Available online at www.sciencedirect.com

ScienceDirect

Applied Numerical Mathematics 59 (2009) 1322–1336

APPLIED NUMERICAL MATHEMATICS

www.elsevier.com/locate/apnum

On the edge of stability analysis

Ercília Sousa

CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal

Available online 20 August 2008

Abstract

The application of high order methods to solve problems with physical boundary conditions in many cases requires a careful treatment near the boundary, where additional numerical boundary schemes have to be introduced. The choice of boundary schemes influences most of the times the stability of the numerical method. The von Neumann analysis does not allow us to define accurately the influence of boundary conditions on the stability of the scheme. The spectral analysis, often called the matrix method, considers the eigenvalues of the matrix iteration of the scheme and although they reflect some of the influence of boundary conditions on the stability, many times eigenvalues fail to capture the transient effects in time-dependent partial differential equations. The Lax stability analysis does provide information on the influence of numerical boundary conditions although in practical situations it is generally not easy to derive the corresponding stability conditions. In this paper we present properties that relates the von Neumann analysis, the spectral analysis and the Lax analysis and show under which circumstances the von Neumann analysis together with the spectral analysis provides sufficient conditions to achieve Lax stability.

Keywords: Lax stability; High-order methods; von Neumann analysis; Matrix method

1. Introduction

The analysis of numerical schemes involves the study of consistency, accuracy, stability and convergence. One of the primary difficulties encountered in applying higher-order finite difference methods is the need for numerical boundary schemes, which preserve the global order of accuracy of the interior scheme and produce a stable scheme.

The conditions of consistency, stability and convergence are related to each other and the precise relation is contained in the fundamental Equivalence Theorem of Lax a proof of which can be found in Richtmyer and Morton [18]. The theorem says that for a well-posed initial value problem and a consistent discretization scheme, stability is the necessary and sufficient condition for convergence.

Stability analysis has been a concern since the beginning of the twentieth century, when in 1928 Courant, Friedrichs and Levy [1] formulated the fundamental CFL condition. Since then, the main question has been how to find stability criteria that lead to conditions that we are able to apply in real world problems.

The von Neumann analysis appeared in the forties. It was developed in Los Alamos by Von Neumann and was considered classified until its brief description by Crank and Nicolson (1947) [2]. This theory was developed for the Cauchy problem and problems with periodic boundary conditions and now is the most well know classical method to determine stability conditions.

E-mail address: ecs@mat.uc.pt.

0168-9274/$30.00 © 2008 IMACS. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.apnum.2008.08.001

The normal mode analysis was initially presented by Godunov and Ryabenkii [4] in 1963. The question they were trying to answer was how to predict the instability based on the eigenvalue distribution of the matrix iteration. The spectral analysis, often called the matrix method, takes in consideration the spectral radius of the matrix iteration. If the matrix is normal, the spectral radius gives accurate information about the matrix norm and this guarantees sufficient and necessary conditions for stability. However, when the matrix iteration is not normal the spectral radius gives no indication of the magnitude of the error for a finite time. It guarantees eventual decay of the solution, but does not control the intermediate growth.

Godunov and Ryabenkii [4,5] introduced a new concept of a family of operators, and a new definition of the spectrum. This theory was later developed by Kreiss (1968) [14] and Osher (1969) [10]. The original Godunov–Ryabenkii theory provided a necessary condition for stability. A sufficient condition was later developed by Gustafsson et al. (1972) [6], henceforth called GKS theory. This theory covers linear, first order hyperbolic systems in one space dimension. Since GKS theory was first presented, related work has been done by Varah [29] for parabolic problems, by Strikwerda [20] for semi-discretized equations, by Michelson [12] for multidimensional problems and by Trefethen [27,28] where a relation between the GKS theory and group velocity is established. Applying the normal mode analysis to high-order finite difference methods often leads to very complex calculations. To overcome some of the difficulty of the theoretical approach, Thuné proposed an algorithm to calculate GKS stability for linear hyperbolic equations [24] and for linear hyperbolic systems [25].

Reddy and Trefethen [17] used the concept of ϵ-pseudoeigenvalues to provide a necessary and sufficient condition for stability of non-normal matrices. The pseudospectra is a more general definition than the spectrum of the family of operators introduced by Godunov and Ryabenkii. In the recent book, Trefethen and Embree [26] remark that the connection between pseudospectra and the stability of initial boundary value problems has gone largely unexploited.

The Lax stability guarantees that the norm of the powers of the matrix iteration is uniformly bounded but we cannot find easily analytical conditions to bound the powers of a matrix.

Recently, D.W. Zingg [31,32] addressed the question: can a method be Cauchy stable and spectral stable, but not Lax stable? They present various examples with different kinds of instabilities, instabilities related with an inflow boundary, an outflow boundary and an interior scheme. When a scheme is not Lax stable one of the conditions, von Neumann or spectral, was violated. This seems also to be the case in the examples presented in [22]. Also Zingg et al. [33] and Rai and Chakravarthy [16] rely on von Neumann analysis and the matrix method to analyze the stability of their scheme.

In this paper we show that under certain circumstances the von Neumann stability of the Cauchy problem and the spectral stability of the problem with non-periodic boundary conditions implies the power matrices are uniformly bounded and therefore ensure Lax stability as predicted experimentally in previous works.

The outline of the paper is as follows. In the next section we introduce the model problem and the general form of the finite difference schemes considered. In Section 3, we give an overview of the stability theories we are discussing. In Section 4, we show two motivation examples and we point out some aspects of the von Neumann analysis showing that the relevance of the von Neumann method to the initial boundary value problems is not yet fully understood. In Section 5, we present the main results. The last section includes some conclusions.

2. The finite difference schemes

The initial boundary problem we consider is a linear equation and is defined on a half-real line:

$$\frac{\partial u}{\partial t} = \mathcal{L}u(x,t), \quad x \geqslant 0,\ t \geqslant 0, \tag{1}$$

where $\mathcal{L}$ is a general differential operator in x. The initial condition and the boundary conditions are given by

$$u(x,0) = f(x), \tag{2}$$

$$u(0,t) = g(t), \qquad \lim_{x\to\infty} u(x,t) = 0. \tag{3}$$

Suppose we have approximations U_j^n to the values $u(x_j,t_n)$ at the mesh points x_j, $j = 0,1,2,\ldots,$ and assume we approximate the problem (1)–(3) by the difference scheme

$$U_j^{n+1} = QU_j^n, \quad j = r, r+1, \ldots, \tag{4}$$

$$Q = \sum_{j=-r}^{p} a_j \mathcal{E}^j, \qquad \mathcal{E}^a U_j^n = U_{j+a}^n, \tag{5}$$

where a_{-r} and a_p are non-zero. The a_j's also depend on parameters Δx and Δt, where Δt is the time step and Δx the space step.

Considering the finite difference scheme (4) we observe that as Q uses r points upstream, the basic approximation cannot be used at $x_0, x_1, x_2, \ldots, x_{r-1}$, so there we need to apply numerical boundary conditions. In our particular case the boundary given by the physical problem is associated only with the point x_0. At the other points the boundary conditions, called numerical boundary conditions, affect the difference scheme. Let us assume that the boundary conditions can be written as

$$U_\beta^{n+1} = \sum_{j=0}^{q} b_{\beta j} U_j^n, \quad \beta = 1, \ldots, r-1, \tag{6}$$

and $b_{\beta j}$ depend also of parameters Δx and Δt.

3. Stability analysis

Nowadays there are many textbooks that describe the three types of stability analysis mentioned below, such as, the classical book by Richtmyer and Morton [18], or some more recent books [7,9,15,19].

3.1. The von Neumann analysis

The von Neumann (Fourier) method is the most well-known classical method to determine necessary and sufficient stability conditions. If we assume periodic boundary conditions the von Neumann analysis is based on the decomposition of the numerical solution into a Fourier sum as

$$U_j^n = \sum_{p=0}^{N-1} \kappa_p^n \mathrm{e}^{\mathrm{i}\xi_p (j\Delta x)},$$

where $\mathrm{i} = \sqrt{-1}$, κ_p^n is the amplification factor of the pth harmonic and $\xi_p = p2\pi/N\Delta x$. The product $\xi_p \Delta x$ is often called the phase angle: $\theta = \xi_p \Delta x$ and covers the domain $[0, 2\pi)$ in steps of $2\pi/N$. The region around $\theta = 0$ corresponds to the low frequencies while the region $\theta = \pi$ is associated with the high-frequencies. In particular, the value $\theta = \pi$ corresponds to the highest frequency resolvable on the mesh, namely the frequency of wavelength $2\Delta x$.

Considering a single mode, $\kappa^n \mathrm{e}^{\mathrm{i}j\theta}$, its time evolution is determined by the same numerical scheme as the complete numerical solution U_j^n. Hence inserting a representation of this form into a numerical scheme we obtain a stability condition by imposing an upper bound to the amplification factor, κ.

The amplification factor is said to satisfy the *von Neumann condition* if there is a constant K such that

$$\left|\kappa(\xi)\right| \leqslant 1 + K\Delta t, \quad \forall \xi \in \mathbb{R}. \tag{7}$$

However, for some problems the presence of the arbitrary constant in (7) is too generous for practical purposes, although being adequate for eventual convergence in the limit $\Delta t \to 0$. In practice, the inequality (7) is substituted by the following stronger condition,

$$\left|\kappa(\xi)\right| \leqslant 1, \quad \forall \xi \in \mathbb{R}, \tag{8}$$

or in terms of the phase angle,

$$\left|\kappa(\theta)\right| \leqslant 1, \quad \forall \theta \in [0, 2\pi). \tag{9}$$

This has been called practical stability by Richtmyer and Morton [18] or strict stability by other authors. In some cases condition (7) allows numerical modes to grow exponentially in time for finite values of Δt. Therefore, the practical, or strict, stability condition (8) is recommended in order to prevent numerical modes from growing faster than the physical modes of the differential equation.

Von Neumann stability. The amplification factor satisfy the practical von Neumann condition if

$$\left|\kappa(\xi)\right| \leqslant 1, \quad \forall \xi \in \mathbb{R}.$$

3.2. The spectral analysis or the matrix method

Let us assume that we are considering an explicit method that can be written in the form of a matrix iteration, where the nodal points are U_j^n, $j = 0, \dots, N-1$.

Introducing the vector $U^n = [U_1^n, \dots, U_{N-1}^n]^T$, the scheme may be written as a matrix equation

$$U^{n+1} = AU^n + \mathbf{v}, \quad n = 0, 1, 2, \dots, \tag{10}$$

where A is an $N \times N$ matrix and $\mathbf{v}$ is a vector that may contain some data from the physical boundary conditions.

Any errors E^n in a calculation based on (10) will grow according to

$$E^{n+1} = AE^n, \quad n = 0, 1, 2, \dots, \tag{11}$$

where $E^n = u^n - U^n$ with u^n, U^n the exact and numerical solutions of (10), respectively, at $t = n\Delta t$.

Given $A \in \mathbb{R}^{N\times N}$ denote the spectral radius of A by $\rho(A)$ and the 2-norm of the matrix A by $\|A\|$. We recall that

$$\|A\| = \rho(A) \quad \text{if } A \in \mathbb{R}^{N\times N} \text{ is normal.}$$

It is well known that for any $A \in \mathbb{R}^{N\times N}$

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

and that

$$A^m \to 0 \text{ as } m \to \infty \quad \text{if and only if} \quad \rho(A) \leqslant 1, \tag{12}$$

where the eigenvalues on the unit circle must be simple.

A simple criterion for regulating the error growth governed by (11) is given by

$$\rho(A) \leqslant 1. \tag{13}$$

When the matrix A is not normal the spectral radius gives no indication of the magnitude of E^n for finite n. In this case a condition of the form $\rho(A) \leqslant 1$ guarantees eventual decay of the solution, but does not control the intermediate growth of the solution. Then, it is easy to understand that the condition (13) is a necessary condition for stability but not always sufficient.

Spectral stability. Let A be the matrix iteration of a numerical method, then the spectral condition is given by

$$\rho(A) \leqslant 1, \tag{14}$$

where the eigenvalues on the unit circle must be simple.

3.3. The Lax analysis

The Lax stability is in some way connected with the spectral stability. The spectral stability ensures that $\|E^n\| \to 0$ when $n \to \infty$. Practical computations are however, performed at finite values of n, and the spectral condition does not ensure that the norm $\|A^n\|$ does not become large at finite values of n before decaying as n goes to infinity. In order to guarantee that $\|A^n\|$ does not become too large we require $\|A^n\|$ to remain uniformly bounded for all values of n.

Lax stability. In order for all $\mathbf{U}^n$ to remain bounded and the scheme, defined by the operator A, to remain stable the infinite set of operators A^n has to be uniformly bounded. That is, a constant K exists, such that,

$$\left\|A^n\right\| \leqslant K, \quad 0 < \Delta t < \tau,\ 0 \leqslant n\Delta t \leqslant T,$$

for fixed values of τ and T and for all n, in particular for $n \to \infty$, $\Delta t \to 0$ with $n\Delta t$ fixed, where K is independent of n, Δt and Δx. This condition implies the definition of some norm in the considered functional space.

4. Motivation examples

In this section two examples are given: a classical example with physical boundary conditions and a more complex example with numerical boundary conditions.

4.1. A classical example

This example is a classical example showing that the eigenvalues of the matrix iteration does not give a sufficient condition for stability.

Suppose we consider the problem (1)–(3), where

$$\mathcal{L}u(x,t) = -\frac{\partial u}{\partial x}, \quad x \in (0,1)$$

with an initial condition and Dirichlet boundary conditions.

We discretize the equation by

$$U_j^{n+1} = U_j^n - \sigma\big(U_j^n - U_{j-1}^n\big), \quad \text{where } \sigma = \frac{\Delta t}{\Delta x}. \tag{15}$$

Then $\mathbf{U}^{n+1} = A_N\mathbf{U}^n + \mathbf{v}$, where $\mathbf{U}^n = [U_1^n, \ldots, U_{N-1}^n]^T$,

$$A_N = \begin{bmatrix} 1-\sigma & & & \\ \sigma & 1-\sigma & & \\ \vdots & \ddots & & \vdots \\ & & \sigma & 1-\sigma \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} \sigma U_0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}$$

The eigenvalues are $\lambda_i = 1-\sigma$, $i = 0, \ldots, N-1$, and $\rho(A_N) < 1$ if and only if $0 < \sigma < 2$.

On the other hand, the von Neumann stability analysis for the interior scheme says that the amplification factor is given by

$$\kappa = 1 - \sigma\big(1 - \mathrm{e}^{-\mathrm{i}k\Delta x}\big).$$

Therefore, for $\theta = k\Delta x$,

$$\kappa = 1 - \sigma + \sigma\cos\theta - i\sigma\sin\theta.$$

We have

$$|\kappa|^2 = 1 - 2\sigma(1-\sigma)(1-\cos\theta).$$

For that reason we have that $|\kappa|^2 \leqslant 1$ if and only if $0 \leqslant \sigma(1-\sigma)(1-\cos\theta) \leqslant 1$ that is $0 < \sigma \leqslant 1$.

In conclusion we have

$$\text{von Neumann:} \quad 0 < \sigma \leqslant 1, \qquad \text{Spectral:} \quad 0 < \sigma < 2.$$

Therefore, the spectral radius suggests stability for $1 < \sigma < 2$, although the scheme is unstable for these values.

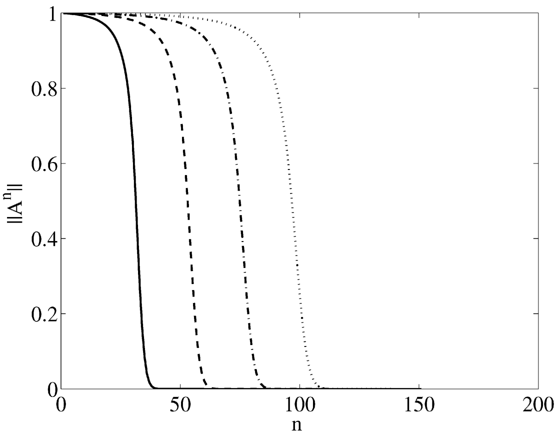
An investigation of the behavior of the norm of the matrix $\|A_N^n\|$ has been performed in [8]. By evaluating

$$M_N(\sigma) = \max_{n \geqslant 1} m_N(\sigma, n),$$

where $m_N(\sigma, n) = \max_{ij} |(A_N^n)_{ij}|$ it is found that for $1 < \sigma < 2$ and $N \geqslant 2$

$$M_N(\sigma) \leqslant \left(\frac{\sigma}{2-\sigma}\right)^{N-1}$$

and therefore it can reach very high values at some intermediate values of n when σ is very close to 2. Note that $m_N(\sigma, n) \leqslant \|A_N\|$, where $\|\cdot\|$ is the 2-norm.



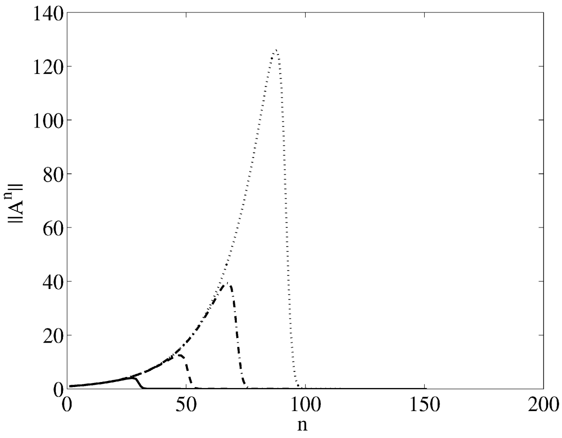


Fig. 1. Norms of powers of the matrix iteration A, of the finite difference scheme (15), for different matrix sizes: $N = 30$, (–), $N = 50$, (– –), $N = 70$, (– · –), $N = 90$ (· · ·). We consider two different values of σ: (a) $\sigma = 0.9$—the scheme for this case is von Neumann stable and spectral stable; (b) $\sigma = 1.03$—the scheme for this case is von Neumann unstable and spectral stable.

Additionally, as $N \to \infty$, that is $\Delta x \to 0$, although $\|A_N^n\|$ tends to zero as $n \to \infty$, there is no uniform bound on the norm if $1 < \sigma \leqslant 2$ as shown in Fig. 1.

This is a very important point for numerical calculations, since if a certain computation for fixed Δx is not accurate enough, one would like to have a better result with a smaller Δx. This is not the case under the condition $\rho(A_N) < 1$ but it seems to be the case for $0 < \sigma \leqslant 1$.

For $0 < \sigma \leqslant 1$, as N becomes larger the eigenvalues of A_N are the same. We can also conclude the stability condition, $0 < \sigma \leqslant 1$, is the intersection of both conditions and is the same as the von Neumann condition.

4.2. An example with numerical boundary conditions

Consider the one-dimensional problem with constant velocity V in the positive x direction and constant diffusion with coefficient $D > 0$:

$$\frac{\partial u}{\partial t} + V\frac{\partial u}{\partial x} = D\frac{\partial^2 u}{\partial x^2}, \quad t > 0,\ x > 0 \tag{16}$$

with the initial condition

$$u(x,0) = f(x), \quad x \geqslant 0 \tag{17}$$

and the boundary conditions

$$u(0,t) = g(t), \quad t > 0, \qquad u(x,t) \to 0, \quad x \to \infty. \tag{18}$$

We discretize at the interior points $j = 2, \ldots, N-1$, using the scheme introduced by Leonard [11],

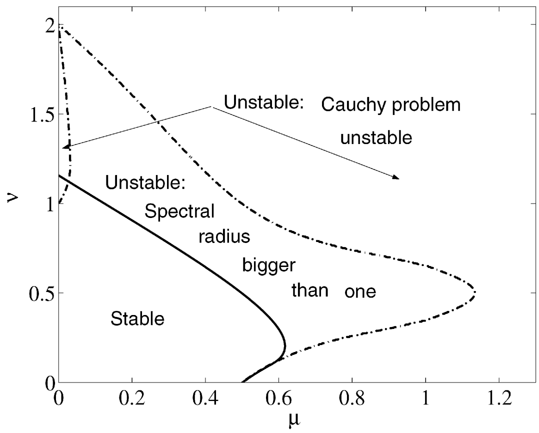
$$U_j^{n+1} = \left\{1 - \nu\Delta_0 + \left(\frac{1}{2}\nu^2 + \mu\right)\delta^2 + \frac{1}{6}\nu\left(1 - \nu^2 - 6\mu\right)\delta^2\Delta_-\right\}U_j^n \tag{19}$$

that interpolates the mesh points U_{j-2}^n, U_{j-1}^n, U_j^n, U_{j+1}^n and where

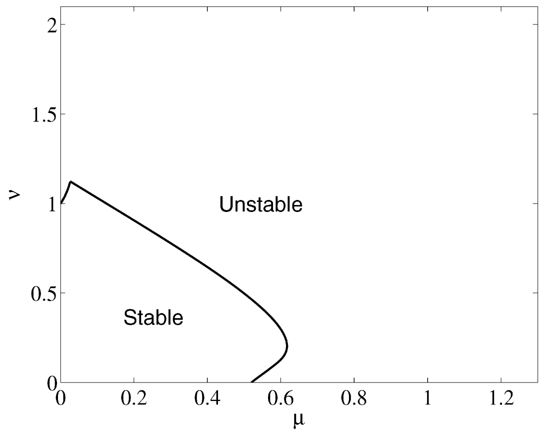
$$\nu = V\frac{\Delta t}{\Delta x} \quad \text{and} \quad \mu = D\frac{\Delta t}{\Delta x^2}.$$

The operators are the usual central, backward and second difference operators

$$\Delta_0 U_j := \frac{1}{2}(U_{j+1} - U_{j-1}), \qquad \Delta_- U_j := U_j - U_{j-1},$$
$$\delta^2 U_j := U_{j+1} - 2U_j + U_{j-1}.$$



(a)



(b)

Fig. 2. Stability region for the finite difference scheme (19) with numerical boundary condition (20): (a) Spectral stability region (–); von Neumann stability region (– · · –); (b) Region where we have simultaneously von Neumann and spectral stability.

At $j = 1$, we use a numerical boundary condition suggested in [22], that interpolates the mesh points U^n_{j-1}, U^n_j, U^n_{j+1} and U^n_{j+2},

$$U_j^{n+1} = \left\{1 - \nu\Delta_0 + \left(\frac{1}{2}\nu^2 + \mu\right)\delta^2 + \frac{1}{6}\nu\left(1 - \nu^2 - 6\mu\right)\delta^2\Delta_+\right\}U_j^n, \tag{20}$$

where $\Delta_+ U_j := U_{j+1} - U_j$.

In Fig. 2 we show the von Neumann condition for the scheme (19), considering we have a Cauchy problem or periodic boundary conditions. We also plot the spectral condition for the matrix iteration A_N that takes in consideration the interior scheme (19) and also the numerical boundary condition (20). The spectral radius of the matrix iteration was computed for $N = 30$. Although, when increasing the matrix size N we observe slight changes in the spectral radius, the region where the eigenvalues are less than one it is not affected since they do not become larger than one. The matrix equation for this case is $\mathbf{U}^{n+1} = A_N\mathbf{U}^n + \mathbf{v}$, where A_N and $\mathbf{v}$ are given by

$$A_N = \begin{bmatrix} b^* & c^* & d^* & 0 & 0 & \ldots & \ldots & 0 & 0 & 0 & 0 \\ b & c & d & 0 & 0 & \ldots & \ldots & 0 & 0 & 0 & 0 \\ a & b & c & d & 0 & \ldots & \ldots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & a & b & c & d \\ 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & 0 & a & b & c \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} a^* U_0 \\ a U_0 \\ 0 \\ \vdots \\ 0 \\ d U_N \end{bmatrix}$$

where

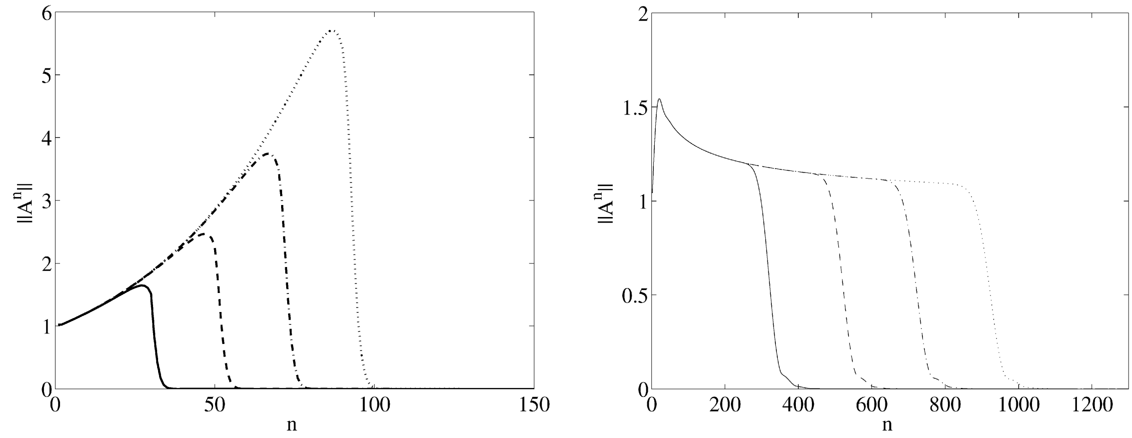
$$\begin{aligned} a^* &= c_1 + c_2 - c_3, & a &= -c_3, \\ b^* &= 1 - 2c_2 + 3c_3, & b &= c_1 + c_2 + 3c_3, \\ c^* &= -c_1 + c_2 - 3c_3, & c &= 1 - 2c_2 - 3c_3, \\ d^* &= c_3, & d &= -c_1 + c_2 + c_3, \end{aligned}$$

for

$$c_1 = \frac{\nu}{2}, \qquad c_2 = \frac{\nu^2}{2} + \mu, \qquad c_3 = \frac{\nu}{6}\left(1 - \nu^2 - 6\mu\right).$$

It was proved in [23], for this example, using normal mode analysis that the stable region is given by the region in Fig. 2(b). This region is the same as the region shown in Fig. 2(a) that represents the intersection of the von Neumann condition and spectral condition.

In Fig. 3 we plot $\|A_N^n\|$ as $n \to \infty$. For $\mu = 0.001$ and $\nu = 1.02$ we are in the spectral stable region but von Neumann unstable and we observe in Fig. 3(a) that the $\max_n \|A_N^n\|$ is increasing as the size of the matrix A_N increases.



(a) (b)

Fig. 3. Norms of powers of the matrix iteration A, for the finite difference scheme (19)–(20), for different matrix sizes: $N = 30$, (–), $N = 50$, (– –), $N = 70$, (– · –), $N = 90$ (· · ·). We consider different values of μ and ν and give the spectral radius for each case: (a) $\mu = 0.001$, $\nu = 1.02$—the scheme for this case is von Neumann unstable and spectral stable. Spectral radius is respectively given by $\rho(A_{30}) = 0.3031$, $\rho(A_{50}) = 0.3053$, $\rho(A_{70}) = 0.5448$, $\rho(A_{90}) = 0.5671$. (b) $\mu = 0.001$, $\nu = 0.1$—the scheme for this case is von Neumann stable and spectral stable. Spectral radius is respectively given by $\rho(A_{30}) = 0.9608$, $\rho(A_{50}) = 0.9608$, $\rho(A_{70}) = 0.9608$, $\rho(A_{90}) = 0.9642$.

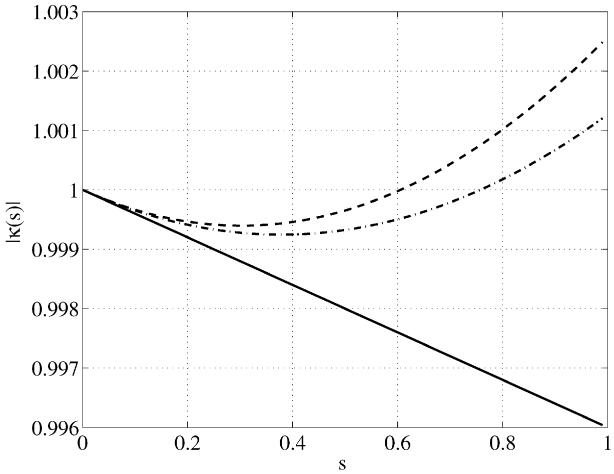


Fig. 4. Absolute value of the amplification factor (21), for the scheme (19) for $\mu = 0.001$ and $\nu = 1$ (–), $\nu = 1.005$ (– · –) $\nu = 1.02$ (– –).

On the other hand, for $\mu = 0.001$ and $\nu = 0.1$ we are in the region where the scheme is spectral and von Neumann stable and it is shown in Fig. 3(b) that the maximum value is always the same as N increases.

The spectral condition guarantees eventual decay of the solution but does not control the intermediate growth of the solution, this being guaranteed by the von Neumann condition, that is, the von Neumann condition seems to assure that as N becomes larger there will not be a strong effect in the value $\max_n \|A_N^n\|$. Note also that if the Cauchy problem is stable, any instabilities must thus be associated with the boundary condition, which is independent of the matrix size.

In Fig. 4 we plot the amplification factor for the scheme (19) given by

$$\begin{aligned}\left|\kappa(s)\right| &= 1 - 8\mu s + 4\left[\left(2\mu + \nu^2\right)^2 - \nu^2 + 2\alpha(1 - 2\nu)\right]s^2 \\ &\quad - 16\alpha\left(2\mu + \nu^2 - \nu - \alpha\right)s^3, \end{aligned} \tag{21}$$

where $\alpha = 2\nu\mu - (\nu/3)(1 - \nu^2)$ and $s = \sin^2(\xi/2)$. In Fig. 5, $\|A_N^n\|$ has a behavior very close to $|\kappa(\pi)|^n$ where $\xi = \pi$ is the value for which the amplification factor takes the biggest value, that is, the norm $\|A_N^n\|$ seems to follow the rise of the amplification factor. Therefore the relevance of the von Neumann method to initial boundary value problems is not yet fully understood.

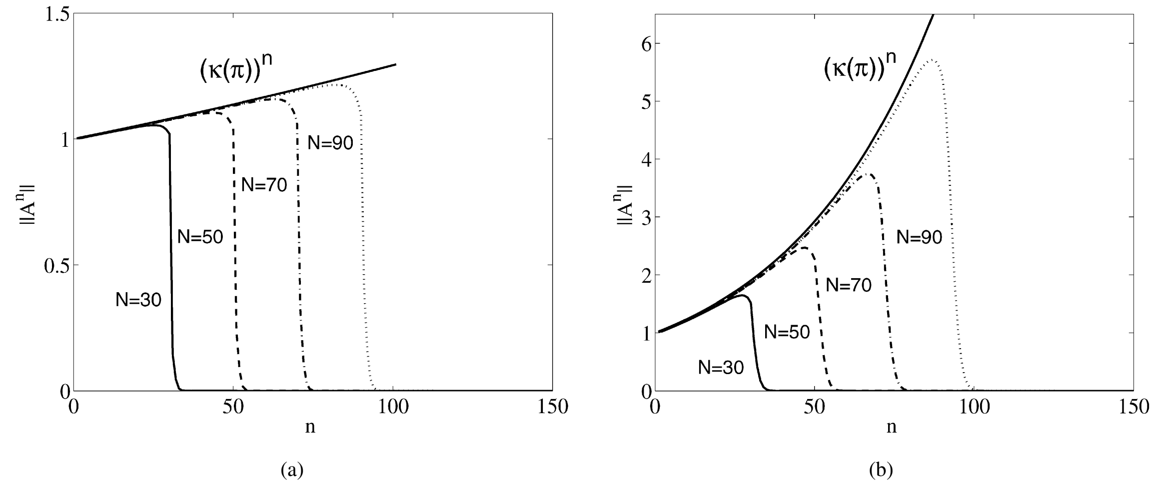


Fig. 5. Norms of powers of the matrix iteration A, for the finite difference scheme (19)–(20), for different matrix sizes: $N = 30$, (–), $N = 50$, (– –), $N = 70$, (– · –), $N = 90$ (· · ·). We consider different values of μ and ν: (a) $\mu = 0.001$, $\nu = 1.005$; (b) $\mu = 0.001$, $\nu = 1.02$. The scheme for both cases is von Neumann unstable and spectral stable. We observe the norm of the matrix iteration A follows the rise of the amplification factor (21) as N increases.

5. Main results

For completeness, we start this section by presenting the results for problems with periodic boundary conditions before showing the results for the case we have numerical boundary schemes. We assume the norm $\|\cdot\|$ of vectors and matrices is the 2-norm.

5.1. Periodic boundary conditions

Consider the difference scheme

$$U_j^{n+1} = QU_j^n, \quad j = r, r+1, \ldots, \tag{22}$$

$$Q = \sum_{j=-r}^{p} a_j \mathcal{E}^j, \qquad \mathcal{E}^a U_j^n = U_{j+a}^n, \tag{23}$$

and assume we have periodic boundary conditions. The matricial form of the numerical scheme is given by

$$\mathbf{U}^{n+1} = Q_N \mathbf{U}^n \tag{24}$$

for $\mathbf{U}^n = [U_0^n, U_1^n, \ldots, U_{N-1}^n]^T$ and Q_N is a circulant $N \times N$ matrix given by

$$Q_N = \begin{bmatrix}
a_0 & & \ldots & a_p & 0 & \ldots & 0 & a_{-r} & \ldots & a_{-1} \\
\vdots & & \ddots & & & \ddots & & \ddots & \ddots & \vdots \\
a_{-r+1} & \ldots & a_0 & \ldots & & a_p & 0 & \ldots & 0 & a_{-r} \\
a_{-r} & \ldots & & a_0 & & \ldots & a_p & 0 & \ldots & 0 \\
\vdots & & \ddots & & \ddots & \ddots & & \ddots & \ddots & \vdots \\
0 & \ldots & 0 & a_{-r} & & \ldots & a_0 & \ldots & a_p & 0 \\
0 & & & & a_{-r} & & & a_0 & & a_p \\
a_p & & & & & & & & & a_{p-1} \\
\vdots & \ddots & & & \ddots & \ddots & & \ddots & \ddots & \vdots \\
a_1 & a_p & 0 & & \ldots & 0 & a_{-r} & & & a_0
\end{bmatrix}. \tag{25}$$

The finite difference scheme (22)–(23) is spectral stable if and only if it is von Neumann stable. This statement comes directly from the well-known result that circulant matrices have eigenvalues which are shared with discrete Fourier modes [30].

Theorem 1. *The finite difference scheme* (22)–(23) *is von Neumann stable if and only if it is Lax stable.*

Proof. A circulant matrix is a normal matrix (see for instance [3]). If the matrix is normal then $\|Q_N\| = \rho(Q_N)$, where $\|\cdot\|$ is the 2-norm. Therefore if the finite difference scheme is von Neumann stable then $\|Q_N\| \leqslant 1$, for all N, since from the previous theorem $\rho(Q_N) \leqslant 1$. Also $\|Q_N^n\| \leqslant \|Q_N\|^n \leqslant 1$ and we conclude the scheme is Lax stable.

Conversely, we shall show that if the von Neumann condition is not satisfied then the scheme is not Lax stable.

We define the *Discrete Fourier Transform*:

$$\hat{U}(\theta) = \frac{1}{N}\sum_{j=0}^{N-1} U_j e^{ij\theta},$$

where $0 \leqslant \theta < 2\pi$, and $\theta = 2k\pi/N$, $k = 0, \dots, N-1$.

The *Rayleigh Energy Theorem* (Discrete Parseval's relation) says [13]

$$\sum_{j=0}^{N-1} |U_j|^2 = N \sum_{k=0}^{N-1} \left|\hat{U}(\theta_k)\right|^2.$$

Let us assume the scheme (22) is not von Neumann stable. Therefore the von Neumann condition is not satisfied, that is, suppose for each $K > 0$, there exists a θ_K, $0 \leqslant \theta_K < 2\pi$ such that

$$\left|\kappa(\theta_K)\right| > 1 + K\Delta t,$$

where $\kappa(\theta) = \sum_{k=-r}^{p} a_k e^{ik\theta}$. This proof follows an idea presented in Sod [21].

We have that $\kappa(\theta)$ is a continuous function of θ. Therefore there exists an interval I_K such that $\theta \in I_K$ and $|\kappa(\theta)| > 1 + K\Delta t$, $\forall \theta \in I_K$.

Consider initial data u_j^0 which is chosen so that the discrete Fourier transform $\hat{U}^0(\theta)$ of U_j^0 is equal to zero outside I_K. Thus, for $0 \leqslant \theta < 2\pi$ with $\hat{U}^0(\theta) \neq 0$, $|\kappa(\theta)| > 1 + K\Delta t$. We have that

$$\hat{U}^n(\theta) = Q\hat{U}^{n-1}(\theta) = \kappa(\theta)\hat{U}^{n-1}(\theta).$$

Therefore, it follows that

$$\hat{U}^n(\theta) = \kappa(\theta)\hat{U}^{n-1}(\theta) = \cdots = \kappa^n(\theta)\hat{U}^0(\theta).$$

Then

$$\hat{U}^n(\theta) > (1 + K\Delta t)^n \hat{U}^0(\theta).$$

Now using the Discrete Parseval's relation we have

$$\begin{aligned}
\|U^n\|^2 = \sum_{j=0}^{N-1} \left|U_j^n\right|^2 &= N \sum_{k=0}^{N-1} \left|\hat{U}^n(\theta_k)\right|^2 \\
&= N \sum_{k=0}^{N-1} \left|\kappa(\theta_k)\right|^{2n} \left|\hat{U}^0(\theta_k)\right|^2 \\
&> (1 + K\Delta t)^{2n} N \sum_{k=0}^{N-1} \left|\hat{U}^0(\theta_k)\right|^2 \\
&= (1 + K\Delta t)^{2n} \left\|U^0\right\|^2.
\end{aligned}$$

We have $U^n = Q^n U^0$ and then

$$\left\|Q^n U^0\right\|^2 > (1 + K\Delta t)^{2n} \left\|U^0\right\|^2$$

or

$$\frac{\|Q^n U^0\|^2}{\|U^0\|^2} > (1 + K\Delta t)^{2n}.$$

But

$$\begin{aligned}\|Q^n\|^2 &= \max_{\|U\| \neq 0} \frac{\|Q^n U\|^2}{\|U\|^2} > \frac{\|Q^n U^0\|^2}{\|U^0\|^2} \\ &> (1 + K\Delta t)^{2n}.\end{aligned}$$

Finally we have

$$\|Q^n\|^2 > (1 + K\Delta t)^{2n},$$

for all constants $K > 0$ and therefore $\|Q^n\|$ is unbounded which implies the scheme is not Lax stable. □

5.2. Numerical boundary conditions

Suppose we have approximations U_j^n as described in Section 2, given by the difference scheme

$$U_j^{n+1} = QU_j^n, \quad j = r, r+1, \ldots, \tag{26}$$

$$Q = \sum_{j=-r}^{p} a_j \mathcal{E}^j, \qquad \mathcal{E}^a U_j^n = U_{j+a}^n \tag{27}$$

and the numerical boundary conditions

$$U_\beta^{n+1} = \sum_{j=0}^{q} b_{\beta j} U_j^n, \quad \beta = 1, \ldots, r-1. \tag{28}$$

We consider $p \leqslant q$. Usually $q = p + r$. Also in what follows $N > p + r$.

Assume that the matricial form of the scheme is

$$\mathbf{U}^{n+1} = A_N \mathbf{U}^n, \tag{29}$$

and the iterative matrix A_N is now given by

$$A_N = \begin{bmatrix} B_{r\times N} \\ Q_{(N-(r+p))\times N} \\ F_{p\times N} \end{bmatrix} \tag{30}$$

where B contains the part of the boundary conditions and Q the part of the interior scheme, that is

$$U_\beta^{n+1} = B_\beta U^n, \quad \beta = 0, \ldots, r-1,$$
$$U_j^{n+1} = Q_j U^n, \quad j = r, \ldots, N-(r+p).$$

Note that B_β represents the β-row and Q_j represents the j-row. The matrix F represents the discrete points computed using (26) and takes in consideration we are assuming the solution $u(x,t)$ is equal to zero as x goes to infinity, that is, we assume that

$$U_N^n = U_{N+1}^n = \cdots = U_{N+p-1}^n = 0.$$

Note that in most applications r and p take the values 1, 2 or 3.

Explicitly the matrix $B_{r\times N}$ is given by

$$B_{r\times N} = \begin{bmatrix} b_{00} & b_{01} & \ldots & b_{0q} & 0 & \ldots & 0 \\ b_{10} & b_{11} & \ldots & b_{1q} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{r-10} & b_{r-11} & \ldots & b_{r-1q} & 0 & \ldots & 0 \end{bmatrix},$$

$Q_{(N-(r+p))\times N}$ is of the form

$$\begin{bmatrix} a_{-r} & a_{-r+1} & \dots & \dots & a_0 & \dots & \dots & a_{p-1} & a_p & 0 & \dots & \dots & 0 \\ 0 & a_{-r} & a_{-r+1} & \dots & \dots & a_0 & \dots & \dots & a_{p-1} & a_p & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & & & \ddots & & & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{-r} & a_{-r+1} & \dots & \dots & a_0 & \dots & & a_{p-1} & a_p & 0 \\ 0 & \dots & \dots & 0 & a_{-r} & a_{-r+1} & \dots & \dots & a_0 & \dots & & a_{p-1} & a_p \end{bmatrix}$$

and the matrix $F_{p\times N}$ is

$$\begin{bmatrix} 0 & \dots & \dots & 0 & a_{-r} & a_{-r+1} & \dots & a_{-1} & a_0 & \dots & \dots & \dots & a_{p-1} \\ 0 & \dots & \dots & \dots & 0 & a_{-r} & a_{-r+1} & \dots & a_{-1} & a_0 & \dots & \dots & a_{p-2} \\ \vdots & & & & & \ddots & \ddots & \ddots & & \ddots & \ddots & & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & a_{-r} & a_{-r+1} & \dots & a_{-1} & a_0 & a_1 \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & a_{-r} & a_{-r+1} & \dots & a_{-1} & a_0 \end{bmatrix}.$$

In what follows we prove properties that relates Lax stability with spectral and von Neumann stability.

Theorem 2. *If A_N is power-bounded for all n, then $\rho(A_N) \leqslant 1$.*

Proof. If A_N is power-bounded for all n, then $\|A_N^n\| \leqslant K$, for $n = 1, 2, \dots$ where K is independent of n, Δt and Δx. We know that $\rho(A_N) \leqslant \|A_N\|$. Then

$$\rho(A_N^n) \leqslant \|A_N^n\| \leqslant K, \quad n = 1, 2, \dots.$$

But $\rho(A_N^n) = \rho^n(A_N)$, hence $\rho^n(A_N) \leqslant K$, $n = 1, \dots$, and therefore $\rho(A_N) \leqslant K^{1/n}$. Then $\lim_{n\to\infty} K^{1/n} = 1$. □

Remark. If a scheme is Lax stable then $\|A^n\| \leqslant K$, for $n = 1, \dots, T/\Delta t$ with $0 < \Delta t < \tau$ and when $\Delta t \to 0$ then $n \to \infty$. Note also that if $\lim_{n\to\infty} A^n = 0$, the eigenvalues of A on the unit circle are simple.

Theorem 3. *If the interior scheme is von Neumann stable then*

$$\|A_N^n\| \leqslant e^{n\|L_N\|}, \quad \text{where } L_N = A_N - Q_N,$$

for A_N and Q_N $N \times N$ matrices given by (30) *and* (25)*, respectively.*

Proof. We prove by induction that

$$\|(Q_N + L_N)^n\| \leqslant e^{n\|L_N\|}.$$

For $n = 1$ is true because

$$\|Q_N + L_N\| \leqslant \|Q_N\| + \|L_N\| \leqslant 1 + \|L_N\| \leqslant e^{\|L_N\|}.$$

It can be proved by induction that

$$(Q_N + L_N)^n = Q_N^n + \sum_{k=0}^{n-1} Q_N^{n-k-1} L_N (Q_N + L_N)^k, \quad n = 1, 2, \dots.$$

Therefore

$$\|(Q_N + L_N)^n\| \leqslant \|Q_N^n\| + \sum_{k=0}^{n-1} \|Q_N^{n-k-1}\| \|L_N\| \|(Q_N + L_N)^k\|.$$

By assumption, the interior scheme is von Neumann stable. Therefore, $\|Q_N^n\| \leqslant 1$ and $\|Q_N^{n-k-1}\| \leqslant 1$, for $k = 0, \dots, n$. By the inductive hypothesis, $\|(Q_N + L_N)^k\|$ is bounded, that is, $\|(Q_N + L_N)^k\| \leqslant e^{k\|L_N\|}$.

We have

$$\begin{aligned}\left\|(Q_N+L_N)^n\right\| &\leqslant 1+\sum_{k=0}^{n-1}\|L_N\|\mathrm{e}^{k\|L_N\|}\\ &\leqslant 1+\|L_N\|\sum_{k=0}^{n-1}\mathrm{e}^{k\|L_N\|}.\end{aligned}$$

Since

$$\mathrm{e}^{n\|L_N\|}\geqslant 1+\|L_N\|\sum_{k=0}^{n-1}\mathrm{e}^{k\|L_N\|},$$

it follows

$$\left\|(Q_N+L_N)^n\right\|\leqslant \mathrm{e}^{n\|L_N\|}. \qquad \square$$

Theorem 4. *Assume the interior scheme is von Neumann stable. If A_N verifies the spectral stability condition* (14)*, that is, $\rho(A_N)\leqslant 1$, where the eigenvalues of A_N on the unit circle must be simple, then $\|A_N^n\|\leqslant \mathrm{e}^{k_N\|L_N\|}$, for all n and where k_N is a constant.*

Proof. We begin by noting that we have the property [26] $\lim_{k\to\infty}\|A_N^k\|^{1/k}=\rho(A_N)$. If A_N verifies the spectral condition (14), we can easily obtain that, there is a k_N such that $\|A_N^k\|\leqslant 1$, for all $k\geqslant k_N$. Therefore, from Theorem 4, follows

$$\left\|A_N^n\right\|\leqslant \mathrm{e}^{k_N\|L_N\|}, \quad \text{for all } n. \qquad \square$$

We have that von Neumann and spectral stability are sufficient conditions for $\|A_N^n\|\leqslant K$ for all n and a fixed N.

How can we prove that $\|A_N^n\|\leqslant K$ for all n and for all $N>p+r$? Note that N is the size of the matrix A_N and is directly related with the space step Δx, such that, $\Delta x\to 0$ as $N\to\infty$.

Remark. If A_N verifies the spectral condition (14) for all $N>p+r$, there exists $c=\max_{N>p+r}k_N$ such that $\|A_N^n\|\leqslant \mathrm{e}^{c\|L_N\|}$, for all $N>p+r$.

Theorem 5. *For L_N such that $L_N=A_N-Q_N$ we have*

$$\|L_N\|=\|L_{N+1}\|, \quad \textit{for all } N>p+r.$$

Proof. Note that

$$L_N=A_N-Q_N=\begin{bmatrix} B_{r\times N}-Q_{r\times N}\\ O_{(N-(r+p))\times N}\\ F_{p\times N}-Q_{p\times N}\end{bmatrix},$$

where $O_{(N-(r+p))\times N}$ is the null matrix or zero matrix. The block matrix $Q_{r\times N}$ represent the first r lines of the matrix Q given by (25) and $Q_{p\times N}$ the last p lines.

Also,

$$L_{N+1}=A_{N+1}-Q_{N+1}=\begin{bmatrix} B_{r\times N+1}-Q_{r\times N+1}\\ O_{(N+1-(r+p))\times N+1}\\ F_{p\times N+1}-Q_{p\times N+1}\end{bmatrix}.$$

The difference between $B_{r\times N}-Q_{r\times N}$ and $B_{r\times N+1}-Q_{r\times N+1}$ is that the latter has an additional column of zeros. The same difference occur between the matrices $F_{p\times N}-Q_{p\times N}$ and $F_{p\times N+1}-Q_{p\times N+1}$. Therefore it is easy to conclude that $\|L_N\|=\|L_{N+1}\|$. Note that usually in practical applications the block matrices B and F does not have more than three lines. $\square$

From the previous results we have the following theorem.

Theorem 6. *If the interior scheme is von Neumann stable and if for all $N > p + r$, A_N verifies the spectral condition* (14) *and $\|L_N\| \leqslant C$, then $\|A_N^n\| \leqslant K$, for all n and for all $N > p + r$.*

The condition $\|L_N\| \leqslant C$ for all $N > p + r$ is satisfied in several practical applications when the von Neumann and spectral conditions are verified. We believe this is the reason we can rely in both conditions to assure Lax stability in the various examples mentioned in literature [16,22,31–33]. For instance if we go back to the first motivation example in Section 4.1 we have

$$L_N = \begin{bmatrix} 0 & \dots & -\sigma \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{bmatrix}$$

and $\|L_N\| = \sigma$. From the von Neumann and spectral analysis we have $0 < \sigma \leqslant 1$ and therefore $\|L_N\| \leqslant 1$, for all $N > p + r$.

6. Conclusion

The aim of this paper is to present some properties that relates the von Neumann analysis, the spectral analysis and the Lax theory in order to give a practical way to verify when a finite difference scheme is stable even if the matrix iteration is not normal. To control the growing of the norm of the power matrices is usually more difficult than to verify the von Neumann condition for the interior scheme and to calculate the maximum eigenvalue of the matrix iteration.

References

[1] R. Courant, K. Friedrichs, H. Levy, Uber die partielle differentialgleichungen der matematischen physik, Math. Annal. 100 (1928) 32–74.
[2] J. Crank, P. Nicolson, A practical method for numerical evaluation of solutions of partial differential equations of heat conduction type, Proc. Cambridge Philos. Soc. 43 (1947) 50–67.
[3] P.J. Davis, Circulant Matrices, Wiley-Interscience, New York, 1979.
[4] S.K. Godunov, V.S. Ryabenkii, Spectral criteria for the stability of boundary problems for non-self-adjoint difference equations, Uspekhi Mat. Nauk 18 (1963) 1–12.
[5] S.K. Godunov, V.S. Ryabenkii, The Theory of Difference Schemes, North-Holland, Amsterdam, 1964.
[6] B. Gustafsson, H.-O. Kreiss, A. Sundstrom, Stability theory of difference approximations for mixed initial boundary value problems, Math. Comp. 26 (1972) 649–686.
[7] B. Gustafsson, H.-O. Kreiss, J. Oliger, The Time Dependent Problems and Difference Methods, Wiley-Interscience, New York, 1995.
[8] A.C. Hindmarsh, P.M. Gresho, D.F. Griffiths, The stability of explicit Euler time-integration for certain finite difference approximations of the multi-dimensional advection-diffusion equation, International Journal for Numerical Methods in Fluids 4 (1984) 853–897.
[9] C. Hirsch, Numerical Computation of Internal and External Flows, Wiley-Interscience, New York, 1988.
[10] S. Osher, Systems of difference equations with general homogeneous boundary conditions, Trans. Amer. Math. Soc. 137 (1969) 177–201.
[11] B.P. Leonard, A stable and accurate convective modeling procedure based on quadratic upstream interpolation, Comput. Methods Appl. Mech. Engrg. 19 (1979) 59–98.
[12] D. Michelson, Stability theory of difference approximations for multidimensional initial-boundary value problems, Math. Comp. 40 (1983) 1–45.
[13] D.W. Kammler, A First Course in Fourier Analysis, Prentice-Hall, New Jersey, 2000.
[14] H.-O. Kreiss, Stability theory for difference approximations of mixed initial boundary value problems I, Math. Comp. 22 (1968) 703–714.
[15] K.W. Morton, D.F. Mayers, Numerical Solution of Partial Differential Equations, Cambridge University Press, Cambridge, 1996.
[16] M.M. Rai, S. Chakravarthy, Conservative high-order-accurate finite difference methods for curvilinear grids, AIAA Paper 93-3380, 1993.
[17] S.C. Reddy, L.N. Trefethen, Stability of the method of lines, Numer. Math. 62 (1992) 235–267.
[18] R.D. Richtmyer, K.W. Morton, Difference Methods for Initial-Value Problems, Wiley-Interscience, New York, 1967.
[19] G.D. Smith, Numerical Solution of Partial Differential Equations: Finite Difference Methods, Oxford University Press, Oxford, 1985.
[20] J. Strikwerda, Initial boundary value problems for the method of lines, J. Comput. Phys. 34 (1980) 94–107.
[21] G.A. Sod, Numerical Methods in Fluid Dynamics: Initial and Initial Boundary-Value Problems, Cambridge University Press, Cambridge, 1988.
[22] E. Sousa, I.J. Sobey, On the influence of numerical boundary conditions, Appl. Numer. Math. 41 (2002) 325–344.
[23] E. Sousa, Stability analysis of difference methods for parabolic initial value problems, J. Sci. Comput. 26 (2006) 45–66.
[24] M. Thuné, Automatic GKS stability analysis, SIAM J. Sci. Stat. Comput. 7 (1986) 959–977.

[25] M. Thuné, A numerical algorithm for stability analysis of difference methods for hyperbolic systems, SIAM J. Sci. Stat. Comput. 11 (1990) 63–81.

[26] L.N. Trefethen, M. Embree, Spectra and Pseudospectra, Princeton University Press, New Jersey, 2005.

[27] L.N. Trefethen, Group velocity interpretation of the stability theory of Gustafsson, Kreiss and Sundstrom, J. Comput. Phys. 49 (1983) 199–217.

[28] L.N. Trefethen, Instability of difference models for hyperbolic initial boundary value problems, Comm. Pure Appl. Math. 37 (1984) 329–367.

[29] J.M. Varah, Stability of difference approximations to the mixed initial boundary value problems for parabolic systems, SIAM J. Numer. Anal. 8 (1971) 598–615.

[30] R. Vichnevestsky, J.B. Bowles, Fourier Analysis of Numerical Approximations of Hyperbolic Equations, SIAM, Philadelphia, 1982.

[31] D.W. Zingg, Aspects of linear stability analysis of higher-order finite difference methods, AIAA Paper 97-1939, 1997.

[32] D.W. Zingg, M. Lederle, On linear stability analysis of high-order finite difference methods, AIAA Paper 2005-5249, 2005.

[33] D.W. Zingg, H. Lomax, H. Jurgens, High-accuracy finite difference schemes for linear wave propagation, SIAM J. Sci. Comput. 17 (1996) 328–346.