \quad t \geqslant 0 \qquad (1b)$$

where the subscripts x and t denote partial differentiation with respect to x and t, respectively, and where the coefficients p_i, q_i, $i = 1,2$, are constants. The usual assumptions of continuity of initial and boundary conditions, and at the corners of R, are made, and ϕ_0, ϕ_1 are bounded as $t \to \infty$.

Following the method used in chapter 2, section 6, we discretise only the space variable x. Each line t = constant in R is divided into N parts, giving N+1 points, x_i, $i = 0,1,\ldots,N$, where $x_i = ih$, and $Nh = 1$. Equation (1) is approximated by the equation

$$\frac{d^2 w_i}{d t^2} = \frac{1}{h^2} \delta_x^2 w_i \, , \qquad (2)$$

where $i = 0,1,\ldots,N$; w_i is the approximation to $u(x_i,t)$; and δ_x is the usual central difference operator.

When $i = 0$, or N, in equation (2), values of w at points outside R, are introduced. These are eliminated by using the boundary conditions (1b). For example, when $i = 0$, the value of w_{-1} is eliminated, by using the equation

$$\frac{w_1 - w_{-1}}{2h} - p_1 w_0 - p_2 \frac{dw_0}{dt} = \phi_0(t) \tag{3}$$

and when $i = N$, w_{N+1} is eliminated, by using the equation

$$\frac{w_{N+1} - w_{N-1}}{2h} + q_1 w_N + q_2 \frac{dw_N}{dt} = \phi_1(t) \tag{4}$$

Substitution of the relations (3) and (4) into equation (2), gives a system of equations for $w_0, w_1, \ldots, w_N$. If we put

$$\underset{\sim}{w} = (w_0, w_1, \ldots, w_N)^T,$$

then the system of difference-differential equations may be written in the form

$$\frac{d^2 \underset{\sim}{w}}{d\, t^2} = -\frac{1}{h^2} U \underset{\sim}{w} - \frac{1}{h} \frac{d}{dt} A \underset{\sim}{w} + \underset{\sim}{l}, \tag{5}$$

where $\underset{\sim}{l}$ is a vector with $N+1$ components, involving the boundary functions ϕ_0 and ϕ_1, and U and A are matrices of order $N+1$, given by

$$(6) \qquad U = \begin{bmatrix} 2(1+\frac{p_1}{N}) & -2 & & & & \\ -1 & 2 & -1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & -1 & 2 & -1 \\ & & & & -2 & 2(1+\frac{q_1}{N}) \end{bmatrix}$$

and

$$A = \begin{bmatrix} 2p_2 & & & & & & \\ & 0 & & & & & \\ & & 0 & & & & \\ & & & \cdot & & & \\ & & & & \cdot & & \\ & & & & & 0 & \\ & & & & & & 2q_2 \end{bmatrix}$$

Let

$$\frac{d\underset{\sim}{w}}{dt} = \underset{\sim}{v}$$

and rewrite the system (5) in the form

$$\frac{d\underset{\sim}{v}}{dt} = -\frac{1}{h^2}\, U \underset{\sim}{w} - \frac{1}{h} A \underset{\sim}{v} + \underset{\sim}{1}.$$

This system is equivalent to the system

$$\frac{d}{dt}\begin{bmatrix} \underset{\sim}{v} \\ \underset{\sim}{w} \end{bmatrix} = -\begin{bmatrix} \frac{1}{h} A & \frac{1}{h^2} U \\ -I & 0 \end{bmatrix}\begin{bmatrix} \underset{\sim}{v} \\ \underset{\sim}{w} \end{bmatrix} + \underset{\sim}{1} \qquad (8)$$

where I is the unit matrix of order N+1.

Most difference approximations to equation (1), subject to (1a) and (1b), are obtained from equation (5), by approximating the time derivatives in some way, to give an implicit or an explicit method. Clearly, no difference approximation to equation (5) will be stable for all t, unless the solutions of (5) are bounded for all t, and for all ϕ_0, ϕ_1. The system (5) has bounded solutions if, and only if, the system (8) has bounded solutions. The condition for this, (e.g. Bellman, [3]), is that the matrix of order 2N+2, given by

$$B = \begin{bmatrix} \frac{1}{h} A & \frac{1}{h^2} U \\ -I & 0 \end{bmatrix}$$

should have all its eigenvalues μ_i, $i = 1,2,\ldots,2N+2$, in $R(\mu_i) \geq 0$.

Elementary row operations, applied to the matrix in the determinantal equation

$$\det \begin{bmatrix} \frac{1}{h} A - \mu I & \frac{1}{h^2} U \\ -I & -\mu I \end{bmatrix} = 0 \tag{9}$$

give the equation

$$\det \begin{bmatrix} \frac{1}{h} A - \mu & \frac{1}{h^2} U \\ 0 & - I(\frac{1}{h} A - \mu I) + \frac{1}{h^2} U \end{bmatrix} = 0$$

i.e.

$$\det(\frac{1}{h} A - \mu I). \det(\mu^2 I - \frac{\mu}{h} \quad + \frac{1}{h^2} U) = 0.$$

If $-\beta$ is substituted for μh, then the condition for boundedness of the solutions of (5) is that the roots of the equation

$$\det(A + \beta I). \det(\beta^2 I + \beta A + U) = 0, \tag{10}$$

should satisfy the condition

$$R(\beta) < 0.$$

The values of β given by (10) are thus the roots of

$$F(\beta) = \det(\beta^2 I + \beta A + U) = 0, \tag{11}$$

since, in the row operations on the matrix in (9), it is assumed that

$$\det(A + \beta I) \neq 0.$$

The determination of conditions on p_i, q_i, $i = 1,2$, so that the roots of (11) satisfy the condition $R(\beta) < 0$, is equivalent to the solution of a standard Routh-Hurwitz problem for the polynomial $F(\beta)$.

4.3 The Routh-Hurwitz Problem.

The Routh-Hurwitz problem [17] for a given polynomial with real coefficients, is to find conditions on the coefficients so that all the roots of the polynomial lie in the negative real half-plane. This problem is easily solved for $F(\beta)$, when no time derivatives occur on the boundary of R, $(x = 0, 1)$, in the conditions (1b). For, in this case $p_2 = q_2 = 0$, and $A = 0$, so that

$$F(\beta) = \det(\beta^2 I + U) = 0. \tag{12}$$

But, in chapter 1, it has already been shown that the eigenvalues λ_i $(i=0, 1, \ldots, N)$ of U are real, and so the roots of (12) are

$$\beta_j^2 = -\lambda_j , \quad j = 0, 1, \ldots, N;$$

i.e.

$$\beta_j = \pm i\sqrt{\lambda_j} ,$$

provided $\lambda_j \geq 0$. If any one of the eigenvalues of U is negative (e.g. suppose $\lambda_0 < 0$), then

$$\beta_0 = \pm \sqrt{-\lambda_0},$$

and so there is a root of (12) in $R(\beta) > 0$. Thus, when p_2, q_2 are both zero, the conditions for equation (5) to have bounded solutions are the conditions for the matrix U to have only non-negative eigenvalues; i.e. (chapter 1)

$$p_1 q_1 + p_1 + q_1 \geq 0,$$
$$p_1 + q_1 \geq 0.$$

The case when p_2 and q_2 are not both zero requires further analysis. If the determintal equation (10) is expanded, the result

$$F(\beta) = [\beta^4 + 4(1+p_2q_2)\beta^2 + 4(p_2q_1+p_1q_2)\beta/N + 4p_1q_1/N^2]\,T_{N-1}(-\beta^2) + 2[(p_1+q_1)/N + \beta(p_2+q_2)][T_N(-\beta^2) - T_{N-2}(-\beta^2)] = 0, \tag{13}$$

is obtained. In equation (13) the polynomials T_m, $m = N-2, N-1, N$, are as defined in chapter 1, i.e.

$$T_m(-\beta^2) = \sum_{r=0}^{m} C_r^{2m-r+1}\,\beta^{2m-2r} . \tag{14}$$

Necessary conditions for a polynomial to have all its roots with non-positive real parts were given by Hurwitz [23]. These conditions are that the coefficients should all have the same sign. Since the coefficient of β^{2N+2}, in (13), is clearly 1, then the coefficients must be positive. Using equations(13) and (14), the following coefficients are easily found.

(i) The constant term, which is $4/N\,(p_1q_1 + p_1 + q_1)$;

(ii) The coefficient of β, which is $4(p_1q_2 + p_2q_1 + p_2 + q_2)$;

(iii) The coefficient of β^{2N+1}, which is $2(p_2 + q_2)$;

(iv) The coefficient of β^{2r}, which is a positive multiple of :

$$(2r+1)2r(2r-1)(2r-2) + 4(1+p_2q_2)(N+r-1)(N-r+1)(2r+1)2r + 4p_1q_1/N^2\,(N+r)(N+r-1)(N-r+1)(N-r) + 4(p_1+q_1)(N+r-1)(N-r+1)(2r+1).$$

In the coefficient of β^{2r}; given in (iv), let $r = aN$, where $0< a <1$, and $1/a$ is a factor of N. For large N, a may be small. Putting $r = aN$ in (iv) gives the expression :

$$N^4(2a+1/N)2a(2a-1/N)(2a-2/N) + 4(1+p_2q_2)N^4(1+a-1/N)(1-a+1/N)2a(2a+1/N) + 4p_1q_1N^2(1+a)(1+a-1/N)(1-a+1/N)(1-a) + 4(p_1+q_1)N^3(1+a-1/N)(1-a+1/N)(2a+1/N),$$

and, for large N, the sign of this will be determined by the sign of

$$a^4 + a^2(1-a^2)(1+p_2q_2) .$$

For large N, a may be taken very small, but still much greater than $1/N$. In this case a^4 is negligible, compared with a^2, and the sign of the coefficient of β^{2r} is determined by the sign of $p_2q_2 + 1$.

Thus, the following set of necessary conditions, for F to have no roots in the positive real half-plane, have been obtained :

$$\text{(i)}\quad p_1q_1 + p_1 + q_1 \geq 0;$$

$$\text{(ii)}\quad p_1q_2 + p_2q_1 + p_2 + q_2 \geq 0;$$

$$\text{(iii)}\quad p_2 + q_2 \geq 0;$$

$$\text{(iv)}\quad 1 + p_2q_2 \geq 0.$$

The number of zeros of $F(\beta)$, in $R(\beta) > 0$, is given by

$$\frac{1}{2\pi}\{\text{ change in } \arg[F(\beta)]\ \}$$

as β goes round the large contour $|\beta| = R$, $R(\beta) \geq 0$, provided F has no zeros on this contour. The semi-circle may be taken large enough, to contain all the zeros of F in the positive real half-plane, and it will be assumed that $p_1q_1 + p_1 + q_1 \neq 0$; i.e. $\beta = 0$ is not a root of $F(\beta) = 0$. In addition, it is assumed that F has no zeros on the imaginary axis. Since the largest power of β in $F(\beta)$ is $2N+2$, then, as β goes from $-\infty i$ to $+\infty i$ around the semi-circle (figure 1), $\arg[F(\beta)]$ increases by $(N+1)\pi$, approximately. On the imaginary axis $\beta = yi$, and

$$F(yi) = a(y) + ib(y),$$

where

$$a(y) = [y^4 - 4(1+p_2q_2)y^2 + 4p_1q_1/N^2]T_{N-1}(y^2) + 2(p_1+q_1)/N[T_N(y^2) - T_{N-2}(y^2)],$$

and

$$b(y) = y[4(p_1q_2+p_2q_1)/N\ T_{N-1}(y^2) + 2(p_2+q_2)\{T_N(y^2) - T_{N-2}(y^2)\}].$$

Clearly, therefore, at $\beta = +\infty i$,

$a(y) < 0$, if N is even, and

$b(y) > 0$.

Thus, if we assume N to be even, (the same argument holds for N odd), then, at ∞i, $F(\beta)$ has negative real part, and positive imaginary part. For $F(\beta)$ to have no roots in $R(\beta) > 0$, it is necessary, and sufficient, that, as β goes from $+ i$ to $- i$, down the imaginary axis, the argument of $F(\beta)$ should decrease by approximately $(N+1)\pi$; i.e. $a(y)$ must have a zero before $b(y)$, and the zeros of a and b must then alternate.

A discussion of the zeros of a and b is complicated, and so the question of whether they alternate or not was answered numerically. (A theoretical analysis is possible, but very tedious, and still leaves some questions unanswered.) The Routh algorithm [17] was used, with $N = 10$, for a set of 24 problems, given in table 1. The number of roots in $R(\beta) > 0$ is given for each value of p_i, q_i, $i = 1,2$, except when a coefficient is negative, in which case this fact is noted. The numerical results verify the necessary conditions (i) - (iv), and suggest that $p_2, q_2 \geq 0$, are also necessary conditions for the zeros of F to be in the positive real half-plane. The last two conditions may be obtained by an argument involving the zeros of a and b.

Thus, to the conditions (i) - (iv) may be added the conditions

$$\text{(v)}\quad p_2 \geq 0,\ q_2 \geq 0.$$

The numerical results of the next section suggest that these conditions are sufficient to ensure boundedness of the solutions of (5), and, therefore, stability of the particular difference method of solution used.

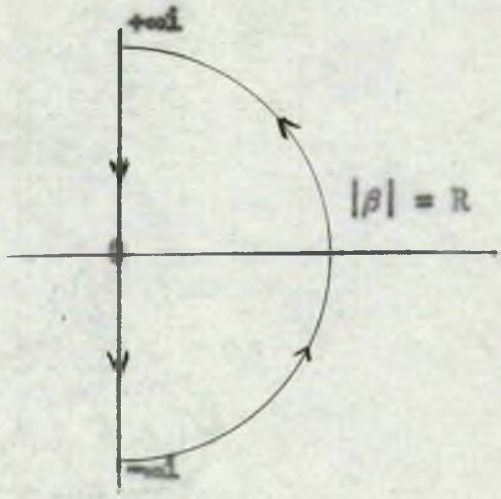

FIGURE 1.

(p_1, q_1) \ (p_2, q_2)	$(\frac{1}{2}, \frac{1}{2})$	$(0, 1)$	$(-1/4, 1)$	$(-1, \frac{1}{2})$
$(0, 0)$	NONE	NONE	8 roots	coeff. < 0.
$(\frac{1}{2}, \frac{1}{2})$	NONE	NONE	6 roots	coeff. < 0.
$(0, 1)$	NONE	NONE	6 roots	coeff. < 0.
$(-\frac{1}{2}, 1)$	NONE	NONE	6 roots	coeff. < 0.
$(-\frac{1}{2}, 2)$	NONE	NONE	coeff. < 0.	coeff. < 0.
$(-1, 1)$	coeff.<0.	coeff.<0.	coeff. < 0.	coeff. < 0.

TABLE 1.

4.4 Numerical Results.

The numerical solutions of the 24 problems considered in the last section, were obtained for equation (1), subject to the initial conditions

$$\left.\begin{aligned} u(x,0) &= \sin \pi x \\ u_t(x,0) &= 0 \end{aligned}\right\} \quad 0 \leqslant x \leqslant 1,$$

and the boundary conditions

$$\left.\begin{aligned} u_x - p_1 u - p_2 u_t &= \pi \cos \pi t, \quad x = 0, \\ u_x + q_1 u + q_2 u_t &= -\pi \cos \pi t, \quad x = 1. \end{aligned}\right\} \quad t \geqslant 0.$$

The theoretical solution of this system is

$$u(x,t) = \sin \pi x \cos \pi t, \qquad \forall\, p_i, q_i, \ i = 1,2.$$

The difference approximation used was the standard von-Neumann approximation, viz. :

$$\delta_t^2 w_j^n = r^2/4\, \delta_x^2 \left(w_j^{n+1} + 2w_j^n + w_j^{n-1} \right),$$

where δ_x and δ_t are the usual central difference operators, in the x and t directions; w_j^m is the numerical approximation to $u(jh, mk)$, $(m = n+1, n, n-1)$; k is the time increment; and $r = k/h$ is kept fixed. The first order time derivative in (5) was approximated by

$$\frac{d}{dt} u(jh, nk) = \frac{1}{2k} \left(w_j^{n+1} - w_j^{n-1} \right).$$

In the computations, N = 10, and r = 1. The calculations were each run for 400 time steps, and the maximum errors were calculated at this stage. Table 2 shows the error after 10 time steps, and the final error, both as percentages of the maximum of the solution.

The results suggest that if (i) - (iv) are not satisfied, instability is pronounced. If these are satisfied and (v) is not, then the instability is still evident, although it is slower to appear.

(p_2, q_2)

(p_1, q_1)	$(\frac{1}{2},\frac{1}{2})$	$(0,1)$	$(-1/2,1)$	$(-1,\frac{1}{2})$
$(0,0)$	2 % 2 %	3 % 4 %	4 % 10 %	10 % 10^{57}
$(\frac{1}{2},\frac{1}{2})$	2 % 2 %	3 % 3 %	4 % 40 %	18 % 10^{57}
$(0,1)$	2 % 2 %	3 % 4 %	4 % 10 %	10 % 10^{31}
$(-\frac{1}{2},1)$	$2\frac{1}{2}$ % $2\frac{1}{2}$ %	4 % 4 %	5 % 24 %	100 % 10^{99}
$(-\frac{1}{2},2)$	$2\frac{1}{2}$ % 3 %	4 % 4 %	5 % 60 %	100 % 10^{99}
$(-1,1)$	$2\frac{1}{2}$ % 10^6	4 % 10^5	5 % 10^{20}	10^3 10^{99}

TABLE 2.

4.5 The Solution Of The Differential System.

In a manner similar to that of chapter 2, the solution of equation (1), subject to (1a), and (1b), may be obtained as the sum of the solutions of the three problems :

(a) (1),(1a),(1b), with $\phi_0 = \phi_1 = 0$;

(b) (1),(1a),(1b), with $f_1 = f_2 = 0$, and $\phi_1 = 0$;

(c) (1),(1a),(1b), with $f_1 = f_2 = 0$, and $\phi_0 = 0$.

When $p_2 = q_2 = 0$, problem (a) is solved as a Sturm-Liouville problem, and (b) and (c) are solved by means of the Laplace Transform method. The method for problems (b) and (c) follows similarly to themethod used in chapter 2 for the heat equation. Problem (a) is now solved as follows :

Let

$$u(x,t) = X(x)\,T(t).$$

Then

$$X'' = -\lambda^2 X\,, \tag{15}$$

and

$$T'' = -\lambda^2 T, \tag{16}$$

where $X(x)$ satisfies the boundary conditions

$$\left.\begin{aligned} X'(0) - p_1 X(0) &= 0 \\ X'(1) + q_1 X(1) &= 0. \end{aligned}\right\} \tag{17}$$

Equations (15) and (17) give a Sturm-Liouville problem for the eigenfunctions $X_\lambda(x)$, with associated eigenvalues λ. As in chapter 2, the eigenvalues are the roots of the equation

$$\cot\lambda = \frac{\lambda^2 - p_1 q_1}{\lambda(p_1 + q_1)} \tag{18}$$

and the solution $u(x,t)$ is then given by

$$u(x,t) = \sum_{\lambda} (A_{\lambda} \cos \lambda t + B_{\lambda} \sin \lambda t)(\lambda \cos \lambda x + p_1 \sin \lambda x), \quad (19)$$

th e values of the constants A_{λ}, and B_{λ}, being obtained from the initial conditions, using the orthogonality properties of the funstions X_{λ}, i.e.

$$f_1(x) = \sum_{\lambda} A_{\lambda} X_{\lambda}(x),$$

and

$$f_2(x) = \sum_{\lambda} B_{\lambda} \lambda X_{\lambda}(x).$$

Thus

$$A_{\lambda} = \int_0^1 f_1(x)\, X_{\lambda}(x)\, dx, \text{ and } B_{\lambda} = \int_0^1 f_2(x)\, X_{\lambda}(x)\, dx/\lambda .$$

The roots λ of equation (18) have been shown to be real, or pure imaginary, in chapter 2. Clearly, if there are any imaginary roots, then $u(x,t)$, given by equation (19), will be unbounded, as $t \to \infty$. Thus, the conditions for the solutions to be bounded are

$$\text{(i)} \quad p_1 q_1 + p_1 + q_1 \geq 0 ;$$

$$\text{(ii)} \quad p_1 + q_1 \geq 0 .$$

In problems (b) and (c), it can be shown that the same conditions are necessary for the solutions to be bounded.

When p_2 and q_2 are not both zero, the Sturm-Liouville method is not applicable. The following technique was developed, by Professor Copson, to solve problem (a).

The solution may be written in the form

$$u(x,t) = f(x+t) + g(t-x),$$

where $f(x)$ and $g(-x)$ are obtained from the initial conditions. The functions f and g satisfy the following differential equations, obtained by substituting $u(x,t)$ in the boundary conditions.

$$(1-p_2)\, f'(t) - (1+p_2)\, g'(t) - p_1\, f(t) - p_1\, g(t) \;=\; 0 \tag{20}$$

$$(1+q_2)\, f'(1+t) - (1-q_2)\, g'(t-1) + q_1 f(1+t) + q_1 g(t-1) \;=\; 0, \tag{21}$$

where ' denotes differentiation with respect to t. The following technique enables us to find the values of f and g on the boundary $x = 0$, and on $x = 1$, and hence the values of u there.

Let

$$F(s) \;=\; \int_0^\infty f(t)\, e^{-st}\, dt,$$

and

$$G(s) \;=\; \int_0^\infty g(t)\, e^{-st}\, dt.$$

Then

$$\int_0^\infty f'(t)\, e^{-st} dt = sF(s) - f(0),$$

provided $f(t)$ stays bounded, as $t \to \infty$; and, similarly,

$$\int_0^\infty g'(t)\, e^{-st} dt = sG(s) - g(0).$$

Equation (20) then gives

$$(1-p_2)(sF(s) - f(0)) - (1+p_2)(sG(s) - g(0)) - p_1F(s) - p_1G(s) \;=\; 0;$$

i.e.

$$[\, s(1-p_2) - p_1]\, F(s) - [\, s(1+p_2) + p_1]\, G(s) \;=\; \alpha(s), \tag{22}$$

where α is a function of s, which may be obtained from the initial conditions.

Rewriting equation (21) in the form :

$$(1+q_2)\, f'(t) - (1-q_2)\, g'(t-2) + q_1\, f(t) + q_1\, g(t-2) \;=\; 0,$$

and using the equation

$$\int_0^\infty e^{-st}\, g(t-2)\, dt \;=\; e^{-2s} \int_0^\infty e^{-st} g(t)\, dt + e^{-2s} \int_{-2}^0 e^{-st} g(t)\, dt$$

$$= e^{-2s} G + e^{-2s} A,$$

then

$$(1+q_2)(sF(s)-f(0)) - (1-q_2)[e^{-2s}(sG(s)-g(0)) + e^{-2s}B] + q_1F(s) + q_1[e^{-2s}G(s) + e^{-2s}A] = 0,$$

where A and B may be found from the initial conditions. Thus

$$[s(1+q_2) + q_1]\, F(s) - [s(1-q_2) - q_1]\, e^{-2s}\, G(s) = \beta(s), \tag{23}$$

where β is a known function of s. Equations (22) and (23), together, give

$$F(s) = \Phi(s)/y(s) \tag{24}$$

and

$$G(s) = \Psi(s)/y(s), \tag{25}$$

where the functions Φ, Ψ, y are easily calculated.

The functions f and g are now obtained by inversion of the Laplace transforms. By contour integration, the functions f and g are given as infinite series in the residues of the integrands, calculated at the poles of F and G (c.f. chapter 2). For f and g to be bounded, as $t \to \infty$, it is necessary, and sufficient, that all the poles of F and G should lie in $R(s) < 0.$; i.e. the roots of

$$y(s) = 0,$$

that is, of

$$e^{2s} = \frac{[s(1-p_2)-p_1][s(1-q_2)-q_1]}{[s(1+p_2)+p_1][s(1+q_2)+q_1]}, \tag{26}$$

should lie in $R(s) < 0$.

If the solution $u(x,t)$ is known on $x = 0$, and $x = 1$, then, by using the characteristics, here $x = \pm t$, the solution is obtained at any point in R, as a linear combination of thevalues on the boundary. If the solution is unbounded on the lines $x = 0$, and $x = 1$, then the solution is unbounded everywhere in R, as $t \to \infty$. Thus, the condition for

$u(x, t)$ to stay bounded, as $t \to \infty$, is $R(s) < 0$, where s is any root of equation (26).

4.6 The Routh-Hurwitz Problem For A Transcendental Equation.

Finding conditions on p_i, q_i, $i = 1,2$, so that $R(s) < 0$, for all roots of equation (26), is a Routh-Hurwitz problem for the transcendental equation. An obvious set of sufficient conditions for the roots of (26) to lie in $R(s) < 0$, is obtained by requiring the function

$$\chi(s) = \frac{[s(1-p_2)-p_1][s(1-q_2)-q_1]}{[s(1+p_2)+p_1][s(1+q_2)+q_1]},$$

to have modulus $\leqslant 1$, in $R(s) > 0$. For, when $R(s) > 0$, then $|e^{2s}| > 1$, and hence, the two functions χ, and e^{2s}, cannot intersect.

The condition $|\chi(s)| < 1$, in $R(s)>0$, is equivalent to the condition

$$4(p_2+q_2)(1+p_2q_2)(s\bar{s})^2 + 2[(1+p_2q_2)(p_1+q_1)+(p_2+q_2)(p_1q_2+p_2q_1)]s\bar{s}(s+\bar{s})+$$

$$+2p_1q_1(p_2+q_2)(s+\bar{s})^2 + 4(p_1^2q_2+q_1^2p_2)s\bar{s} + 2p_1q_1(p_1+q_1)(s+\bar{s}) \geqslant 0.$$

Since $R(s) > 0$, then $s+\bar{s} > 0$, and so a set of sufficient conditions is obtained by requiring the coefficients of $(s\bar{s})^2$, $s\bar{s}(s+\bar{s})$, etc., to be non-negative. Thus, it is sufficient if the following conditions are satisfied :

(i) $(p_2+q_2)(1+p_2q_2) \geqslant 0$;

(ii) $(1+p_2q_2)(p_1+q_1)+(p_2+q_2)(p_1q_2+p_2q_1) \geqslant 0$;

(iii) $p_1q_1(p_2+q_2) \geqslant 0$;

(iv) $p_1^2q_2+q_1^2p_2 \geqslant 0$;

(v) $p_1q_1(p_1+q_1) \geqslant 0$.

By (i), (p_2+q_2) has the same sign as $(1+p_2q_2)$; which, by (iii), has the same sign as p_1q_1; and which, in turn, has the same sign as (p_1+q_1), by (v). Thus, $(1+p_2q_2)(p_1+q_1) \geqslant 0$, and so (ii) may be replaced by the sufficient condition

$$(p_2+q_2)(p_1q_2+p_2q_1) \geq 0.$$

Thus, for equation (26) to have all its roots in $R(s) < 0$, it is sufficient that

$$(p_1+q_1),\ (p_2+q_2),\ p_1q_1,\ (1+p_2q_2),\ \text{and}\ (p_1q_2+p_2q_1),$$

should all have the same sign, and that

$$p_1^2q_2 + p_2q_1^2 \geq 0.$$

These conditions are strict ones, since, if $p_2 = q_2 = 0$, they reduce to p_1, and $q_1 \geq 0$, which are not necessary (see section 4.5).

In [4], Bellman and Cooke have considered the Hurwitz problem, for the equation (26), and have obtained sufficient conditions for the roots of equations of the form

$$H(s,e^s) = 0,$$

to lie in $R(s) < 0$, when H is a polynomial in s and e^s. In order to apply the results of [4], we write equation (26) in the form

$$[(p_2+q_2)s + (p_1+q_1)]\cosh s + [(1+p_2q_2)s^2 + (p_1q_2+p_2q_1)s + p_1q_1]\frac{\sinh s}{s} = 0. \quad (27)$$

The roots of the equation

$$(a_1s^2 + a_2s + a_3)\cosh s + (b_1s^2 + b_2s + b_3)\sinh s = 0,$$

have been discussed by V.N. Capyrin, [51], for the case when a_3 is not zero. This case is the one which occurs in (27). However, in [4], the following theorem is proved : (p. 443) :

In order that all the zeros of the function $H(s,e^s)$, a polynomial in s and e^s, should lie in $R(s) < 0$, it is sufficient that, if

$$H(ia,e^{ia}) = F(a) + iG(a),$$

then either

(i) all the zeros of $F(a)$ are real, and for each zero, a_0, of

$F(a)$, the condition

$$F'(a_0).G(a_0) \leq 0,$$

is satisfied; or

(ii) all the zeros of $G(a)$ are real, and for each zero, b_0, of $G(a)$, the condition

$$F(b_0).G'(b_0) \geq 0,$$

is satisfied.

Putting $s = ia$, in (27), we obtain

$$G(a) = (p_2+q_2)a \cos a + (p_1q_2+p_2q_1) \sin a \text{ , and}$$

$$F(a) = (p_1+q_1) \cos a + \frac{\sin a}{a}[p_1q_1 - (1+p_2q_2)a^2].$$

Then

$$G(a) = 0, \text{ when}$$

$$\tan a = \frac{-(p_2+q_2)a}{p_1q_2+p_2q_1} \qquad (28)$$

In section 2.8, it was shown that all the roots of

$$\tan \lambda = \lambda/(-p),$$

were real if, and only if,

$$p \geq -1.$$

Thus, the roots of equation (28) are real if, and only if,

$$\frac{+(p_1q_2+p_2q_1)}{-(p_2 + q_2)} \leq 1. \qquad (29)$$

The condition

$$G'(a).F(a) \geq 0,$$

at roots of $G(a) = 0$, gives the condition

$$[(p_2+q_2)(p_1q_2+p_2q_1)+(p_1q_2+p_2q_1)^2+(p_2+q_2)^2a^2][p_1^2q_2+p_2q_1^2+(1+p_2q_2)(p_2+q_2)a^2] \geq 0. \qquad (30)$$

Since $a = 0$, is a root of (28), and the other roots are $O(n\pi)$, then (30) gives the conditions

$$(p_2+q_2)(1+p_2q_2) > 0, \text{ and}$$

$$(p_1q_2+p_2q_1+p_2+q_2)(p_1q_2+p_2q_1)(p_1^2q_2+p_2q_1^2) > 0.$$

Thus, a set of sufficient conditions for the roots of (27) to lie in the negative real half-plane, is

$$(1) \qquad \frac{p_1q_2+p_2q_1}{p_2+q_2} \quad > \quad -1;$$

$$(2) \qquad (p_2+q_2)(1+p_2q_2) \quad > \quad 0;$$

$$(3) \quad (p_1q_2+p_2q_1+p_2+q_2)(p_1q_2+p_2q_1)(p_1^2q_2+p_2q_1^2) \quad > \quad 0.$$

Clearly, conditions (i) - (v), above, imply these conditions.

Recalling the necessary conditions for stability of the difference-differential system of section 4.3, i.e.

$$\begin{aligned} p_1q_1 + p_1 + q_1 &> 0; \\ p_1q_2 + p_2q_1 + p_2 + q_2 &> 0; \\ p_2 + q_2 &> 0; \\ 1 + p_2q_2 &> 0, \end{aligned} \qquad \text{(c)}$$

it is seen that the second and third conditions together imply condition (1), above. However, it is not the case that the sufficient conditions for stability of the <u>differential system</u>, imply the necessary conditions for stability of the <u>difference-differential system</u>, or vice-versa. For there are values of p_i, q_i, $i = 1,2$, satisfying conditions (1),(2),(3), but not (c); e.g.

$$p_1 = q_1 = 0, \quad (p_2+q_2)(1+p_2q_2) > 0;$$

and there are values which satisfy (c), but not (1),(2),(3); e.g.

$$p_1 = 0, \; q_1 = 1, \; p_2 = -1/4, \; q_2 = 1.$$

However, it should be noted that, if the further conditions

$$p_2 \geq 0, \ q_2 \geq 0,$$

are added to (c), then (1),(2),(3), are satisfied. The numerical results suggested that p_2, q_2 non-negative were necessary conditions.

It was noted that, when $p_2 = q_2 = 0$, the differential system had only bounded solutions for the same problems for which the difference-differential system was stable. But, as the above analysis shows, the difference-differential system, when p_2 and q_2 are not both zero, may be unstable, although the differential system has only bounded solutions. This is in contrast to the position for the heat equation, for which it was shown that the differential and difference-differential systems had only bounded solutions for the same problems.

As might be expected, the difference in stability properties, between the differential and difference-differential systems, disappears as $N \to \infty$. For, if, in the stability polynomial (13), we put

$$\beta^2 = 4 \sinh^2 s/(2N),$$

then

$$\beta^2 = -2(1 - \cos\theta),$$

where

$$\theta = \frac{s}{i\,N},$$

and so, using the properties of $T_N(-\beta^2)$ given in chapter 1,

$$T_{N-1}(\beta^2) = \frac{\sin N\theta}{\sin\theta} = \frac{\sinh s}{\sinh s/N}$$

and

$$T_N(-\beta^2) - T_{N-2}(-\beta^2) = 2 \cos N\theta = 2 \cosh s.$$

In addition, as $N \to \infty$, $N\beta \to s$, and $\beta \to 0$. Thus, rewriting (13) as the equation

$$4 \cosh s \left[(p_1+q_1) + s\frac{\sinh s/(2N)}{s/(2N)}(p_2+q_2)\right] +$$

$$\frac{\sinh s}{N \sinh s/N}\left[4p_1q_1+4N\beta(p_2q_1+p_1q_2) + 4(1+p_2q_2)N^2\beta^2 + N^2\beta^4\right]=0,$$

and, letting $N \to \infty$, we obtain the equation

$$\cosh s\,(p_1+q_1+ s\{p_2+q_2\}) + \frac{\sinh s}{s}(p_1q_1+ \{p_1q_2+p_2q_1\}s + (1+p_2q_2)s^2)=0,$$

which is equation (27). The condition $R(\beta) < 0$, in the limit as N tends to infinity, is thus equivalent to $R(s) < 0$. As $N \to \infty$, therefore, the stability properties of the difference-differential and differential systems are identical.

This makes the following fact seem more strange. No matter how large N is taken, the polynomial (13) still has roots in $R(\beta) > 0$, for values of p_i, q_i, $i = 1,2$, for which all roots of (27) are in $R(s) < 0$. A partial explanation of this fact is found in the paper [24], of K.E. Iverson. In this paper, polynomial approximations to the function e^s are considered, of order N, for $N < 23$. It is shown that, for all such N, the polynomials (which were the partial sums of the exponential series) have zeros in $R(s) > 0$, despite the fact that, in this region, $|e^s| > 1$. In addition, it is suggested that the partial sums of e^s will have roots in $R(s) > 0$, for all N however large. The equation (13) is, in fact, a polynomial approximation to the exponential function in equation (26). The occurrence of roots in $R(s) > 0$, when equation (26) has no roots there, may, therefore, be explained by the above remarks.

4.7 Conclusion.

The preceding study of the effect of the boundary conditions on solutions of the wave equation, clearly leaves some questions unanswered. Necessary and sufficient conditions, for the solutions of the differential system to be bounded, have yet to be obtained. It has been shown, however, that, if time derivatives appear in the boundary conditions, problems will arise which give bounded solutions of the differential system, but which cannot be solved by any discretisation method, of the type which uses the space derivative replacement discussed in this chapter. This is in sharp contrast to the position when time derivatives do not appear in the boundary conditions; then the class of boundary conditions, for which the differential equation has bounded solutions, coincides with the class of boundary conditions, for which the difference-differential system is stable.

5 Concluding Remarks.

In this thesis, initial-boundary value problems, and numerical approximations to them, have been discussed. It has been shown that the stability of a difference approximation to such a problem is affected as much by the boundary conditions, as by the method of approximation to the differential operator which is used. For the heat equation, and for a class of wave operators, the instability observed in the difference approximations, was shown to be caused by the problems' being improperly posed, in the sense defined in chapter 0. As far as these problems show, we can say that, if a difference approximation is unconditionally stable for one properly posed problem, then it is unconditionally stable for all boundary conditions for which the problem is properly posed. That this conclusion is not invariably correct, was shown by the example of the wave equation with boundary conditions involving time (i.e. tangential) derivatives - there exist properly posed problems for which the difference approximations are unconditionally unstable.

It has also been noted that the instability observed, is an asymptotic instability, i.e. as $t \to \infty$. For this reason, such instability becomes more important in the iterative solutions of equations arising from numerical approximations to, say, Laplace's equation. It was demonstrated, in chapter 3, that, for a certain class of boundary conditions, all the commonly used iterative methods, for the numerical solution of Laplace's equation, break down.

A fruitful area for further research, would be the extension of the results of this thesis, to equations with variable coefficients, and also to non-rectangular regions. The former extension is essentially similar in form to the case of constant coefficients, but, for non-rectangular regions, nothing has yet been attempted.

REFERENCES.

[1] Albasiny, E.L.: On the numerical solution of a cylindrical heat-conduction problem. Quarterly J. of Mech. and Appl. Math. 13, 374-384 (1960).

[2] Batten, G.W. : Second order correct boundary conditions for the numerical solution of the mixed boundary problem for parabolic equations. Math. of Comp. 17, 405-413 (1963).

[3] Bellman, R. : Stability theory of differential equations. Mc. Graw-Hill Inc., New York (1953).

[4] Bellman, R., and K.L. Cooke : Differential-difference equations. Academic Press, New York, (1963).

[5] Buchanan, M.L. : A necessary and sufficient condition for stability of difference schemes for second-order initial value problems. J. Soc. Indust. Appl. Math. 11, 474-501 (1963).

[6] Buchanan, M.L. : A necessary and sufficient condition for stability of difference schemes for initial value problems. J. Soc. Indust. Appl. Math. 11, 919-935 (1963).

[7] Brunings, J. : Discussion of iterative methods for solving certain mixed boundary value problems. Space Laberatory Report NN-71 (1957).

[8] Buchanan, M.L. : A condition for uniform Jordanisation of a family of triangular matrices. Num. Math. 6, 335-345 (1964).

[9] Copson, E.T., and P. Keast : On a boundary value problem for the equation of heat. J. Inst. Math. Applic. 2, 358-363 (1966).

[10] Courant, R., K.O. Friedrichs, and H. Lewy : Uber die partiellen differenzengleichungen der mathematischen physik. Math. Ann. 100, (1928). Translation by P. Fox, AEC Research and development Report, NYO-7689, Institute of Mathematical Sciences, New York University, (9156).

[11] Courant, R. and D. Hilbert : Methods of mathematical physics. Vol. 1, Interscience, New York (1953).

[12] Douglas, J. : On the relation between stability and convergence in the numerical solution of linear parabolic and hyperbolic differential equations. J. Soc. Indust. Appl. Math. 4, 20-37 (1956).

[13] Douglas, J., and C.M. Pearcy : On the convergence of alternating direction procedures, in the presence of singular operators. Numer. Math. 5, 175-184 (1963).

[14] Faddeeva, V.N. : Computational methods of linear algebra. Translated by C.D. Benster, Dover, New York (1959).

[15] Friedman, A. : Partial differential equations of parabolic type. Prentice-Hall, Englewood Cliff, New Jersey, (1964).

[16] Frjazinov, I.V. : On a difference approximation of the boundary conditions for the third boundary value problem. Z. Vycisl. Mat. Fiz. 4, 1106-1112 (1964).

[17] Gantmacher, F.R. : The theory of matrices. Chelsea Publishing Co., New York (1959).

[18] Garabedian, P.R. : Partial differential equations. J. Wiley and Sons, New York (1964).

[19] Gary, J. : A generalisation of the Lax-Richtmeyer theorem on finite difference schemes. J. S.I.A.M., Numer. Anal. 3, 467-473 (1966).

[20] Godunov, S.K., and V.S. Ryabenki : Theory of difference schemes. North Holland Publishing Co., Amsterdam, (1964).

[21] Gunn, J.F. : The solution of elliptic difference equations by semi-explicit iterative techniques. J. S.I.A.M., Numer. Anal. Ser.B, 2, 23-45 (1964).

[22] Halmos, P.R. : Finite dimensional vector spaces. D. van Nostrand Co., Princeton (1958).

[23] Hurwitz, A. : Uber die Bedingungen unter welchen eine Gleichung nur Wurzeln mit negativen reelen Theilen besitzt. Math. Ann., 46, 273-284 (1895).

[24] Iverson, K.E. : The zeros of the partial sums of e^z. M.T.A.C., 12, 165-168 (1958).

[25] Keast, P., and A.R. Mitchell : On the instability of the Crank-Nicolson formula under derivative boundary conditions. Comp. J. 9, 110-114 (1966).

[26] Keast, P., and A.R. Mitchell : Finite difference solution of the third boundary problem in elliptic and parabolic equations. Num. Math. (1967), (to appear).

[27] Kreiss, H.O. : Uber die losung des Cauchyproblems fur lineare partielle differentialgleichungen mit hilfe von differenzengleichungen. Acta Math. 101, 179-199 (1959).

[28] Kreiss, H.O. : Uber matrizen die beschranke halbgruppen erzengen. Math. Scand. 7, 71-80 (1959).

[29] Kreiss, H.O. : Uber die Losung von anfangsrandwertaufgaben fur partielle differentialgleichungen mit hilfe von differenzengleichungen. Trans. Royal Inst. of Tech., Stockholm, No. 166, (1960).

[30] Kreiss, H.O. : Uber die stabilitatsdefinition fur differenzengleichungen die partielle differentialgleichungen approximieren. B.I.T. 2, 152-181 (1962).

[31] Kreiss, H.O. : Uber implizite differenzmethoden fur partielle differentialgleichungen. Num. Math. 5, 24-47 (1963).

[32] Kreiss, H.O. : On difference approximations of the dissipative type for hyperbolic differential equations. Comm. Pure Appl. Math. 17, 335-353 (1964).

[33] Lax, P.D., and R.D. Richtmeyer : Survey of the stability of linear finite difference equations. Comm. Pure Appl. Math. 9, 267-293 (1956).

[34] Lebedev, V.I., and E.G. D'Jakonov : The application of difference schemes with splitting operator for the solution of the third boundary value problem for an equation of parabolic type. Sibirsk Mat. Z. 6, 108-113 (1965).

[35] Lowan, A.N. : The operator approach to problems of stability and convergence. Scripta Mathematica, New York (1957).

[36] Martin, D.W., and G.J. Tee : Iterative methods for linear equations with symmetric positive definite matrix. Comp. J., 4, (1961).

[37] O'Brien, G.G., M.A. Hyman, and S. Kaplan : A study of the numerical solution of partial differential equations. J. Math. Phys. 29, 223-251 (1951).

[38] Parker, I.B., and J. Crank : Persistent discretisation errors in partial differential equations of parabolic type. Comp. J. 7, 163-167 (1964).

[39] Peaceman, D.W., and H.H. Rachford : The numerical solution of parabolic and elliptic equations. J. Soc. Indust. Appl. Math. 3, 28-41 (1955).

[40] Pearson, C.E. : Impulsive end condition for diffusion equation. Sperry-Rand research centre report SRRC-RR-64-29 (1964).

[41] Richtmeyer, R.D. : Difference methods for initial value problems. Interscience Tracts in Pure and Applied Mathematics, Vol. 4, Interscience, New York (1957).

[42] Rose, M.E. : On the integration of non-linear parabolic differential equations by finite difference methods. Quart. Appl. Math. 14, 237-248 (1956).

[43] Rutherford, D.E. : Some continuant determinants arising in Physics and Chemistry. Proc. Roy. Soc. Edin. 62, 229-236 (1945).

[44] Rutherford, D.E. : Some continuant determinants arising in Physics and Chemistry - II. Proc. Roy. Soc. Edin., 63, 232-261 (1952).

[45] Schechter, S. : Quasi-tridiagonal matrices and type insensitive difference equations. Quart. Appl. Math., 18, 285-295 (1960).

[46] Todd, J. : A direct approach to the problem of stability in the numerical solution of partial differential equations. Comm. Pure Appl. Math., 9, 597-612 (1956).

[47] Varga, R.S. : Matrix iterative analysis. Prentice-Hall, Englewood Cliff, New Jersey (1962).

[48] Wachspress, E.L. : Extended application of alternating direction implicit iteration model problem theory. J. S.I.A.M., 11, 994-1015 (1963).

[49] Wachspress, E.L. : Iterative solution of elliptic systems. Prentice-Hall, Englewood Cliff, New Jersey (1966).

[50] Walsh, J. : Numerical Analysis : An introduction. Academic Press, London (1966).

[51] Capyrin, V.N. : On the problem of Hurwitz for transcendental equations. Prikl. Mat. Meh., 12, 301-328 (1948).