

DANIELE FUNARO

**Analysis of the Du Fort-Frankel method
for linear systems**

M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique, tome 19, n° 3 (1985), p. 429-441

http://www.numdam.org/item?id=M2AN_1985__19_3_429_0

© AFCET, 1985, tous droits réservés.

L'accès aux archives de la revue « M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

ANALYSIS OF THE DU FORT-FRANKEL METHOD FOR LINEAR SYSTEMS (*)

Daniele FUNARO ⁽¹⁾

Communicated by F. BREZZI

Résumé. — *On étudie le schéma de Du Fort-Frankel comme étant une méthode itérative pour la résolution du système linéaire $Au = f$, où A est une matrice complexe $N \times N$. On donne des conditions de stabilité en fonction des valeurs propres de A . Enfin, on détermine les paramètres qui donnent la convergence la plus rapide et on présente une estimation d'erreur.*

Abstract. — *The Du Fort-Frankel scheme is studied as an iterative method to solve the linear system $Au = f$, where A is a complex $N \times N$ matrix. Stability conditions are given in terms of the eigenvalues of A . An error estimate is presented. In order to provide a faster convergence, the optimal choice of the parameters is analysed.*

INTRODUCTION

The Du Fort-Frankel method (DFF) is, under suitable assumptions, a stable and convergent method to discretize in time differential equations of parabolic type (see for instance [5] or [7]). In this paper such a method will be studied as an iterative two-steps procedure to converge to the solution $u \in \mathbb{C}^N$ of the linear system $Au = f$ where A is a $N \times N$ complex matrix and $f \in \mathbb{C}^N$. Since u can be seen as the steady state solution of the equation $v_i = Av - f$, the sequence of approximations $\{u^n\}_{n \in \mathbb{N}}$ in \mathbb{C}^N is obtained from the relation :

$$\frac{u^{n+1} - u^{n-1}}{2 \Delta t} = Au^n - f - \sigma(u^{n+1} - 2u^n + u^{n-1})$$

where $\sigma, \Delta t > 0$ are considered as fixed parameters.

In the first part of the paper a characterization of the domain of stability in the plane $(\sigma, \Delta t)$ is provided, in terms of the eigenvalues of A . These are supposed to have negative real parts. An estimate of the error $u^n - u$ is given, in terms of the spectral radius $r = r(\sigma, \Delta t)$ of the amplification matrix.

(*) Received in May 1984.

⁽¹⁾ Dipartimento di Matematica, Università degli Studi, Strada Nuova, 65, 27100 PAVIA (Italia).

A further section of the paper is devoted to the study of the dependence of r on the parameters. In particular, if the eigenvalues of A are real, the pair $(\sigma^*, \Delta t^*)$ which minimizes r , is determined in the domain of stability. Explicit analytic formulas of the *optimal parameters* σ^* and Δt^* are produced. A remarkable improvement in the rate of convergence, can be observed when these parameters are used. This makes the DFF scheme competitive with other iterative methods.

Frequently A is the approximation of an elliptic differential operator, obtained through finite difference, finite element or spectral methods. The simultaneous use of the DFF method and spectral techniques was previously studied in Gottlieb and Gustafsson [2] and Gottlieb and Lustman [3].

Since the Chebyshev spectral approximation of an elliptic operator leads to full and ill-conditioned matrices, the convergence of the DFF method may be rather slow even if the *optimal parameters* are used.

However an impressive acceleration of the convergence is achieved if the DFF method described in this paper, is applied after preconditioning the matrices (see for instance [1]).

1. THE APPROXIMATION SCHEME

We are concerned with the problem of computing the solution $u \in \mathbb{C}^N$ ($N \geq 2$) of the system :

$$Au = f \quad (1.1)$$

where $f \in \mathbb{C}^N$ and $A : \mathbb{C}^N \rightarrow \mathbb{C}^N$ is a linear operator expressed as a $N \times N$ matrix with respect to the canonical base.

We shall use the DFF method as an iterative procedure to converge to the solution of (1.1).

For some fixed $u_0, u_1 \in \mathbb{C}^N$ we shall consider the following finite difference scheme :

$$\left. \begin{aligned} \frac{u^{n+1} - u^{n-1}}{2 \Delta t} &= Au^n - f - \sigma(u^{n+1} - 2u^n + u^{n-1}), \sigma \text{ and } \Delta t > 0 \\ u^1 &= u_1 \\ u^0 &= u_0 \end{aligned} \right\} \quad (1.2)$$

The aim is twofold. First we want to find conditions on $A, \Delta t$ and σ in order to obtain stability for (1.2). Moreover we shall estimate the error between u^{n+1} and u in the $l^{(2)}$ norm. Thus we shall write (1.2) explicitly as :

$$u^{n+1} - u = \Gamma(u^n - u) + \gamma(u^{n-1} - u) \quad \forall n \geq 1, \quad (1.3)$$

where

$$\Gamma = \frac{2 \Delta t}{1 + 2 \sigma \Delta t} (A + 2 \sigma I) \quad \text{and} \quad \gamma = \frac{1 - 2 \sigma \Delta t}{1 + 2 \sigma \Delta t}.$$

Equivalently we can write :

$$v^{n+1} = Gv^n = G^n v^1 \quad \forall n \in \mathbb{N} \quad (1.4)$$

where :

$$v^n = \begin{pmatrix} u^n - u \\ u^{n-1} - u \end{pmatrix} \quad \forall n \in \mathbb{N} \quad \text{and}$$

$$G = \begin{pmatrix} \Gamma & \gamma I \\ I & 0 \end{pmatrix} \quad \text{is a complex } 2N \times 2N \text{ matrix.}$$

The stability for (1.2) will be discussed through the derived formulation (1.4).

2. SOME PRELIMINARY RESULTS

In order to get an estimate on the eigenvalues of G we begin our investigation with some lemmas.

LEMMA 2.1 : *For every eigenvalue μ of Γ the two roots of the equation : $\lambda^2 - \mu\lambda - \gamma = 0$ are eigenvalues of G .*

Proof : We start by observing that if $\gamma = 0$ we have $\det(G - \lambda I) = (-\lambda)^N \det(\Gamma - \lambda I) = 0$ which implies $\lambda = 0$ or $\lambda = \mu$.

Let's suppose now $\gamma \neq 0$ (so $\lambda \neq 0$). We can write :

$$\begin{aligned} G - \lambda I &= \begin{pmatrix} \Gamma - \lambda I & \gamma I \\ I & -\lambda I \end{pmatrix} = \\ &= \begin{pmatrix} I & \Gamma - \lambda I + \frac{\gamma}{\lambda} I \\ -\frac{\lambda}{\gamma} I & 0 \end{pmatrix} \begin{pmatrix} -\frac{\gamma}{\lambda} I & \gamma I \\ I & 0 \end{pmatrix}, \end{aligned} \quad (2.1)$$

hence :

$$\begin{aligned} \det(G - \lambda I) &= \det\left(-\frac{\lambda}{\gamma} I\right) \det\left(\Gamma - \lambda I + \frac{\gamma}{\lambda} I\right) \det(\gamma I) = \\ &= (-\lambda)^N \det\left(\Gamma - \left(\lambda - \frac{\gamma}{\lambda}\right) I\right). \end{aligned} \quad (2.2)$$

So (2.2) shows that λ is an eigenvalue of G if and only if $\lambda = \frac{\gamma}{\lambda}$ is an eigenvalue of Γ . The proof is so completed. ■

We define now :

$$R_M = \max |\operatorname{Re} \xi|, \quad R_m = \min |\operatorname{Re} \xi| \quad \text{and} \quad I_M = \max |\operatorname{Im} \xi|$$

when ξ runs among all the eigenvalues of A .

Then we can show (the result, in the case $I_M = 0$, has already been obtained in [2]) :

LEMMA 2.2 : If $\operatorname{Re} \xi < 0$ for every eigenvalue ξ of A and if :

$$\left. \begin{aligned} \frac{R_M(1 - \sqrt{1 - \Delta t^2 I_M^2})}{2 \Delta t^2 I_M^2} < \sigma < \frac{R_m(1 + \sqrt{1 - \Delta t^2 I_M^2})}{2 \Delta t^2 I_M^2} \\ \text{with } \Delta t \leq \frac{2\sqrt{R_M R_m}}{I_M(R_M + R_m)} \leq \frac{1}{I_M} \\ \text{if } I_M \neq 0 \\ \sigma > \frac{R_M}{4} \quad \text{if } I_M = 0 \end{aligned} \right\} \quad (2.3)$$

then every eigenvalue λ of G satisfies : $|\lambda| < 1$.

Proof. — The hypotheses (2.3) imply the inequality :

$$4 |\operatorname{Re} \xi| \sigma > \operatorname{Re}^2 \xi + \frac{(1 - \gamma)^2}{(1 + \gamma)^2} \operatorname{Im}^2 \xi, \quad \forall \xi \text{ eigenvalue of } A. \quad (2.4)$$

Moreover we note that each eigenvalue μ of Γ can be obtained from an eigenvalue ξ of A by the equality :

$$\mu = (1 - \gamma) \left(\frac{\xi}{2\sigma} + 1 \right). \quad (2.5)$$

Therefore, taking into account (2.4), we get :

$$-2 \left(\frac{\operatorname{Re} \mu}{1 - \gamma} - 1 \right) > \left(\frac{\operatorname{Re} \mu}{1 - \gamma} - 1 \right)^2 + \left(\frac{\operatorname{Im} \mu}{1 + \gamma} \right)^2, \quad (2.6)$$

and we easily arrive to the inequality :

$$\frac{\operatorname{Re}^2 \mu}{(1 - \gamma)^2} + \frac{\operatorname{Im}^2 \mu}{(1 + \gamma)^2} < 1. \quad (2.7)$$

Now, if $\lambda = 0$ or $|\lambda| = \sqrt{|\gamma|} < 1$, the lemma is trivial. On the contrary, setting $\lambda = \rho(\cos \theta + i \sin \theta)$, we must have by lemma 2.1 :

$$\left(\rho - \frac{\gamma}{\rho}\right) \cos \theta = \Re \mu \quad \text{and} \quad \left(\rho + \frac{\gamma}{\rho}\right) \sin \theta = \Im \mu; \quad (2.8)$$

hence, since $\sin^2 \theta + \cos^2 \theta = 1$, one has :

$$\frac{\rho^2 \Re^2 \mu}{(\rho^2 - \gamma)^2} + \frac{\rho^2 \Im^2 \mu}{(\rho^2 + \gamma)^2} = 1. \quad (2.9)$$

Finally we get, by subtracting (2.7) from (2.9) :

$$(\rho^2 - \gamma^2)(1 - \rho^2) \left[\frac{\Re^2 \mu}{(\rho^2 - \gamma)^2 (1 - \gamma)^2} + \frac{\Im^2 \mu}{(\rho^2 + \gamma)^2 (1 + \gamma)^2} \right] > 0. \quad (2.10)$$

Therefore we have $\gamma^2 < \rho^2 = |\lambda|^2 < 1$ and the lemma is proved. ■

Let ξ_i ($i = 1, 2, \dots, N$) be the eigenvalues of A . In the following we shall denote by μ_i ($i = 1, 2, \dots, N$) the eigenvalues of Γ related to ξ_i by (2.5), and by $\lambda_{i,1}, \lambda_{i,2}$ ($i = 1, 2, \dots, N$) the two correspondent eigenvalues of G obtained through the equation : $\lambda^2 - \mu_i \lambda - \gamma = 0$.

We also define :

$$D(x) = \begin{pmatrix} x_1 & & & 0 \\ & \ddots & & \\ & & x_i & \\ & & & \ddots \\ 0 & & & & x_N \end{pmatrix}$$

where $x = (x_1, \dots, x_i, \dots, x_N)$ and $x_i \in \mathbb{C}$ ($i = 1, 2, \dots, N$).

Another basic result is the following lemma.

LEMMA 2.3 : Assume $\Delta t \neq \frac{1}{\sqrt{-\xi(\xi + 4\sigma)}}$ for every eigenvalue ξ of A ; then, if A admits a diagonal form, G can also be expressed in a diagonal form.

Proof : We can write :

$$A = V D(\xi) V^{-1} \quad (2.11)$$

where $V = (v_1, \dots, v_i, \dots, v_N)$ is a $N \times N$ matrix and v_i ($i = 1, 2, \dots, N$) is the

eigenvector of A correspondent to the eigenvalue ξ_i ($i = 1, 2, \dots, N$). We notice that $\{v_i\}_{1 \leq i \leq N}$ form a basis for \mathbb{C}^N .

The condition on Δt is equivalent to assume that $\lambda_{i,1} \neq \lambda_{i,2}$ for every $i = 1, 2, \dots, N$; actually :

$$|\lambda_{i,1} - \lambda_{i,2}|^2 = |\mu_i^2 + 4\gamma| = \frac{4}{(1 + 2\sigma\Delta t)^2} |1 + \Delta t^2 \xi_i(\xi_i + 4\sigma)| \neq 0. \quad (2.12)$$

Hence, the eigenvectors of G which form a basis of \mathbb{C}^{2N} are :

$$w_{i,1} = \begin{pmatrix} \lambda_{i,1} v_i \\ v_i \end{pmatrix} \quad \text{and} \quad w_{i,2} = \begin{pmatrix} \lambda_{i,2} v_i \\ v_i \end{pmatrix} \quad i = 1, 2, \dots, N,$$

since :

$$Gw_{i,1} = \begin{pmatrix} \lambda_{i,1} \left(\mu_i + \frac{\gamma}{\lambda_{i,1}} \right) v_i \\ \lambda_{i,1} v_i \end{pmatrix} = \begin{pmatrix} \lambda_{i,1}^2 v_i \\ \lambda_{i,1} v_i \end{pmatrix} = \lambda_{i,1} w_{i,1} \quad i = 1, 2, \dots, N. \quad (2.13)$$

Similarly we can argue for $w_{i,2}$. So it is possible to find a diagonal form for G . ■

By the last proposition we are allowed to write :

$$G = W \begin{pmatrix} D(\lambda_1) & 0 \\ 0 & D(\lambda_2) \end{pmatrix} W^{-1} \quad (2.14)$$

with $\lambda_1 = (\lambda_{1,1}, \dots, \lambda_{i,1}, \dots, \lambda_{N,1})$ and $\lambda_2 = (\lambda_{1,2}, \dots, \lambda_{i,2}, \dots, \lambda_{N,2})$. In particular for W we have the explicit expression :

$$W = \begin{pmatrix} V & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} D(\lambda_1) & D(\lambda_2) \\ I & I \end{pmatrix}. \quad (2.15)$$

3. STABILITY CONDITIONS

We shall denote by $r(M)$ the spectral radius of a given matrix M . Furthermore $\|M\|$ will denote the $l^{(2)}$ -norm of the matrix M . We recall that if $r(G) < 1$, there exists a constant $K > 0$ such that $\|G^n\| < K$, $\forall n \in \mathbb{N}$ (see for instance [5, p. 63]). In general we cannot determine explicitly K in terms of G . Nevertheless the results of the previous section can be used to give an explicit expression for K .

THEOREM 3.1 : *Assume that the matrix A admits a diagonal form. Then there exists a positive constant c such that, for every choice of σ and Δt which implies*

$r(G) < 1$, we have :

$$\|G^n\| \leq cnr^{n-1}(G) \quad \forall n \geq 1. \quad (3.1)$$

Proof : We start by supposing that G admits a diagonal form.

Hence, by (2.14) and (2.15), we get :

$$G^n = \begin{pmatrix} V & 0 \\ 0 & V \end{pmatrix} \underbrace{\begin{pmatrix} D\left(\frac{\lambda_1^{n+1} - \lambda_2^{n+1}}{\lambda_1 - \lambda_2}\right) D\left(-\lambda_1 \lambda_2 \frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2}\right) \\ D\left(\frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2}\right) D\left(-\lambda_1 \lambda_2 \frac{\lambda_1^{n-1} - \lambda_2^{n-1}}{\lambda_1 - \lambda_2}\right) \end{pmatrix}}_{Q_n} \begin{pmatrix} V^{-1} & 0 \\ 0 & V^{-1} \end{pmatrix}, \quad (3.2)$$

and consequently :

$$\|G^n\| \leq \|V\| \cdot \|V^{-1}\| \cdot \|Q_n\| = c_1 [r(Q_n Q_n^*)]^{1/2}, \quad c_1 \in \mathbb{R}. \quad (3.3)$$

A straightforward computation shows that the eigenvalues of $Q_n Q_n^*$ are the $2N$ roots $\alpha_i^+, \alpha_i^-, i = 1, 2, \dots, N$ of the equations :

$$\alpha_i^2 - I_T^i \alpha_i + I_D^i = 0 \quad i = 1, 2, \dots, N, \quad (3.4)$$

where we have defined :

$$I_T^i = \frac{1}{|\lambda_{i,1} - \lambda_{i,2}|^2} \left[|\lambda_{i,1}^{n+1} - \lambda_{i,2}^{n+1}|^2 + (|\lambda_{i,1}|^2 |\lambda_{i,2}|^2 + 1) |\lambda_{i,1}^n - \lambda_{i,2}^n|^2 + |\lambda_{i,1}|^2 |\lambda_{i,2}|^2 |\lambda_{i,1}^{n-1} - \lambda_{i,2}^{n-1}|^2 \right] \quad i = 1, 2, \dots, N,$$

and

$$I_D^i = \frac{|\lambda_{i,1}|^2 |\lambda_{i,2}|^2}{|\lambda_{i,1} - \lambda_{i,2}|^4} \left| (\lambda_{i,1}^{n+1} - \lambda_{i,2}^{n+1})(\lambda_{i,1}^{n-1} - \lambda_{i,2}^{n-1}) - (\lambda_{i,1}^n - \lambda_{i,2}^n)^2 \right|^2 \quad i = 1, 2, \dots, N.$$

Thus we have for $i = 1, 2, \dots, N$:

$$\begin{aligned} I_T^i &= \frac{1}{|\lambda_{i,1} - \lambda_{i,2}|^2} \left[(1 + |\lambda_{i,1}|^2)(1 + |\lambda_{i,2}|^2) |\lambda_{i,1}^n - \lambda_{i,2}^n|^2 + \right. \\ &\quad + (|\lambda_{i,1}|^2 - \lambda_{i,1} \bar{\lambda}_{i,2} - \bar{\lambda}_{i,1} \lambda_{i,2} + |\lambda_{i,2}|^2) \lambda_{i,1}^n \bar{\lambda}_{i,2}^n + \\ &\quad \left. + (|\lambda_{i,1}|^2 - \lambda_{i,1} \bar{\lambda}_{i,2} - \bar{\lambda}_{i,1} \lambda_{i,2} + |\lambda_{i,2}|^2) \bar{\lambda}_{i,1}^n \lambda_{i,2}^n \right] = \\ &= (1 + |\lambda_{i,1}|^2)(1 + |\lambda_{i,2}|^2) \frac{|\lambda_{i,1}^n - \lambda_{i,2}^n|^2}{|\lambda_{i,1} - \lambda_{i,2}|^2} + 2 \operatorname{Re} \lambda_{i,1}^n \bar{\lambda}_{i,2}^n, \quad (3.5) \end{aligned}$$

and

$$I_D^i = \frac{|\lambda_{i,1}|^2 |\lambda_{i,2}|^2}{|\lambda_{i,1} - \lambda_{i,2}|^4} |\lambda_{i,1}^{n-1}|^2 |\lambda_{i,2}^{n-1}|^2 |\lambda_{i,1} - \lambda_{i,2}|^4 = |\lambda_{i,1}^n|^2 |\lambda_{i,2}^n|^2. \quad (3.6)$$

Hence, for every $i = 1, 2, \dots, N$ we have the estimates :

$$\begin{aligned} |I_T^i| &\leq (1 + r^2(G))^2 \left(\sum_{j=0}^{n-1} |\lambda_{i,1}|^{n-j-1} |\lambda_{i,2}|^j \right)^2 + 2r^{2n}(G) \\ &\leq 4 \left(\sum_{j=0}^{n-1} r^{n-1}(G) \right)^2 + 2r^{2n}(G) \leq c_2 n^2 r^{2n-2}(G), \quad c_2 \in \mathbb{R}^+ \end{aligned} \quad (3.7)$$

and

$$|I_D^i| \leq r^{4n}(G). \quad (3.8)$$

Whence :

$$|\alpha_i^\pm| \leq \frac{|I_T^i| + \sqrt{|I_T^i|^2 + 4|I_D^i|}}{2} \leq c_3 n^2 r^{2n-2}(G) \quad i = 1, 2, \dots, N, \quad c_3 \in \mathbb{R}^+. \quad (3.9)$$

Finally the inequality (3.1) follows from (3.3).

Let's suppose now that G does not admit a diagonal form. By lemma 2.3 this can be possible only if : $\Delta t = 1/\sqrt{-\xi(\xi + 4\sigma)}$ where ξ is an eigenvalue of A .

Now, this relation between Δt and σ defines a curve in the plane $(\sigma, \Delta t)$. Since $\|G^n\|$ and $r(G)$ are continuous functions of σ and Δt , the inequality (3.1) in the general case is consequence of a standard density argument. ■

Remark 1 : The constant c in (3.1) depends in general on the order of A .

Remark 2 : In general, it is not possible to have a better estimate for $\|G^n\|$ than the one presented in (3.1). The difficulties arise when we try to find a bound to the term $\|W^{-1}\|$ (W is defined by (2.15)) or, that is equivalent, to the terms $1/|\lambda_{i,1} - \lambda_{i,2}|$, $i = 1, 2, \dots, N$.

By a Taylor expansion with respect to the variable Δt we have :

$$1/|\lambda_{i,1} - \lambda_{i,2}| = \frac{1}{2} |1 + 2\sigma \Delta t + o(\Delta t)| \quad i = 1, 2, \dots, N.$$

Therefore if Δt is sufficiently close to 0 we are able to bound $\|W^{-1}\|$. In this way we can get : $\|G^n\| \leq kr^n(G)$, $k \in \mathbb{R}$.

In particular this implies that the DFF method used in the approximation of the solution $u(t)$ of the equation $u_t = Au - f$, in the finite interval $]0, T[$

is stable and convergent for $\Delta t \rightarrow 0$ and σ fixed. Actually stability derives from [7] (p. 84; condition 1) and convergence is a classical consequence of stability and consistency.

4. A CONVERGENCE THEOREM

THEOREM 4.1 : *Let u be the solution of the problem (1.1) and let $\{u^n\}_{n \in \mathbb{N}}$ be the sequence generated by (1.2) for some $u_0, u_1 \in \mathbb{C}^N$.*

If σ and Δt satisfy the stability conditions (2.3) and if (3.1) holds, then we have :

$$\|u^{n+1} - u\| \leq cnr^{n-1}(G) \sqrt{\|u_0 - u\|^2 + \|u_1 - u\|^2}, \quad \forall n \geq 1. \quad (4.1)$$

Proof : In terms of norms the relation (1.4) yields :

$$\|v^{n+1}\| \leq \|G^n\| \|v^1\|, \quad \forall n \geq 1. \quad (4.3)$$

Hence (4.1) follows directly from (4.3) and (3.1). ■

5. INFLUENCE OF THE PARAMETERS IN THE CONVERGENCE BEHAVIOR

This section is devoted to the study of the convergence behavior in dependence of σ and Δt . More precisely we are interested in minimizing $r(G)$ in order to achieve the fastest convergence rate. Numerical tests performed over a great number of problems, show that an appropriate pair $(\sigma^*, \Delta t^*)$ determined experimentally always leads to a much faster convergence than that obtained with any other choice of these parameters. We shall prove theoretically, in the case in which all the eigenvalues of A are real and strictly negative, that there exists a unique pair $(\sigma^*, \Delta t^*)$ which minimizes $r(G)$. Moreover we shall give an explicit formula for σ^* and Δt^* in terms of the smallest and the largest of the eigenvalues. For this purpose we define the functions :

$$F^+(\xi, \sigma, \Delta t) = \frac{\Delta t(\xi + 2\sigma) + \sqrt{1 + \Delta t^2 \xi(\xi + 4\sigma)}}{1 + 2\sigma \Delta t}$$

and

$$F^-(\xi, \sigma, \Delta t) = \frac{\Delta t(\xi + 2\sigma) - \sqrt{1 + \Delta t^2 \xi(\xi + 4\sigma)}}{1 + 2\sigma \Delta t}$$

and we set :

$$\Lambda(\sigma, \Delta t) = \max_{1 \leq i \leq N} \{ |F^+(\xi_i, \sigma, \Delta t)|, |F^-(\xi_i, \sigma, \Delta t)| \}. \quad (5.1)$$

Lemma 2.1 and the relation (2.5) imply, for every $\sigma, \Delta t > 0$: $r(G) = \Lambda(\sigma, \Delta t)$. We shall suppose : $\xi_i \in \mathbb{R}^-$, $i = 1, 2, \dots, N$ and therefore, with notations

already used, we can write :

$$R_M = \max_{1 \leq i \leq N} \{ -\xi_i \} \quad \text{and} \quad R_m = \min_{1 \leq i \leq N} \{ -\xi_i \}.$$

Our aim is to determine a useful expression for

$$r^* = \inf_{\Delta t > 0} \inf_{\sigma > R_M/4} \Lambda(\sigma, \Delta t). \quad (5.2)$$

The next proposition answers this question.

PROPOSITION 5.1 : *Let $\sigma^* = (R_M + R_m)/4$ ($> R_M/4$) and $\Delta t^* = 1/\sqrt{R_M R_m}$. Then we have :*

$$r^* = \Lambda(\sigma^*, \Delta t^*). \quad (5.3)$$

Proof : We start by observing that, for every $\sigma > 0$ and $\Delta t > 0$, one has :

$$\max \{ |F^+(\xi, \sigma, \Delta t)|, |F^-(\xi, \sigma, \Delta t)| \} = \begin{cases} |F^+(\xi, \sigma, \Delta t)| & \text{if } \xi > -2\sigma \\ |F^-(\xi, \sigma, \Delta t)| & \text{if } \xi < -2\sigma. \end{cases} \quad (5.4)$$

Moreover $|F^+(\xi, \sigma, \Delta t)|$ is an increasing function in the variable ξ if $\xi > -2\sigma$ and $|F^-(\xi, \sigma, \Delta t)|$ is a decreasing function in the variable ξ if $\xi < -2\sigma$, so that :

$$\Lambda(\sigma, \Delta t) = \max \{ |F^+(-R_m, \sigma, \Delta t)|, |F^-(-R_M, \sigma, \Delta t)| \}. \quad (5.5)$$

A diligent investigation shows that $\forall \Delta t > 0$:

$$|F^+(-R_m, \sigma, \Delta t)| \geq |F^-(-R_M, \sigma, \Delta t)| \quad \text{if } \sigma \geq \sigma^*. \quad (5.6)$$

So we are allowed to write $\forall \Delta t > 0$:

$$\begin{aligned} \inf_{\sigma > R_M/4} \Lambda(\sigma, \Delta t) &= \\ &= \min \left\{ \inf_{\sigma > \sigma^*} |F^+(-R_m, \sigma, \Delta t)|, \inf_{R_M/4 < \sigma < \sigma^*} |F^-(-R_M, \sigma, \Delta t)| \right\}. \end{aligned} \quad (5.7)$$

Considering that $-\gamma = -\gamma(\sigma, \Delta t)$ is the product between the two roots of the equation $\lambda^2 - \mu\lambda - \gamma = 0$, by lemma 2.1 we must have :

$$\begin{aligned} |F^+(-R_m, \sigma, \Delta t)| |F^-(-R_m, \sigma, \Delta t)| &= |\gamma(\sigma, \Delta t)| = \\ &= |F^+(-R_M, \sigma, \Delta t)| |F^-(-R_M, \sigma, \Delta t)|. \end{aligned} \quad (5.8)$$

This shows that $\sqrt{|\gamma(\sigma, \Delta t)|}$ is an inferior bound for $|F^+(-R_m, \sigma, \Delta t)|$ and $|F^-(-R_m, \sigma, \Delta t)|$. Now the equality: $|F^+(-R_m, \sigma, \Delta t)| = \sqrt{|\gamma(\sigma, \Delta t)|}$ implies $\sigma = \frac{1 + \Delta t^2 R_m^2}{4 \Delta t^2 R_m}$, $\forall \Delta t > 0$ so that the condition $\sigma \geq \sigma^*$ implies $\Delta t \leq \Delta t^*$. At this point by (5.7) it is not difficult to prove:

$$\begin{aligned} \inf_{\substack{\sigma > \sigma^* \\ \Delta t > 0}} \Lambda(\sigma, \Delta t) &= \inf_{0 < \Delta t < \Delta t^*} \left| \gamma\left(\frac{1 + \Delta t^2 R_m^2}{4 \Delta t^2 R_m}, \Delta t\right) \right|^{1/2} = \\ &= \inf_{0 < \Delta t < \Delta t^*} \left| \frac{R_m \Delta t - 1}{R_m \Delta t + 1} \right| = \left| \frac{R_m \Delta t^* - 1}{R_m \Delta t^* + 1} \right| = \Lambda(\sigma^*, \Delta t^*), \quad (5.9) \end{aligned}$$

since

$$\sigma^* = \frac{1 + \Delta t^{*2} R_m^2}{4 \Delta t^{*2} R_m}.$$

Similarly we can argue for $|F^-(-R_m, \sigma, \Delta t)|$ so that we have:

$$\inf_{\substack{R_M/4 < \sigma < \sigma^* \\ \Delta t > 0}} \Lambda(\sigma, \Delta t) = \Lambda(\sigma^*, \Delta t^*). \quad (5.10)$$

Thus (5.3) is a trivial consequence of (5.9) and (5.10). ■

By the previous proposition we can find an explicit expression for r^* as follows:

$$r^* = \sqrt{\gamma^*} = \sqrt{\frac{1 - 2 \sigma^* \Delta t^*}{1 + 2 \sigma^* \Delta t^*}} = \frac{\sqrt{R_M} - \sqrt{R_m}}{\sqrt{R_M} + \sqrt{R_m}} = \frac{\sqrt{R_M/R_m} - 1}{\sqrt{R_M/R_m} + 1}. \quad (5.11)$$

This shows the dependence of r^* on the ratio R_M/R_m .

The matrix G , corresponding to the pair $(\sigma^*, \Delta t^*)$, does not admit a diagonal form. This means that for an index i we have: $\lambda_{i,1} = \lambda_{i,2}$ so that, recalling the *remark 2*, in this case (3.1) cannot be improved.

Finally we have to notice that, for some vector v in the plane $(\sigma, \Delta t)$, the derivative $\partial \Lambda / \partial v$ at the point $(\sigma^*, \Delta t^*)$, attains the value ∞ . For this reason the computation of the optimal parameters must be carried out with high accuracy.

6. FINAL REMARKS

Applications of the previous algorithm to the solution of linear systems, arising from the discretization of elliptic partial differential equations, are particularly suggested when spectral methods are used. The matrix deriving from a spectral approximation is generally full and heavily ill-conditioned. Therefore direct methods are impracticable, while the efficiency of an iterative method relies on the dependence of the convergence factor on the condition number of the matrix. For the matrix arising from a N -frequency Fourier approximation of a second order elliptic operator with periodic boundary conditions, the ratio R_M/R_m between the largest and the smallest eigenvalue is $O(N^2)$. Hence by (5.11), the minimal radius r^* for the DFF method is asymptotic with $\frac{N-1}{N+1}$. Instead, the minimal radius for the Richardson method [6] is asymptotic with $\frac{N^2-1}{N^2+1}$.

The improvement is even more dramatic for a Chebyshev approximation for the same operator with non-periodic boundary conditions, since R_M/R_m is $O(N^4)$.

However the convergence factor may still be too close to 1 to allow a reasonable speed of convergence in practical applications. It has been recently pointed out (see [6]) that the spectral matrix can be efficiently preconditioned by a low order finite difference matrix. Hence the DFF method discussed here, can be applied as well to the preconditioned matrix. Clearly the same analysis applies, provided the eigenvalues of the preconditioned matrix satisfy the hypotheses of reality.

This approach has been followed in [1], in which the DFF method with the optimal parameters obtained in the present paper, has been successfully applied, in conjunction with different preconditioning techniques, to Chebyshev approximations of second order operators. The values of the optimal parameters, under the hypotheses of reality of the eigenvalues, have also been used in [1] for preconditioned matrices with some complex eigenvalue. The experimental results show that the behavior predicted by the present theory also applies in these cases.

REFERENCES

- [1] C. CANUTO and A. QUARTERONI, *Preconditioned Minimal Residual Methods for Chebyshev Spectral Calculations*, ICASE Report 83-28.
- [2] D. GOTTLIEB and B. GUSTAFSSON, *Generalized Du Fort-Frankel Methods for Parabolic Initial Boundary Value Problems*, SIAM J. Numer. Anal., 13 (1976), pp. 129-144.
- [3] D. GOTTLIEB and L. LUSTMAN, *The Du Fort-Frankel Chebyshev Method for Parabolic Initial Boundary Value Problems*, ICASE Report No. 83-84, December 4, 1980.
- [4] D. GOTTLIEB and S. A. ORSZAG, *Numerical Analysis of Spectral Methods-Theory and Applications*, Regional Conference Series in Applied Mathematics, SIAM (Philadelphia) (1977).
- [5] E. ISAACSON and H. B. KELLER, *Analysis of Numerical Methods*, John Wiley & Sons, New York (1966).
- [6] S. A. ORSZAG, *Spectral Methods for Problems in Complex Geometries*, J. Comput. Phys., 37 (1980), pp. 70-92.
- [7] RICHTMYER and MORTON, *Difference Methods for Initial Value Problems*, John Wiley & Sons, New York (1967).