COMPUTER SCIENCE'S

REINCARNATION OF FINITE DIFFERENCES

by

Dana L. Ulery

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Applied Science (Computer Science).

June, 1976

COMPUTER SCIENCE'S

REINCARNATION OF FINITE DIFFERENCES

by

Dana L. Ulery

Approved: Hatem M. Khalil
Professor in charge of thesis on
behalf of the Advisory Committee

Approved: John W. Gries
Professor in charge of thesis on
behalf of the Advisory Committee

Approved:
Chairman of the Department of Statistics
and Computer Science

Approved:
Dean of the College of Graduate Studies

ACKNOWLEDGMENTS

Embodied in this work are the ideas, encouragements and enthusiasms of several people. Foremost among these is my advisor, Dr. Hatem M. Khalil, under whose guidance this work was undertaken and completed. His initial encouragement and confidence was a major factor in my decision to pursue this degree. He has never failed to be there when needed, always ready to share his uncanny ability to uncover the essence of a problem and determine a path leading to progress. It has been my unusual good fortune to find in him teacher, collaborator, and true friend, so that work has indeed been pleasure.

I am also most indebted to my coadvisor and friend, Dr. John H. Giese, and my husband, Dr. Harris E. Ulery. With patience and good humor, Dr. Giese has spent long hours in discussion, offering many suggestions and carefully checking details. He deserves credit for the technique used for the Method of Lines. In addition to great patience, despite late dinners and all sorts of other inconveniences, I am grateful to my husband for frequent and encouraging discussions concerning my work. His talent

for modifying standard techniques to fit new situations has provided me with fresh approaches to my problem on many occasions. I must also thank the members of my committee, and in particular Arthur E. Hoerl.

I wish to thank the Mathlab Group at M.I.T. for permission to use the MACSYMA system. The Mathlab Group is supported by the Advanced Research Agency (ARPA), Department of Defence, under Office of Naval Research Contract N0014-70-A-0362-0001.

This thesis is dedicated to my family: my mother Meriam; my husband Harris; and my dear children, Brad and Terrie. Without their support and cooperation this work could not have been done. I can only hope that they are as aware of my devotion to them as I am aware of their devotion to me.

TABLE OF CONTENTS

1. INTRODUCTION

Many scientific problems are formulated as partial differential equations. Since very few of these can be solved analytically, various techniques have been devised for obtaining approximate solutions [1,7,9,10]. Among the large number of numerical methods proposed for solving partial differential equations, the method of finite differences has particular importance because of its universal applicability to both linear and nonlinear problems. While the repetitive nature of the method makes it particularly well-suited for digital computations, time and space complexities are often quite large. The search for efficient finite difference methods has, therefore, been intensive over the past decades, resulting in the development of many schemes. However, until recently there has been no unified approach for generating and testing new difference schemes. Each formula has had to be considered individually with respect to properties such as accuracy, consistency, stability and convergence.

Recently, Khalil [5] proposed an algorithm for generating consistent families of difference approximations

that depend on several parameters. This was illustrated by deriving and analyzing a two-parameter, eighteen-point, two-level family of high-order approximations to the two-dimensional equation of heat flow in a polygonal region with Dirichlet boundary conditions. Later, Giese and Khalil [6] used inverse Vandermonde matrices to formalize the approach and derived a twelve-parameter family of eighteen-point, two-level approximations for the same equation. With this unifying approach, families can be analyzed in terms of their parameters to determine methods of increased efficiency.

A natural extension of their work is to generalize the approach to handle equations in arbitrary regions and with various boundary conditions. Embedding these generalizations in a multiparameter family of difference methods will undoubtedly be profitable, since it enables us to study the effects of various boundary conditions on the choice of parameters. Moreover, the algorithm thus developed will form the nucleus of software for solving partial differential equations.

The desired generalization was explored using the Vandermonde matrix formalization. This led to a numeric approach in which each specific case had to be handled individually, resulting in an increase of complexity in the

generation of families. Although the increase in complexity for a specific case is not significant, the number of cases to be considered grows rapidly with the number of spatial coordinates.

It therefore seemed advantageous to vary the approach in such a way as to minimize the complexity but still retain the technique of automatically generating consistent multiparameter families. A symbolic approach, based on a net of variable geometry, was selected for this purpose. In addition to generation parameters, explicit spatial displacement parameters are used to achieve non-uniformity of step size along the spatial axes. This device permits algorithmic generation of multiparameter families of difference schemes applicable without modification both in the interior and along the boundaries of arbitrary regions, thus minimizing complexity. In addition, such an approach permits apriori stability analysis which yields valuable information regarding the choice of generation parameters. Finally, this approach provides a unifying principle for independently developed methods: all the well-known two-level approximations, as well as the five and twelve-parameter families of Giese and Khalil, are embedded in the approximations derived here.

In Section 1 we develop and analyze a multiparameter

family of non-uniform difference methods for the one-dimensional equation of heat flow with Dirichlet boundary conditions.

Section 2 deals with the extension of the method to higher dimensions with Dirichlet boundary conditions. As a specific case, a ten-point, two-level multiparameter family is studied in detail.

Section 3 deals with the extensions of the proposed method to derivative boundary conditions.

2. A NON-UNIFORM MULTIPARAMETER FAMILY OF DIFFERENCE ANALOGUES TO $u_t = u_{xx}$

The simplest technique for solving a partial differential equation numerically is the finite difference method. Our aim is to automatically generate a non-uniform analogue to the heat operator which will serve as a nucleus in the development of computer software for the solution of this class of problems. To this end we seek the most general analogue. Consequently, our approach is to construct families of methods which depend on several parameters: spatial displacement parameters (SDP) and generation parameters (GP). Assignment of values to these parameters yields various methods which may be used as circumstances dictate.

To illustrate this idea we shall consider in this section the one-dimensional heat equation

$$u_t = u_{xx} \tag{2.1}$$

in an arbitrary region R of the (x,t) space with boundary ∂R, subject to initial conditions

$$u(x,0) = f(x), x \in R$$

and Dirichlet boundary conditions

$$u(x,t) = g(x,t), \quad x \in \partial R.$$

To solve this problem using finite differences, we cover the region $R \times \partial R$ by a lattice of discrete points and approximate $u(x,t)$ by the difference operator

$$\begin{aligned} L[u(x,t)] = {} & Au(x-\alpha_n h, t-\tfrac{1}{2}k) + Bu(x,t-\tfrac{1}{2}k) \\ & + Cu(x+\beta_n h, t-\tfrac{1}{2}k) + Du(x-\alpha_{n+1} h, t+\tfrac{1}{2}k) \\ & + Eu(x,t+\tfrac{1}{2}k) + Fu(x+\beta_{n+1} h, t+\tfrac{1}{2}k) \end{aligned} \tag{2.2}$$

where

$$h = \Delta x > 0, \quad k = \Delta t > 0, \quad \text{and} \quad 0 < \alpha, \beta \leq 1 . \tag{2.3}$$

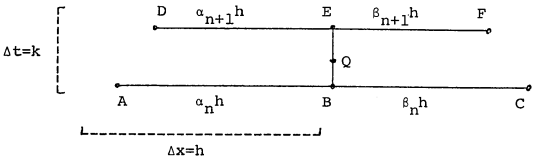
Here α_n, β_n, α_{n+1}, β_{n+1} denote the left and right SDP at the lower and upper time levels respectively.

Unlike familiar difference operators, (2.2) generates a non-uniform lattice.

2.1 The Fundamental Stencil

The set of coefficients $\mathcal{C} = \{A,B,C,D,E,F\}$ and center point $Q = (x,t)$ define the stencil $S(Q)$ associated with the difference operator (2.2). We say that this stencil is *conformable*, in that nodes $A,C,D,F \in \mathcal{C}$ can be positioned to lie directly on ∂R. The vehicle permitting this

conformability is the set A of SDP, $\{\alpha_n, \beta_n, \alpha_{n+1}, \beta_{n+1}\}$. The stencil is represented by the diagram below:



To compute the coefficients C of the difference operator (2.2), we expand L[u(x,t)] about the central point (x,t) as a Taylor's series in powers of h and k to obtain

$$L[u(x,t)] = \sum_{p,q=0}^{\infty} [(-\alpha_n)^p (-\tfrac{1}{2})^q A + (-\tfrac{1}{2})^q B + (\beta_n)^p (-\tfrac{1}{2})^q C + (-\alpha_{n+1})^p (\tfrac{1}{2})^q D + (\tfrac{1}{2})^q E + (\beta_{n+1})^p (\tfrac{1}{2})^q F] \cdot \frac{h^p k^q}{p!q!} \cdot \frac{\partial^{p+q} u}{\partial x^p \partial t^q} \qquad (2.4)$$

We select (p,q) pairs to include the explicit terms of the heat operator and force the u and u_x terms in (2.4) to vanish. The four conditions thus imposed for the (p,q) pairs (0,0), (1,0), (2,0) and (0,1) will restrict the choices of the coefficient set C to a two-GP, four-SDP

family. To uniquely define C, we adjoin two additional (p,q) pairs, (1,1) and (2,1), which contain the two desired generation parameters, ξ and η. This yields the system of linear symbolic equations:

$$Q\,x = f \tag{2.5}$$

where

$$Q = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ -\alpha_n & 0 & \beta_n & -\alpha_{n+1} & 0 & \beta_{n+1} \\ \alpha_n^2 & 0 & \beta_n^2 & \alpha_{n+1}^2 & 0 & \beta_{n+1}^2 \\ -\tfrac{1}{2} & -\tfrac{1}{2} & -\tfrac{1}{2} & \tfrac{1}{2} & \tfrac{1}{2} & \tfrac{1}{2} \\ \tfrac{1}{2}\alpha_n & 0 & -\tfrac{1}{2}\beta_n & -\tfrac{1}{2}\alpha_{n+1} & 0 & \tfrac{1}{2}\beta_{n+1} \\ -\tfrac{1}{2}\alpha_n^2 & 0 & -\tfrac{1}{2}\beta_n^2 & \tfrac{1}{2}\alpha_{n+1}^2 & 0 & \tfrac{1}{2}\beta_{n+1}^2 \end{bmatrix}$$

$$x = (A \quad B \quad C \quad D \quad E \quad F)^T,$$

$$f = (0 \quad 0 \quad -2h^{-2} \quad k^{-1} \quad \xi \quad \eta)^T$$

Thus, from (2.1), (2,2), and (2.5)

$$L[u(x,t)] = u_t - u_{xx} + \xi h u_{tx} + \eta h^2 k u_{txx} + 0(h^3) + 0(k^2) \tag{2.6}$$

We are, of course, ultimately concerned with solving the difference equations generated by $L[u(x,t)]$ by computer. We could store system (2.5) and numerically determine the values of the set of coefficients C for a given (ξ,η) pair and each new value assignment to the SDP

set A. This would involve the numerical solution of a 6×6 system of equations twice/sweep in the one-dimensional case; in the higher dimensional cases, the number of solutions required per sweep would increase greatly. We could design much more efficient software if we could solve for the coefficients explicitly. Moreover, an explicit solution would allow us to analytically investigate the effect of our GP on the stability, local accuracy, etc. of our difference methods. The task of solving the system of equations (2.5) by hand would be too tedious and error-prone. Thus we have a good candidate for solution by an algebraic symbol manipulating system. The problem was solved symbolically by the MACSYMA [8] system, yielding the following solution:

$$\begin{bmatrix} A \\ B \\ C \\ D \\ E \\ F \end{bmatrix} = \begin{bmatrix} -[\eta-\beta_n\xi+h^{-2}]/[\alpha_n\beta_n+\alpha_n^2] \\ [\eta+\xi(\alpha_n-\beta_n)-k^{-1}\alpha_n\beta_n+h^{-2}]/\alpha_n\beta_n \\ -[\eta+\alpha_n\xi+h^{-2}]/[\alpha_n\beta_n+\beta_n^2] \\ [\eta-\beta_{n+1}\xi-h^{-2}]/[\alpha_{n+1}\beta_{n+1}+\alpha_{n+1}^2] \\ -[\eta+\xi(\alpha_{n+1}-\beta_{n+1})-k^{-1}\alpha_{n+1}\beta_{n+1}-h^{-2}]/\alpha_{n+1}\beta_{n+1} \\ [\eta+\alpha_{n+1}\xi-h^{-2}]/[\alpha_{n+1}\beta_{n+1}+\beta_{n+1}^2] \end{bmatrix} \qquad (2.7)$$

At present, the inversion of symbolic matrices of even this order of magnitude can be troublesome. The manipulation

can be simplified by taking advantage of the fact that the matrix can be partitioned in the form

$$Q = \begin{bmatrix} V_1 & V_2 \\ -\frac{1}{2}V_1 & \frac{1}{2}V_2 \end{bmatrix}$$

where V_1 and V_2 are Vandermonde matrices of order 3, and whose inverses are easily computed [6]. Then

$$Q^{-1} = \begin{bmatrix} \frac{1}{2}V_1^{-1} & -V_1^{-1} \\ \frac{1}{2}V_2^{-1} & V_2^{-1} \end{bmatrix}$$

2.2 Matrix Formulation

Let $u_i^n = u(x_i,t_n)$, where $0 \leq i \leq N+1$, $n \geq 0$. We approximate u_i^n by solutions U_i^n of the linear system

$$M_{n+1}U^{n+1} = M_nU^n + Z^n \qquad (2.8)$$

where $U^s = (U_1^s, U_2^s, \ldots, U_N^s)^T$ and $s=n, n+1$.

M_{n+1} and M_n are square tridiagonal matrices of order N whose elements are functions of the SDP and GP, and are given by

$$M_{n+1} = \begin{bmatrix} E & E & & & \\ D & E & F & & \\ & D & E & F & \\ & & D & E & \\ & & & & \end{bmatrix} \qquad M_n = -\begin{bmatrix} B & C & & & \\ A & B & C & & \\ & A & B & C & \\ & & A & B & \\ & & & & \end{bmatrix} \tag{2.9}$$

Z^n is a vector of N components involving the boundary conditions.

We can express (2.9) more conveniently as the weighted sum of simpler matrices I, W_n, W_{n+1}, Y_n, and Y_{n+1}:

$$\begin{aligned} M_{n+1} &= k^{-1}I + \tfrac{1}{2}(h^{-2}-\eta)W_{n+1} - \tfrac{1}{2}\xi Y_{n+1} \\ M_n &= k^{-1}I - \tfrac{1}{2}(h^{-2}+\eta)W_n - \tfrac{1}{2}\xi Y_n \end{aligned} \tag{2.10}$$

Let s=n, n+1, and $1 \le i, j \le N$. Then the elements in the i^{th} row and j^{th} column of W_s and Y_s are:

$$W_s = \begin{cases} 2/\alpha_s\beta_s & \text{if } i = j \\ -2/[\alpha_s(\alpha_s+\beta_s)] & \text{if } i-j = 1 \\ -2/[\beta_s(\alpha_s+\beta_s)] & \text{if } j-i = 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.11}$$

$$Y_s = \begin{cases} [2(\alpha_s-\beta_s)]/\alpha_s\beta_s & \text{if } i = j \\ 2\beta_s/[\alpha_s(\alpha_s+\beta_s)] & \text{if } i-j = 1 \\ -2\alpha_s/[\beta_s(\alpha_s+\beta_s)] & \text{if } j-i = 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.12}$$

I denotes the N×N unit matrix.

As usual, in the interior of the region R, we will work with a uniform net. Moreover, for $N \geq 2$, it is reasonable to assign the value one to β_n and β_{n+1} in row 1 (at the left-hand boundary) and to α_n and α_{n+1} in row N (the right-hand boundary). Then (2.11) and (2.12) have the forms

$$W_s = \begin{bmatrix} 2/\alpha_s & -2/(1+\alpha_s) & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots \\ & & & -2/(1+\beta_s) & 2/\beta_s \end{bmatrix} \tag{2.13}$$

$$Y_s = \begin{bmatrix} 2(\alpha_s-1)/\alpha_s & 2\alpha_s/(1+\alpha_s) & & & \\ 1 & 0 & -1 & & \\ & 1 & 0 & -1 & \\ & & \ddots & \ddots & \ddots \\ & & & 2\beta_s/(\beta_s+1) & 2(1-\beta_s)/\beta_s \end{bmatrix} \tag{2.14}$$

For future use, we shall need some information about the eigenvalues of W_s and Y_s. First, we note that $\tilde{W}_s = H_s^{-1} W_s H_s$ is symmetric, where

$$H_s = \begin{cases} \sqrt{2/(1+\alpha_s)} & \text{if } i = j = 1 \\ \sqrt{2/(1+\beta_s)} & \text{if } i = j = N \\ \delta_{ij} & \text{otherwise} \end{cases} \qquad (2.15)$$

s=n, n+1 and δ_{ij} denotes the Kronecker delta of rank 2. Hence, the eigenvalues of $\tilde{W}_s$, and therefore those of W_s, are real.

The matrix W_s is irreducible, with diagonal dominance assured for rows 2 through N-1. Strict inequality occurs for rows 1 and N when

$$\left|\frac{2}{\alpha_s}\right| > \left|\frac{-2}{1+\alpha_s}\right| \quad \text{and} \quad \left|\frac{2}{\beta_s}\right| > \left|\frac{-2}{1+\beta_s}\right| , \qquad (2.16)$$

respectively. By (2.3), these conditions are always met. Thus, the eigenvalues μ_s of W_s are always positive. Using Gerschgoren's Theorem [11] to derive an upper bound, we find

$$0 < \mu_s \leq \mu_s^* \qquad (2.17)$$

$$\text{where } \mu_s^* = \max\left[4, \frac{2(2\alpha_s+1)}{\alpha_s(\alpha_s+1)}, \frac{2(2\beta_s+1)}{\beta_s(\beta_s+1)}\right] . \qquad (2.18)$$

The matrix Y_s is asymmetric, with eigenvalues occurring as imaginary pairs and zero.

2.3 Stability

If M_{n+1} is nonsingular, (2.8) implies

$$U^{n+1} = M_{n+1}^{-1} M_n U^n + M_{n+1}^{-1} Z^n \quad . \tag{2.19}$$

A necessary and sufficient condition for stability is that the spectral radius ρ of the amplification matrix $M_{n+1}^{-1} M_n$ be less than or equal to unity [3].

Because of the asymmetry of the matrix Y_s, it is quite difficult to obtain a good estimate for $\rho\,(M_{n+1}^{-1} M_n)$. We can, however, obtain a reasonable estimate by setting $\xi=0$. This has the effect of setting A=C and D=F in the interior of R. Let us further simplify the analysis by letting $\sigma = \min[\alpha_s, \beta_s]$, and let W, Y, H, and μ^* denote W_s, Y_s, H_s, and μ_s^*, respectively, with α_s and β_s replaced by σ.

Then

$$M_{n+1}^{-1} M_n = [k^{-1} I + \tfrac{1}{2}(h^{-2} - \eta) W]^{-1} [k^{-1} I - \tfrac{1}{2}(h^{-2} + \eta) W] \tag{2.20}$$

By straightforward substitution we find

$$\Lambda = \frac{k^{-1} - \tfrac{1}{2}(h^{-2} + \eta)\mu}{k^{-1} + \tfrac{1}{2}(h^{-2} - \eta)\mu} \tag{2.21}$$

where $\Lambda \epsilon M$, the set of eigenvalues of $M_{n+1}^{-1} M_n$ and $\mu \epsilon W$, the

set of eigenvalues of W.

Moreover, using (2.16) we see that

$$\begin{aligned}(\widetilde{M_{n+1}^{-1}M_n}) &= H^{-1}(M_{n+1}^{-1}M_n)H \\ &= H^{-1}[k^{-1}I+\tfrac{1}{2}(h^{-2}-\eta)W]^{-1}[k^{-1}I-\tfrac{1}{2}(h^{-2}+\eta)W]H \\ &= \{H^{-1}[k^{-1}I+\tfrac{1}{2}(h^{-2}-\eta)W]^{-1}H\}\{H^{-1}[k^{-1}I-\tfrac{1}{2}(h^{-2}+\eta)W]H\} \\ &= \{H^{-1}[k^{-1}I+\tfrac{1}{2}(h^{-2}-\eta)W]H\}^{-1}\{H^{-1}[k^{-1}I-\tfrac{1}{2}(h^{-2}+\eta)W]H\} \\ &= [k^{-1}I+\tfrac{1}{2}(h^{-2}-\eta)\tilde{W}]^{-1}[k^{-1}I-\tfrac{1}{2}(h^{-2}+\eta)\tilde{W}]. \qquad (2.22)\end{aligned}$$

Since matrices $[k^{-1}I+\frac{1}{2}(h^{-2}-\eta)\tilde{W}]^{-1}$ and $[k^{-1}I-\frac{1}{2}(h^{-2}+\eta)\tilde{W}]$ are symmetric and commute, $(\widetilde{M_{n+1}^{-1}M_n})$ is symmetric. Thus $(M_{n+1}^{-1}M_n)$ is similar to a symmetric matrix so that

$$\rho(M_{n+1}^{-1}M_n) \equiv \rho(\widetilde{M_{n+1}^{-1}M_n}) \leq 1 \qquad (2.23)$$

is a necessary and sufficient condition for stability. We insist then that $|\Lambda| \leq 1$ for all $\Lambda \in M$, which leads to the following conditions:

$$\text{(i)} \quad \Lambda \leq 1 \qquad \text{if } -h^{-2}\mu \leq 0, \qquad (2.24)$$

$$\text{(ii)} \quad \Lambda \geq -1 \qquad \text{if } \eta\mu \leq 2k^{-1}. \qquad (2.25)$$

Condition (i) is trivially satisfied since h and μ are always positive. From (2.17) we find that a necessary and sufficient condition for the unconditional stability of the family defined by (2.2), (2.7), and $\xi=0$, is that η and k

satisfy the following condition:

$$\eta \leq \frac{2}{k\mu^*} , \tag{2.26}$$

where

$$\mu^* = \begin{cases} \dfrac{2(2\sigma+1)}{\sigma(\sigma+1)} & \text{for } 0 < \sigma < \sqrt{1/2} \\ 4 & \text{for } \sqrt{1/2} \leq \sigma \leq 1 \end{cases}$$

In the limit, this condition takes the form

$$\lim_{\sigma \to 0} \eta \leq 0 , \tag{2.27}$$

which assures unconditional stability independent of the values of the SDP.

If we restrict the region to the usual rectangular shape, we can set all the SDP to unity to achieve a uniform net throughout R. The necessary and sufficient condition for stability for arbitrary η and k with ξ=0, then, is given by

$$\eta \leq \tfrac{1}{2}k^{-1} . \tag{2.28}$$

Since $M_{n+1}^{-1}M_n$ is a continuous rational function of the GP, we can expect that its eigenvalues will also be continuous functions of the GP. This being the case, we can generalize our results to the extent of predicting that methods satisfying the stability criterion (2.26) will also be unconditionally stable for sufficiently small

values of $|\xi|$. By Lax's equivalence theorem [3], immediately applicable for the single equation (2.1), stability of our difference methods is equivalent to the convergence of U_i^n to $u(x,t)$ in R.

2.4 Generation of Standard Two-Level Difference Operators

The familiar two-level approximations [9,10] displayed in Table I, can be obtained from our method by specializing the generation parameters ξ and η after setting all SDP to unity. The local accuracy and stability criteria of schemes in which $\xi=0$ is immediately apparent from (2.6) and (2.28).

TABLE I: Familiar Two-Level Difference Operators

Scheme	ξ	η	A	B	C	D	E	F	Stability Conditions
Classical Explicit	0	h^{-2}	$-h^{-2}$	$2h^{-2}-k^{-1}$	$-h^{-2}$	0	k^{-1}	0	$kh^{-2} \leq \frac{1}{2}$
Classical Implicit	0	$-h^{-2}$	0	$-k^{-1}$	0	$-h^{-2}$	$2h^{-2}+k^{-1}$	$-h^{-2}$	none
Crank-Nicolson	0	0	$-\frac{1}{2}h^{-2}$	$-k^{-1}+h^{-2}$	$-\frac{1}{2}h^{-2}$	$-\frac{1}{2}h^{-2}$	$k^{-1}+h^{-2}$	$-\frac{1}{2}h^{-2}$	none
One-Parameter	0	$h^{-2}(1-2\theta)$	$(\theta-1)h^{-2}$	$-k^{-1}+2(1-\theta)h^{-2}$	$(\theta-1)h^{-2}$	$-\theta h^{-2}$	$k^{-1}+2\theta h^{-2}$	$-\theta h^{-2}$	$kh^{-2}(1-2\theta) \leq \frac{1}{2}$
Asym-metric	h^{-2}	0	0	$-k^{-1}+h^{-2}$	$-h^{-2}$	$-h^{-2}$	$k^{-1}+h^{-2}$	0	none, if used on alternate time steps
	$-h^{-2}$	0	$-h^{-2}$	$-k^{-1}+h^{-2}$	0	0	$k^{-1}+h^{-2}$	$-h^{-2}$	

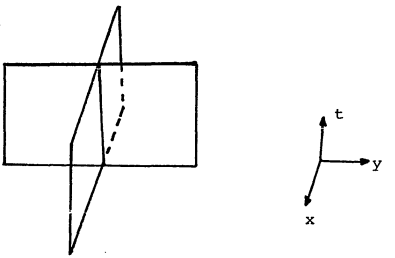
3. EQUATIONS OF HIGHER DIMENSIONS

In this section we consider a technique for using the conformable stencil developed for the one-dimensional heat equation to build lattices for higher order heat equations subject to Dirichlet boundary conditions in an arbitrary region. Since most of the additional complications arise immediately with the addition of one more space dimension, we content ourselves with considering only the two-dimensional case in detail. The further extension to three or more space dimensions appears to be straightforward.

3.1 Extension of Method to Higher Dimensions

A natural extension of the approach developed for the one-dimensional case can be thought of in geometric terms as the intersection of two conformable stencils. The stencil for the one-dimensional heat equation consisted of six points in the xt-plane; in two dimensions it is natural to combine a six-point stencil in a plane normal to the y-axis with another six-point stencil in a plane normal to the x-axis, insisting that the centered points coincide.

This results in a ten-point stencil, illustrated below:



The difference operator associated with this stencil can be formed as the algebraic sum of two judiciously modified one-dimensional operators. Let $L_x[u]$ and $L_y[u]$ be the operators associated with the stencils in the planes normal to the x- and y-axes, respectively. Instead of approximating the standard heat equation, we approximate $\frac{1}{2}u_t - u_{xx}$ by $L_x[u]$ and $\frac{1}{2}u_t - u_{yy}$ by $L_y[u]$. We can easily specify these operators by making a slight change in the right-hand vector of the system of equations defining the stencil coefficients (2.5). The component k^{-1} in the vector f is replaced by $\frac{1}{2}k^{-1}$. The effect of this replacement is a simple change of variable in the solution vector given by (2.7). The difference operator associated with the ten-point stencil, $L_{10}[u]$, can then be written as

$$L_{10}[u] = L_x[u] + L_y[u]$$

$$= u_t - u_{xx} - u_{yy} + \xi_x hku_{tx} + \eta_x h^2 ku_{txx}$$

$$+ \xi_y hku_{ty} + \eta_y h^2 ku_{tyy} + 0(h^3) + 0(k^2)$$

$$= \sum A_{a,b}(x \pm ah,\ y \pm bh,\ t \pm \tfrac{1}{2}k) \qquad (3.1)$$

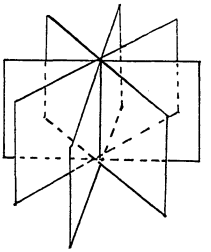
where $a = 0,\ \alpha_n^x,\ \alpha_{n+1}^x,\ \beta_n^x,\ \beta_{n+1}^x$

$b = 0,\ \alpha_n^y,\ \alpha_{n+1}^y,\ \beta_n^y,\ \beta_{n+1}^y$

$0 < \alpha,\ \beta \leq 1.$

Thus, $L_{10}[u]$ defines a two-level, ten-point, four-GP, eight-SDP family of difference approximations to the two-dimensional heat equation whose coefficients, $A_{a,b}$, are explicitly known.

Extending this approach further, we combine two pairs of mutually perpendicular conformable stencils--one pair oriented coincident with the coordinate axes, the second pair rotated 45° about the t-axis--to produce an eighteen-point stencil:



Let $L_x^{\phi}[u]$ and $L_y^{\phi}[u]$ be the operators associated with the rotated stencil-pair. Since each pair of stencils approximates the two-dimensional heat-equation, we define our eighteen-point operator as

$$L_{18}[u] = \tfrac{1}{2}[L_x+L_x^{\phi}+L_y+L_y^{\phi}] \ . \qquad (3.2)$$

The operator $L_{18}[u]$ defines a two-level, eighteen-point, eight-GP, sixteen-SDP family of difference approximations.

In general, the intersection of n conformable stencils in the (x,y,t) space leads to a variety of two-level, (4n+2)-point, 2n-GP, 4n-SDP family of finite difference approximations to the two-dimensional heat operator. The coefficients of the difference operators can be obtained in explicit form, after some manipulation, from (2.7).

3.2 Analysis of a Ten-point Multiparameter Family for $u_t = u_{xx} + u_{yy}$

We consider the two-dimensional heat equation

$$u_t = u_{xx} + u_{yy} \tag{3.3}$$

in an arbitrary region R of the (x,y,t) space with boundary ∂R, subject to initial conditions

$$u(x,y,0) = f(x,y), \quad x,y \in R$$

and the boundary conditions

$$u(x,y,t) = g(x,y,t), \quad x,y \in \partial R.$$

3.2.1 Matrix Formulation

Let $u^n_{i,j} = u(x_i, y_j, t_n)$, where $i,j,n \geq 0$. We approximate $u^n_{i,j}$ by solutions $U^n_{i,j}$ of the linear system

$$M_{n+1} U^{n+1} = M_n U^n + Z^n . \tag{3.4}$$

U^n and U^{n+1} are N-dimension vectors; Z^n is an N-dimension vector involving boundary conditions; and M_{n+1} and M_n are square block tridiagonal matrices of order N, whose elements are the coefficients of the difference operator (3.1).

We anticipate from Section 2.2 that the matrices

M_{n+1} and M_n can be expressed as the weighted sum of simpler matrices, I_N, $\bar{\bar{W}}_n$, $\bar{\bar{W}}_{n+1}$, $\bar{\bar{Y}}_n$, and $\bar{\bar{Y}}_{n+1}$, each block tridiagonal and of order N. Let $\bar{\bar{W}}_s(I,J;i,j)$ and $\bar{\bar{Y}}_s(I,J;i,j)$ denote the i^{th} row and j^{th} column of the (I,J) block of $\bar{\bar{W}}_s$ and $\bar{\bar{Y}}_s$, respectively, where s=n, n+1. Then the entries of $\bar{\bar{W}}_s$ and $\bar{\bar{Y}}_s$ are shown in the table below:

		$\bar{\bar{W}}_s$	$\bar{\bar{Y}}_s$
I=J	i=j	$\frac{2}{\alpha_s^x\beta_s^x} + \frac{2}{\alpha_s^y\beta_s^y}$	$\frac{2(\alpha_s^x+\beta_s^x)}{\alpha_s^x\beta_s^x} + \frac{2(\alpha_s^y+\beta_s^y)}{\alpha_s^y\beta_s^y}$
	i-j=1	$-\frac{2}{\alpha_s^y(\alpha_s^y+\beta_s^y)}$	$\frac{2\beta_s^y}{\alpha_s^y(\alpha_s^y+\beta_s^y)}$
	j-i=1	$-\frac{2}{\beta_s^y(\alpha_s^y+\beta_s^y)}$	$\frac{2\alpha_s^y}{\beta_s^y(\alpha_s^y+\beta_s^y)}$
	else	0	0
I-J=1	i=j	$-\frac{2}{\alpha_s^x(\alpha_s^x+\beta_s^x)}$	$\frac{2\beta_s^x}{\alpha_s^x(\alpha_s^x+\beta_s^x)}$
	else	0	0
J-I=1	i=j	$-\frac{2}{\beta_s^x(\alpha_s^x+\beta_s^x)}$	$\frac{2\alpha_s^x}{\beta_s^x(\alpha_s^x+\beta_s^x)}$
	else	0	0

Separating components contributed by the operators $L_x[u]$

and $L_y[u]$, we can express $\underline{\overline{W}}_s$ and $\underline{\overline{Y}}_s$ as

$$\begin{aligned} \underline{\overline{W}}_s &= W_s^x + W_s^y \\ \underline{\overline{Y}}_s &= Y_s^x + Y_s^y \end{aligned} \tag{3.5}$$

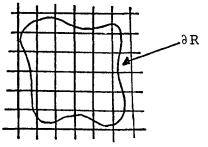
We can now express M_{n+1} and M_n as

$$\begin{aligned} M_{n+1} &= k^{-1}I_N + \tfrac{1}{2}(h^{-2}-\eta_x)W_{n+1}^x - \tfrac{1}{2}\xi_x Y_{n+1}^x + \tfrac{1}{2}(h^{-2}-\eta_y)W_{n+1}^x \\ &\quad - \tfrac{1}{2}\xi_y Y_{n+1}^y \end{aligned} \tag{3.6}$$

$$M_n = k^{-1}I_N - \tfrac{1}{2}(h^{-2}+\eta_x)W_n^x - \xi_x Y_n^x - \tfrac{1}{2}(h^{-2}+\eta_y)W_n^y - \tfrac{1}{2}\xi_y Y_n^y .$$

3.2.2 Stability Analysis

For purposes of analysis, we shall assume that the step size is chosen in such a way that each block in M_{n+1} and M_n is of equal size, i.e., ∂R has the form illustrated below:



Furthermore, let $\sigma = \min[\alpha,\beta]$, and let W_x, W_y, Y_x, and Y_y denote W_s^x, W_s^y, Y_s^x, and Y_s^y, respectively, with α_s^x, α_s^y, β_s^x, β_s^y replaced by σ.

We define two $\sqrt{N}$-order matrices W and Y of the form

$$W = \frac{1}{\sigma^2}\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \qquad Y = \frac{1}{\sigma}\begin{bmatrix} 0 & -1 & & & \\ 1 & 0 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & -1 \\ & & & 1 & 0 \end{bmatrix} \tag{3.7}$$

Then

$$\begin{aligned} W_x &= W \otimes I \\ W_y &= I \otimes W \\ Y_x &= Y \otimes I \\ Y_y &= I \otimes Y \end{aligned} \tag{3.8}$$

where I denotes the unit matrix of order $\sqrt{N}$ and the symbol $\otimes$ signifies the Kronecker product.

A necessary and sufficient condition for stability is that

$$\rho(M_{n+1}^{-1} M_n) \leq 1 . \tag{3.9}$$

As in the one-dimensional case, it is necessary to set $\xi_x = \xi_y = 0$ in order to obtain a reasonable estimate for ρ. Then

$$\begin{aligned} M_{n+1} &= k^{-1} I_N + \tfrac{1}{2}(h^{-2} - \eta_x) W_x + \tfrac{1}{2}(h^{-2} - \eta_y) W_y \\ M_n &= k^{-1} I_N - \tfrac{1}{2}(h^{-2} + \eta_x) W_x - \tfrac{1}{2}(h^{-2} + \eta_y) W_y . \end{aligned} \tag{3.10}$$

Since

$$W_x W_y = (W \otimes I)(I \otimes W) = W \otimes W = W_y W_x \qquad (3.11)$$

the matrices possess a common set of orthonormal eigenvectors ν with the corresponding eigenvalues given by

$$W_x \nu = \mu \nu$$

$$W_y \nu = \lambda \nu \ .$$

It is clear from (3.11) that

$$M_{n+1} M_n = M_n M_{n+1} \ , \qquad (3.12)$$

which implies that

$$M_n M_{n+1}^{-1} = M_{n+1}^{-1} M_n \ . \qquad (3.13)$$

Hence, using Frobenius' lemma [11], we can express the eigenvalues Λ of $M_{n+1}^{-1} M_n$ in the form

$$\Lambda = \frac{k^{-1} - \tfrac{1}{2}(h^{-2} + \eta_x)\mu - \tfrac{1}{2}(h^{-2} + \eta_y)\lambda}{k^{-1} + \tfrac{1}{2}(h^{-2} - \eta_x)\mu + \tfrac{1}{2}(h^{-2} - \eta_y)\lambda} \ . \qquad (3.14)$$

Both M_{n+1}^{-1} and M_n are symmetric, since W is symmetric. From (3.13), we see that $M_{n+1}^{-1} M_n$ is also symmetric, so that the condition

$$|\Lambda| \leq 1 \qquad \text{for all } \Lambda, \mu, \text{ and } \lambda \qquad (3.15)$$

is a necessary and sufficient condition for the stability of (3.4). This leads to the following conditions:

$$\text{(i)} \quad \Lambda \leq 1 \quad \text{if} \quad -h^{-2}(\mu+\lambda) \leq 0 \tag{3.16}$$

$$\text{(ii)} \quad \Lambda \geq -1 \quad \text{if} \quad \mu\eta_x + \lambda\eta_y \leq 2k^{-1} . \tag{3.17}$$

Using Gerschgorin's Theorem, we find

$$0 < \mu, \lambda \leq 4/\sigma^2 . \tag{3.18}$$

Since h, μ, and λ are always positive, condition (i) is trivially satisfied. The most restrictive bound for condition (ii) is then

$$\eta_x + \eta_y \leq \frac{\sigma^2}{2k} . \tag{3.19}$$

In the limit, this condition takes the form

$$\lim_{\sigma \to 0} \eta_x + \eta_y \leq 0 , \tag{3.20}$$

which assures unconditional stability independent of the values of α and β. Again, by the argument of continuity, we can expect condition (3.19) to be necessary and sufficient for the unconditional stability of methods specified by sufficiently small values of $|\xi_x|$ and $|\xi_y|$.

3.2.3 Special Cases

3.2.3.1 Generation of Standard Two-Level Difference Operators

The familiar two-level approximations, displayed in Table II, are special cases of the subclass of the family (3.6) with $\sigma = 1$. The local accuracy and stability criteria of each scheme is immediately apparent from (3.1) and (3.19).

3.2.3.2 Split-Formula Schemes

We now consider a variant of (3.10) to derive a multiparameter family of ADI methods. Let

$$M_{n+1} = [k^{-1}I_N + \tfrac{1}{2}(h^{-2} - \eta_x)W_x][k^{-1}I_N + \tfrac{1}{2}(h^{-2} - \eta_y)W_y]$$
$$M_n = [k^{-1}I_N - \tfrac{1}{2}(h^{-2} + \eta_x)W_x][k^{-1}I_N - \tfrac{1}{2}(h^{-2} + \eta_y)W_y] \quad . \tag{3.21}$$

The split formulas

$$[k^{-1}I_N + \tfrac{1}{2}(h^{-2} - \eta_x)W_x]U^{n+1*} = [k^{-1}I_N - \tfrac{1}{2}(h^{-2} + \eta_y)W_y]U^n$$
$$[k^{-1}I_N + \tfrac{1}{2}(h^{-2} - \eta_y)W_y]U^{n+1} = [k^{-1}I_N - \tfrac{1}{2}(h^{-2} + \eta_x)W_x]U^{n+1*} \tag{3.22}$$

are obtained from (3.21).

Since the matrices W_x and W_y are symmetric and commute, (3.11), it can be shown that each of the factors comprising M_{n+1}^{-1} and M_n are symmetric and possess a common

TABLE II: Familiar Cases

Scheme	ξ_x	ξ_y	η_x	η_y	M_{n+1}	M_n	Stability
Explicit	0	0	h^{-2}	h^{-2}	$k^{-1}I_N$	$k^{-1}I_N - h^{-2}(W_x+W_y)$	$kh^{-2} \leq \frac{1}{4}$
Crank-Nicholson	0	0	0	0	$k^{-1}I_N + \frac{1}{2}h^{-2}(W_x+W_y)$	$k^{-1}I_N - \frac{1}{2}h^{-2}(W_x+W_y)$	none

set of orthogonal eigenvectors. Hence, we can express the eigenvalues Λ of $M_{n+1}^{-1}M_n$, from (3.21), in the form

$$\Lambda = \frac{[k^{-1}-\frac{1}{2}(h^{-2}+\eta_x)\mu][k^{-1}-\frac{1}{2}(h^{-2}+\eta_y)\lambda]}{[k^{-1}+\frac{1}{2}(h^{-2}-\eta_x)\mu][k^{-1}+\frac{1}{2}(h^{-2}-\eta_y)\lambda]} \tag{3.22}$$

where μ and λ are the eigenvalues of W_x and W_y, respectively. From (3.18), $0 < \mu,\lambda \leq 4/\sigma^2$. A necessary and sufficient condition for the stability of this family of ADI methods is that

$$|\Lambda| \leq 1 \quad \text{for all } \Lambda,\ \mu,\ \lambda\ . \tag{3.23}$$

This leads to the following conditions:

$$\text{(i)} \quad \Lambda \leq 1 \text{ if } \eta_x + \eta_y \leq \frac{\sigma^2}{k} \tag{3.24}$$

$$\text{(ii)} \quad \Lambda \geq 1 \text{ if } \eta_x + \eta_y - \frac{2k}{\sigma^2}\eta_x\eta_y \leq \tfrac{1}{2}k^{-1}\sigma^2 + \frac{2k}{h^4\sigma^2}\ . \tag{3.25}$$

Note that if we choose $\eta_x \leq 0$ and $\eta_y \leq 0$, the method will be unconditionally stable, independent of boundary irregularities.

Special cases are:

Peaceman-Rachford Method: $\sigma = 1$, $\eta_x = \eta_y = 0$

Mitchell-Fairweather Method: $\sigma = 1$, $\eta_x = \eta_y = -1/6$.

From (3.24) and (3.25) it is clear that both these methods are unconditionally stable.

3.3 An Eighteen-Point Multiparameter Family for $u_t = u_{xx} + u_{yy}$

Extending the techniques demonstrated for the ten-point operator to the eighteen-point operator (3.2), we define

$$\begin{aligned} \overline{W} &= 3W_x + W_y + W_x^{\phi} + W_y^{\phi} \\ \overline{Y} &= Y_x + Y_y + Y_x^{\phi} + Y_y^{\phi} \end{aligned} \tag{3.26}$$

where

$$\begin{aligned} W_x &= W \otimes I \\ W_y &= I \otimes W \\ W_x^{\phi} &= E \otimes W \\ W_y^{\phi} &= E^* \otimes W \\ Y_x &= Y \otimes I \\ Y_y &= I \otimes Y \\ Y_x^{\phi} &= E \otimes Y \\ Y_y^{\phi} &= E^* \otimes Y \end{aligned}$$

Matrices W, Y, and I are defined by (3.7). E is the forward shift operator defined by the $\sqrt{N}$-order matrix

$$E = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \sim & & & 0 \end{bmatrix}$$

E^* is the backward shift operator defined by the $\sqrt{N}$-order matrix

$$E^* = \begin{bmatrix} 0 & & & \\ & \ddots & & \\ 1 & \ddots & \ddots & \\ \sim & & 1 & 0 \end{bmatrix} .$$

The difference scheme is then defined by the matrices

$$M_{n+1} = k^{-1}I_N + 3/2(h^{-2}-\eta_x)W_x - \tfrac{1}{2}\xi_x Y_x + \tfrac{1}{2}(h^{-2}-\eta_y)W_y$$
$$- \tfrac{1}{2}\xi_y Y_y + \tfrac{1}{2}(h^{-2}-\eta_x^{\phi})W_x^{\phi} - \tfrac{1}{2}\xi_x^{\phi}Y_x^{\phi} + \tfrac{1}{2}(h^{-2}-\eta_y^{\phi})W_y^{\phi}$$
$$- \tfrac{1}{2}\xi_y^{\phi}Y_y^{\phi}$$

$$M_n = k^{-1}I_N - 3/2(h^{-2}+\eta_x)W_x - \tfrac{1}{2}\xi_x Y_x - \tfrac{1}{2}(h^{-2}+\eta_y)W_y$$
$$- \tfrac{1}{2}\xi_y Y_y - \tfrac{1}{2}(h^{-2}+\eta_x^{\phi})W_x^{\phi} - \tfrac{1}{2}\xi_x^{\phi}Y_x^{\phi} - \tfrac{1}{2}(h^{-2}+\eta_y^{\phi})W_y^{\phi}$$
$$- \tfrac{1}{2}\xi_y^{\phi}Y_y^{\phi} .$$

3.4 Introduction of Extra Parameters

The following technique can be used to increase the number of GP associated with operators for the two-

dimensional heat equation. Let a two-level, ten-point conformable stencil $L[u(x,y,t)]$ be the fundamental operator, where

$$L[u] = L_x[u] + L_y[u] + \theta_x u_x + \theta_y u_y + \theta_{xy} u_{xy} , \tag{3.28}$$

where $L_x[u]$ and $L_y[u]$ are the operators introduced in Section 3.1 and θ_x, θ_y, and θ_{xy} are additional GP. If we choose to approximate the heat equation with a ten-point operator of $O(h^3)$, as in (3.1), we require that

$$\theta_x u_x = \theta_y u_y = \theta_{xy} u_{xy} = 0$$

and no new GP have been added. However, if we choose to combine at least two of these new fundamental operators to construct an eighteen-point operator, we have

$$\begin{aligned} L_{18}[u] &= \tfrac{1}{2}\{L[u] + L^{\phi}[u]\} \\ &= \tfrac{1}{2}\{L_x[u] + L_y[u] + L_x^{\phi}[u] + L_y^{\phi}[u] + (\theta_x+\theta_x^{\phi})u_x \\ &\quad + (\theta_y+\theta_y^{\phi})u_y + (\theta_{xy}+\theta_{xy}^{\phi})u_{xy}\} , \end{aligned} \tag{3.29}$$

in which six additional GP have been introduced. Imposing the three constraints necessary to assure that our operator is $O(h^3)$ leaves us with an operator with three more GP than we had for $L_{18}[u]$. This new operator thus has a total of eleven GP, and is equivalent to the eleven-parameter operator discussed by Giese and Khalil [6].

In general, combining n fundamental operators $L[u]$ results in a two-level, (8n+2)-point, 8n - SDP, (7n-3)-GP family of difference operators. Moreover, it is obvious that judicious selection of fundamental operators can lead to multiparameter families with varying characteristics, such as three-level families.

4. EXTENSION TO DERIVATIVE BOUNDARY CONDITIONS

The multiparameter family of difference methods introduced in Section 2 for the solution of the first initial boundary value problem for the one-dimensional heat equation will be extended in this section to the third initial boundary value problem in an arbitrary region. The Neumann, or second, initial boundary value problem is considered as a particular case.

Consider the one-dimensional heat equation

$$u_t = u_{xx} \tag{4.1}$$

in an arbitrary region R of the (x,t) space with boundary ∂R, subject to initial conditions

$$u(x,0) = f(x), \quad x \in R \tag{4.2}$$

and known boundary conditions

$$p(x)u(x,t) + \frac{\partial u(x,t)}{\partial x} = g(x,t), \quad x \in \partial R \tag{4.3}$$

4.1 Difference Analogue

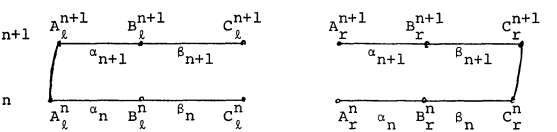
Since our fundamental operator (2.2) can achieve

local accuracy $O(h^3)$ for certain parameter selections, we would like the approximation to the normal derivative on the boundary to have the same order of accuracy. To accomplish this, we must use at least three points to approximate the derivative at point Q on ∂R for one time level. We overlay the conformable stencil so that the model conforms to the irregular boundary (either left or right) at both time levels. At time s=n, n+1 we choose to use points $u^s_{i-\alpha_s}$, u^s_i, and $u^s_{i+\beta_s}$ to approximate $\partial u^s_{i-\alpha_s}/\partial x$. We thus define the difference operators $L^s_\ell[u]$ and $L^s_r[u]$ to approximate the normal derivative on the left and right boundaries, respectively, as

$$L^s_d[u] = A^s_d u^s_{i-\alpha_s} + B^s_d u^s_i + C^s_d u^s_{i+\beta_s} \qquad (4.4)$$

where d = ℓ,r.

These ideas are illustrated below



Using the technique described in Section 2, we expand the operators $L_{\ell}^{s}[u]$ in a Taylor's Series about the point $(x-\alpha_s h,t)$ and select terms which will include the first derivative and assure an error $0(h^3)$. This leads to the system of linear symbolic equations

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & \alpha_s & \alpha_s+\beta_s \\ 0 & \tfrac{1}{2}\alpha_s^2 & \tfrac{1}{2}(\alpha_s+\beta_s)^2 \end{bmatrix} \begin{bmatrix} A_{\ell}^{s} \\ B_{\ell}^{s} \\ C_{\ell}^{s} \end{bmatrix} = \begin{bmatrix} 0 \\ h^{-1} \\ 0 \end{bmatrix} \tag{4.5}$$

which uniquely determines the coefficients as

$$\begin{bmatrix} A_{\ell}^{s} \\ B_{\ell}^{s} \\ C_{\ell}^{s} \end{bmatrix} = \begin{bmatrix} \dfrac{\alpha_s^2 - (\alpha_s+\beta_s)^2}{\alpha_s\beta_s(\alpha_s+\beta_s)}\, h^{-1} \\ \dfrac{\alpha_s + \beta_s}{\alpha_s\beta_s}\, h^{-1} \\ -\dfrac{\alpha_s^2}{\alpha_s\beta_s(\alpha_s+\beta_s)}\, h^{-1} \end{bmatrix} \tag{4.6}$$

Similarly, at the right-hand boundary we expand about the point $(x+\beta_s h,t)$, which leads to the system

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & -\alpha_s & -(\alpha_s+\beta_s) \\ 0 & \tfrac{1}{2}\alpha_s^2 & \tfrac{1}{2}(\alpha_s+\beta_s)^2 \end{bmatrix} \begin{bmatrix} C_{r}^{s} \\ B_{r}^{s} \\ A_{r}^{s} \end{bmatrix} = \begin{bmatrix} 0 \\ h^{-1} \\ 0 \end{bmatrix} \tag{4.7}$$

and the solution

$$\begin{bmatrix} A_r^s \\ B_r^s \\ C_r^s \end{bmatrix} = \begin{bmatrix} \dfrac{\alpha_s^2}{\alpha_s\beta_s(\alpha_s+\beta_s)}\, h^{-1} \\ -\dfrac{(\alpha_s+\beta_s)}{\alpha_s\beta_s}\, h^{-1} \\ -\dfrac{[\alpha_s^2-(\alpha_s+\beta_s)^2]}{\alpha_s\beta_s(\alpha_s+\beta_s)}\, h^{-1} \end{bmatrix} \tag{4.8}$$

4.2 Matrix Formulation

Replacing the normal derivative in equation (4.3) by the appropriate operator (4.4) determines grid points on the boundaries in terms of the two interior stencil points:

$$\begin{aligned} u^s_{i-\alpha_s} &= \frac{g^s_{i-\alpha_s} - B^s_\ell u^s_i - C^s_\ell u^s_{i+\beta_s}}{P^s_{i-\alpha_s} + A^s_\ell} \\ u^s_{i+\beta_s} &= \frac{g^s_{i+\beta_s} - B^s_r u^s_i - C^s_r u^s_{i-\alpha_s}}{P^s_{i+\beta_s} + A^s_r} \end{aligned} \tag{4.9}$$

Substituting these values into our basic operator L[u], (2.2), yields the linear system

$$M_{n+1}U^{n+1} = M_nU^n + Z^n \tag{4.10}$$

whose solutions U_i^n approximate u_i^n. U^n and U^{n+1} are

solution vectors of order N. Z^n is a vector of order N involving boundary conditions, whose i^{th} element is given by

$$Z^n = -\begin{cases} \dfrac{A\, g^n_{i-\alpha_n}}{p^n_{i-\alpha_n} + A^n_\ell} + \dfrac{D\, g^{n+1}_{i-\alpha_{n+1}}}{p^{n+1}_{i-\alpha_{n+1}} + A^{n+1}_\ell} & \text{if } i = 1 \\ \dfrac{C\, g^n_{i+\beta_n}}{p^n_{i+\beta_s} + C^n_r} + \dfrac{F\, g^{n+1}_{i+\beta_{n+1}}}{p^{n+1}_{i+\beta_s} + C^{n+1}_r} & \text{if } i = N \\ 0 & \text{otherwise .} \end{cases} \tag{4.11}$$

The elements in the i^{th} row and j^{th} column ($1 \le i,j \le N$) of the tridiagonal matrices M_n and M_{n+1} are:

$$M_n = -\begin{cases} B - \dfrac{A\, B^n_\ell}{p^n_{i-\alpha_n} + A^n_\ell} & \text{if } i = j = 1 \\ C - \dfrac{A\, C^n_\ell}{p^n_{i-\alpha_n} + A^n_\ell} & \text{if } i = 1,\ j = 2 \\ \left.\begin{array}{ll} A & \text{if } j = i-1 \\ B & \text{if } j = i \\ C & \text{if } j = i+1 \end{array}\right\} & 2 \le i \le N-1 \\ A - \dfrac{C\, A^n_r}{p^n_{i+\beta_n} + C^n_r} & \text{if } i = N,\ j = N-1 \end{cases} \tag{4.12}$$

$$
\begin{cases}
B - \dfrac{C\, B_r^n}{p_{i+\beta_n}^n + C_r^n} & \text{if } i = j = N \\
0 & \text{otherwise}
\end{cases}
$$

$$
M_{n+1} = \begin{cases}
E - \dfrac{D\, B_\ell^{n+1}}{p_{i-\alpha_{n+1}}^{n+1} + A_\ell^{n+1}} & \text{if } i = j = 1 \\
F - \dfrac{D\, C_\ell^{n+1}}{p_{i-\alpha_{n+1}}^{n+1} + A_\ell^{n+1}} & \text{if } i = 1,\ j = 2 \\
D & \text{if } j = i-1 \\
E & \text{if } j = i \quad \Big\} \ 2 \le i \le N-1 \\
F & \text{if } j = i+1 \\
D - \dfrac{F\, A_r^{n+1}}{p_{i+\beta_{n+1}}^{n+1} + C_r^{n+1}} & \text{if } i = N,\ j = N-1 \\
E - \dfrac{F\, B_r^{n+1}}{p_{i+\beta_{n+1}}^{n+1} + C_r^{n+1}} & \text{if } i = j = N \\
0 & \text{otherwise}
\end{cases}
\tag{4.13}
$$

As in Section 2, we assign the value one to β_s in row 1 (the left boundary), α_s in row N (the right boundary), and both α_s and β_s in rows 2 to N-1. Thus, the net is uniform in the interior of region R. As usual, we split M_n and M_{n+1} so that they have the form

$$M_{n+1} = k^{-1}I + \tfrac{1}{2}(h^{-2}{}_{n})W_{n+1} - \tfrac{1}{2}\xi Y_{n+1}$$

$$M_{n} = k^{-1}I - \tfrac{1}{2}(h^{-2}{}_{n})W_{n} - \tfrac{1}{2}\xi Y_{n} \quad . \tag{4.14}$$

Matrices W_s and Y_s have the same form as in the case of Dirichlet boundaries (2.13, 2.14), except for the elements in the first and last rows, given as

$$W_s = \begin{cases} 2(y_s-\alpha_s)/\alpha_s(y_s-2\alpha_s-1) & \text{if } i = j = 1 \\ 2(-y_s+\alpha_s+1)/(\alpha_s+1)(y_s-2\alpha_s-1) & \text{if } i = 1,\ j = 2 \\ -2(q_s+\beta_s+1)/(\beta_s+1)(q_s+2\beta_s+1) & \text{if } i=N,\ j=N-1 \\ 2(q_s+\beta_s)/\beta_s(q_s+2\beta_s+1) & \text{if } i = j = N \end{cases} \tag{4.15}$$

$$Y_s = \begin{cases} [2y_s(\alpha_s-1)-4\alpha_s^2]/\alpha_s(y_s-2\alpha_s-1) & \text{if } i = j = 1 \\ 2[-\alpha_s y_s+2\alpha_s(\alpha_s+1)]/[(\alpha_s+1)(y_s-2\alpha_s-1)] & \text{if } i = 1,\ j = 2 \\ 2\beta_s(q_s+2\beta_s+2)/[(1+\beta_s)(q_s+2\beta_s+1)] & \text{if } i=N,\ j=N-1 \\ 2[q_s(1-\beta_s)-2\beta_s^2]/\beta_s(q_s+2\beta_s+1) & \text{if } i = j = N \end{cases} \tag{4.16}$$

where

$$y_s = p^{s}_{i-\alpha_s}\ h\ \alpha_s(\alpha_s+1)$$

$$q_s = p^{s}_{i+\beta_s}\ h\ \beta_s(\beta_s+1)$$

and s=n, n+1.

Let us now consider the properties of W_s for use in further analysis. It is clear that, as before, W_s is a real, irreducible matrix and similar to a symmetric matrix

$\tilde{W}_s = H_s^{-1} W_s H_s$, where

$$H_s = \begin{cases} \sqrt{W_s(1,2)} & \text{If } i = j = 1 \\ \sqrt{W_s(N,N-1)} & \text{if } i = j = N \\ \delta_{ij} & \text{otherwise} \end{cases} \qquad (4.17)$$

s=n, n+1 and δ_{ij} denotes the Kronecker delta of rank 2.

Thus the eigenvalues μ_s of W_s are real.

Diagonal dominance is assured for rows 2 through N-1. Diagonal dominance occurs for rows 1 and N when

$$|W_s(1,1)| \geq |W_s(1,2)| \qquad (4.18)$$

and

$$|W_s(N,N)| \geq |W_s(N,N-1)| \qquad (4.19)$$

When strict inequality is achieved for at least one of these rows, with equality achieved for all others, the eigenvalues of W_s are always greater than zero. An upper bound μ_s^* can be computed using Gerschgorin's Theorem.

4.3 Stability

As before, we simplify the analysis by setting $\xi=0$ and $\sigma = \min[\alpha_s, \beta_s]$, and let W and μ^* denote W_s and μ_s^*, respectively. From (2.23) the necessary and sufficient

condition for stability is that $|\Lambda| \leq 1$ for all $\Lambda \epsilon M$, the set of eigenvalues of $M_{n+1}^{-1} M_n$, where Λ is given by (2.21). This restriction leads to conditions (2.24) and (2.25).

4.3.1 The Neumann Problem

Let us consider the special case where $p(x) = 0$ for all x. This constitutes the Neumann Problem. For this case, we rewrite the conditions for diagonal dominance, (4.18) and (4.19), as

$$\left|\frac{2}{2\sigma+1}\right| \geq \left|\frac{-2}{2\sigma+1}\right| \quad \text{and} \quad \left|\frac{2}{2\sigma+1}\right| \geq \left|\frac{-2}{2\sigma+1}\right| \tag{4.20}$$

Thus we have diagonal dominance in all rows, but in no row do we have strict diagonal dominance. Thus, the eigenvalues μ of W satisfy,

$$0 \leq \mu \leq 4 \tag{4.21}$$

(note: since $0 < \sigma \leq 1$, $4 = \max\left[4, \frac{4}{2\sigma+1}\right]$).

Consider what happens to Λ when $\mu = 0$; from equation (2.21) we find that $\Lambda = 1$. In this case, the difference family is stable but shows the presence of persistent error [7]. For $0 < \mu \leq 4$, stability depends upon our choice of η. By (2.24) and (2.25), the stability condition is that

$$\eta \leq \tfrac{1}{2}k^{-1} \quad . \tag{4.22}$$

In the Neumann Problem then, stability is unaffected by the values of the SDP; moreover the restriction on η in an arbitrary region is exactly the same as that found for a rectilinear region with Dirichlet boundaries (2.28).

4.3.2 The Third Boundary Value Problem in a Rectilinear Region

In the special case where the region is rectilinear, all SDP are set equal to unity and the region is covered by a uniform net. It is this special case that is generally referred to in studies of the third boundary value problem. Under these circumstances, the conditions for diagonal dominance (4.18) and (4.19) become:

$$|2p_0h-1| \geq |1-p_0h| \tag{4.23}$$

and

$$|2p_{N+1}h+1| \geq |1+p_{N+1}h| \quad . \tag{4.24}$$

From (4.23) and (4.24), W will have strict diagonal dominance in at least one row except when

$$0 \leq p_0h \leq 2/3 \quad \text{and/or} \quad -2/3 \leq p_{N+1}h \leq 0 \ . \tag{4.25}$$

Thus, for all values of p_0h and $p_{N+1}h$ outside the intervals (4.25), the eigenvalues μ of W lie within the interval

$$0 < \mu \leq \mu^* \tag{4.26}$$

where

$$\mu^* = \max\left[4,\ 4\left|\frac{2p_0h-1}{2p_0h-3}\right|,\ 4\left|\frac{2p_{N+1}h+1}{2p_{N+1}h+3}\right|\right] \quad .$$

Note that μ^* is unbounded at $p_0h = 3/2$ and $p_{N+1}h = -3/2$.

By conditions (2.24) and (2.25), the difference operator (4.14) will be stable when

$$\eta' \leq \frac{2}{k\mu^*} \tag{4.27}$$

and p_0h and $p_{N+1}h$ lie outside the intervals (4.25). For negative values of p_0 and positive values of p_{N+1}, the proper choice of η can define an unconditionally stable scheme. Otherwise, assuming that η satisfies (4.27), stability is conditional as indicated below:

$$\begin{aligned}
&\text{if } p_0 > 0 \text{ and } p_{N+1} > 0,\ h \leq \frac{2}{3p_0} \\
&\text{if } p_0 < 0 \text{ and } p_{N+1} < 0,\ h \leq \left|\frac{2}{3p_{N+1}}\right| \\
&\text{if } p_0 > 0 \text{ and } p_{N+1} < 0,\ h \leq \min\left[\frac{2}{3p_0},\ \left|\frac{2}{3p_{N+1}}\right|\right] .
\end{aligned} \tag{4.28}$$

4.3.3 The Third Boundary Value Problem in an Arbitrary Region

In the general case where the shape of the region is unrestricted, it is more difficult to determine directly conditions for diagonal dominance of W (4.18) and

(4.19). We define P_i as the set of all $p_i h$, where i=0, N+1, and S as the set of minimum values of the SDP, σ. We wish to determine the relations

$$f_i : P_i \to S \tag{4.29}$$

which defines the value pairs $(p_i h, \sigma)$ for which conditions (4.18) and (4.19) are NOT satisfied. From (4.15) it is apparent that f_i are 'multi-valued functions' and thus not amenable to analytic solution. We therefore take a heuristic approach. Using an iterative technique, we find the restricted regions Rf_0 and Rf_{N+1} within which conditions (4.17) and (4.19), respectively, are not satisfied. The regions are shown in Figures 1 and 2.

The stability conditions (2.24) and (2.25) are thus satisfied when

$$\eta \leq \frac{2}{k\mu^*} \quad \text{and} \quad (p_i h, \sigma) \notin Rf_i, \quad i=0, N+1 \tag{4.30}$$

where $\mu^* = \max [4, |W(1,1)| + |W(1,2)|, |W(N,N)| + |W(N,N-1)|]$.

In the limit as $\sigma \to 0$, μ^* becomes unbounded. However, the choice $\eta \leq 0$ assures stability for this class of difference operators outside the restricted regions Rf_i.

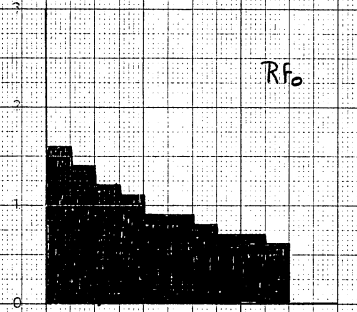


Figure 1

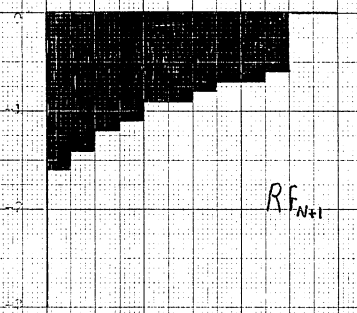


Figure 2

5. THE METHOD OF LINES

We have been concerned with determining discrete approximations $U(x_j,t_n)$ to solutions $u(x_j,t_n)$ of the heat equation at the points (x_j,t_n) of a discrete lattice over an arbitrary region $R\times\partial R$. We restrict our attention in this section to the one-dimensional heat equation in a rectangular region. By taking the limit, $k\to 0$, we can generate semi-discrete approximations $U(x_j,t)$ to $u(x_j,t)$ on the lines $x=x_j$ in terms of ordinary derivatives of U. This process is called the method of lines.

We set all displacement parameters to unity and cover the region $R\times\partial R$ by a uniform lattice. As before, we approximate $u(x,t)$ by the difference operator

$$L[U(x,t)] = 0$$

defined by (2.2) and (2.6). Let us split our parameters η and ξ so that

$$\begin{aligned} \eta &= \eta_1 + \eta_2 k^{-1} \\ \xi &= \xi_1 + \xi_2 k^{-1} \ . \end{aligned} \tag{5.1}$$

Rewriting our coefficients (2.7) in terms of these new parameters, we get

$$\begin{bmatrix} A \\ B \\ C \\ D \\ E \\ F \end{bmatrix} = \begin{bmatrix} -\tfrac{1}{2}(\eta_1+\eta_2k^{-1}-\xi_1-\xi_2k^{-1}+h^{-2}) \\ (\eta_1+\eta_2k^{-1}-k^{-1}+h^{-2}) \\ -\tfrac{1}{2}(\eta_1+\eta_2k^{-1}+\xi_1+\xi_2k^{-1}+h^{-2}) \\ \tfrac{1}{2}(\eta_1+\eta_2k^{-1}-\xi_1-\xi_2k^{-1}-h^{-2}) \\ -(\eta_1+\eta_2k^{-1}-k^{-1}-h^{-2}) \\ \tfrac{1}{2}(\eta_1+\eta_2k^{-1}+\xi_1+\xi_2k^{-1}-h^{-2}) \end{bmatrix} \tag{5.2}$$

We now write L[U] in terms of the coefficients defined by (5.2):

$$\begin{aligned} L[U(x_j,t_n)] &= \tfrac{1}{2}(\eta_1k+\eta_2-\xi_1k-\xi_2)(1/k)(U^{n+1}_{j-1}-U^n_{j-1}) \\ &-\tfrac{1}{2}h^{-2}(U^{n+1}_{j-1}+U^n_{j-1})-(\eta_1k+\eta_2-1)(1/k)(U^{n+1}_j-U^n_j) \\ &+h^{-2}(U^{n+1}_j+U^n_j)+\tfrac{1}{2}(\eta_1k+\eta_2+\xi_1k+\xi_2)(1/k) \\ &(U^{n+1}_{j+1}-U^n_{j+1})-\tfrac{1}{2}h^{-2}(U^{n+1}_{j+1}+U^n_{j+1}) = 0 \end{aligned} \tag{5.3}$$

Now

$$\lim_{k\to 0} L[U] = 0$$

yields

$$\tfrac{1}{2}(\eta_2-\xi_2)\,\frac{dU(x-h,t)}{dt} - (\eta_2-1)\,\frac{dU(x,t)}{dt} + \tfrac{1}{2}(\eta_2+\xi_2)\,\frac{dU(x+h,t)}{dt} - h^{-2}[U(x-h,t)-2U(x,t)+U(x+h,t)] = 0 \,. \tag{5.4}$$

For $\eta_2 = \xi_2 = 0$,

$$\frac{dU(x,t)}{dt} = h^{-2}[U(x-h,t)-2U(x,t)+U(x+h,t)] \,, \tag{5.5}$$

the equation most commonly used to formulate the method of lines for the one-dimensional heat equation. For $\eta_2 = \pm\xi_2$, we obtain a system of equations that can readily be solved for $\frac{dU(x,t)}{dt}$:

$$\begin{aligned}(1-\eta_2)\,\frac{dU(x,t)}{dt} &= h^{-2}[U(x-h,t)-2U(x,t)+U(x+h,t)] \\ &\quad - \eta_2\,\frac{dU(x+h,t)}{dt} \\ (1-\eta_2)\,\frac{dU(x,t)}{dt} &= h^{-2}[U(x-h,t)-2U(x,t)+U(x+h,t)] \\ &\quad + \eta_2\,\frac{dU(x-h,t)}{dt}\end{aligned} \tag{5.6}$$

These equations are analogues of asymmetrical difference operators which become explicit, in effect, for boundary value problems. Finally, when $\eta_2 \neq \pm\xi_2$, (5.4) yields a system of linear equations for $dU(x_j,t)$ which has a tri-diagonal coefficient matrix.

6. CONCLUSIONS

The use of a conformable stencil to algorithmically generate families of difference approximations to the first, second and third boundary-value problems for the heat equation has led to some useful results:

1. a procedure for constructing multiparameter families of difference schemes which contain boundary information for regions of arbitrary shape;
2. a simple method for determining the conditions for stability of each family symbolically, in terms of generative parameters and spatial displacement parameters;
3. a unified account of existing two-level difference approximations to the one- and two-dimensional heat operator.

We have demonstrated these techniques in detail for:

1. a two-level, six-point, two-GP, four-SDP family of approximations to the one-dimensional heat operator with Dirichlet boundary conditions in an

arbitrary region;

2. a two-level, ten-point, four-GP, eight-SDP family of approximations to the two-dimensional heat operator with Dirichlet boundaries in an arbitrary region;
3. a two-level, ten-point, four-GP, eight-SDP family of ADI methods for the two-dimensional heat operator with Dirichlet boundary conditions in an arbitrary region;
4. a two-level, six-point, two-GP, four-SDP family of approximations to the one-dimensional heat operator with derivative boundary conditions in an arbitrary region.

Necessary and sufficient conditions for the stability of each of these families was determined symbolically in terms of their parameters; it was shown that proper choices of the generative parameters assured stability independent of boundary conditions. The well-known corresponding difference schemes were shown to be a subclass of these multiparameter families.

A slight modification of the difference scheme for the one-dimensional heat operator in a rectangular region was shown to produce a family for the method of lines.

Two factors make the technique demonstrated here particularly suitable to software design. First, the coefficients are given as explicit symbolic expressions. Secondly, the basic stencil is used as a building block to construct finite-difference families in terms of simply expressed matrices. Thus we have developed a fairly simple algorithm to serve as the core for software to solve partial differential equations.

Finally, the ease with which stability analysis can be performed should encourage exploration of these families to discover optimal schemes.

BIBLIOGRAPHY

1. Douglas, J., "A Survey of Numerical Methods for Parabolic Differential Equations," Advances in Computers," Vol. 2, Academic Press, New York, N. Y., 1961.

2. Giese, J. H. and H. M. Khalil, "A General Method for Generating Consistent Difference Operators," Egyptian Comp. J., Vol. 1, No. 2, 1973.

3. Isaacson, E. and H. B. Keller, Analysis of Numerical Methods, John Wiley and Sons, New York, N. Y., 1967.

4. Keast, P. and A. R. Mitchell, "Finite Difference Solution of the Third Boundary Problem in Elliptic and Parabolic Equations," Numer. Math., Vol. 10, 1967.

5. Khalil, H. M., "A Two-Parameter Family of Approximations to the Two-Dimensional Heat Equation," Ph.D. Dissertation, University of Delaware, Newark, Delaware, 1969.

6. Khalil, H. M. and J. H. Giese, "A Two-Parameter Family of Two-Level Higher Order Difference Methods for the Two-Dimensional Heat Equation," J.I.M.A., Vol. 15, 1975.

7. Mitchell, A. R., Computational Methods in Partial Differential Equations, John Wiley and Sons, New York, N. Y., 1969.

8. Moses, J., "MACSYMA--The Fifth Year," Proc. of Eurosam, Royal Inst. of Tech., Sweden, 1974.

9. Richtmyer, R. D. and K. W. Morton, Difference Methods for Initial-Value Problems, Interscience Press, New York, N. Y., 1967.

10. Saul'yev, V. K., Integration of Equations of Parabolic Type by the Method of Nets, The MacMillan Co., New York, N. Y., 1964.

11. Varga, R. S., Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, N. J., 1962.