SEMIGROUP STABILITY OF FINITE DIFFERENCE SCHEMES FOR MULTIDIMENSIONAL HYPERBOLIC INITIAL BOUNDARY VALUE PROBLEMS

Jean-François Coulombel

CNRS, Université Lille 1 and Team Project SIMPAF of INRIA Lille - Nord Europe
Laboratoire Paul Painlevé, Bâtiment M2, Cité Scientifique
59655 Villeneuve d'Ascq Cedex, France

Antoine Gloria

Team Project SIMPAF of INRIA Lille - Nord Europe
Park Plazza, 40 avenue Halley, 59655 Villeneuve d'Ascq Cedex, France

Abstract. We develop a simple energy method to prove the stability of finite difference schemes for multidimensional hyperbolic initial boundary value problems. In particular we extend to several space dimensions a crucial result by Goldberg and Tadmor. This allows us to give two conditions on the discretized operator that ensure that stability estimates for zero initial data imply a semigroup stability estimate for general initial data. We then apply this criterion to several numerical schemes in two space dimensions.

1. Introduction

The aim of this article is to prove stability estimates for finite difference discretizations of hyperbolic initial boundary value problems. A general theory to derive such estimates has been developed in [4] for one-dimensional problems and later extended in [7] to multidimensional problems. The analysis for the discretized equations is similar to the theory in [5] for the continuous problem (namely for hyperbolic systems of partial differential equations), and relies on the so-called normal modes analysis. Due to the fact that the method uses a Laplace transform in time, the estimates in [4, 5, 7] are restricted to zero initial data. A natural question is then to show that problems that are stable for zero initial data are also stable for non-zero initial data. Making the space of "suitable" initial data precise is part of the question.

For continuous problems, this question was solved in [8]. (We also refer to [1, chapter 4] for a complete description of the results.) For discretized problems, the question was solved in [13] where the author proves stability for non-zero initial data in one space dimension. The proof in [13] relies on a crucial Lemma (Lemma 2.3 in [2]), that we shall refer to as Goldberg-Tadmor's Lemma. Goldberg-Tadmor's Lemma shows that the Dirichlet boundary conditions yield stable problems for discretized *scalar* equations and zero initial data. Consequently, it is not clear whether Goldberg-Tadmor's Lemma, and therefore the results of [13], extends to hyperbolic systems in several space dimensions (because systems usually do not decouple into a collection of scalar equations). In this article, we develop a simple energy method with which we recover Goldberg-Tadmor's Lemma and the stability results of [13] and that is flexible enough to handle discretized multidimensional systems. Opposite to the original proof in [2], our new proof of Goldberg-Tadmor's Lemma covers the case of non-zero initial data in ℓ^2. Once we have obtained a stability estimate for Dirichlet boundary conditions and non-zero initial data, it is almost straightforward to show that discretizations that are stable for zero initial data are also stable for non-zero initial data. Note that, although this is an old problem, Theorem 3.1 below seems to be the first general stability result for discretized initial boundary problems in several space dimensions.

1991 *Mathematics Subject Classification.* Primary: 65M12; Secondary: 65M06, 35L50.

Key words and phrases. Hyperbolic systems, boundary value problems, finite difference schemes, stability.

Research of the first author was supported by the Agence Nationale de la Recherche, contract ANR-08-JCJC-0132-01.

We first develop our approach on one-dimensional problems. The one-dimensional analysis is extended to multidimensional problems in Section 3 using a partial Fourier transform in the tangential space variables. Then in Section 4, we give examples of discretizations to which our analysis applies.

Notation. In all this paper, we let $\mathscr{M}_{d,D}(\mathbb{K})$ denote the set of $d \times D$ matrices with entries in $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and we use the notation $\mathscr{M}_D(\mathbb{K})$ when $d = D$. If $M \in \mathscr{M}_D(\mathbb{C})$, $\mathrm{sp}(M)$ denotes the spectrum of M, $\rho(M)$ denotes the spectral radius of M, while M^* denotes the conjugate transpose of M. We let I denote the identity matrix, without mentioning the dimension. The norm of a vector $x \in \mathbb{C}^D$ is $|x| := (x^*\, x)^{1/2}$. Eventually, we let ℓ^2 denote the set of square integrable sequences, without mentioning the indices of the sequences (sequences may be valued in $\mathbb{C}^d$ for some integer d).

The letter C denotes a constant that may vary from line to line or within the same line. The dependence of the constants on the various parameters is made precise throughout the text.

2. One-dimensional problems

For one-dimensional problems, we introduce the following notation for norms on $\ell^2(\mathbb{Z})$. Let $\Delta x > 0$ be a space step. For all integers $m_1 \le m_2$, we set

$$\|u\|^2_{m_1,m_2} := \Delta x \sum_{j=m_1}^{m_2} |u_j|^2$$

to denote the ℓ^2-norm on the interval $[m_1, m_2]$ (m_1 may equal $-\infty$ and m_2 may equal $+\infty$). The corresponding scalar product is denoted by $(\cdot,\cdot)_{m_1,m_2}$.

2.1. Main result in one space dimension.

We consider a hyperbolic initial boundary value problem in one space dimension:

$$\begin{cases} \partial_t u + A\,\partial_x u = F(t,x)\,, & (t,x) \in \mathbb{R}^+ \times \mathbb{R}^+\,, \\ B\,u(t,0) = g(t)\,, & t \in \mathbb{R}^+\,, \\ u(0,x) = f(x)\,, & x \in \mathbb{R}^+\,, \end{cases} \tag{1}$$

where $A \in \mathscr{M}_D(\mathbb{R})$ is diagonalizable with real eigenvalues, and $B \in \mathscr{M}_{D_+,D}(\mathbb{R})$ with D_+ the number of positive eigenvalues of A (counted with their multiplicity). We assume that the boundary is noncharacteristic, that is $0 \notin \mathrm{sp}(A)$. Problem (1) is well-posed in any suitable sense if and only if:

$$\mathbb{R}^D = \text{ Ker } B \oplus E_+(A)\,,$$

where $E_+(A)$ is the unstable eigenspace of A (associated with positive eigenvalues of A). In that case, the solution u to (1) belongs to $\mathscr{C}(\mathbb{R}^+; L^2(\mathbb{R}^+))$ and its trace on $\{x = 0\}$ is well-defined and belongs to $\mathrm{e}^{\gamma\,t}\, L^2(\mathbb{R}^+)$ for all $\gamma > 0$. Moreover, for every parameter $\gamma > 0$, u satisfies the energy estimate:

$$\begin{aligned} \sup_{t\ge 0} \mathrm{e}^{-2\,\gamma\,t}\, \|u(t,\cdot)\|^2_{L^2(\mathbb{R}^+)} + \gamma \int_0^{+\infty} \mathrm{e}^{-2\,\gamma\,t}\, \|u(t,\cdot)\|^2_{L^2(\mathbb{R}^+)}\, \mathrm{d}t + \int_0^{+\infty} \mathrm{e}^{-2\,\gamma\,t}\, |u(t,0)|^2\, \mathrm{d}t \\ \le C \left(\|f\|^2_{L^2(\mathbb{R}^+)} + \frac{1}{\gamma} \int_0^{+\infty} \mathrm{e}^{-2\,\gamma\,t}\, \|F(t,\cdot)\|^2_{L^2(\mathbb{R}^+)}\, \mathrm{d}t + \int_0^{+\infty} \mathrm{e}^{-2\,\gamma\,t}\, |g(t)|^2\, \mathrm{d}t \right)\,, \end{aligned} \tag{2}$$

where the constant C is independent of γ, f, F, g. The estimate (2) can be localized on any finite time interval $[0, T]$ because the solutions to (1) satisfy a causality principle ("future does not affect the past").

We now introduce the finite difference approximation of (1). Let $\Delta x, \Delta t > 0$ denote the space and time steps, where the ratio $\lambda = \Delta t/\Delta x$ is a fixed positive constant, and let p, q, r be some integers. The solution u to (1) is approximated by a sequence (U_j^n) defined for $n \in \mathbb{N}$, and $j \in 1 - r + \mathbb{N}$. For $j = 1 - r, \dots, 0$, U_j^n approximates the trace $u(n\,\Delta t, 0)$ on the boundary $\{x = 0\}$, and possibly the trace of normal derivatives. The boundary meshes $[j\,\Delta x, (j+1)\,\Delta x[$, $j = 1 - r, \dots, 0$, shrink to $\{0\}$

as Δx tends to 0. Hence the "formal" limit problem as Δx tends to 0 is set on the half-line $\mathbb{R}^+$. We consider one-step finite difference approximations of (1) that read[1]:

$$\begin{cases} U_j^{n+1} = Q\,U_j^n + \Delta t\,F_j^n\,, & j \ge 1\,, \quad n \ge 0\,, \\ U_j^{n+1} = B_{j,-1}\,U_1^{n+1} + B_{j,0}\,U_1^n + g_j^{n+1}\,, & j = 1-r,\dots,0\,, \quad n \ge 0\,, \\ U_j^0 = f_j\,, & j \ge 1-r\,, \end{cases} \tag{3}$$

where the operators Q, $B_{j,-1}, B_{j,0}$ are given by:

$$Q := \sum_{\ell=-r}^{p} A_\ell\,T^\ell\,, \quad B_{j,\sigma} := \sum_{\ell=0}^{q} B_{\ell,j,\sigma}\,T^\ell\,, \qquad T^\ell\,U_k^m := U_{k+\ell}^m\,. \tag{4}$$

In (4), all matrices $A_\ell, B_{\ell,j,\sigma}$ belong to $\mathscr{M}_D(\mathbb{R})$ and depend on λ, A, B but not on Δt (or equivalently Δx). We recall the following definition from [4]:

Definition 2.1 (Strong stability [4])**.** *The finite difference approximation (3) is said to be strongly stable if there exists a constant C such that for all $\gamma > 0$ and all $\Delta t \in]0,1]$, the solution (U_j^n) of (3) with $f = 0$ satisfies the estimate:*

$$\frac{\gamma}{\gamma\,\Delta t + 1} \sum_{n \ge 0} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,\|U^n\|_{1-r,+\infty}^2 + \sum_{n \ge 0} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=1-r}^{0} |U_j^n|^2$$
$$\le C\,\left\{ \frac{\gamma\,\Delta t + 1}{\gamma} \sum_{n \ge 0} \Delta t\,\mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\,\|F^n\|_{1,+\infty}^2 + \sum_{n \ge 1} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=1-r}^{0} |g_j^n|^2 \right\}\,.$$

The estimate in Definition 2.1 is the discrete counterpart of the energy estimate (2) for the "continuous" problem (1) in the case of zero initial data (and when one does not require to control the $L_t^\infty(L_x^2)$ norm of the solution). We recall that strong stability in the sense of Definition 2.1 is usually proved by performing a Laplace transform with respect to the time variable. The energy estimate for the resolvent equation is then obtained by using symmetrizers whose construction relies on the so-called *uniform Kreiss-Lopatinskii condition* (non-existence of unstable nor weakly unstable normal modes). We refer to [4] for some results in this direction. In this paper, we shall consider that the scheme (3) is strongly stable, and we wish to prove an energy estimate for (3) in the case of non-zero initial data. In view of (2), the most obvious space of initial data for (3) is ℓ^2. Let us now introduce our main assumptions, and then state our result.

For $\ell = -r,\dots,p$, and $z \in \mathbb{C} \setminus \{0\}$, let us define the matrices:

$$\mathbb{A}_\ell(z) := \delta_{\ell 0}\,I - \frac{1}{z}\,A_\ell\,, \tag{5}$$

where $\delta_{\ell_1 \ell_2}$ is the Kronecker symbol. We make the following assumption[2]:

Assumption 2.1. *The matrix $\mathbb{A}_p(z)$ is invertible for all $z \in \mathbb{C}$ with $|z| \ge 1$.*

Our second crucial assumption is the following:

Assumption 2.2. *The operator Q satisfies $\|Q\,v\|_{-\infty,+\infty} \le \|v\|_{-\infty,+\infty}$ for all $v \in \ell^2$.*

Our main result is stated as follows:

Theorem 2.1. *Let Assumptions 2.1 and 2.2 be satisfied, and assume that the scheme (3) is strongly stable in the sense of Definition 2.1. Then there exists a constant C such that for all $\gamma > 0$ and all*

[1]We do not focus here on the construction of such approximations and refer to [3] for some examples that enter this framework, see also Section 4.

[2]Assumption 2.1 is similar to Assumption 5.5 in [4].

$\Delta t \in]0,1]$, *the solution U to* (3) *satisfies the estimate:*

$$\sup_{n\geq 0} \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,\|U^n\|^2_{1-r,+\infty} + \frac{\gamma}{\gamma\,\Delta t+1}\sum_{n\geq 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,\|U^n\|^2_{1-r,+\infty} + \sum_{n\geq 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{j=1-r}^{0}|U_j^n|^2$$
$$\leq C\left\{\|f\|^2_{1-r,+\infty} + \frac{\gamma\,\Delta t+1}{\gamma}\sum_{n\geq 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\,\|F^n\|^2_{1,+\infty} + \sum_{n\geq 1}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{j=1-r}^{0}|g_j^n|^2\right\}. \tag{6}$$

The method of proof is inspired from [13] with however some important modifications. More precisely, we shall introduce an auxiliary discretized problem where we modify the boundary operators $B_{j,-1}, B_{j,0}$. The auxiliary problem is chosen in such a way that even for non-zero initial data, the solution can be estimated by applying the energy method (in some sense, the auxiliary boundary conditions are "strictly dissipative"). Our auxiliary problem is not the same as in [13]. As a matter of fact, we shall show directly that the Dirichlet boundary conditions are strictly dissipative. As announced in the introduction, this is an improved version of Goldberg-Tadmor's Lemma. Our new proof of Goldberg-Tadmor's Lemma relies on the energy method and can therefore be extended to multidimensional problems even if the equation is not scalar ($D \geq 2$), see Section 3. Once we have the estimate for the auxiliary boundary conditions, the end of the proof follows [13], see also the arguments in [1, chapter 4] for the continuous problem.

2.2. **A refined version of Goldberg-Tadmor's Lemma.** In this paragraph, we consider the following auxiliary discretization where the (non-homogeneous) Dirichlet conditions are enforced at the boundary:

$$\begin{cases} V_j^{n+1} = Q\,V_j^n + \Delta t\,F_j^n\,, & j\geq 1\,,\quad n\geq 0\,,\\ V_j^{n+1} = g_j^{n+1}\,, & j = 1-r,\dots,0\,,\quad n\geq 0\,,\\ V_j^0 = f_j\,, & j\geq 1-r\,. \end{cases} \tag{7}$$

The aim of this paragraph is to prove the following:

Theorem 2.2. *Let Assumptions 2.1 and 2.2 be satisfied. Then there exists a constant C such that for all $\gamma > 0$ and all $\Delta t \in]0,1]$, the solution V to* (7) *satisfies the estimate:*

$$\sup_{n\geq 0} \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,\|V^n\|^2_{1-r,+\infty} + \frac{\gamma}{\gamma\,\Delta t+1}\sum_{n\geq 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,\|V^n\|^2_{1-r,+\infty}$$
$$+\sum_{n\geq 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{j=1-r}^{\max(p,q+1)}|V_j^n|^2 \leq C\left\{\|f\|^2_{1-r,+\infty} + \frac{\gamma\,\Delta t+1}{\gamma}\sum_{n\geq 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\,\|F^n\|^2_{1,+\infty}\right.$$
$$\left.+\sum_{n\geq 1}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{j=1-r}^{0}|g_j^n|^2\right\}. \tag{8}$$

In particular, the discretization (7) *is strongly stable in the sense of Definition 2.1.*

The key point in the proof of Theorem 2.2 is Assumption 2.2, which can be understood as a symmetry assumption on the matrices A_ℓ. We emphasize that Assumption 2.2 can be extended to multidimensional systems, as we shall see in section 3, while the "scalar" assumption in [2] does not extend to general multidimensional systems.

The proof of Theorem 2.2 is split in several steps. We first observe that the solution V to (7) depends linearly on the source terms (f,g,F). It is therefore sufficient to prove (8) in the case $F=0$ (no source term in the interior equation) and in the case $(f,g)=0$ (zero initial data, and homogeneous boundary conditions). It turns out that we use slightly different arguments for the two cases. We begin with the case $F=0$, and then treat the case $(f,g)=0$.

We point out that in (8), we estimate the (weighted) ℓ^2-norm in time of the trace $(V_j^n)_{n\geq 0}$, for all j from $1-r$ to $\max(p,q+1)$. As a matter of fact, it would have been sufficient for the proof of Theorem 2.1 to have this type of estimate up to $j=q+1$. However, in the proof below, we shall first obtain an estimate up to $j=p$ (see Corollary 2.1 below). This is the reason why we have stated (8) in this way. We start the proof of Theorem 2.2 proper with a series of preliminary results:

Lemma 2.1. *Let Assumptions 2.1 and 2.2 be satisfied. Then there exists a constant C such that for all $\gamma>0$ and all $\Delta t\in\,]0,1]$, the solution V to (7) with $F=0$ satisfies the estimate:*

$$\begin{aligned}\mathrm{e}^{2\,\gamma\,\Delta t}\sup_{n\ge 1}\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,\|V^n\|^2_{1,+\infty}+\gamma\sum_{n\ge 1}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,\|V^n\|^2_{1,+\infty}+\sum_{n\ge 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{j=1-r}^{p}|V_j^n|^2\\ \le C\left\{\|f\|^2_{1-r,+\infty}+\sum_{n\ge 1}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{j=1-r}^{0}|g_j^n|^2\right\}.\qquad(9)\end{aligned}$$

Proof of Lemma 2.1. We decompose the operator Q as:

$$Q:=I+\widetilde{Q}\,.$$

Then Assumption 2.2 is equivalent to the inequality:

$$\forall\, w\in\ell^2\,,\quad 2\,\big(w,\widetilde{Q}\,w\big)_{-\infty,+\infty}+\|\widetilde{Q}\,w\|^2_{-\infty,+\infty}\le 0\,.\qquad(10)$$

We first use the relation $V_j^{n+1}=(I+\widetilde{Q})\,V_j^n$ for $j\ge 1$ (recall that $F=0$), and derive:

$$\|V^{n+1}\|^2_{1,+\infty}-\|V^n\|^2_{1,+\infty}=2\,\big(V^n,\widetilde{Q}\,V^n\big)_{1,+\infty}+\|\widetilde{Q}\,V^n\|^2_{1,+\infty}\,.\qquad(11)$$

For a fixed integer n, we introduce the sequence (W_j) such that $W_j=V_j^n$ for $j\ge 1-r$ and $W_j=0$ for $j\le -r$. Due to the structure of the operator $\widetilde{Q}$, see (4), we have $\widetilde{Q}\,W_j=0$ if $j\le -r-p$, and $\widetilde{Q}\,W_j=\widetilde{Q}\,V_j^n$ if $j\ge 1$. Using (10), we thus get:

$$\begin{aligned}0&\ge 2\,\big(W,\widetilde{Q}\,W\big)_{-\infty,+\infty}+\|\widetilde{Q}\,W\|^2_{-\infty,+\infty}\\ &=2\,\big(V^n,\widetilde{Q}\,W\big)_{1-r,0}+2\,\big(V^n,\widetilde{Q}\,V^n\big)_{1,+\infty}+\|\widetilde{Q}\,W\|^2_{1-r-p,-r}+\|\widetilde{Q}\,W\|^2_{1-r,0}+\|\widetilde{Q}\,V^n\|^2_{1,+\infty}\\ &=2\,\big(V^n,\widetilde{Q}\,V^n\big)_{1,+\infty}+\|\widetilde{Q}\,V^n\|^2_{1,+\infty}+\|V^n+\widetilde{Q}\,W\|^2_{1-r,0}+\|\widetilde{Q}\,W\|^2_{1-r-p,-r}-\|V^n\|^2_{1-r,0}\,.\qquad(12)\end{aligned}$$

We insert (12) into (11) and obtain:

$$\|V^{n+1}\|^2_{1,+\infty}-\|V^n\|^2_{1,+\infty}+\|\widetilde{Q}\,W\|^2_{1-r-p,-r}+\|V^n+\widetilde{Q}\,W\|^2_{1-r,0}\le\|V^n\|^2_{1-r,0}\,.\qquad(13)$$

At this point, two situations may occur depending on p. Let us first consider the case $p\ge 1$. Then, by Assumption 2.1, A_p is an invertible matrix. We have the following:

Lemma 2.2. *Let $p\ge 1$ and let A_p be invertible. Then there exists a constant $c>0$ that does not depend on Δt nor on V^n such that the following estimate holds:*

$$\|\widetilde{Q}\,W\|^2_{1-r-p,-r}+\|V^n+\widetilde{Q}\,W\|^2_{1-r,0}\ge c\,\|V^n\|^2_{1-r,p}\,.$$

Let us assume that Lemma 2.2 holds and go back to (13). We have:

$$\|V^{n+1}\|^2_{1,+\infty}-\|V^n\|^2_{1,+\infty}+c\,\Delta x\sum_{j=1-r}^{p}|V_j^n|^2\le\Delta x\sum_{j=1-r}^{0}|V_j^n|^2\,.\qquad(14)$$

The end of the proof consists in integrating (14) over $\mathbb{N}$, see a similar calculation in the continous case in [1, page 95]. Let $\gamma>0$, and for the sake of clarity, let us introduce the following notation:

$$\mathscr{V}_n:=\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,\|V^n\|^2_{1,+\infty}\,,\quad\mathscr{B}_n:=\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{j=1-r}^{p}|V_j^n|^2\,,\quad\mathscr{G}_n:=\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{j=1-r}^{0}|V_j^n|^2\,.$$

We multiply (14) by $\exp(-2\,\gamma\,n\,\Delta t)$:

$$\mathrm{e}^{2\,\gamma\,\Delta t}\,\mathscr{V}_{n+1}-\mathscr{V}_n+\frac{c}{\lambda}\,\Delta t\,\mathscr{B}_n\le\frac{1}{\lambda}\,\Delta t\,\mathscr{G}_n\,.$$

Summing this inequality from 0 to N yields:

$$e^{2\gamma\Delta t}\mathscr{V}_{N+1}+\frac{e^{2\gamma\Delta t}-1}{\Delta t}\sum_1^N\Delta t\,\mathscr{V}_n+\frac{c}{\lambda}\sum_0^N\Delta t\,\mathscr{B}_n\le\mathscr{V}_0+\frac{1}{\lambda}\sum_0^N\Delta t\,\mathscr{G}_n\le\mathscr{V}_0+\frac{1}{\lambda}\sum_{n\ge 0}\Delta t\,\mathscr{G}_n\,.$$

Letting N tend to $+\infty$, we have proved:

$$e^{2\gamma\Delta t}\sup_{n\ge1}\mathscr{V}_n+\gamma\sum_{n\ge1}\Delta t\,\mathscr{V}_n+\sum_{n\ge0}\Delta t\,\mathscr{B}_n\le C\left(\mathscr{V}_0+\Delta x\,\mathscr{G}_0+\sum_{n\ge1}\Delta t\,\mathscr{G}_n\right),\tag{15}$$

and the right-hand side of (15) is directly estimated by the right-hand side of (9), see the definition above for $\mathscr{G}_n$ and use (7). The constant C in (15) is independent of γ and Δt and we have therefore completed the proof of (9) in the case $p\ge1$.

It remains to treat the case $p=0$ for which Lemma 2.2 does not hold anymore. In this case, we go back to (13) and simply ignore the nonnegative "boundary terms" on the left-hand side:

$$\|V^{n+1}\|^2_{1,+\infty}-\|V^n\|^2_{1,+\infty}\le\|V^n\|^2_{1-r,0}\,.$$

Then we proceed as above (with the same notation) and derive the weighted-in-time estimate:

$$e^{2\gamma\Delta t}\sup_{n\ge1}\mathscr{V}_n+\gamma\sum_{n\ge1}\Delta t\,\mathscr{V}_n\le C\left(\mathscr{V}_0+\Delta x\,\mathscr{G}_0+\sum_{n\ge1}\Delta t\,\mathscr{G}_n\right).$$

In the case $p=0$, the term:

$$\sum_{n\ge0}\Delta t\,e^{-2\gamma n\Delta t}\sum_{j=1-r}^{p}|V_j^n|^2\,,$$

in the left-hand side of (9) is directly estimated by the right-hand side of (9) so the proof of Lemma 2.1 is complete (provided that we prove Lemma 2.2, which is done below). □

Proof of Lemma 2.2. Proving Lemma 2.2 is equivalent to proving that the following quadratic form (that is independent on n):

$$(V^n_{1-r},\dots,V^n_p)\longmapsto\sum_{j=1-r-p}^{-r}|\widetilde{Q}\,W_j|^2+\sum_{j=1-r}^{0}|V_j^n+\widetilde{Q}\,W_j|^2\,,\tag{16}$$

is positive definite. (Recall that W denotes the extension of V^n by zero for $j\le -r$.) The quadratic form (16) is clearly nonnegative. Let us therefore consider some vector $(V^n_{1-r},\dots,V^n_p)$ that satisfies:

$$\forall\,j=1-r-p,\dots,-r,\quad \widetilde{Q}\,W_j=0\,,\qquad \forall\,j=1-r,\dots,0,\quad V_j^n+\widetilde{Q}\,W_j=0\,.\tag{17}$$

We first show by induction on j that $V_j^n=0$ for all $j=1-r,\dots,p-r$. Let us recall that $p\ge1$, so we can write $\widetilde{Q}=Q-I$ under the form:

$$\widetilde{Q}=A_p\,T^p+\sum_{\ell=-r}^{p-1}\widetilde{A}_\ell\,T^\ell\,.$$

In particular, we have $\widetilde{Q}\,W_{1-r-p}=A_p\,V^n_{1-r}$, so $V^n_{1-r}=0$ because A_p is invertible. For $j=1-r-p,\dots,-r$, $\widetilde{Q}\,W_j$ equals $A_p\,V^n_{j+p}$ plus a linear combination of the V^n_ℓ, $\ell<j+p$. Since the first term V^n_{1-r} is zero, we can proceed by induction and we thus get $V^n_{1-r}=\dots=V^n_{p-r}=0$.

We now use the second set of equalities in (17). In particular, we have $V^n_{1-r}+\widetilde{Q}\,W_{1-r}=\widetilde{Q}\,W_{1-r}=A_p\,V^n_{1-r+p}$. We therefore get $V^n_{1-r+p}=0$, and the rest of the proof follows from another induction argument. We have therefore shown that (17) implies $(V^n_{1-r},\dots,V^n_p)=0$, so the quadratic form (16) is positive definite. The proof of Lemma 2.2 is complete. □

We next turn to the case $(f,g)=0$ and the interior source F is arbitrary. We have:

Lemma 2.3. *Let Assumptions 2.1 and 2.2 be satisfied. Then there exists a constant C such that for all $\gamma>0$ and all $\Delta t \in]0,1]$, the solution (V_j^n) to (7) with $(f,g)=0$ satisfies the estimate:*

$$\sup_{n\ge 1} \mathrm{e}^{-2\gamma\, n\,\Delta t}\,\|V^n\|_{1,+\infty}^2 + \frac{\gamma}{\gamma\,\Delta t+1} \sum_{n\ge 1} \Delta t\, \mathrm{e}^{-2\gamma\, n\,\Delta t}\,\|V^n\|_{1,+\infty}^2 \\ + \mathrm{e}^{-2\gamma\,\Delta t} \sum_{n\ge 1} \Delta t\, \mathrm{e}^{-2\gamma\, n\,\Delta t} \sum_{j=1}^{p} |V_j^n|^2 \le C\, \frac{\gamma\,\Delta t+1}{\gamma} \sum_{n\ge 0} \Delta t\, \mathrm{e}^{-2\gamma\,(n+1)\,\Delta t}\,\|F^n\|_{1,+\infty}^2 \,. \quad (18)$$

Proof of Lemma 2.3. Following the proof of Lemma 2.1, we decompose the operator Q as $Q=I+\widetilde{Q}$. Then we use the relation $V_j^{n+1}=Q\,V_j^n+\Delta t\,F_j^n$ for $j\ge 1$ and derive:

$$\|V^{n+1}\|_{1,+\infty}^2 - \|V^n\|_{1,+\infty}^2 \\ = 2\left(V^n,\widetilde{Q}\,V^n\right)_{1,+\infty} + \|\widetilde{Q}\,V^n\|_{1,+\infty}^2 + 2\left(Q\,V^n,F^n\right)_{1,+\infty} + \Delta t^2\,\|F^n\|_{1,+\infty}^2 \,.$$

Let us first assume that p is positive, so that Lemma 2.2 holds. Proceeding as in the proof of Lemma 2.1, we obtain the inequality (recall that here we have homogeneous boundary conditions):

$$\|V^{n+1}\|_{1,+\infty}^2 - \|V^n\|_{1,+\infty}^2 + c\,\Delta x \sum_{j=1}^{p} |V_j^n|^2 \le 2\,\Delta t\,\|V^n\|_{1,+\infty}\,\|F^n\|_{1,+\infty} + \Delta t^2\,\|F^n\|_{1,+\infty}^2 \,. \quad (19)$$

For the sake of clarity, we now introduce the notation:

$$\mathscr{V}_n := \mathrm{e}^{-2\gamma\, n\,\Delta t}\,\|V^n\|_{1,+\infty}^2 \,,\quad \mathscr{B}_n := \mathrm{e}^{-2\gamma\, n\,\Delta t} \sum_{j=1}^{p} |V_j^n|^2 \,,\quad \mathscr{F}_n := \mathrm{e}^{-2\gamma\,(n+1)\,\Delta t}\,\|F^n\|_{1,+\infty}^2 \,.$$

We multiply (19) by $\exp(-2\gamma\,(n+1)\,\Delta t)$ and get:

$$\mathscr{V}_{n+1} - \mathrm{e}^{-2\gamma\,\Delta t}\,\mathscr{V}_n + \frac{c}{\lambda}\,\mathrm{e}^{-2\gamma\,\Delta t}\,\Delta t\,\mathscr{B}_n \le 2\,\Delta t\,\mathrm{e}^{-\gamma\,\Delta t}\,\mathscr{F}_n^{1/2}\,\mathscr{V}_n^{1/2} + \Delta t^2\,\mathscr{F}_n \,.$$

Summing this inequality from 0 to N and recalling that the initial data is zero, we obtain:

$$\mathscr{V}_{N+1} + \frac{1-\mathrm{e}^{-2\gamma\,\Delta t}}{\Delta t} \sum_{1}^{N} \Delta t\,\mathscr{V}_n + \frac{c}{\lambda}\,\mathrm{e}^{-2\gamma\,\Delta t} \sum_{0}^{N} \Delta t\,\mathscr{B}_n \le \Delta t \sum_{0}^{N} \Delta t\,\mathscr{F}_n + C \sum_{0}^{N} \sqrt{\Delta t\,\mathscr{F}_n}\,\sqrt{\Delta t\,\mathscr{V}_n} \,.$$

We apply Young's inequality for the last term on the right-hand side and we end up with:

$$\mathscr{V}_{N+1} + \frac{1-\mathrm{e}^{-2\gamma\,\Delta t}}{2\,\Delta t} \sum_{0}^{N} \Delta t\,\mathscr{V}_n + \frac{c}{\lambda}\,\mathrm{e}^{-2\gamma\,\Delta t} \sum_{0}^{N} \Delta t\,\mathscr{B}_n \le C\,\frac{\Delta t}{1-\mathrm{e}^{-2\gamma\,\Delta t}} \sum_{0}^{N} \Delta t\,\mathscr{F}_n \,.$$

Letting N tend to $+\infty$, we have proved (recall that the initial data is zero):

$$\sup_{n\ge 1} \mathscr{V}_n + \frac{\gamma}{\gamma\,\Delta t+1} \sum_{n\ge 1} \Delta t\,\mathscr{V}_n + \mathrm{e}^{-2\gamma\,\Delta t} \sum_{n\ge 1} \Delta t\,\mathscr{B}_n \le C\,\frac{\gamma\,\Delta t+1}{\gamma} \sum_{n\ge 0} \Delta t\,\mathscr{F}_n \,.$$

The constant C is independent of γ and Δt.

The case $p=0$ is dealt with in an entirely similar way. In this case, the term:

$$\sum_{n\ge 1} \Delta t\, \mathrm{e}^{-2\gamma\, n\,\Delta t} \sum_{j=1}^{p} |V_j^n|^2$$

that appears on the left-hand side of (18) vanishes. The proof of Lemma 2.3 is thus complete. □

If we compare Lemma 2.1 to Lemma 2.3, we observe that in (18) the estimate for the trace $(V_j^n)_{n\ge 0}$, $j=1,\dots,p$, involves a factor $\exp(-2\gamma\,\Delta t)$ that deteriorates the estimate when $\gamma\,\Delta t$ is large. We are going to derive an additional estimate that will enable us to get rid of this factor. This is done in:

Lemma 2.4. *Let Assumptions 2.1 and 2.2 be satisfied. Then there exists a constant C such that for all $\gamma > 0$ and all $\Delta t \in]0,1]$ verifying $\gamma\,\Delta t \geq 1$, the solution (V_j^n) to (7) with $(f,g)=0$ satisfies the estimate:*

$$\sum_{n\geq 1} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=1}^{p} |V_j^n|^2 \leq C\, \frac{\gamma\,\Delta t+1}{\gamma} \sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\, \|F^n\|_{1,+\infty}^2 \,. \tag{20}$$

Proof of Lemma 2.4. Let $j \in \{1,\dots,p\}$. We use the equation $V_j^{n+1} = Q\,V_j^n + \Delta t\, F_j^n$, and derive:

$$|V_j^{n+1}|^2 \leq 2\,|Q\,V_j^n|^2 + 2\,\Delta t^2\,|F_j^n|^2 \leq C\left(\frac{1}{\Delta t}\,\|V^n\|_{1,+\infty}^2 + \Delta t\,\|F^n\|_{1,+\infty}^2\right).$$

We multiply this inequality by $\exp(-2\,\gamma\,(n+1)\,\Delta t)$ and sum with respect to $n \geq 0$:

$$\sum_{n\geq 1} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\, |V_j^n|^2 \leq C\left(\frac{\mathrm{e}^{-2\,\gamma\,\Delta t}}{\Delta t} \sum_{n\geq 1} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\, \|V^n\|_{1,+\infty}^2 + \Delta t \sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\, \|F^n\|_{1,+\infty}^2\right).$$

We now use Lemma 2.3 to estimate the first term on the right-hand side of the inequality. We obtain:

$$\sum_{n\geq 1} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\, |V_j^n|^2 \leq C\left(\mathrm{e}^{-2\,\gamma\,\Delta t}\,\frac{\gamma\,\Delta t+1}{\gamma\,\Delta t}+1\right)\frac{\gamma\,\Delta t+1}{\gamma} \sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\, \|F^n\|_{1,+\infty}^2 \,,$$

and the result follows. □

The proof of Lemma 2.4 is surprisingly simple, but unfortunately it does cover the small values of $\gamma\,\Delta t$. If we collect Lemma 2.1, Lemma 2.3 and Lemma 2.4, we obtain:

Corollary 2.1. *Let Assumptions 2.1 and 2.2 be satisfied. Then there exists a constant C such that for all $\gamma > 0$ and all $\Delta t \in]0,1]$, the solution V to (7) satisfies the estimate:*

$$\sup_{n\geq 0} \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\, \|V^n\|_{1-r,+\infty}^2 + \frac{\gamma}{\gamma\,\Delta t+1} \sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\, \|V^n\|_{1-r,+\infty}^2 + \sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=1-r}^{p} |V_j^n|^2$$
$$\leq C\left\{\|f\|_{1-r,+\infty}^2 + \frac{\gamma\,\Delta t+1}{\gamma} \sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\, \|F^n\|_{1,+\infty}^2 + \sum_{n\geq 1} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=1-r}^{0} |g_j^n|^2\right\}. \tag{21}$$

Proof of Corollary 2.1. First of all, Lemma 2.1 shows that (21) holds when the interior source term F vanishes. Indeed, (9) implies the weaker inequality:

$$\sup_{n\geq 1} \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\, \|V^n\|_{1,+\infty}^2 + \frac{\gamma}{\gamma\,\Delta t+1} \sum_{n\geq 1} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\, \|V^n\|_{1,+\infty}^2 + \sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=1-r}^{p} |V_j^n|^2$$
$$\leq C\left\{\|f\|_{1-r,+\infty}^2 + \sum_{n\geq 1} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=1-r}^{0} |g_j^n|^2\right\}. \tag{22}$$

To obtain (21) in the case $F=0$, it remains to argue that one may replace $\sup_{n\geq 1} \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,\|V^n\|_{1,+\infty}^2$ in (22) by $\sup_{n\geq 1} \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,\|V^n\|_{1-r,+\infty}^2$, and add the term

$$\frac{\gamma}{\gamma\Delta t+1} \sum_{n\geq 1} \Delta t e^{-2\gamma n\Delta t} \|V^n\|_{1-r,0}^2 = \frac{1}{\lambda}\,\frac{\gamma\Delta t}{\gamma\Delta t+1} \sum_{n\geq 1} \Delta t\, e^{-2\gamma n\Delta t} \sum_{j=1-r}^{0} |V_j^n|^2$$

to the left handside. To this aim, it is enough to note that

$$\|V^n\|_{1-r,0}^2 = \sum_{j=1-r}^{0} \Delta x \, |V_j^n|^2 = \lambda \, \Delta t \sum_{j=1-r}^{0} |g_j^n|^2 ,$$

which precisely appears in the right handside of (22). Changing C accordingly proves the claim.

When the initial data and the boundary source term vanish, the estimate (21) is obtained by combining Lemma 2.3 (when $\gamma \, \Delta t \in]0,1]$) and Lemma 2.4 (when $\gamma \, \Delta t \geq 1$). We need not recover the boundary terms in the supremum as above since these boundary terms vanish here. □

If we compare the result of Corollary 2.1 with [13, Theorem 3.2], we get a better information when p is greater than 2, since we get additional trace estimates. Another advantage of our approach is that in [13], the author uses Goldberg-Tadmor's Lemma after introducing his auxiliary boundary operator. Hence, in some sense he needs two auxiliary problems. Here we shall only deal with the original discretization (3) and the auxiliary discretization (7).

Corollary 2.1 gives the result of Theorem 2.2 when $p > q$. For $q \geq p$, we need some additional trace estimates. These estimates can be obtained by adapting the method described in [13, page 85]. As a matter of fact, we use simpler arguments than in [13], which are only based on the energy method. In particular, we nowhere refer to the results of [4].

End of the proof of Theorem 2.2. From now on, we consider the case $q \geq p$ since for $q < p$, Corollary 2.1 gives the result of Theorem 2.2. Once again, the proof of (9) is slightly different according to the value of p. Let us first assume $p \geq 1$. As in [13], we define the sequence $W_j^n := V_{j+1}^n$ for $n \geq 0$ and $j \geq 1-r$. Then (W_j^n) solves the system:

$$\begin{cases} W_j^{n+1} = Q \, W_j^n + \Delta t \, F_{j+1}^n \, , & j \geq 1 \, , \quad n \geq 0 \, , \\ W_j^{n+1} = g_{j+1}^{n+1} \, , & j = 1-r, \dots, -1 \, , \quad n \geq 0 \, , \\ W_0^{n+1} = V_1^{n+1} \, , & n \geq 0 \, , \\ W_j^0 = f_{j+1} \, , & j \geq 1-r \, . \end{cases}$$

We can apply Corollary 2.1 to W and obtain:

$$\sum_{n \geq 0} \Delta t \, \mathrm{e}^{-2\gamma n \Delta t} \, |W_p^n|^2 \leq C \left\{ \|f\|_{2-r,+\infty}^2 + \frac{\gamma \, \Delta t + 1}{\gamma} \sum_{n \geq 0} \Delta t \, \mathrm{e}^{-2\gamma (n+1) \Delta t} \, \|F^n\|_{2,+\infty}^2 \right.$$
$$\left. + \sum_{n \geq 1} \Delta t \, \mathrm{e}^{-2\gamma n \Delta t} \sum_{j=2-r}^{0} |g_j^n|^2 + \sum_{n \geq 1} \Delta t \, \mathrm{e}^{-2\gamma n \Delta t} \, |V_1^n|^2 \right\} .$$

We use Corollary 2.1 again to estimate the last term on the right-hand (this is possible because $p \geq 1$), and we get:

$$\sum_{n \geq 0} \Delta t \, \mathrm{e}^{-2\gamma n \Delta t} \, |V_{p+1}^n|^2 \leq C \left\{ \|f\|_{1-r,+\infty}^2 + \frac{\gamma \, \Delta t + 1}{\gamma} \sum_{n \geq 0} \Delta t \, \mathrm{e}^{-2\gamma (n+1) \Delta t} \, \|F^n\|_{1,+\infty}^2 \right.$$
$$\left. + \sum_{n \geq 1} \Delta t \, \mathrm{e}^{-2\gamma n \Delta t} \sum_{j=1-r}^{0} |g_j^n|^2 \right\} .$$

We have therefore derived a trace estimate for $(V_{p+1}^n)_{n\ge 0}$. A straightforward induction argument gives:

$$\sum_{n\ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=p+1}^{q+1} |V_j^n|^2 \le C \left\{ \|f\|_{1-r,+\infty}^2 + \frac{\gamma\,\Delta t+1}{\gamma} \sum_{n\ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\, \|F^n\|_{1,+\infty}^2 + \sum_{n\ge 1} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=1-r}^{0} |g_j^n|^2 \right\} . \tag{23}$$

Combining (21) and (23), we obtain (9) for the case $p \ge 1$.

We now consider the case $p = 0$, for which Corollary 2.1 does not give any trace estimate of $(V_j^n)_{n\ge 0}$ with $j \ge 1$. Hence, we can not use the shift argument of [13] as above. Using Assumption 2.1, we know that the spectral radius of A_0 is strictly less than 1. Consequently, there exists a positive definite symmetric matrix H and there exists a positive number ε_0 such that if we consider the new Euclidean norm on $\mathbb{R}^D$:

$$\forall\, X \in \mathbb{R}^D\,, \quad |X|_H := \sqrt{\langle X; H\, X\rangle}\,,$$

then we have:

$$\forall\, X \in \mathbb{R}^D\,, \quad |A_0\, X|_H \le \sqrt{1-2\,\varepsilon_0}\, |X|_H\,.$$

We start from the relation:

$$V_1^{n+1} = A_0\, V_1^n + \sum_{\ell=-r}^{-1} A_\ell\, V_{1+\ell}^n + \Delta t\, F_1^n = A_0\, V_1^n + \underbrace{\sum_{j=1-r}^{0} A_{j-1}\, g_j^n + \Delta t\, F_1^n}_{=:X^n}\,,$$

where we use the notation $g_j^0 := f_j$ for $j = 1-r, \dots, 0$. Then we derive:

$$\begin{aligned} |V_1^{n+1}|_H^2 &= |A_0\, V_1^n|_H^2 + 2\,\langle A_0\, V_1^n; H\, X^n\rangle + |X^n|_H^2 \\ &\le (1-2\,\varepsilon_0)\, |V_1^n|_H^2 + 2\,\langle A_0\, V_1^n; H\, X^n\rangle + |X^n|_H^2 \le (1-\varepsilon_0)\, |V_1^n|_H^2 + (1+\varepsilon_0^{-1})\, |X^n|_H^2\,. \end{aligned}$$

Using the definition of X^n, the latter inequality gives:

$$|V_1^{n+1}|_H^2 - |V_1^n|_H^2 + \varepsilon_0\, |V_1^n|_H^2 \le C \left(\Delta t\, \|F^n\|_{1,+\infty}^2 + \sum_{j=1-r}^{0} |g_j^n|^2 \right) .$$

Using the same summation process as earlier, we obtain:

$$\begin{aligned} &\left\{ (1-\mathrm{e}^{-2\,\gamma\,\Delta t}) + \varepsilon_0\, \mathrm{e}^{-2\,\gamma\,\Delta t} \right\} \sum_{n\ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\, |V_1^n|_H^2 \\ &\le C \left\{ \|f\|_{1-r,+\infty}^2 + \frac{\gamma\,\Delta t+1}{\gamma} \sum_{n\ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\, \|F^n\|_{1,+\infty}^2 + \sum_{n\ge 1} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=1-r}^{0} |g_j^n|^2 \right\} . \end{aligned}$$

The norm $|\cdot|_H$ and the standard Euclidean norm are equivalent, so we get:

$$\begin{aligned} &\sum_{n\ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\, |V_1^n|^2 \\ &\le C \left\{ \|f\|_{1-r,+\infty}^2 + \frac{\gamma\,\Delta t+1}{\gamma} \sum_{n\ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\, \|F^n\|_{1,+\infty}^2 + \sum_{n\ge 1} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=1-r}^{0} |g_j^n|^2 \right\} , \end{aligned} \tag{24}$$

with a constant C that does not depend on γ nor on Δt. The proof of (9) follows from an induction argument where we apply the above method to recover the estimate for the trace $(V_j^n)_{n\ge 0}$, $j = 2, \dots, q+1$. The proof of Theorem 2.2 is now complete. □

2.3. **Proof of Theorem 2.1.** We decompose the solution U to (3) as $U=V+W$, where V satisfies:

$$\begin{cases} V_j^{n+1} = Q\,V_j^n + \Delta t\,F_j^n\,, & j\ge 1\,, \quad n\ge 0\,,\\ V_j^{n+1} = g_j^{n+1}\,, & j=1-r,\dots,0\,, \quad n\ge 0\,,\\ V_j^0 = f_j\,, & j\ge 1-r\,, \end{cases} \tag{25}$$

and W satisfies:

$$\begin{cases} W_j^{n+1} = Q\,W_j^n\,, & j\ge 1\,, \quad n\ge 0\,,\\ W_j^{n+1} = B_{j,-1}\,W_1^{n+1} + B_{j,0}\,W_1^n + \widetilde{g}_j^{n+1}\,, & j=1-r,\dots,0\,, \quad n\ge 0\,,\\ W_j^0 = 0\,, & j\ge 1-r\,. \end{cases} \tag{26}$$

The source term $\widetilde{g}$ in (26) is defined by:

$$\forall\, j=1-r,\dots,0\,, \quad \forall\, n\ge 1\,, \quad \widetilde{g}_j^n := B_{j,-1}\,V_1^n + B_{j,0}\,V_1^{n-1}\,. \tag{27}$$

The estimate of V is given by Theorem 2.2. Moreover, the discretization (3) is strongly stable in the sense of Definition 2.1, so W satisfies:

$$\frac{\gamma}{\gamma\,\Delta t+1} \sum_{n\ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\, n\,\Delta t}\, \|W^n\|_{1-r,+\infty}^2 + \sum_{n\ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\, n\,\Delta t} \sum_{j=1-r}^{0} |W_j^n|^2 \le C \sum_{n\ge 1} \Delta t\, \mathrm{e}^{-2\,\gamma\, n\,\Delta t} \sum_{j=1-r}^{0} |\widetilde{g}_j^n|^2\,.$$

Here, we use the fact that the initial data for (26) is zero. The estimate of $\widetilde{g}_j^n$ is straightforward using the definition (27) and (9). In the end, we obtain:

$$\frac{\gamma}{\gamma\,\Delta t+1} \sum_{n\ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\, n\,\Delta t}\, \|W^n\|_{1-r,+\infty}^2 + \sum_{n\ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\, n\,\Delta t} \sum_{j=1-r}^{0} |W_j^n|^2$$
$$\le C \left\{ \|f\|_{1-r,+\infty}^2 + \frac{\gamma\,\Delta t+1}{\gamma} \sum_{n\ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\, \|F^n\|_{1,+\infty}^2 + \sum_{n\ge 1} \Delta t\, \mathrm{e}^{-2\,\gamma\, n\,\Delta t} \sum_{j=1-r}^{0} |g_j^n|^2 \right\}. \tag{28}$$

If we combine the estimate (28) for W and the estimate (9) for V, we see that the final task in the proof of Theorem 2.1 is to control the $\ell_n^\infty(\ell_j^2)$ norm of W. We recall the following result that is the analogue of [13, Lemma 3.1]:

Lemma 2.5. *Let Assumptions 2.1 and 2.2 be satisfied, and assume that the discretization (3) is strongly stable. Then there exists a constant C that does not depend on the data $\widetilde{g}$ such that for all $\gamma>0$ and all $\Delta t\in\,]0,1]$, the solution W to (26) satisfies:*

$$\sum_{n\ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\, n\,\Delta t} \sum_{j=1}^{q+1} |W_j^n|^2 \le C \sum_{n\ge 1} \Delta t\, \mathrm{e}^{-2\,\gamma\, n\,\Delta t} \sum_{j=1-r}^{0} |\widetilde{g}_j^n|^2\,.$$

We give a proof of Lemma 2.5 in Appendix A. As a matter of fact, Lemma 2.5 was stated without a proof in [13], although the result does not appear exactly in this form in [4]. We find it useful to give a complete and detailed proof here. Using Lemma 2.5, let us rewrite (26) as:

$$\begin{cases} W_j^{n+1} = Q\,W_j^n\,, & j\ge 1\,, \quad n\ge 0\,,\\ W_j^{n+1} = G_j^{n+1}\,, & j=1-r,\dots,0\,, \quad n\ge 0\,,\\ W_j^0 = 0\,, & j\ge 1-r\,, \end{cases} \tag{29}$$

with an obvious definition for the source term G. We apply Theorem 2.2 to (29) and obtain:

$$\sup_{n\ge 0} \mathrm{e}^{-2\,\gamma\, n\,\Delta t}\, \|W^n\|_{1-r,+\infty}^2 \le C \sum_{n\ge 1} \Delta t\, \mathrm{e}^{-2\,\gamma\, n\,\Delta t} \sum_{j=1-r}^{0} |G_j^n|^2\,.$$

We now combine Lemma 2.5 and our previous estimate of $\widetilde{g}$ (see the argument above to get (28)), to derive:

$$\sum_{n\geq 1}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{j=1-r}^{0}|G_j^n|^2$$
$$\leq C\left\{\|f\|_{1-r,+\infty}^2+\frac{\gamma\,\Delta t+1}{\gamma}\sum_{n\geq 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\,\|F^n\|_{1,+\infty}^2+\sum_{n\geq 1}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{j=1-r}^{0}|g_j^n|^2\right\}.$$

We thus get the expected estimate for the $\ell^\infty_n(\ell^2_j)$ norm of W, and the proof of Theorem 2.1 is complete.

3. Multidimensional problems

For multidimensional problems, we need further notation for norms on $\ell^2(\mathbb{Z}^d)$. Let $\Delta x_i>0$ for $i=1,\dots,d$ be d space steps. For all integers $m_1\leq m_2$, we set:

$$\|u\|_{m_1,m_2}^2:=\Delta x_1\sum_{j_1=m_1}^{m_2}\left(\prod_{k=2}^{d}\Delta x_k\right)\sum_{i=2}^{d}\sum_{j_i\in\mathbb{Z}}|u_{j_1,\dots,j_d}|^2\,,$$

to denote the ℓ^2-norm on the set $[m_1,m_2]\times\mathbb{Z}^{d-1}$ (m_1 may equal $-\infty$ and m_2 may equal $+\infty$).

We shall also make use of the $\ell^2(\mathbb{Z}^{d-1})$-norm that we denote by $\|\cdot\|$: for all $v\in\ell^2(\mathbb{Z}^{d-1})$,

$$\|v\|^2:=\left(\prod_{k=2}^{d}\Delta x_k\right)\sum_{i=2}^{d}\sum_{j_i\in\mathbb{Z}}|v_{j_2,\dots,j_d}|^2\,.$$

3.1. **Main result in several space dimensions.** We consider the hyperbolic initial boundary value problem corresponding to (1) in several space dimensions $d>1$, that is:

$$\begin{cases}\partial_t u+\displaystyle\sum_{i=1}^{d}A_i\,\partial_{x_i}u=F(t,x)\,, & (t,x)\in\mathbb{R}^+\times\mathbb{R}^d_+\,,\\ B\,u(t,(0,x'))=g(t,x')\,, & t\in\mathbb{R}^+,\ x'\in\mathbb{R}^{d-1},\\ u(0,x)=f(x)\,, & x\in\mathbb{R}^d_+\,,\end{cases}\tag{30}$$

where $\mathbb{R}^d_+:=\mathbb{R}^+\times\mathbb{R}^{d-1}$, the matrices $A_i\in\mathscr{M}_D(\mathbb{R})$ are such that the symbol $\mathbb{R}^d\ni\xi\mapsto A(\xi):=\sum_{i=1}^d\xi_iA_i$ is *uniformly* diagonalizable (see [1, Theorem 1.3]) in $\mathbb{R}$, and $B\in\mathscr{M}_{D_+,D}(\mathbb{R})$ with D_+ the number of positive eigenvalues of A_1. We assume that the boundary is noncharacteristic, that is $0\notin\mathrm{sp}(A_1)$. Problem (30) is strongly well-posed in $L^2(\mathbb{R}^d_+)$ if and only if the matrices $\{A_i\}$ and B satisfy the so-called *uniform Kreiss-Lopatinskii condition*. In that case, the solution u to (30) belongs to $\mathscr{C}(\mathbb{R}^+;L^2(\mathbb{R}^d_+))$ and its trace on $\{x_1=0\}$ is well-defined and belongs to $\mathrm{e}^{\gamma\,t}\,L^2(\mathbb{R}^+;L^2(\mathbb{R}^{d-1}))$ for all $\gamma>0$. Moreover, for all $\gamma>0$, u satisfies the energy estimate:

$$\sup_{t\geq 0}\mathrm{e}^{-2\,\gamma\,t}\,\|u(t,\cdot)\|_{L^2(\mathbb{R}^d_+)}^2+\gamma\int_0^{+\infty}\mathrm{e}^{-2\,\gamma\,t}\,\|u(t,\cdot)\|_{L^2(\mathbb{R}^d_+)}^2\,\mathrm{d}t+\int_0^{+\infty}\mathrm{e}^{-2\,\gamma\,t}\,\|u(t,(0,\cdot))\|_{L^2(\mathbb{R}^{d-1})}^2\,\mathrm{d}t$$
$$\leq C\left(\|f\|_{L^2(\mathbb{R}^d_+)}^2+\frac{1}{\gamma}\int_0^{+\infty}\mathrm{e}^{-2\,\gamma\,t}\,\|F(t,\cdot)\|_{L^2(\mathbb{R}^d_+)}^2\,\mathrm{d}t+\int_0^{+\infty}\mathrm{e}^{-2\,\gamma\,t}\,\|g(t,\cdot)\|_{L^2(\mathbb{R}^{d-1})}^2\,\mathrm{d}t\right)\,,\tag{31}$$

where the constant C is independent of γ,f,F,g.

As for the one-dimensional case, we introduce the finite difference approximation of (30). We denote by $\Delta x:=\{\Delta x_i\}_{i=1,\dots,d}$ and Δt the space and time steps related by the fixed ratios $\lambda_i=\Delta t/\Delta x_i$. For all $j\in\mathbb{Z}^d$, we set $j=(j_1,j')$ with $j'=(j_2,\dots,j_d)$. We let $p,q,r\in\mathbb{N}^d$ be some multi-integers, and define p_1,q_1,r_1,p',q',r' according to the above notation. The solution u to (30) is approximated by a sequence $(U_j^n)=(U_{j_1,j'}^n)$ for $n\in\mathbb{N}$, $j_1\in 1-r_1+\mathbb{N}$, and $j'\in\mathbb{Z}^{d-1}$. For $j_1=1-r_1,\dots,0$, $U_{j_1,\cdot}^n$ approximates the trace $u(n\Delta t,0,\cdot)$ on the boundary $\{x_1=0\}$, and possibly the trace of normal derivatives. We consider one-step finite difference approximations of (30) that read:

$$\begin{cases} U_{j_1,j'}^{n+1} = Q\,U_{j_1,j'}^n + \Delta t\,F_{j_1,j'}^n\,, & j_1 \geq 1\,,\ j' \in \mathbb{Z}^{d-1}\,, \quad n \geq 0\,,\\ U_{j_1,j'}^{n+1} = B_{j_1,-1}\,U_{1,j'}^{n+1} + B_{j_1,0}\,U_{1,j'}^n + g_{j_1,j'}^{n+1}\,, & j_1 = 1-r_1,\dots,0\,,\ j' \in \mathbb{Z}^{d-1}\,, \quad n \geq 0\,,\\ U_{j_1,j'}^0 = f_{j_1,j'}\,, & j_1 \geq 1-r_1\,,\ j' \in \mathbb{Z}^{d-1}\,, \end{cases} \tag{32}$$

where the operators Q, $B_{j_1,-1}, B_{j_1,0}$ are given by:

$$\begin{aligned} &Q := \sum_{\ell_1=-r_1}^{p_1} \Big(\sum_{\ell'=-r'}^{p'} A_{\ell_1,\ell'}\,{T'}^{\ell'} \Big)\, T^{\ell_1}\,,\\ &B_{j_1,\sigma} := \sum_{\ell_1=0}^{q_1} \Big(\sum_{\ell'=-q'}^{q'} B_{\ell_1,\ell',j_1,\sigma}\,{T'}^{\ell'} \Big)\, T^{\ell_1}\,,\\ &T^{\ell_1}\,U_{k_1,k'}^m := U_{k_1+\ell_1,k'}^m\,,\\ &{T'}^{\ell'}\,U_{k_1,k'}^m := U_{k_1,k'+\ell'}^m\,. \end{aligned} \tag{33}$$

In (33), all matrices $A_\ell, B_{\ell,j_1,\sigma}$ belong to $\mathscr{M}_D(\mathbb{R})$ and depend on $\{\lambda_i, A_i\}_{i=1,\dots,d}, B$ but not on Δt (or equivalently not on Δx). For multidimensional problems, the notion of strong stability now reads:

Definition 3.1 (Strong stability [7]). *The finite difference approximation (32) is said to be strongly stable if there exists a constant C such that for all $\gamma > 0$ and all $\Delta t \in]0,1]$, the solution (U_j^n) of (3) with $f = 0$ satisfies the estimate:*

$$\begin{aligned} \frac{\gamma}{\gamma\,\Delta t + 1} \sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\, \| U^n \|_{1-r_1,+\infty}^2 + \sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j_1=1-r_1}^{0} \| U_{j_1,\cdot}^n \|^2 \\ \leq C \left\{ \frac{\gamma\,\Delta t + 1}{\gamma} \sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\, \| F^n \|_{1,+\infty}^2 + \sum_{n\geq 1} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j_1=1-r_1}^{0} \| g_{j_1,\cdot}^n \|^2 \right\}. \end{aligned}$$

We are in position to introduce the hypotheses corresponding to Assumptions 2.1 and 2.2 in the multidimensional case. For $\ell_1 = -r_1,\dots,p_1$, and $z \in \mathbb{C}\setminus\{0\}$, let us define the linear mappings:

$$\begin{aligned} \mathbb{A}_{\ell_1}(z) :\quad \ell^2(\mathbb{Z}^{d-1}) &\;\rightarrow\; \ell^2(\mathbb{Z}^{d-1}) \\ w &\;\mapsto\; \delta_{\ell_1 0}\, w - \frac{1}{z} \sum_{\ell'=-r'}^{p'} A_{\ell_1,\ell'} {T'}^{\ell'}\, w\,. \end{aligned} \tag{34}$$

We make the following first assumption:

Assumption 3.1. *The mapping $\mathbb{A}_{p_1}(z)$ is coercive on $\ell^2(\mathbb{Z}^{d-1})$ for all $z \in \mathbb{C}$ with $|z| \geq 1$. More precisely, there exists a constant $c > 0$ such that for all $z \in \mathbb{C}$ with $|z| \geq 1$ and for all $w \in \ell^2(\mathbb{Z}^{d-1})$, we have:*

$$\| \mathbb{A}_{p_1}(z) w \| \geq \frac{c}{|z|^\nu}\, \| w \|\,, \quad \nu := \begin{cases} 1 & \text{if } p_1 > 0,\\ 0 & \text{otherwise.} \end{cases}$$

Remark 3.1. *If $p_1 = 0$, then Assumption 3.1 amounts to assuming that $\mathbb{A}_0(z)$ is an isomorphism on $\ell^2(\mathbb{Z}^{d-1})$ for all $|z| \geq 1$. For $p_1 > 0$, Assumption 3.1 is slightly weaker than assuming that $\mathbb{A}_{p_1}(z)$ is an isomorphism since the fulfillment of Assumption 3.1 does not necessarily imply the surjectivity of $\mathbb{A}_{p_1}(z)$ on $\ell^2(\mathbb{Z}^{d-1})$.*

Our second assumption is unchanged:

Assumption 3.2. *The operator Q satisfies $\| Q\, v \|_{-\infty,+\infty} \leq \| v \|_{-\infty,+\infty}$ for all $v \in \ell^2(\mathbb{Z}^d)$.*

Our main stability result corresponding to Theorem 2.1 then reads:

Theorem 3.1. *Let Assumptions 3.1 and 3.2 be satisfied, and assume that the scheme (32) is strongly stable in the sense of Definition 3.1. Then there exists a constant C such that for all $\gamma > 0$ and all $\Delta t \in]0,1]$, the solution U to (32) satisfies the estimate:*

$$\sup_{n\geq 0} \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\, |||U^n|||^2_{1-r_1,+\infty} + \frac{\gamma}{\gamma\,\Delta t+1} \sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\, |||U^n|||^2_{1-r_1,+\infty}$$
$$+ \sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j_1=1-r_1}^{0} \|U^n_{j_1,\cdot}\|^2 \leq C \left\{ |||f|||^2_{1-r_1,+\infty} + \frac{\gamma\,\Delta t+1}{\gamma} \sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\, |||F^n|||^2_{1,+\infty} \right.$$
$$\left. + \sum_{n\geq 1} \Delta t\, \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j_1=1-r_1}^{0} \|g^n_{j_1,\cdot}\|^2 \right\} .$$

Our proof of Theorem 3.1 relies on the application of the one-dimensional analysis. In particular, we shall see below why Assumptions 3.1 and 3.2 are the natural extensions of Assumptions 2.1 and 2.2 in several space dimensions.

Proof of Theorem 3.1. Let us start from the scheme (32) that defines the sequence (U^n_j). For each pair (j_1,n) with $j_1 \geq 1-r_1$ and $n \geq 0$, we define the piecewise constant function $V^n_{j_1}(x')$ on $\mathbb{R}^{d-1}$ such that $V^n_{j_1}(x')$ equals $U^n_{j_1,j'}$ on the mesh element with index j'. In a similar way, we define the functions $F^n_{j_1}, g^n_{j_1}, f_{j_1}$. Applying Fourier transform with respect to the tangential variables x' in (32), we obtain:

$$\begin{cases} \widehat{V^{n+1}_{j_1}}(\xi') = Q(\theta')\,\widehat{V^n_{j_1}}(\xi') + \Delta t\,\widehat{F^n_{j_1}}(\xi')\,, & j_1 \geq 1\,, \quad n\geq 0\,, \\ \widehat{V^{n+1}_{j_1}}(\xi') = B_{j_1,-1}(\theta')\,\widehat{V^{n+1}_1}(\xi') + B_{j_1,0}(\theta')\,\widehat{V^n_1}(\xi') + \widehat{g^{n+1}_{j_1}}(\xi')\,, & j_1 = 1-r_1,\dots,0\,, \quad n\geq 0\,, \\ \widehat{V^0_{j_1}}(\xi') = \widehat{f_{j_1}}(\xi')\,, & j_1 \geq 1-r_1\,, \end{cases} \tag{35}$$

where $\xi' = (\xi_2,\dots,\xi_d)$ denotes the frequency variables, θ' is a short notation for $(\xi_2\,\Delta x_2,\dots,\xi_d\,\Delta x_d)$, and the operators $Q(\theta'), B_{j_1,-1}(\theta'), B_{j_1,0}(\theta')$ are defined by:

$$Q(\theta') := \sum_{\ell_1=-r_1}^{p_1} \Big(\sum_{\ell'=-r'}^{p'} A_{\ell_1,\ell'}\, \mathrm{e}^{i\,\ell'\cdot\theta'} \Big)\, T^{\ell_1}\,,$$
$$B_{j_1,\sigma}(\theta') := \sum_{\ell_1=0}^{q_1} \Big(\sum_{\ell'=-q'}^{q'} B_{\ell_1,\ell',j_1,\sigma}\, \mathrm{e}^{i\,\ell'\cdot\theta'} \Big)\, T^{\ell_1}\,.$$

We are reduced to a collection of one-dimensional problems parametrized by the frequencies ξ'. The end of the proof is based on the following observation:

Lemma 3.1. *Let Assumptions 3.1 and 3.2 be satisfied, and assume that the scheme (32) is strongly stable in the sense of Definition 3.1. Then for all $\xi' \in \mathbb{R}^{d-1}$, the scheme (35) satisfies Assumptions 2.1 and 2.2, and is strongly stable in the sense of Definition 2.1. Moreover, the estimate of Theorem 2.1 holds with a constant C that is independent of ξ'.*

The proof of Lemma 3.1 is performed below. With the help of Lemma 3.1, let us now complete the proof of Theorem 3.1. We can apply Theorem 2.1 to the scheme (35) with a constant C that is independent of ξ'. We thus obtain:

$$\sup_{n\ge 0} \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,\|\widehat{V^n}(\xi')\|^2_{1-r_1,+\infty} + \frac{\gamma}{\gamma\,\Delta t+1}\,\sum_{n\ge 0} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,\|\widehat{V^n}(\xi')\|^2_{1-r_1,+\infty}$$

$$+\sum_{n\ge 0} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j_1=1-r_1}^{0} |\widehat{V^n_{j_1}}(\xi')|^2 \le C\,\Bigg\{ \|\widehat{f}(\xi')\|^2_{1-r,+\infty} + \sum_{n\ge 1} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j_1=1-r_1}^{0} |\widehat{g^n_{j_1}}(\xi')|^2$$

$$+\frac{\gamma\,\Delta t+1}{\gamma} \sum_{n\ge 0} \Delta t\,\mathrm{e}^{-2\,\gamma\,(n+1)\,\Delta t}\,\|\widehat{F^n}(\xi')\|^2_{1,+\infty} \Bigg\}\,. \tag{36}$$

Fubini's and Plancherel's Theorems yield

$$\begin{aligned}\sup_{n\ge 0} \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,|\!|\!|U^n|\!|\!|^2_{1-r_1,+\infty} &= \sup_{n\ge 0} \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \int_{\mathbb{R}^{d-1}} \|V^n(x')\|^2_{1-r_1,+\infty}\,\mathrm{d}x' \\ &= \frac{1}{(2\,\pi)^{d-1}}\,\sup_{n\ge 0} \mathrm{e}^{-2\,\gamma\,n\,\Delta t} \int_{\mathbb{R}^{d-1}} \|\widehat{V^n}(\xi')\|^2_{1-r_1,+\infty}\,\mathrm{d}\xi' \\ &\le \frac{1}{(2\,\pi)^{d-1}} \int_{\mathbb{R}^{d-1}} \sup_{n\ge 0} \mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,\|\widehat{V^n}(\xi')\|^2_{1-r_1,+\infty}\,\mathrm{d}\xi'\,,\end{aligned}$$

after switching the order of the supremum and of the integral. We then integrate (36) with respect to $\xi' \in \mathbb{R}^{d-1}$, use Fubini's and Plancherel's Theorem. Combined with the latter inequality, this proves Theorem 3.1. (Observe that it is crucial here that the right handside of (6) only contains ℓ^2 type norms so that we may apply Plancherel's Theorem, while the ℓ^∞-norm only appears on the left handside of (6).) The proof of Theorem 3.1 is therefore complete provided we prove Lemma 3.1, which is done below. □

Proof of Lemma 3.1. Let us first show that Assumption 2.2 is satisfied. We know that the norm of Q as an operator on $\ell^2(\mathbb{Z}^d)$ is not greater than 1. Using Plancherel's Theorem, this property is equivalent to the fact that the norm of the symbol $\widehat{Q}$ of Q is not greater than 1. Decomposing frequencies $\theta \in \mathbb{R}^d$ as $\theta = (\theta_1, \theta')$, the symbol $\widehat{Q}$ is given by:

$$\widehat{Q}(\theta) = \sum_{\ell_1=-r_1}^{p_1} \Big(\sum_{\ell'=-r'}^{p'} A_{\ell_1,\ell'}\,\mathrm{e}^{i\,\ell'\cdot\theta'} \Big)\,\mathrm{e}^{i\,\ell_1\,\theta_1}\,.$$

In other words, the symbol of the operator $Q(\theta')$ is nothing but $\widehat{Q}(\cdot,\theta')$. This shows that the norm of $Q(\theta')$ as an operator on $\ell^2(\mathbb{Z})$ is not greater than 1, and Assumption 2.2 is satisfied.

Let us turn to Assumption 2.1. With slight abuse of notation, if $w \in \ell^2(\mathbb{Z}^{d-1})$, we still denote w the piecewise constant function defined on $\mathbb{R}^{d-1}$, and whose value on the mesh element with index j' equals $w_{j'}$. Using Plancherel's Theorem, we have:

$$\|\mathbb{A}_{p_1}(z)w\|^2 = \frac{1}{(2\,\pi)^{d-1}} \int_{\mathbb{R}^{d-1}} |(\widehat{\mathbb{A}_{p_1}(z)w})(\xi')|^2\,\mathrm{d}\xi' = \frac{1}{(2\,\pi)^{d-1}} \int_{\mathbb{R}^{d-1}} |\mathbb{A}_{p_1}(z,\theta')\,\widehat{w}(\xi')|^2\,\mathrm{d}\xi'\,,$$

where θ' is again a short notation for $(\xi_2\,\Delta x_2, \dots, \xi_d\,\Delta x_d)$, and where the matrices $\mathbb{A}_{p_1}(z,\theta')$ are defined by:

$$\mathbb{A}_{p_1}(z,\theta') := \delta_{p_1 0}\,I - \frac{1}{z} \sum_{\ell'=-r'}^{p'} A_{p_1,\ell'}\,\mathrm{e}^{i\,\ell'\cdot\theta'}\,, \tag{37}$$

see (34). The operator $\mathbb{A}_{p_1}$ is coercive if and only if the matrices $\mathbb{A}_{p_1}(z,\theta')$ are invertible for all $z \in \mathbb{C}$ with $|z| \ge 1$ and all $\theta' \in \mathbb{R}^{d-1}$. This proves that Assumption 2.1 is satisfied for all $\theta' \in \mathbb{R}^{d-1}$.

Next we argue that the constants in Theorem 2.1 are independent of ξ'. A close look at the proof shows that the only places where the constants may depend on ξ' are Lemma 2.2 and (24). On the one hand, Assumption 3.1 provides a lower bound for the coercivity constants of $\mathbb{A}_{p_1}(z,\theta')$ which is uniform in θ', so that the constant in Lemma 2.2 is uniform in θ'. On the other hand, for $p_1 = 0$, we claim that the norm $|\cdot|_{H(\theta')}$ and the standard Euclidian norm $|\cdot|$ are equivalent uniformly in θ' so that the constant in (24) can be chosen independent of θ'. To this aim, it is enough to prove that

$$\sup_{\theta' \in \mathbb{R}^{d-1}} |\mathbb{A}_0(1,\theta')| \quad < \quad \infty \,, \tag{38}$$

$$\sup_{\theta' \in \mathbb{R}^{d-1}} \rho(\mathbb{A}_0(1,\theta')) \quad < \quad 1 \,. \tag{39}$$

Since $\theta' \mapsto \mathbb{A}_0(1,\theta')$ is continuous and periodic on $\mathbb{R}^{d-1}$, (38) is trivial, the supremum in (39) is attained and the bound follows from Assumption 3.1 and (37).

The only remaining task is to prove that the scheme (35) is strongly stable in the sense of Definition 2.1. This is done, as above, by applying Plancherel's Theorem to the strong stability estimate in Definition 3.1. We omit the details. □

3.2. **A multidimensional version of Goldberg-Tadmor's lemma.** The same argument as above (Fourier transform in the tangential variables and application of the one-dimensional results of Section 2) also applies to the case of nonhomogeneous Dirichlet boundary conditions. More precisely, let us consider the following auxiliary problem:

$$\begin{cases} V_{j_1,j'}^{n+1} = Q\, V_{j_1,j'}^n + \Delta t\, F_{j_1,j'}^n \,, & j_1 \ge 1 \,,\, j' \in \mathbb{Z}^{d-1} \,, \quad n \ge 0 \,, \\ V_{j_1,j'}^{n+1} = g_{j_1,j'}^{n+1} \,, & j_1 = 1-r_1, \dots, 0 \,,\, j' \in \mathbb{Z}^{d-1} \,, \quad n \ge 0 \,, \\ V_{j_1,j'}^0 = f_{j_1,j'} \,, & j_1 \ge 1-r_1 \,,\, j' \in \mathbb{Z}^{d-1} \,. \end{cases} \tag{40}$$

The multidimensional version of Theorem 2.2 reads:

Theorem 3.2. *Let Assumptions 3.1 and 3.2 be satisfied. Then there exists a constant C such that for all $\gamma > 0$ and all $\Delta t \in]0,1]$, the solution V to (40) satisfies the estimate:*

$$\sup_{n \ge 0} \mathrm{e}^{-2\,\gamma\, n\, \Delta t} \, ||| V^n |||_{1-r_1,+\infty}^2 + \frac{\gamma}{\gamma\, \Delta t + 1} \sum_{n \ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\, n\, \Delta t} \, ||| V^n |||_{1-r_1,+\infty}^2$$
$$+ \sum_{n \ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\, n\, \Delta t} \sum_{j_1=1-r_1}^{\max(p_1,q_1+1)} \| V_{j_1,\cdot}^n \|^2 \le C \Bigg\{ ||| f |||_{1-r_1,+\infty}^2 + \sum_{n \ge 1} \Delta t\, \mathrm{e}^{-2\,\gamma\, n\, \Delta t} \sum_{j_1=1-r_1}^{0} \| g_{j_1,\cdot}^n \|^2$$
$$+ \frac{\gamma\, \Delta t + 1}{\gamma} \sum_{n \ge 0} \Delta t\, \mathrm{e}^{-2\,\gamma\, (n+1)\, \Delta t} \, ||| F^n |||_{1,+\infty}^2 \Bigg\} \,.$$

In particular, the discretization (40) is strongly stable in the sense of Definition 3.1.

We do not detail the proof of Theorem 3.2 since the arguments are entirely similar to the arguments used in the proof of Theorem 3.1 for passing from one-dimensional results to multidimensional results.

4. Examples and comments

4.1. **Examples and comments in one space dimension.** For one-dimensional problems, the matrices A_ℓ in the finite difference operator Q are usually polynomials of the matrix $\lambda\, A$ where A is the matrix of the hyperbolic operator in (1) and $\lambda = \Delta t / \Delta x$. If we assume furthermore that A is symmetric, then there exists an orthogonal matrix that diagonalizes simultaneously all the matrices A_ℓ. In this case, Assumption 2.2 is exactly equivalent to the ℓ^2-stability of the finite difference approximation, which is itself equivalent to the well-known von Neumann condition (see [3, chapter 5]):

$$\forall\, \kappa \in \mathbb{S}^1 \,, \quad \rho \left(\sum_{\ell=-r}^{p} \kappa^\ell\, A_\ell \right) \le 1 \,.$$

We recall that ρ denotes the spectral radius of a square matrix. Assumption 2.2 is therefore very reasonable and not restrictive in one space dimension. Assumption 2.1 is also satisfied in numerous situations and is rather easy to check.

We now make a few comments on the proof of Theorem 2.2. The proof in the case $F = 0$ follows from Lemma 2.1 and is done in a single way, whatever the values of γ and Δt. However, we have seen that the case $F \neq 0$ requires two different approaches depending on the value of $\gamma\,\Delta t$. Here, we report on a simple numerical test which shows that the two regimes $\gamma\,\Delta t \leq 1$ and $\gamma\,\Delta t \geq 1$ are different when $F \neq 0$. We consider the scalar transport equation:

$$\begin{cases} \partial_t u - \partial_x u = F(t,x)\,, & (t,x) \in \mathbb{R}^+ \times \mathbb{R}^+\,, \\ u(0,x) = f(x)\,, & x \in \mathbb{R}^+\,, \end{cases}$$

where no boundary condition is required. We discretize the equation with the Lax-Friedrichs scheme and homogeneous Dirichlet boundary condition[3]:

$$\begin{cases} v_j^{n+1} = (v_{j-1}^n + v_{j+1}^n)/2 + \lambda\,(v_{j+1}^n - v_{j-1}^n)/2 + \Delta t\,F(n\,\Delta t, j\,\Delta x)\,, & j \geq 1\,, \quad n \geq 0\,, \\ v_0^{n+1} = 0\,, & n \geq 0\,, \\ v_j^0 = f(j\,\Delta x)\,, & j \geq 0\,, \end{cases} \tag{41}$$

or with the Lax-Wendroff scheme and homogeneous Dirichlet boundary condition:

$$\begin{cases} v_j^{n+1} = v_j^n + \lambda\,(v_{j+1}^n - v_{j-1}^n)/2 + \lambda^2\,(v_{j+1}^n + v_{j-1}^n - 2\,v_j^n)/2 + \Delta t\,F(n\,\Delta t, j\,\Delta x)\,, & j \geq 1\,, n \geq 0\,, \\ v_0^{n+1} = 0\,, & n \geq 0\,, \\ v_j^0 = f(j\,\Delta x)\,, & j \geq 0\,. \end{cases} \tag{42}$$

The CFL parameter λ is 0.9 in both situations, which ensures that Theorem 2.2 holds.

We plot the the ratio between the norm of the trace:

$$\sum_{n \geq 0} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,|v_1^n|^2\,,$$

and the right-hand side of (8) as a function of γ for the following cases:

(1) $f(x) = 1$ for $x \in [0,1]$, $F = 0$. The simulation is performed on the space interval $[0,2]$ with 1000 grid points. The parameter γ ranges from 10^{-2} to 10^2.
(2) $f = 0$, $F(t,x) = 1$ for $t \geq 0$ and $x \in [0,1]$. The simulation is performed on the space interval $[0,2]$ with 1000 grid points. The parameter γ ranges from 10^{-2} to 10^2.

The results are plotted in Figure 1 for the Lax-Friedrichs scheme (41), and in Figure 2 for the Lax-Wendroff scheme (42). The observation is the following: the ratio between the norm of the trace and the norm of the source term depends monotonically on γ when $f \neq 0$ and $F = 0$, while it does not depend monotonically on γ when $f = 0$ and $F \neq 0$. This seems to indicate that in the case $F \neq 0$, the estimate (8) for small values of γ does not follow from the same arguments as for large values of γ. This is the reason why we believe that Lemmas 2.3 and 2.4 are both useful.

4.2. **Examples and comments in several space dimensions.** Let us first comment on our multidimensional version of Golberg-Tadmor's lemma. In one dimension, the commutation assumption by Golberg and Tadmor is very natural, and essentially, it is "equivalent" to Assumption 2.2. The original work by Golberg and Tadmor [2] also covers the multidimensional case provided the matrices $\{A_\ell\}$ in (33) commute. This assumption is very restrictive for $d > 1$ and amounts more or less to consider d uncoupled scalar equations in (30). Assumption 3.2, however, can hold independently of the fact that the matrices $\{A_\ell\}$ do commute or not. Theorem 3.2 is therefore a true generalization of Golberg-Tadmor's lemma.

In the remaining part of this paragraph, we consider $d = 2$ and give several examples of discretizations to which Theorem 3.1 applies.

[3] Observe that the homogeneous Dirichlet condition is not consistent in the L^∞-norm with the continuous problem for which no boundary condition is required. However, we are concerned here with stability estimates, and consistency is a different issue.

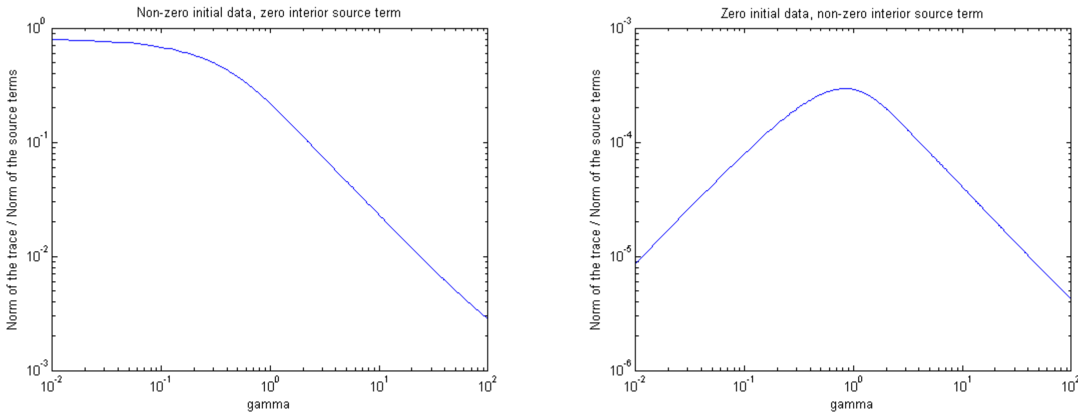

FIGURE 1. Lax-Friedrichs scheme (41): ratio between the norm of the trace and the norm of the source terms. Non-zero initial data and zero interior source term (left). Zero initial data and non-zero interior source term (right).

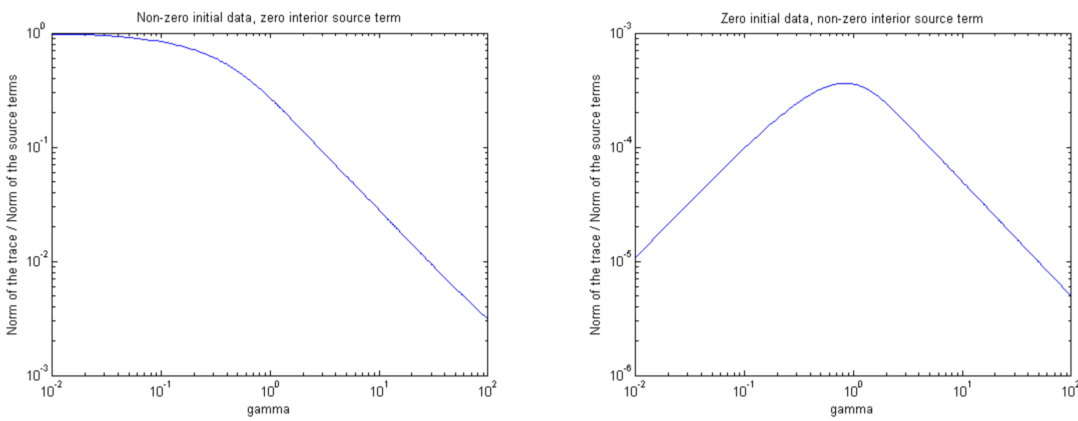

FIGURE 2. Lax-Wendroff scheme (42): ratio between the norm of the trace and the norm of the source terms. Non-zero initial data and zero interior source term (left). Zero initial data and non-zero interior source term (right).

We consider the following two-dimensional problem:

$$\begin{cases} \partial_t u + A_1\,\partial_{x_1} u + A_2\,\partial_{x_2} u = F(t,x)\,, & (t,x) \in \mathbb{R}^+ \times \mathbb{R}^2_+\,, \\ B\,u(t,0,x') = g(t,x')\,, & t \in \mathbb{R}^+,\ x' \in \mathbb{R}, \\ u(0,x) = f(x)\,, & x \in \mathbb{R}^2_+\,, \end{cases} \tag{43}$$

where A_1, A_2 and B are matrices, and A_1 and A_2 are symmetric. We assume that $f \in L^2(\mathbb{R}^2_+)$, and that there exists $\gamma > 0$ such that $g \in e^{\gamma t} L^2(\mathbb{R}^+; L^2(\mathbb{R}))$, and $F \in e^{\gamma t} L^2(\mathbb{R}^+; L^2(\mathbb{R}^2))$. We let f_{j_1,j_2}, $F^n_{j_1,j_2}$ and $g^n_{j_1,j_2}$ be discrete approximations of f, F and g at time $n\Delta t$, and points $(j_1\Delta x_1, j_2\Delta x_2)$, where Δt, Δx_1 and Δx_2 are the time and space steps related through the fixed ratios $\lambda_1 = \Delta x_1/\Delta t$ and $\lambda_2 = \Delta x_2/\Delta t$. As before, $U^n_{j_1,j_2}$ denotes the approximation of the solution u to (43) at time $n\Delta t$, and points $(j_1\Delta x_1, j_2\Delta x_2)$. As for the one-dimensional examples (41) and (42), we address stability issues, not consistency, and we replace the boundary condition in (43) by Dirichlet boundary conditions for the numerical scheme.

4.2.1. *Lax-Friedrichs' scheme.* The Lax-Friedrichs' scheme with homogeneous Dirichlet boundary condition is as follows:

$$\begin{cases} U^{n+1}_{j_1,j_2} &= Q_{LF}\, U^n_{j_1,j_2} + \Delta t\, F^n_{j_1,j_2}\,, & \text{for } j_1 \geq 1\,,\ j_2 \in \mathbb{Z}\,, \quad n \geq 0\,, \\ U^{n+1}_{j_1,j_2} &= g^{n+1}_{j_1,j_2}\,, & \text{for } j_1 \in \{-1,0\}\,,\ j_2 \in \mathbb{Z}\,, \quad n \geq 0\,, \\ U^0_{j_1,j_2} &= f_{j_1,j_2}\,, & \text{for } j_1 \geq -1,\ j_2 \in \mathbb{Z}\,, \end{cases} \tag{44}$$

with the operator Q_{LF} given by:

$$Q_{LF} := \frac{1}{4}\,(T_1^{-1} + T_1 + T_2^{-1} + T_2) - \frac{\lambda_1}{2}\,A_1\,(T_1 - T_1^{-1}) - \frac{\lambda_2}{2}\,A_2\,(T_2 - T_2^{-1})\,.$$

In particular, it is of the form (32) with $p_1 = p_2 = r_1 = r_2 = 1$, $q_1 = q_2 = 0$, and obvious definitions for the matrices A_{j_1,j_2}. Assumptions 3.1 and 3.2 are translated in terms of admissible zones for the CFL parameters (λ_1, λ_2).

Let us begin with Assumption 3.1. In this case, $p_1 = 1$, and the mapping:

$$\mathbb{A}_1(z) : w \,\mapsto\, -\frac{1}{z}\,\left(\frac{1}{4}\,w - \frac{\lambda_1}{2}\,A_1\,w\right) = -\frac{1}{4\,z}\,(I - 2\,\lambda_1\,A_1)\,w$$

has to be coercive on $\ell^2(\mathbb{Z})$ for all $|z| \geq 1$. Since A_1 is diagonalizable, a sufficient condition for Assumption 3.1 to be satisfied is:

$$\lambda_1\,\rho(A_1) \,<\, 1/2\,. \tag{45}$$

Assumption 3.2 is particularly simple in this case, because the symbol of the discretized operator Q_{LF} is a normal matrix. Hence it is diagonalizable in an orthonormal basis and its Hermitian norm coincides with its spectral radius. We refer to [11] for the following sufficient condition for Assumption 3.2 to hold:

$$\forall\,\theta \in [0, 2\,\pi]\,, \quad \rho\big(\lambda_1\,\cos\theta\,A_1 + \lambda_2\,\sin\theta\,A_2\big) \,\leq\, \frac{1}{\sqrt{2}}\,.$$

As shown in [10], the latter condition holds provided that we have[4]:

$$\lambda_1^2\,A_1^2 + \lambda_2^2\,A_2 \,\leq\, \frac{1}{2}\,I\,. \tag{46}$$

If conditions (45) and (46) hold, Theorem 3.2 shows that the solution to (44) satisfies the estimate of Theorem 3.2.

4.2.2. *Modified Lax-Wendroff's scheme I.* Let us now address a variant of the Lax-Wendroff scheme introduced by Wendroff in [12], whose ℓ^2-stability has been further studied by Vaillancourt in [11]. With the notation of Section 3, the scheme is as follows:

$$Q_{LW1} := I - \frac{1}{8}\,\Big(\lambda_1\,A_1\,(T_1 - T_1^{-1}) + \lambda_2\,A_2\,(T_2 - T_2^{-1})\Big) \\ \Big(T_1 + T_1^{-1} + T_2 + T_2^{-1} - \lambda_1\,A_1\,(T_1 - T_1^{-1}) - \lambda_2\,A_2\,(T_2 - T_2^{-1})\Big)\,. \tag{47}$$

For this scheme, $p_1 = 2$ and for all $j_2 \in \mathbb{Z}$,

$$A_{2,j_2} = \frac{1}{8}\,\lambda_1\,A_1\,(I + \lambda_1\,A_1)\,\delta_{0j_2}\,.$$

We first determine the CFL parameters (λ_1, λ_2) for which Assumption 3.2 is satisfied. The symbol of the discretized operator Q_{LW1} is given by:

$$G(\xi) := I - i\,J(\xi)\,\left(c(\xi)\,I - i\,\frac{J(\xi)}{2}\right) = I - \frac{J(\xi)^2}{2} - i\,c(\xi)\,J(\xi)\,,$$

where $J(\xi) := \sin\xi_1\,\lambda_1\,A_1 + \sin\xi_2\,\lambda_2\,A_2$, and $c(\xi) := (\cos\xi_1 + \cos\xi_2)/2$. In particular, $G(\xi)$ is a normal matrix. Hence, $G(\xi)$ is diagonalizable in an orthonormal basis and the scheme is ℓ^2-stable if and only if Assumption 3.2 is satisfied. This stability condition is equivalent to the fulfillment of the von Neumann condition, which reads $G(\xi)\,G^*(\xi) \leq I$, that is:

$$J(\xi)^2 \,\leq\, 4\,(1 - c(\xi))\,I\,.$$

[4] Condition (46) is sufficient for stability, but it is also necessary when the matrices A_1 and A_2 commute.

Following [11], one notes that

$$1-c(\xi)^2=\frac{1}{2}\,(\sin^2\xi_1+\sin^2\xi_2)+\frac{1}{4}\,(\cos\xi_1-\cos\xi_2)^2\,.$$

Hence, Assumption 3.2 holds provided

$$(\sin\xi_1\,\lambda_1\,A_1+\sin\xi_2\,\lambda_2\,A_2)^2\ \le\ 2\,(\sin^2\xi_1+\sin^2\xi_2)\,I\,.$$

This condition is equivalent to

$$\forall\,\theta\in[0,2\,\pi]\,,\quad (\lambda_1\,\cos\theta\,A_1+\lambda_2\,\sin\theta\,A_2)^2\ \le\ 2\,I\,.$$

Therefore, Q_{LW1} satisfies Assumption 3.2 provided

$$\lambda_1^2\,A_1^2+\lambda_2^2\,A_2^2\le 2\,I\,. \tag{48}$$

Let us now turn to Assumption 3.1. In this case, for all $z\in\mathbb{C}^*$, we have:

$$\mathbb{A}_2(z):w\mapsto\frac{1}{8\,z}\,\lambda_1\,A_1\,(I+\lambda_1\,A_1)\,w\,,$$

and $\mathbb{A}_2(z)$ is nothing but a simple multiplication. The mapping $\mathbb{A}_2(z)$ is coercive if:

$$\det A_1\neq 0\quad\text{and}\quad \lambda_1\,\rho_1(A_1)\ <\ 1\,. \tag{49}$$

Consequently Theorem 3.2 shows stability for the modified Lax-Wendroff scheme (47) with Dirichlet boundary conditions provided (48) and (49) hold.

4.2.3. *Modified Lax-Wendroff's scheme II.* Our last example is another modification of the Lax-Wendroff scheme introduced by Lax and Wendroff in [6]. The discretized operator is as follows:

$$Q_{LW2}:=I-\frac{1}{2}\Big(\lambda_1\,A_1\,(T_1-T_1^{-1})+\lambda_2\,A_2\,(T_2-T_2^{-1})\Big)+\frac{1}{8}\Big(\lambda_1\,A_1\,(T_1-T_1^{-1})+\lambda_2\,A_2\,(T_2-T_2^{-1})\Big)^2-L\,, \tag{50}$$

where:

$$L:=\frac{1}{8}\,\big(T_1+T_1^{-1}+T_2+T_2^{-1}-4\,I\big)\,\big(\lambda_1^2\,A_1^2\,(T_1+T_1^{-1}-2\,I)+\lambda_2^2\,A_2^2\,(T_2+T_2^{-1}-2\,I)\big)\,.$$

Although it may seem at first glance that the scheme involves 5 points in each direction, there are cancellations for the terms T_j^2,T_j^{-2} and we have $r_1=r_2=p_1=p_2=1$.

The ℓ^2-stablity of Q_{LW2} has been characterized by Lax and Wendroff in [6] (see also Turkel [10]). However there is a small gap between the general concept of ℓ^2-stablity for which the operators Q_{LW2}^n are bounded uniformly in $n\in\mathbb{N}$, and Assumption 3.2 where we require the norm of Q_{LW2} to be not greater than 1. The latter property is called strong ℓ^2-stability by Tadmor in [9]. The results of [6] and [9] show that the discretized operator Q_{LW2} in (50) is ℓ^2-stable if:

$$\lambda_1^2\,A_1^2+\lambda_2^2\,A_2^2\ \le\ \frac{1}{2}\,I\,,$$

while Assumption 3.2 is satisfied under the slightly more restrictive condition:

$$\lambda_1^4\,A_1^4+\lambda_2^4\,A_2^4\ \le\ \frac{1}{8}\,I\,. \tag{51}$$

Let us now address Assumption 3.1. Unlike the previous examples, $\mathbb{A}_1(z)$ is not a multiplication anylonger. Here, for all $z\in\mathbb{C}^*$, we have:

$$\mathbb{A}_1(z):w\mapsto\frac{1}{2\,z}\Big[\lambda_1\,A_1+\lambda_1^2\,A_1^2+\frac{\lambda_1\lambda_2}{2}\,(A_1\,A_2+A_2\,A_1)\,(T_2-T_2^{-1}) +\frac{1}{4}\,(\lambda_1^2\,A_1^2+\lambda_2^2\,A_2^2)\,(T_2+T_2^{-1}-2\,I)\Big]\,w\,,$$

and Assumption 3.1 is satisfied if and only if $\mathbb{A}_1(1/2)$ is coercive. The symbol of $\mathbb{A}_1(1/2)$ is given by:

$$\widehat{\mathscr{A}}(\xi_2):=\lambda_1\,A_1\,(I+\lambda_1\,A_1)-(\lambda_1^2\,A_1^2+\lambda_2\,A_2^2)\,\sin^2\frac{\xi_2}{2}+i\,\frac{\lambda_1\,\lambda_2}{2}\,(A_1\,A_2+A_2\,A_1)\,\sin\xi_2\,.$$

Using Plancherel's Theorem, it is rather easy to prove that the operator $\mathbb{A}_1(1/2)$ is coercive if and only if its symbol is invertible for all ξ_2.

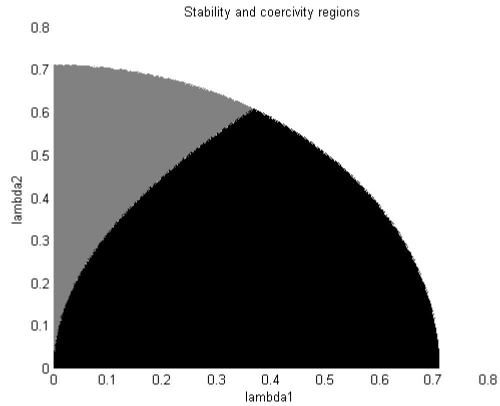

FIGURE 3. Strong stability (grey and black) and coercivity (black) regions for the modified Lax-Wendroff scheme II in the plane (λ_1, λ_2) for the specific example.

Let $C_0 > 0$ denote an arbitrary constant. Then for $\lambda_2 \le C_0\,\lambda_1$, we have:

$$\widehat{\mathscr{A}}(\xi_2) = \lambda_1\,A_1 + O(\lambda_1^2)\,,$$

uniformly in $\xi_2 \in \mathbb{R}$. Hence, Assumption 3.1 holds if $\det A_1 \ne 0$ and λ_1 is small enough (the second CFL parameter λ_2 is subject to the restriction $\lambda_2 \le C_0\,\lambda_1$). In this case, Theorem 3.2 shows the stability of the scheme (50) with Dirichlet boundary conditions and general initial data.

Let us take a closer look at Assumption 3.1 on one specific example. To this aim, we consider:

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,.$$

These matrices satisfy $A_1^2 = A_2^2 = I$ and $A_1\,A_2 + A_2\,A_1 = 0$, so that the symbol $\widehat{\mathscr{A}}$ above reduces to:

$$\widehat{\mathscr{A}}(\xi_2) = \begin{pmatrix} \lambda_1 + \lambda_1^2 - (\lambda_1^2+\lambda_2^2)\,\sin^2\dfrac{\xi_2}{2} & 0 \\ 0 & -\lambda_1 + \lambda_1^2 - (\lambda_1^2+\lambda_2^2)\,\sin^2\dfrac{\xi_2}{2} \end{pmatrix}\,.$$

Let $X := \sin^2(\xi_2/2) \in [0,1]$. The symbol $\widehat{\mathscr{A}}(\xi_2)$ is non-invertible if and only if $\lambda_1+\lambda_2^2-(\lambda_1^2+\lambda_2^2)\,X = 0$, that is:

$$X = \frac{\lambda_1+\lambda_1^2}{\lambda_1^2+\lambda_2^2}\,.$$

Since $X \in [0,1]$, this may only happen for $\lambda_1 \le \lambda_2^2$. Hence, Assumption 3.1 is satisfied if and only if

$$\lambda_2 \,<\, \sqrt{\lambda_1}\,. \tag{52}$$

Theorem 3.2 then shows that the scheme (50) with Dirichlet boundary conditions and general initial data is stable provided (51) and (52) hold. In this case, Assumption 3.1 does have an impact on the stability region, as illustrated on Figure 3. In particular, the condition (52) does not only involve λ_1 but also λ_2. The strong stability region (Assumption 3.2 satisfied) in the plane (λ_1,λ_2) for Q_{LW2} corresponds to the grey and black disk, whereas the black region corresponds to the fulfillment of both (51) and (52).

Appendix A. Proof of Lemma 2.5

Let us recall that the matrices $\mathbb{A}_\ell$'s are defined in (5). Using Assumption 2.1, we can define the following matrix $\mathbb{M}(z)$ that depends holomorphically on z, with $|z| \geq 1$:

$$\forall z \in \mathbb{C}, \quad |z| \geq 1, \quad \mathbb{M}(z) := \begin{pmatrix} -\mathbb{A}_p(z)^{-1}\,\mathbb{A}_{p-1}(z) & \dots & \dots & -\mathbb{A}_p(z)^{-1}\,\mathbb{A}_{-r}(z) \\ I & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 0 & 0 & I & 0 \end{pmatrix} \in \mathscr{M}_{D\,(p+r)}(\mathbb{C})\,. \tag{53}$$

We recall the following result that can be found in [4, page 659]:

Lemma A.1 ([4]). *Let Assumptions 2.1 and 2.2 be satisfied. Then for all $z \in \mathbb{C}$ with $|z| > 1$, the matrix $\mathbb{M}(z)$ defined by (53) has no eigenvalue on the unit circle $\mathbb{S}^1$. We let $\mathbb{E}^s(z)$ denote the generalized eigenspace associated with those eigenvalues of $\mathbb{M}(z)$ that belong to the unit disk.*

The vector space $\mathbb{E}^s(z)$ is made of all vectors of the form $(w_p, \dots, w_{1-r})$, where each w_j belongs to $\mathbb{C}^D$, and such that the sequence $(w_j)_{j\geq 1-r}$ defined by:

$$z\,w_j - Q\,w_j = 0\,, \quad j \geq 1\,, \tag{54}$$

belongs to ℓ^2 (the sequence is even exponentially decreasing). Observe that (54) defines the sequence $(w_j)_{j\geq 1-r}$ in a unique way thanks to Assumption 2.1. The proof of Lemma 2.5 relies on the following preliminary result that is entirely similar to the analysis in [1, page 110]:

Lemma A.2. *Let Assumptions 2.1 and 2.2 be satisfied. Then there exists a constant $C > 0$ such that for all $z \in \mathbb{C}$ with $|z| > 1$, and for all $W = (w_p, \dots, w_{1-r}) \in \mathbb{E}^s(z)$, the sequence $(w_j)_{j\geq 1-r}$ defined by (54) satisfies:*

$$\sum_{j=1}^{q+1} |w_j|^2 \leq C \sum_{j=1-r}^{0} |w_j|^2\,.$$

Proof of Lemma A.2. Let $N \in \mathbb{N}$, and let $z \in \mathbb{C}$ with $|z| > 1$. We consider an element $W = (w_p, \dots, w_{1-r})$ of $\mathbb{E}^s(z)$ and the sequence $(w_j)_{j\geq 1-r} \in \ell^2$ defined by (54). Then we define the source terms g by:

$$g_j^n := \begin{cases} z^n\,w_j\,, & \text{if } 1 \leq n \leq N \text{ and } 1-r \leq j \leq 0, \\ 0\,, & \text{otherwise.} \end{cases}$$

We let V denote the solution to the numerical scheme:

$$\begin{cases} V_j^{n+1} = Q\,V_j^n\,, & j \geq 1\,, \quad n \geq 0\,, \\ V_j^{n+1} = g_j^{n+1}\,, & j = 1-r, \dots, 0\,, \quad n \geq 0\,, \\ V_j^0 = w_j\,, & j \geq 1-r\,. \end{cases}$$

Since the sequence (w_j) satisfies the induction relation (54), we have $V_j^n = z^n\,w_j$ for all $n \leq N$ and $j \geq 1-r$. We apply Theorem 2.2 to V and obtain the following inequalities for all $\gamma > 0$:

$$\begin{aligned} \sum_{n=0}^{N} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=1}^{q+1} |V_j^n|^2 \leq \sum_{n\geq 0} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=1-r}^{\max(p,q+1)} |V_j^n|^2 \\ \leq C \left\{ \|w\|_{1-r,+\infty}^2 + \sum_{n\geq 1}^{N} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j=1-r}^{0} |g_j^n|^2 \right\}. \end{aligned}$$

Using the expression of V and g for $n \leq N$, we find:

$$\sum_{j=1}^{q+1} |w_j|^2 \leq C\,\|w\|_{1-r,+\infty}^2 \left(\Delta t \sum_{n=0}^{N} (|z|\,\mathrm{e}^{-\gamma\,\Delta t})^{2\,n} \right)^{-1} + C \sum_{j=1-r}^{0} |w_j|^2\,.$$

We choose $\gamma > 0$ such that $|z| > \exp(\gamma\,\Delta t)$, then we let N tend to infinity, and the result follows. ☐

Let us now go back to the proof of Lemma 2.5. We consider the solution W to (26), and introduce the piecewise constant functions:

$$\forall\, j \ge 1-r\,,\quad W_j(t) := \begin{cases} W_j^n\,, & \text{if } n \ge 1 \text{ and } t \in [n\,\Delta t, (n+1)\,\Delta t[,\\ 0\,, & \text{if } t < \Delta t. \end{cases}$$

The definition of $\widetilde{g}_j(t)$ is similar. We apply the Laplace transform to (26) and obtain:

$$\begin{cases} z\,\widehat{W}_j = Q\,\widehat{W}_j\,, & j \ge 1\,,\\ z\,\widehat{W}_j = (z\,B_{j,-1} + B_{j,0})\,\widehat{W}_1 + z\,\widehat{g}_j\,, & j = 1-r,\dots,0\,, \end{cases} \tag{55}$$

where z is a short notation for $\exp((\gamma + i\,\omega)\,\Delta t)$, with $\gamma + i\,\omega$ the point where we evaluate the Laplace transform. Using Theorem 4.2 in [4], we know that $(\widehat{W}_j)_{j\ge 1-r}$ satisfies the estimate[5]:

$$\frac{|z|-1}{|z|}\sum_{j\ge 1-r}|\widehat{W}_j|^2 + \sum_{j=1-r}^{0}|\widehat{W}_j|^2 \le C\sum_{j=-r+1}^{0}|\widehat{g}_j|^2\,, \tag{56}$$

with a constant C that is uniform with respect to z.

Since we have $(\widehat{W}_j)_{j\ge 1-r} \in \ell^2$, we know that $(\widehat{W}_p,\dots,\widehat{W}_{1-r}) \in \mathbb{E}^s(z)$. Consequently, we can apply Lemma A.2 and obtain:

$$\sum_{j=1}^{q+1}|\widehat{W}_j|^2 \le C\sum_{j=-r+1}^{0}|\widehat{W}_j|^2 \le C\sum_{j=-r+1}^{0}|\widehat{g}_j|^2\,, \tag{57}$$

where we use (56) to derive the last inequality. The conclusion of Lemma 2.5 follows after integrating (57) with respect to $\omega \in \mathbb{R}$ and using Plancherel's Theorem. The proof of Lemma 2.5 is now complete.

Let us observe that the estimate (56), that is obtained by using the strong stability of the discretization (3), is not sufficient to derive (57) since the control of the "interior" terms $\widehat{W}_1,\dots,\widehat{W}_{q+1}$ deteriorates as $|z|$ tends to 1. The key point is to combine (56) with the estimate of Lemma A.2. The latter is a typical estimate in the frequency variables when the *uniform Kreiss-Lopatinskii condition* holds, see e.g. [1, chapter 4]. A major advantage of our approach is to bypass the verification of the uniform Kreiss-Lopatinskii condition for (7) in the frequency variables (this was the approach followed in [2] and it seems difficult to carry out for systems). Lemma A.2 appears here a direct and easy consequence of the energy method.

References

[1] S. Benzoni-Gavage, D. Serre. *Multidimensional hyperbolic partial differential equations*. Oxford University Press, 2007. First-order systems and applications.

[2] M. Goldberg, E. Tadmor. Scheme-independent stability criteria for difference approximations of hyperbolic initial-boundary value problems. II. *Math. Comp.*, 36(154):603–626, 1981.

[3] B. Gustafsson, H.-O. Kreiss, and J. Oliger. *Time dependent problems and difference methods*. John Wiley & Sons, 1995.

[4] B. Gustafsson, H.-O. Kreiss, and A. Sundström. Stability theory of difference approximations for mixed initial boundary value problems. II. *Math. Comp.*, 26(119):649–686, 1972.

[5] H.-O. Kreiss. Initial boundary value problems for hyperbolic systems. *Comm. Pure Appl. Math.*, 23:277–298, 1970.

[6] P. D. Lax and B. Wendroff. Difference schemes for hyperbolic equations with high order of accuracy. *Comm. Pure Appl. Math.*, 17:381–398, 1964.

[7] D. Michelson. Stability theory of difference approximations for multidimensional initial-boundary value problems. *Math. Comp.*, 40(161):1–45, 1983.

[8] J. Rauch. $\mathscr{L}^2$ is a continuable initial condition for Kreiss' mixed problems. *Comm. Pure Appl. Math.*, 25:265–285, 1972.

[9] E. Tadmor. Complex symmetric matrices with strongly stable iterates. *Linear Algebra Appl.*, 78:65–77, 1986.

[10] E. Turkel. Symmetric hyperbolic difference schemes and matrix problems. *Linear Algebra and Appl.*, 16(2):109–129, 1977.

[11] R. Vaillancourt. On the stability of Friedrichs' scheme and the modified Lax-Wendroff scheme. *Math. Comp.*, 24:767–770, 1970.

[12] B. Wendroff. Well-posed problems and stable difference operators. *SIAM J. Numer. Anal.*, 5:71–82, 1968.

[5] The estimate (56) is the analogue in the frequency variables of the strong stability estimate in definition 2.1.

[13] L. Wu. The semigroup stability of the difference approximations for initial-boundary value problems. *Math. Comp.*, 64(209):71–88, 1995.

E-mail address: jean-francois.coulombel@math.univ-lille1.fr
E-mail address: antoine.gloria@inria.fr