NASA CR-132653

GENERALIZED DU FORT-FRANKEL METHODS FOR PARABOLIC INITIAL-BOUNDARY-VALUE PROBLEMS

David Gottlieb
Department of Mathematics
Massachusetts Institute of Technology

Bertil Gustafsson
ICASE and Dept. of Computer Science
University of Uppsala

ABSTRACT

The Du Fort-Frankel difference scheme is generalized to difference operators of arbitrary high order accuracy in space and to arbitrary order of the parabolic differential operator. Spectral methods can also be used to approximate the spatial part of the differential operator. The scheme is explicit, and it is unconditionally stable for the initial value problem. Stable boundary conditions are given for two different fourth order accurate space approximations.

(NASA-CR-132653) GENERALIZED DU FORT-FRANKEL METHODS FOR PARABOLIC INITIAL BOUNDARY VALUE PROBLEMS (Massachusetts Inst. of Tech.) 32 p HC $3.75 CSCL 12A

N75-23209
Unclas
G3/64 20739

This paper was prepared as a result of work performed under NASA Grant NGR 47-102-001 while the authors were in residence at ICASE, NASA Langley Research Center, Hampton, VA 23665.

I. INTRODUCTION

The Du Fort-Frankel difference scheme for solving parabolic equations, see e.g. [6], has the advantage of being explicit and yet unconditionally stable. The consistency requires that Δt goes to zero faster than Δx does, but this requirement is in practice not too severe if the coefficient of the second derivative is small.

An analysis of a semidiscretized parabolic model problem, performed along the same lines as in [2], [3] for hyperbolic problems, shows that higher order accurate approximations are more efficient except for very low requirements on the accuracy of the results. Therefore, in this paper the Du Fort-Frankel scheme will be generalized to difference operators of arbitrary high order accuracy in space and to arbitrary order of the parabolic differential operator. The number of space dimensions is also arbitrary and so is the number of equations in the system. The scheme is explicit and unconditionally stable. For a system with ℓ differential equations we also avoid the solution of an $\ell \times \ell$ system of equations for each gridpoint which would result from a straightforward formulation of the scheme. In addition to finite differences, spectral methods and finite element methods can also be used to approximate the spatial part of the differential operator in our scheme.

As for the original Du Fort-Frankel scheme, consistency imposes a restriction on Δt in relation to Δx. However, for the type of applications that we have in mind, like the viscous Navier-Stokes equations, the dominating truncation error comes from the space discretization. Furthermore, when the time dependent equations are used to obtain a steady state solution, the truncation error from the time discretization is of no importance, assuming that the scheme converges for $t \to \infty$.

The generalization to higher order accurate approximations in space was given by Swarz [8] for the scalar equation $u_t = \sigma u_{xx}$ with periodic boundary conditions. He studied the efficiency for different orders of accuracy and found e.g. that 12th order accurate operators are optimal in a certain sense for a relative precision of 10^{-2}, and even higher order for higher precision. In real applications with non periodic boundary conditions, we think that 4th or 6th order operators are more realistic.

In section 2 the scheme will be presented for a sequence of differential equations of increasing generality. In section 3 the stability proofs are given, and in section 4 the so called Fourier method is treated. In section 5 the stability of the mixed initial boundary value problem is proven for two different 4th order accurate space approximations. Section 6 contains a presentation of some numerical experiments that were done for the Burger's equation.

2. THE DU FORT-FRANKEL METHOD FOR FINITE DIFFERENCE SCHEMES

In order to illustrate the idea of the original Du Fort-Frankel scheme we start from the simple equation:

(2.1) $u_t = \sigma u_{xx}$, $\sigma > 0$.

With $u_j^n = u(j\Delta x, n\Delta t)$, it is well known that the scheme

$$\frac{u_j^{n+1}-u_j^{n-1}}{2\Delta t} = \frac{\sigma}{(\Delta x)^2} (u_{j+1}^n - 2u_j^n + u_{j-1}^n)$$

is unconditionally unstable. However, if we replace u_j^n by $\frac{1}{2}(u_j^{n+1}+u_j^{n-1})$, the scheme becomes unconditionally stable. For higher order approximations to u_{xx} it is not enough to replace u_j^n by some average. We adopt, therefore, another approach (see also Swartz [8]).

Let $(\Delta x)^{-2} D_{2p}^2$ be a 2p th order approximation to $\frac{\partial^2}{\partial x^2}$; then the generalized Du Fort-Frankel scheme will be

$$(2.2) \quad \frac{u_j^{n+1}-u_j^{n-1}}{2\Delta t} = \frac{\sigma}{(\Delta x)^2} D_{2p}^2 u_j^n - \frac{\gamma\sigma}{(\Delta x)^2} (u_j^{n+1} - 2u_j^n + u_j^{n-1})$$

where γ is to be chosen such that the scheme is unconditionally stable. The second term on the right hand side is a stabilizer, and it is an approximation to $\gamma\, \sigma \left(\frac{\Delta t}{\Delta x}\right)^2 u_{tt}$. Therefore, consistency requires that $\frac{\Delta t}{\Delta x} \to 0$; moreover, in order to minimize the truncation error we would like to choose γ as small as possible. It should be noted that the operator D_{2p}^2 is not necessarily a difference operator; we can use any method of approximation such as spectral methods [3], [5], and finite element methods [7]. For difference approximations, γ is given by

$$(2.3) \quad \gamma > \gamma_0 \equiv \frac{1}{4} \max_{|\xi| \le \pi} |\hat{D}_{2p}^2(\xi)|$$

where

$$(2.4) \quad \hat{D}_{2p}^2(\xi) = e^{-ikx} D_{2p}^2 \, e^{ikx}, \quad \xi = k\Delta x \, .$$

In the next section we will show that (2.3) yields unconditional stability. The original Du Fort-Frankel scheme corresponds to $\gamma=\gamma_0$. However, in this case the stability is not clear for variable coefficients.

We would like to consider in detail some difference approximations for D_{2p}^2, and the first one is the symmetric explicit operator.

Let $D_+u_j^n=u_{j+1}^n-u_j^n$, $D_-u_j^n=u_j^n-u_{j-1}^n$. Then

$$(2.5) \quad D_{2p}^2=D_+D_- \sum_{j=0}^{p-1} (-1)^j \frac{(j!)^2}{(2j+1)!(j+1)} (D_+D_-)^j$$

From (2.5) we deduce the following formula for the Fourier transform

$$\hat{D}^2_{2p} = -4\sin^2\frac{\xi}{2}\sum_{j=0}^{p-1}\frac{4^j(j!)^2}{(2j+1)!(j+1)}\sin^{2j}\frac{\xi}{2}.$$

It is obvious that $\hat{D}^2_{2p}(\xi)$ takes its maximum at $\xi = \pm\pi$. Therefore

$$(2.6)\quad \gamma_0 = \sum_{j=0}^{p-1}\frac{4^j(j!)^2}{(2j+1)!(j+1)},$$

and for every γ such that $\gamma \geq \gamma_0$, (2.2) is unconditionally stable. The operators D_2^2 and D_4^2 are of special interest from the practical point of view. Formula (2.5) yields

$$(2.7a)\quad D_2^2 = D_+D_-$$

$$(2.7b)\quad D_4^2 = D_+D_-(I - \frac{1}{12}D_+D_-) \equiv Q_1$$

For the first operator we have $\gamma_0 = 1$, and for the second one $\gamma_0 = \frac{4}{3}$. Another operator which is of importance is the fourth order implicit operator given by the following formula:

$$(2.8)\quad \frac{\partial^2 u}{\partial x^2} \doteq \frac{1}{(\Delta x)^2}\cdot\frac{D_+D_-}{1+\frac{1}{12}D_+D_-}u_j^n \equiv \frac{1}{(\Delta x)^2}Q_2u_j^n.$$

It was shown by Kreiss [5] that this operator is more accurate than the explicit operator defined in (2.7b). It is easily verified that for this case $\gamma_0 = \frac{3}{2}$.

We consider now a parabolic system of equations of the form

$$(2.9)\quad u_t = A(x,t,u)u_{xx},$$

where u is an ℓ component vector and A is an $\ell\times\ell$ matrix. The condition for uniform parabolicity is:

$$(2.10)\quad \text{Real } \lambda(A) \geq \delta_1 > 0$$

for each eigenvalue $\lambda(A)$ of A. There are two ways to extend the method defined by (2.2). The first is:

$$\frac{u_j^{n+1}-u_j^{n-1}}{2\Delta t} = \frac{1}{(\Delta x)^2} A_j^n D_{2p}^2 u_j^n - \frac{\gamma}{(\Delta x)^2} A_j^n (u_j^{n+1}-2u_j^n+u_j^{n-1}) ,$$

which requires the solution of an $\ell\times\ell$ system of equations in every time step, and therefore, this method is not desired. A better method can be obtained by:

$$(2.11)\quad \frac{u_j^{n+1}-u_j^{n-1}}{2\Delta t} = \frac{1}{(\Delta x)^2} A_j^n D_{2p}^2 u_j^n - \frac{\gamma}{(\Delta x)^2} \rho(A_j^n)(u_j^{n+1}-2u_j^n+u_j^{n-1})$$

where

$$(2.12)\quad \rho(A) = \max |\lambda(A)| ,$$

and λ is given by (2.3). (Here we assume that $\rho(A_j^n)$ is known explicitly, or that a good upper bound is known. The spectral radius should not be computed numerically at each timestep.) A stability analysis for certain classes of the matrix A will be presented in the next section.

We will now discuss the two dimensional case for systems:

$$(2.13)\quad u_t = A(x,t,u)u_{xx} + B(x,t,u)u_{xy} + C(x,t,u)u_{yy}.$$

The equation is said to be uniformly parabolic in the sense of Petrovskii if there exists a constant δ_2 independent of x,t and u such that:

(2.14) Real $\lambda\ (A\omega_1^2+B\omega_1\omega_2 + C\omega_2^2)\ \geq \delta_2 > 0$ for all real ω_1,ω_2 with $\omega_1^2+\omega_2^2=1$. The Du Fort-Frankel scheme for (2.13) is:

$$\frac{u_{j\ell}^{n+1}-u_{j\ell}^{n-1}}{2\Delta t} = \frac{1}{(\Delta x)^2} A_{j\ell}^n (D_{2p}^2)_x u_{j\ell}^n + \frac{1}{(\Delta y)^2} C_{j\ell}^n (D_{2p}^2)_y u_{j\ell}^n$$

$$(2.15)\qquad + \frac{1}{\Delta x\Delta y} B_{j\ell}^n (D_{2p}^1)_x (D_{2p}^1)_y u_{j\ell}^n - \gamma\left(\frac{\rho(A_{j\ell}^n)}{(\Delta x)^2} + \frac{\rho(C_{j\ell}^n)}{(\Delta y)^2}\right)$$

$$\cdot (u_{j\ell}^{n+1} - 2u_{j\ell}^n + u_{j\ell}^{n-1}) ,$$

where $(D)_x$ means a difference operator in the x direction and $(D)_y$ is a difference operator in the y direction. D_{2p}^1 is any approximation to the first derivative accurate up to 2pth order. It should be noted that the stabilizing term, the last term in (2.15), is independent of B . This means that (2.15) is a very simple explicit method that can be extended easily to more than two space dimensions. We can determine again γ by (2.3).

The last problem to be treated is a general parabolic differential equation of order $2m$ in s space dimensions:

$$(2.16) \qquad u_t = \sum_{|\nu|=2m} A_\nu(X,t)\ \partial_1^{\nu_1}\cdots\partial_s^{\nu_s}u\ , \qquad X = (x_1,\ldots,x_s)$$

$$\nu = (\nu_1,\ldots,\nu_s)\ ,\quad \nu_i \geq 0\ ,\quad |\nu| = \sum_{i=1}^{s}\nu_i\ ,\quad \partial_j = \frac{\partial}{\partial x_j}$$

The equation is said to be parabolic if there is a constant δ_3 such that all the eigenvalues of

$$\sum_{|\nu|=2m} A_\nu(X,t)(i\omega_1)^{\nu_1}\ \ldots\ (i\omega_s)^{\nu_s}$$

satisfy

$$(2.17) \qquad \text{Real } \lambda \leq -\delta_3 < 0$$

for all $\omega = (\omega_1,\ldots,\omega_s)$, ω_i real and $\sum \omega_i^2 = 1$. The scheme for (2.16) will be

$$(2.18) \qquad \frac{u_\nu^{n+1} - u_\nu^{n-1}}{2\Delta t} = \sum_{|\nu|=2m} A_\nu^n \frac{(D_{2p}^{\nu_1})_{x_1}}{(\Delta x_1)^{\nu_1}} \cdots \frac{(D_{2p}^{\nu_s})_{x_s}}{(\Delta x_s)^{\nu_s}} u_\nu^n$$

$$+ (-1)^m\gamma \sum_{\mu=1}^{s}\frac{\rho(A_\mu^n)}{(\Delta x_\mu)^{2m}}(u_\nu^{n+1} - 2u_\nu^n + u_\nu^{n-1})\ ,$$

where in the second sum all the terms with mixed derivatives are excluded.

3. Stability for the initial value problem.

In this section it will be shown that the scheme presented in the first section is unconditionally stable for the linear pure initial value problem. We shall make use of the stability theory developed by Widlund [10] and we will assume that the reader is familiar with that paper.

We start with the following lemma:

Lemma 3.1. Consider the equation:

$$\lambda^2 - 1 = -2y\lambda - x(\lambda-1)^2 \tag{3.1}$$

where

$$x > \frac{y}{2} \geq 0 . \tag{3.2}$$

Then the roots λ_+ and λ_- of (3.1) satisfy

$$|\lambda_\pm| \leq 1 , \tag{3.3}$$

where the equality sign holds only if $y = 0$, and then only for λ_+.

Proof: The roots of (3.1) are

$$\lambda_\pm = \frac{1}{1+x} [(x-y) \pm \sqrt{1 - y(2x-y)}] .$$

If $1-y(2x-y) \geq 0$ then by (3.2) we get

$$\sqrt{1 - y(2x-y)} \leq 1 \, , \quad |x-y| \leq x \, ,$$

and therefore $|\lambda_{\pm}| \leq 1$. Equality holds in (3.3) only if $y=0$, in this case

$$\lambda_{+} = 1 \quad \text{and} \quad |\lambda_{-}| = \left|\frac{x-1}{x+1}\right| \, .$$

If now $1 - y(2x-y) < 0$, then λ_{+} and λ_{-} are complex and therefore

$$|\lambda_{\pm}|^2 = \left|\frac{x-1}{x+1}\right| < 1 \, .$$

This completes the proof.

We discuss first the method defined by (2.2). The Fourier transform of (2.2) is exactly (3.1) with

$$y = \frac{\Delta t}{(\Delta x)^2} \sigma \hat{D}^2_{2p} \, , \quad x = 2\nu\sigma \frac{\Delta t}{(\Delta x)^2} \quad .$$

Equation (3.2) yields the condition (2.3) and by lemma 2.1 only one of the roots lies on the unit circle, which is sufficient for stability. Moreover, since $\hat{D}^2_{2p}(\xi) = 0$ only if $\xi = 0$, we have also:

(3.4) $$|\lambda| \leq 1 - \delta|\xi|^2 \quad .$$

The bound on $|\lambda|$ in (3.4) is important if lower order terms are included in (2.1). It should be noted that with $\nu = \gamma_0$, as in the original Du-Fort-Frankel scheme, (3.4) is not fulfilled.

In order to investigate the stability of (2.11) we make several assumptions that simplify the analysis. We will assume that A is independent of u and has real eigenvalues. The second assumption applies to a large class of problems such as the compressible viscous Navier Stokes equations. (We recognize, though, that the Navier Stokes equations are not uniformly parabolic since the continuity equation does not contain second derivatives).

We shall show now that the scheme (2.11) is a parabolic difference scheme in the sense of Widlund [10]. Rewrite first, the Fourier transform of (2.11) to get:

$$\widehat{\begin{bmatrix} u^{n+1} \\ u^{n} \end{bmatrix}} = G \widehat{\begin{bmatrix} u^{n} \\ u^{n-1} \end{bmatrix}} \tag{3.5}$$

where

$$G = \begin{bmatrix} \dfrac{2\frac{\Delta t}{(\Delta x)^2} A\hat{D}^2_{2p} + 4\nu\rho(A)I\frac{\Delta t}{(\Delta x)^2}}{1 + 2\nu\rho(A)\frac{\Delta t}{(\Delta x)^2}} & \dfrac{1 - 2\gamma\frac{\Delta t}{(\Delta x)^2}\rho(A)I}{1 - 2\gamma\frac{\Delta t}{(\Delta x)^2}\rho(A)} \\ I & 0 \end{bmatrix}$$

The eigenvalues z of G satisfy the equation (3.1) with $\lambda = z$ and

$$y = -\frac{\Delta t}{(\Delta x)^2}\lambda(A)\hat{D}^2_{2p}(\xi \tag{3.6}$$

and

(3.7) $$x = 2\gamma \frac{\Delta t}{(\Delta x)^2}\, \rho(A) \, .$$

Again if (2.3) is satisfied we get:

$$|z| \leq 1-\delta|\xi|^2 \quad , \quad |\xi| \leq \pi \, .$$

It remains to check the root condition of Widlund [10], that is we have to check that the eigenvalues of

(3.8) $$H = \begin{bmatrix} \frac{2x}{1+x} & \frac{1-x}{1+x} \\ 1 & 0 \end{bmatrix}$$

are not outside the unit circle and are simple on the unit circle. But the eigenvalues of H are $\frac{x\pm1}{x+1}$ and therefore, the root condition is satisfied. This completes the proof that (2.11) is unconditionally stable.

It should be noted that H defined in (3.8) satisfies a stronger condition than the root condition since only one of the roots lies on the unit circle. In fact the conditions of theorem 1.1 in [10] are satisfied. Thus the scheme is strongly parabolic in the sense of Varah [9], and we shall make use of this fact when treating the boundary conditions.

We proceed by analyzing the scheme (2.15). We make the additional assumption that the eigenvalues of

$$A\omega_1^2 + B\omega_1\omega_2 + C\omega_2^2$$

are real. This assumption is valid for the Navier Stokes equations, for example. Condition (ii) of theorem 1.1 in [10] is satisfied because now H is given by (3.8) with

$$x = 2\gamma\Delta t \left(\frac{\rho(A)}{(\Delta x)^2} + \frac{\rho(C)}{(\Delta y)^2} \right) \quad ,$$

and therefore, only one root lies on the unit circle. It remains to prove that the eigenvalues of the Fourier transform of (2.15) do not lie outside the unit circle. These eigenvalues are given, again, by (3.1) with

$$y = -\lambda \left(\frac{\Delta t}{(\Delta x)^2} A (\hat{D}^2_{2p})_x + \frac{\Delta t}{(\Delta y)^2} C (\hat{D}^2_{2p})_y + \frac{\Delta t}{\Delta x \Delta y} B (\hat{D}^1_{2p})_x (\hat{D}^1_{2p})_y \right)$$

and

$$x = 2\gamma\Delta t \left(\frac{\rho(A)}{(\Delta x)^2} + \frac{\rho(C)}{(\Delta y)^2} \right) \quad ,$$

Note that $(\hat{D}^2_{2p})_x$ and $(\hat{D}^2_{2p})_y$ are negative and $(\hat{D}^1_{2p})_x$ and $(\hat{D}^1_{2p})_y$ are purely imaginary. For usual difference approximations y is positive.

Condition (3.2) implies that γ must be chosen such that

$$(3.10) \qquad - \frac{\Delta t}{(\Delta x)^2} A (\hat{D}^2_{2p})_x - \frac{\Delta t}{(\Delta y)^2} C (\hat{D}^2_{2p})_y - \frac{\Delta t}{\Delta x \Delta y} B (\hat{D}^1_{2p})_x (\hat{D}^1_{2p})_y$$

$$\leq 4\gamma\Delta t \left(\frac{\rho(A)}{(\Delta x)^2} + \frac{\rho(C)}{(\Delta y)^2} \right) \quad .$$

By the parabolicity condition (2.14) it is clear that (3.10) is satisfied provided that:

$$(3.11)\qquad \gamma > \gamma_0 = \max_{|\xi|\le\pi} \frac{|\hat{D}^2_{2p}(\xi)| + |\hat{D}^1_{2p}(\xi)|^2}{4} \quad .$$

For the approximations (2.7) one can prove that (3.10) is satisfied for

$$\gamma > \gamma_0 = \frac{1}{4} \max_{|\xi|\le\pi} |\hat{D}^2_{2p}(\xi)|$$

A similar analysis holds for (2.18) and therefore we have completed the stability proof of the generalized Du Fort-Frankel methods.

We will now discuss briefly the effect of lower order terms in the equation. The advantage of the Du Fort-Frankel scheme is that it can be combined in a natural way with the Leap-Frog scheme if the equations under consideration have lower order terms.

As an illustration we discuss the equation:

$$(3.12)\qquad u_t = Au_x + \sigma u_{xx} \quad .$$

The scheme will be

$$(3.13)\qquad \frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t} = AD^1_{2p}u^n_j + \sigma D^2_{2p}u^n_j - \gamma\sigma(u_j^{n+1} - 2u_j^n + u_j^{n-1}) \quad .$$

The effect of the term AD^1_{2p} on the amplification matrix will be $\delta i\xi f(\xi)$ in the upper left corner, where δ is very small. Therefore, due to the discussion in [6,Sec.5.3], the scheme remains stable. However, if A depends on u and σ is very small, we need a term that will make the scheme dissipative. Kreiss and Oliger [5] suggested the form

$$u^{n+1} = (I + \frac{\epsilon}{64}(D_+D_-)^3)u^{n-1} + 2\frac{\Delta t}{\Delta x}AD_o(I - \frac{1}{6}D_+D_-)u^n ,$$

where

$$D_o u_j^n = \frac{u_{j+1}^n - u_{j-1}^n}{2} .$$

A trivial calculation shows that a sufficient condition for stability is

$$\frac{\Delta t A}{\Delta x} \leq \frac{1}{2}\sqrt{\frac{1-\epsilon}{0.4706}} \quad , \quad \epsilon < 1 .$$

4. The Fourier method for periodic boundary conditions.

Let N be a natural number, $\Delta x = \frac{1}{2N+1}$ and $x_j = j\Delta x$, $j = 0,1,\ldots,2N$. Consider a function $u(x)$ such that $u(x+1) = u(x)$. An accurate method of approximating u_{xx} at $x = x_j$ is to interpolate $u(x_j)$ by the trigonometric polynomial

$$(4.1) \qquad u(x_j) = \sum_{\omega=-N}^{N} \hat{u}(\omega)\exp(2\pi i\omega x_j) ,$$

where

$$\hat{u}(\omega) = \Delta x \sum_{j=0}^{2N} u(x_j)\exp(-2\pi i\omega x_j) ,$$

and to differentiate this polynomial to get

$$(4.2) \qquad \left.\frac{\partial^2 u}{\partial x^2}\right|_{x=x_j} = \sum_{\omega=-N}^{N} -4\pi^2\omega^2\hat{u}(\omega)\exp(2\pi i\omega x_j) .$$

This approximation can be achieved by two Fast Fourier Transforms and N complex multiplications.

With

$$\vec{u}_N = (u(x_o),\ldots,u(x_{2N}))^T ,$$
$$\vec{v}_N = (\frac{\partial^2 u(x_o)}{\partial x^2},\ldots,\frac{\partial^2 u(x_{2N})}{\partial x^2})^T$$

the above method can be written as $\vec{v}_N = T^*\Lambda T\vec{u}_N \equiv S_N\vec{u}_N$, where

(4.3) $$\Lambda = \text{diag}(-4\pi^2N^2, -4\pi^2(N-1)^2, \ldots , -4\pi^2,0,-4\pi^2,\ldots, -4\pi^2N^2)$$

and

(4.4) $$T_{k\ell} = \sqrt{\Delta x}\,\exp\,(-2\pi i(k-N)x_\ell) .$$

It is obvious that T is a unitary matrix and therefore we have

<u>Lemma 4.1</u>: S_N is Hermitian and its eigenvalues are all negative except one which is zero.

Consider now the equation

(4.5) $$u_t = \sigma u_{xx} ,$$
$$u(x,0) = f(x) \equiv \sum_{\omega=-N}^{N} \hat{f}(\omega)\exp(2\pi i\omega x) ,$$
$$u(0,t) = u(1,t) .$$

We approximate (4.5) by

(4.6) $$\frac{\vec{u}_N^{n+1} - \vec{u}_N^{n-1}}{2\Delta t} = \sigma S_N\vec{u}_N^n - \sigma\gamma N^2(\vec{u}_N^{n+1} - 2\vec{u}_N^n + \vec{u}_N^{n-1}) ,$$

which can be written

$$(4.7) \qquad \begin{bmatrix} \vec{v}_N^{\,n+1} \\ \vec{v}_N^{\,n} \end{bmatrix} = G_N \begin{bmatrix} \vec{v}_N^{\,n} \\ \vec{v}_N^{\,n-1} \end{bmatrix}$$

where

$$(4.8) \qquad G_N = \begin{bmatrix} -\frac{\sigma\Delta t}{1+x}(S_N+2\gamma N^2 I) & \frac{1-x}{1+x} I \\ I & 0 \end{bmatrix}$$

The eigenvalues of G_N are given by equation (3.1) with

$$y = y(\omega) = 4\Delta t\sigma\omega^2\pi^2 , \qquad x = 2\gamma\sigma\Delta t\omega \quad .$$

If now $\gamma > \pi^2$, then by Lemma 3.1 all the eigenvalues of G_N except a simple one, lies inside the unit circle. In order to prove that

$$(4.9) \qquad \| G_N^n \| \le K$$

where K is not a function of either n or N , we shall prove that G can be diagonalized by a similarity transformation which is independent of N . Define first

$$(4.10) \qquad \hat{T} = \begin{bmatrix} T & 0 \\ 0 & T \end{bmatrix} \quad .$$

Then

$$\hat{T}G\hat{T}^{*} = \begin{bmatrix} \frac{2(y(-N)-x)}{1+x} & & 0 & \frac{1-x}{1+x} I \\ & \ddots & & \\ 0 & & \frac{2(y(N)-x)}{1+x} & \\ & I & & 0 \end{bmatrix} = \hat{G} \ .$$

The eigenvectors of $\hat{G}$ are of the form

$$\begin{matrix} \lambda_1^{\pm} & 0 & & 0 \\ 0 & \lambda_2^{\pm} & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \\ 0 & 0 & \cdots & \lambda_{2N+1}^{\pm} \\ 1 & 0 & & 0 \\ 0 & 1 & & \cdot \\ \cdot & 0 & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \\ 0 & 0 & & 1 \end{matrix}$$

where $\lambda_i^{\pm}$ are solutions of (3.1) with $y = y(\omega_i)$, $= -N, -N+1, \ldots, N$.

If

$$R = \begin{bmatrix} L^{+} & L^{-} \\ I & I \end{bmatrix} ,$$

where

$$L_{\pm} = \mathrm{diag}\,(\lambda_1^{\pm}, \lambda_2^{\pm}, \ldots, \lambda_{2N+1}^{\pm}) ,$$

we have

$$R^{-1} = \begin{bmatrix} (L^+ - L^-)^{-1} & -L^-(L^+ - L^-)^{-1} \\ -(L^+ - L^-)^{-1} & L^+(L^+ - L^-)^{-1} \end{bmatrix} .$$

$\| R \|$ and $\| R^{-1} \|$ are bounded independently of N and $R^{-1}\hat{G}R$ is diagonal, which completes the proof.

The generalization of this method to variable coefficients is trivial. Moreover, because of the parabolicity there are no stability problems. This is not the case for hyperbolic equations, see [3].

5. Boundary Conditions

In this section we will treat only 4th order approximations to the scalar equation (2.1). The results can also be applied to systems, but the boundary conditions must then be stated in terms of those variables corresponding to the diagonalized system (if such a form exists). We will always consider the quarter space problem $0 \le x < \infty$, $t \ge 0$ and the theory by Varah [9] will be used. We begin with the Dirichlet boundary condition

$$u(0,t) = g(t)$$

The stability proof will be applied to the more general scheme

$$(5.1) \qquad \sum_{k=-1}^{r} (\alpha_k I + \beta_k D_4^{\,2}) u_j^{\,n-k} = 0$$

where $D_4^{\,2}$ is one of the 4th order accurate operators treated in Sec. 2. (5.1) is assumed to be consistent with (2.1) and stable for the initial value problem. Furthermore we assume strong parabolicity, which was shown to be fulfilled by the Du Fort-Frankel scheme in Sec. 3. For the explicit operator Q_1 defined by (2.7b), the numerical boundary conditions will be

$$(5.2) \qquad u_0^{\,n} = g^n$$

$$(5.3) \qquad \sum_{k=-1}^{r} [\alpha_k I + \beta_k D_+ D_- (I - \tfrac{1}{12} D_+^{\,2})] u_1^{\,n-k} = 0$$

In (5.3) the centered difference operator used for inner points is substituted by a non-centered 3rd order accurate operator. This lower order accuracy at the point $x = \Delta x$ should not effect the overall accuracy, (see Sec. 6). Connected with (5.1) is the resolvent equation

$$(5.4) \quad \tau(z)\hat{u}_j - D_+D_-(I - \frac{1}{12} D_+D_-)\hat{u}_j = 0, \quad j = 2,3,\ldots$$

where

$$\tau(z) = \sum_{k=-1}^{r} \alpha_k z^{r-k} \Big/ \sum_{k=-1}^{r} \beta_k z^{r-k}, \quad |z| > 1,$$

and the characteristic equation

$$(5.5) \quad \tau(z) - f(\kappa) = 0$$

where

$$f(\kappa) = \frac{(\kappa-1)^2}{\kappa} \left(1 - \frac{(\kappa-1)^2}{12\kappa}\right)$$

(5.5) has for $|z| > 1$ two roots κ_1, κ_2 inside the unit circle (see [9]). We can now prove

Theorem 5.1. Assume that at least one of the roots, $\kappa_2(z)$ say, satisfies $\kappa_2(z) \neq 1$, $\kappa_2(z) \neq \kappa_0$ for $|z| \geq 1$, where κ_0 is the root to

$$(5.6) \quad 4\kappa^4 - 9\kappa^3 - 5\kappa^2 - 15\kappa + 1 = 0,$$

which is inside the unit circle. Then (5.1) with $D_4{}^2 \equiv Q_1$ is stable with boundary conditions (5.2), (5.3).

Proof.

The homogenous boundary conditions for $\hat{u}_j$ are

(5.7) $\hat{u}_0 = 0$

(5.8) $\tau(z)\hat{u}_1 + D_+D_-(I - \frac{1}{12} D_+^2)\hat{u}_1 = 0$

Looking for non trivial solutions in $\ell_2(0,\infty)$ (i.e., $\sum_{j=0}^{\infty} |\hat{u}_j|^2 \Delta x < \infty$) for $|z| > 1$, we write $\hat{u}_j$ in the form

(5.9) $\hat{u}_j = a(\kappa_1^j - \kappa_2^j), \quad \kappa_1 \neq \kappa_2$

where κ_1, κ_2 are defined by (5.5).

The condition for the existence of non trivial solutions is obtained by inserting (5.9) into (5.8):

(5.10) $$\tau(z)(\kappa_1-\kappa_2) - [(\kappa_1-1)^2(1 - \frac{(\kappa_1-1)^2}{12}) - (\kappa_2-1)^2(1 - \frac{(\kappa_2-1)^2}{12})] = 0$$

By combining (5.10) with (5.5) it can be shown that κ_1 must fulfill

(5.11) $$\frac{(\kappa_1-1)^5}{\kappa_1} - \frac{(\kappa_2-1)^5}{\kappa_2} = 0$$

together with one of the equations

(5.12) $$\frac{(\kappa_1-1)^2}{\kappa_1} + \frac{(\kappa_2-1)^2}{\kappa_2} = 12$$

or

(5.13) $$\frac{(\kappa_1-1)^2}{\kappa_1} - \frac{(\kappa_2-1)^2}{\kappa_2} = 0$$

We begin with the system (5.11), (5.12). By defining

$c = \frac{\kappa_1 - 1}{\kappa_2 - 1}$ we get $\kappa_1 = c^5\kappa_2$, and from (5.11)

$\kappa_2^{-1} = -c(c+1)(c^2+1)$.

(5.12) then gives the final polynomial for c:

$$c^{10} + c^9 + 2c^8 + 3c^7 + 16c^6 + 3c^5 + 16c^4 + 3c^3 + 2c^2 + c + 1 = 0$$

The roots were obtained by a computer program. They are all such that at least one corresponding κ_i is outside the unit circle, which is a contradiction

An easy calculation shows that the system (5.11), (5.13) implies that $\kappa_1 = \kappa_2$, and in this case the form (5.9) of the solution does not hold. For double roots of (5.5) the form is $\hat{u}_j = aj\kappa^j$. By inserting it into (5.8) and using (5.5), we obtain immediately the condition

$$4\kappa^5 - 13\kappa^4 + 4\kappa^3 - 10\kappa^2 + 16\kappa - 1 = 0$$

for a nontrivial solution. $\kappa=1$ is one root, but is ruled out by our assumption. The remaining deflated polynomial is (5.6), and the theorem is proved. (It should be noted that the assumptions in the theorem could be weakened, since Varah's stability condition permits non-trivial solutions of a certain type.)

Corollary. The Du Fort-Frankel scheme is stable with boundary conditions (5.2), (5.3).

Proof. $\kappa=1$ is actually a double root of (5.5) for $z=1$ but a perturbation calculation shows that only one of them is inside the unit circle for $|z| > 1$. Therefore the first assumption of the theorem is fulfilled.

The root κ_0 of (5.6) with $|\kappa_0| < 1$ is real and located in the interval [0.06, 0.07]. This corresponds to a value of $f(\kappa)$ which is less than the stability limit 16/3, and, therefore $|z| < 1$ (see Sec. 3.).

It is easily verified that the assumptions in the theorem are fulfilled by, for example, the following schemes

$$u^{n+1} = \mu Q_1 u^n,$$

$$\mu Q_1 u^{n+1} = u^n,$$

$$(I - \theta\mu Q_1)u^{n+1} = (I + (1-\theta)\mu Q_1)u^n,$$

$$((1+\theta)I - \mu Q_1)u^{n+1} + (1-2\theta)u^n + \theta u^{n-1} = 0,$$

where it is assumed that $\mu = \frac{\sigma\Delta t}{(\Delta x)^2}$ and θ are in the stability intervals for the initial value problem. The theorem can also be generalized to several space dimensions in the sense that after a Fourier transformation over all space variables except x, the stability conditions of Varah [9] still are fulfilled uniformly in the dual variables

$\xi_2, \xi_3, \ldots$. This depends on the fact that the ξ_1 drop out of our calculations as z and μ do, and the final polynomial will be independent of all parameters.

Next we will treat the operator Q_2 defined by (2.8) One boundary condition for the difference scheme obviously must be (5.2). This condition will also be sufficient to define the solution if we multiply (5.1) by $(I + \frac{1}{12} D_+D_-)$ and solve for u^{n+1} directly. In this case stability follows immediately, since the solution to the resolvent equation has the form $\hat{u}_j = a\kappa^j$, and it cannot satisfy $\hat{u}_0 = 0$ if $a \neq 0$. However, this procedure will become inconvenient for a real problem, where the coefficients depend on x, t and very likely even on u.

In a practical application for an explicit scheme, the second derivative $(\Delta x)^{-2} S_j^n$ is computed from u_j^n by solving the tridiagonal system,

$$(I + \frac{1}{12} D_+D_-)S_j^n = D_+D_-u_j^n, \quad j=1,2,\ldots$$

For an implicit scheme, the system

$$\left.\begin{aligned} &(I + \frac{1}{12} D_+D_-)S_j^{n+1} - D_+D_-u_j^{n+1} = 0 \\ &\alpha_0 u_j^{n+1} + \beta_0 S_j^{n+1} = - \sum_{k=0}^{r} (\alpha_k u_j^{n-k} + \beta_k S_j^{n-k}) \end{aligned}\right\} \quad j=1,2,\ldots$$

is solved for u_j^{n+1}, S_j^{n+1}. In both cases, a boundary condition is needed for S_o. One way of providing that is to express S_o in terms of u_j, and this is also the simplest way for an explicit scheme.

By using a 3rd order accurate one sided formula, we get the boundary condition

$$(5.14) \quad S_0^n = (35u_0^n - 104u_1^n + 114u_2^n - 56u_3^n + 11u_4^n)/12.$$

Theorem 5.2. The scheme (5.1) with $D_4^2 \equiv Q_2$ is stable with the boundary conditions (5.2), (5.14).

Proof.

The resolvent equations for $\hat{u}_j$, $\hat{S}_j$ are

$$\left.\begin{aligned} \tau(z)\hat{u}_j &= \hat{S}_j \\ (I + \tfrac{1}{12} D_+D_-)\hat{S}_j &= D_+D_-\hat{u}_j \end{aligned}\right\} \qquad u=1,2,\ldots$$

with the homogenous boundary conditions

$$\hat{u}_0 = 0, \quad \hat{S}_0 = \sum_{j=0}^{4} e_j\hat{u}_j,$$

e_j being defined by (5.14). After eliminating $\hat{S}_j$, these equations can also be written

$$(5.15) \quad \tau(I + \tfrac{1}{12} D_+D_-)\hat{u}_j = D_+D_-\hat{u}_j, \qquad j=2,3,\ldots$$

$$(5.16) \quad \hat{u}_0 = 0$$

(5.17) $\tau\hat{u}_1 + \frac{1}{12}(\tau\hat{u}_2 - 2\tau\hat{u}_1 + \sum_{j=1}^{4} e_j\hat{u}_j) = \hat{u}_2 - 2\hat{u}_1.$

The solution in $\ell_2(0,\infty)$ to this system has for $|z| > 1$ the form

(5.18) $\hat{u}_j = a(\kappa^j - \delta_{0j})$ (δ_{ij} is the Kronecker symbol)

where κ satisfies

(5.19) $\tau(z)(\kappa + \frac{(\kappa-1)^2}{12}) - (\kappa-1)^2 = 0$

When using (5.18), (5.19) in (5.17) we obtain the condition for a non trivial solution:

$$(\kappa^2 + 10\kappa + 1) \sum_{j=1}^{4} e_j\kappa^{j-1} + 144 = 0$$

For our choice of e_j this equation is

(5.20) $11\kappa^5 + 54\kappa^4 - 435\kappa^3 + 980\kappa^2 - 926\kappa + 1624 = 0$

which has no root inside the unit circle, and the theorem is proved.

We consider next the boundary condition

$$u_x(0,t) + bu(0,t) = g(t).$$

For both of the operators Q_1 and Q_2, $u_x(0)$ is approximated by a one-sided 4th order accurate formula

(5.21) $(-25u_0 + 48u_1 - 36u_2 + 16u_3 - 3u_4)/12\Delta x + bu_0 = g.$

In connection with Q_1 the second boundary condition will be (5.3). The stability of the scheme has not been verified theoretically. However, numerical experiments show no evidence of instability (see Sec. 6.).

For the implicit operator Q_2, the second boundary condition will be (5.14), and in this case we can prove the following

Theorem 5.3.

The scheme (5.1) with $D_4{}^2 \equiv Q_2$ is stable with the boundary conditions (5.14), (5.21).

Proof.

The proof follows the same lines as the proof of Theorem 5.2, and we do not give the details here. Keeping the parameter a defined in (5.18), the final equation corresponding to (5.20) now is

$$(5.22)\quad P(\kappa)a \equiv (602\kappa^5 + 2876\kappa^4 - 24064\kappa^3 + 56764\kappa^2 - 71546\kappa + 35368)a = 0,$$

which has 4 roots outside the unit circle, and one root at $\kappa=1$. By the assumption of strong parabolicity, the corresponding value of z obtained from (5.19) is 1, and for this z-value, $\kappa=1$ is a double root of (5.19). Therefore $\kappa = 1 + \mathcal{O}(\sqrt{|z-1|})$. The stability condition by Varah is that the parameter a can be estimated from $P(\kappa)a = g$ by $|a| \leq |g| / \sqrt{|z-1|}$. This estimate is valid for our case since $\kappa=1$ is a simple root of $P(\kappa)$.

6. Numerical Experiments

The generalized Du Fort-Frankel scheme was tested for the Burgers equation

(6.1) $u_t + (u-1/2)u_x = \sigma u_{xx}, \quad -5 \le x \le 5, \quad 0 \le t,$

using the two 4th order accurate operators Q_1, Q_2 defined in Sec. 2. The initial function was

$u(x,0) = 1 - (x-5)/10,$

and the boundary conditions were

(6.2) $u(-5,t) = 1$

(6.3) $u(5,t) = 0$

The problem (6.1), (6.2), (6.3) has the steady state solution

(6.4) $v(x) = \frac{1}{2}(1 - \tanh \frac{x}{4\sigma})$.

The time integration was stopped when the condition

(6.5) $\max_j |u_j^{n+1} - u_j^n| \le \Delta t \cdot 10^{-5}$

was fulfilled. The error $\varepsilon = \max_j |u_j^n - v(x_j)|$ is listed in table 6.1. Parameter values used were $\sigma = 1/8$, $\Delta x = \Delta t = 0.2$, $\gamma = 4/3$. $\hat{t}$ denotes the time when (6.5) was first satisfied.

	Q_1	Q_2
ε	$10.6\ 10^{-4}$	$3.4\ 10^{-4}$
$\hat{t}$	26.6	26.4

Table 6.1 The error ε and steady state time $\hat{t}$ for the different operators.

Our scheme was also run with $\Delta x = \Delta t = 0.1$, and the error was found to be 16 times smaller in accordance with the 4th order accuracy.

With the same initial function but with the boundary condition $u_x(-5) = 0$ instead of (6.2), we obtained the steady state solution $u(x,\infty) \equiv 0$ for both operators Q_1, Q_2. In both cases the scheme was not showing any signs of instability, which possibly could have occurred from the boundary condition at $x = -5$, when using the operator Q_1.

REFERENCES:

1. Gustafsson, B., Kreiss, H.O. and Sundström, A., Stability theory for mixed initial boundary value problems. II. Math. Comp. 26, 649-686, 1972.

2. Kreiss, H. O. and Oliger, J., Comparison of Accurate Methods for the Integration of Hyperbolic Equations. Tellus, 24, 199-215, 1972.

3. Kreiss, H.O. and Oliger, J., Methods for the approximate solution of time dependent problems. GARP publication series No. 10, 1973.

4. Oliger, J., Fourth order difference methods for the initial boundary-value problem for hyperbolic equations. Math. Comp. 28, 15-25, 1974.

5. Orszag, S. and Israeli, M., Numerical simulation of viscous incompressible flows. Annual Reviews of Fluid Mechanics, Vol. 6, 1974.

6. Richtmyer, R.D. and Morton, K.W., Difference methods for initial value problems. Interscience, 1967.

7. Strang, W.G. and Fix, G.J., Analysis of the finite element method. Prentice Hall, 1973.

8. Swartz, B.K., The construction and comparison of finite difference analougs of some finite element schemes. Los Alamos Scientific Laboratory, Report LAUR-74-771, 1974.

9. Varah, J.M., Stability of difference approximations to the mixed initial boundary value problems for parabolic systems. SIAM J. Numer. Anal. 8, 598-615, 1971.

10. Widlund, O.B., Stability of parabolic difference schemes in the maximum norm Numer. Math. 8, 186-202,1966.