École Polytechnique
Université Paris-Saclay

Journal de l'École polytechnique
Mathématiques

Jean-François Coulombel

The Leray-Gårding method for finite difference schemes

Tome 2 (2015), p. 297-331.

<http://jep.cedram.org/item?id=JEP_2015__2__297_0>

Journal de l'École polytechnique
Mathématiques
Tome 2, 2015, p. 297–331 DOI: 10.5802/jep.25

THE LERAY-GÅRDING METHOD FOR FINITE DIFFERENCE SCHEMES

by Jean-François Coulombel

Abstract. — In the fifties, Leray and Gårding have developed a multiplier technique for deriving a priori estimates for solutions to scalar hyperbolic equations. The existence of such a multiplier is the starting point of the argument by Rauch [23] for the derivation of semigroup estimates for hyperbolic initial boundary value problems. In this article, we explain how this multiplier technique can be adapted to the framework of finite difference approximations of transport equations. The technique applies to numerical schemes with arbitrarily many time levels. The existence and properties of the multiplier enable us to derive *optimal* semigroup estimates for fully discrete hyperbolic initial boundary value problems.

Résumé (La méthode de Leray et Gårding pour les schémas aux différences finies)

Dans les années 1950, Leray et Gårding ont développé une technique de multiplicateur pour obtenir des estimations a priori de solutions d'équations hyperboliques scalaires. L'existence d'un multiplicateur est le point de départ du travail de Rauch [23] pour montrer des estimations de semi-groupe pour les problèmes aux limites hyperboliques. Dans cet article, nous expliquons comment cette technique de multiplicateur peut être adaptée au cadre des schémas aux différences finies pour les équations de transport. Ce travail s'applique à des schémas numériques multi-pas en temps. L'existence et les propriétés du multiplicateur nous permettent d'obtenir des estimations de semi-groupe optimales pour des versions totalement discrètes des problèmes aux limites hyperboliques.

Contents

Mathematical subject classification (2010). — 65M06, 65M12, 35L03, 35L04.
Keywords. — Hyperbolic equations, difference approximations, stability, boundary conditions, semigroup.

The research of the author was supported by the ANR project BoND, ANR-13-BS01-0009-01.

Throughout this article, we use the notation

$$\mathscr{U} := \{\zeta \in \mathbb{C}, |\zeta| > 1\}, \qquad \overline{\mathscr{U}} := \{\zeta \in \mathbb{C}, |\zeta| \geqslant 1\},$$
$$\mathbb{D} := \{\zeta \in \mathbb{C}, |\zeta| < 1\}, \qquad \mathbb{S}^1 := \{\zeta \in \mathbb{C}, |\zeta| = 1\}, \qquad \overline{\mathbb{D}} := \mathbb{D} \cup \mathbb{S}^1.$$

We let $\mathscr{M}_n(\mathbb{K})$ denote the set of $n \times n$ matrices with entries in $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. If $M \in \mathscr{M}_n(\mathbb{C})$, M^* denotes the conjugate transpose of M. We let I denote the identity matrix or the identity operator when it acts on an infinite dimensional space. We use the same notation $x^* y$ for the Hermitian product of two vectors $x, y \in \mathbb{C}^n$ and for the Euclidean product of two vectors $x, y \in \mathbb{R}^n$. The norm of a vector $x \in \mathbb{C}^n$ is $|x| := (x^* x)^{1/2}$. The induced matrix norm on $\mathscr{M}_n(\mathbb{C})$ is denoted $\| \cdot \|$.

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.

In what follows, we let $d \geqslant 1$ denote a fixed integer, which will stand for the dimension of the space domain we are considering. We also use the space ℓ^2 of square integrable sequences. Sequences may be valued in $\mathbb{C}^k$ for some integer k. Some sequences will be indexed by $\mathbb{Z}^{d-1}$ while some will be indexed by $\mathbb{Z}^d$ or a subset of $\mathbb{Z}^d$. We thus introduce some specific notation for the norms. Let $\Delta x_i > 0$ for $i = 1, \dots, d$ be d space steps. We shall make use of the $\ell^2(\mathbb{Z}^{d-1})$-norm that we define as follows: for all $v \in \ell^2(\mathbb{Z}^{d-1})$,

$$\|v\|^2_{\ell^2(\mathbb{Z}^{d-1})} := \Big(\prod_{k=2}^{d} \Delta x_k \Big) \sum_{i=2}^{d} \sum_{j_i \in \mathbb{Z}} |v_{j_2,\dots,j_d}|^2 .$$

The corresponding scalar product is denoted $\langle \cdot , \cdot \rangle_{\ell^2(\mathbb{Z}^{d-1})}$. Then for all integers $m_1 \leqslant m_2$, we set

$$|||u|||^2_{m_1,m_2} := \Delta x_1 \sum_{j_1=m_1}^{m_2} \|u_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})},$$

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 $+\infty$). The corresponding scalar product is denoted $\langle \cdot , \cdot \rangle_{m_1,m_2}$. Other notation throughout the text is meant to be self-explanatory.

1. Introduction

1.1. The context. — The goal of this article is to derive semigroup estimates for finite difference approximations of hyperbolic initial boundary value problems. Up to now, the only available general stability theory for such numerical schemes is due to Gustafsson, Kreiss and Sundström [13]. It relies on a Laplace transform with respect to the time variable. For various technical reasons, the corresponding stability estimates are restricted to zero initial data. A long standing problem in this line of research is, starting from the GKS stability estimates, which are *resolvent* type estimates, to incorporate nonzero initial data and to derive *semigroup* estimates, see, e.g., the discussion by Trefethen in [28, §4] and the conjecture by Kreiss and Wu in [17]. This problem is delicate for the following reason: the validity of the GKS stability estimate is known to be equivalent to a *slightly stronger version* of the resolvent estimate

$$\sup_{z\in\mathscr{U}} (|z|-1)\big\|(zI-T)^{-1}\big\|_{\mathscr{L}(\ell^2(\mathbb{N}))} < +\infty, \tag{1.1}$$

where T is some bounded operator on $\ell^2(\mathbb{N})$ that incorporates both the discretization of the hyperbolic equation and the numerical boundary conditions. Deriving an optimal semigroup estimate amounts to showing that T is power bounded:

$$\sup_{n\geqslant 1} \|T^n\|_{\mathscr{L}(\ell^2(\mathbb{N}))} < +\infty.$$

In finite dimension, the equivalence between power boundedness of T and the resolvent condition (1.1) is known as the Kreiss matrix Theorem, but the analogous equivalence is known to fail in general in infinite dimension. Worse, even the strong resolvent condition

$$\sup_{n\geqslant 1}\sup_{z\in\mathscr{U}} (|z|-1)^n\big\|(zI-T)^{-n}\big\|_{\mathscr{L}(\ell^2(\mathbb{N}))} < +\infty,$$

does not imply in general that T is power bounded, see, e.g., the review [26] or [29, Chap. 18] for details and historical comments.

Optimal semigroup estimates have nevertheless been derived for some discretized hyperbolic initial boundary value problems. The very first results in this direction date back to Kreiss and Osher [16, 22, 21], even though these works precede [13] but the main results are exactly of the form we discuss. The first general derivation of semigroup estimates starting from GKS stability is due to Wu [31], whose analysis deals with numerical schemes with two time levels and scalar equations (as in [16, 22, 21]). The results in [31] were extended by Gloria and the author in [7] to *systems* in arbitrary space dimension, but the arguments in [7] are still restricted to numerical schemes with two time levels. The present article gives, as far as we are aware of, the first systematic derivation of semigroup estimates for fully discrete hyperbolic initial boundary value problems in the case of numerical schemes with arbitrarily many time levels. It generalizes the arguments of [31, 7] and provides new insight for the construction of 'dissipative' (sometimes called 'absorbing') numerical boundary conditions for discretized evolution equations. Let us observe that the leap-frog scheme, with some very specific homogeneous boundary conditions, has been dealt with by Thomas [27] by using a *multiplier* technique. It is precisely this technique which we aim at developing in a systematic fashion for numerical schemes with arbitrarily many time levels. In particular, we shall explain why the somehow magical multiplier $u_j^{n+2}+u_j^n$ for the leap-frog scheme, see, e.g., [24], follows from a general theory that is the analogue of the Leray–Gårding method [18, 10] for hyperbolic partial differential equations.

1.2. The main result. — We first set a few notations. We let $\Delta x_1,\dots,\Delta x_d,\Delta t>0$ denote space and time steps where the ratios, the so-called Courant-Friedrichs-Lewy parameters, $\lambda_i := \Delta t/\Delta x_i$, $i=1,\dots,d$, are fixed positive constants. We keep $\Delta t\in(0,1]$ as a small parameter and let the space steps $\Delta x_1,\dots,\Delta x_d$ vary accordingly. The ℓ^2-norms with respect to the space variables have been previously defined and thus depend on Δt and the CFL parameters through the cell volume

($\Delta x_2 \cdots \Delta x_d$ on $\mathbb{Z}^{d-1}$, and $\Delta x_1 \cdots \Delta x_d$ on $\mathbb{Z}^d$). We always identify a sequence w indexed by either $\mathbb{N}$ (for time), $\mathbb{Z}^{d-1}$ or $\mathbb{Z}^d$ (for space), with the corresponding step function. For instance, for a sequence $(w^n)_{n\in\mathbb{N}}$, the step function w reads:

$$w(t) := w^n, \quad \forall\, t \in [n\Delta t, (n+1)\Delta t), \quad \forall\, n \in \mathbb{N}.$$

In particular, we shall feel free to take Fourier and/or Laplace transforms of sequences.

For all $j \in \mathbb{Z}^d$, we set $j = (j_1, j')$ with $j' := (j_2, \dots, j_d) \in \mathbb{Z}^{d-1}$. We let $p, q, r \in \mathbb{N}^d$ denote some fixed multi-integers, and define $p_1, q_1, r_1, p', q', r'$ according to the above notation. We also let $s \in \mathbb{N}$ denote some fixed integer. We consider a recurrence relation of the form:

$$\begin{cases} \displaystyle\sum_{\sigma=0}^{s+1} Q_\sigma u_j^{n+\sigma} = \Delta t F_j^{n+s+1}, & j \in \mathbb{Z}^d, \quad j_1 \geqslant 1, \quad n \geqslant 0, \\ u_j^{n+s+1} + \displaystyle\sum_{\sigma=0}^{s+1} B_{j_1,\sigma} u_{1,j'}^{n+\sigma} = g_j^{n+s+1}, & j \in \mathbb{Z}^d, \quad j_1 = 1 - r_1, \dots, 0, \quad n \geqslant 0, \\ u_j^n = f_j^n, & j \in \mathbb{Z}^d, \quad j_1 \geqslant 1 - r_1, \quad n = 0, \dots, s, \end{cases} \tag{1.2}$$

where the operators Q_σ and $B_{j_1,\sigma}$ are given by:

$$Q_\sigma := \sum_{\ell_1=-r_1}^{p_1} \sum_{\ell'=-r'}^{p'} a_{\ell,\sigma} \boldsymbol{S}^\ell, \quad B_{j_1,\sigma} := \sum_{\ell_1=0}^{q_1} \sum_{\ell'=-q'}^{q'} b_{\ell,j_1,\sigma} \boldsymbol{S}^\ell. \tag{1.3}$$

Let us comment a little on (1.2), (1.3). First of all, in (1.3), the coefficients $a_{\ell,\sigma}, b_{\ell,j_1,\sigma}$ are *real numbers* and are independent of the small parameter Δt (they may depend on the CFL parameters though), while $\boldsymbol{S}$ denotes the shift operator on the space grid: $(\boldsymbol{S}^\ell v)_j := v_{j+\ell}$ for $j, \ell \in \mathbb{Z}^d$. We have also used the short notation

$$\sum_{\ell'=-r'}^{p'} := \sum_{i=2}^{d} \sum_{\ell_i=-r_i}^{p_i}, \qquad \sum_{\ell'=-q'}^{q'} := \sum_{i=2}^{d} \sum_{\ell_i=-q_i}^{q_i}.$$

This means that each operator $Q_\sigma, B_{j_1,\sigma}$ in (1.2) acts on sequences indexed by the 'spatial variable' $j \in \mathbb{Z}^d$. In particular, the 'time variable' n enters as a parameter when we apply these operators and, for instance, $Q_\sigma u_j^{n+\sigma}$ is a short notation for the application of the operator Q_σ to the sequence $u^{n+\sigma}$, this resulting sequence being evaluated at the space index j.

The numerical scheme (1.2) should then be understood as follows: one starts with ℓ^2 initial data $(f_j^0), \dots, (f_j^s)$ defined for $j_1 \geqslant 1 - r_1$, $j' \in \mathbb{Z}^{d-1}$, some given interior source term (F_j^n) defined for $j_1 \geqslant 1$, $j' \in \mathbb{Z}^{d-1}$, $n \geqslant s+1$, and some given boundary source term (g_j^n) defined for $j_1 = 1 - r_1, \dots, 0$, $j' \in \mathbb{Z}^{d-1}$, $n \geqslant s+1$. The (space) cells associated with $j_1 \geqslant 1$ correspond to the *interior domain* (the discretized counterpart of the half-space $\{x_1 > 0\}$ in $\mathbb{R}^d$), while those associated with $j_1 = 1 - r_1, \dots, 0$ represent the *discrete boundary* (the discretized counterpart of the hyperplane $\{x_1 = 0\}$ in $\mathbb{R}^d$). Assuming that the solution (u_j^n) has been defined up to some time index $n + s$ for all $j_1 \geqslant 1 - r_1$, $j' \in \mathbb{Z}^{d-1}$, with $n \geqslant 0$, then the first and second equations in (1.2) should uniquely determine u_j^{n+s+1} for $j_1 \geqslant 1-r_1$, $j' \in \mathbb{Z}^{d-1}$. More precisely, the equation

$$\sum_{\sigma=0}^{s+1} Q_\sigma u_j^{n+\sigma} = \Delta t F_j^{n+s+1},$$

is meant to give an update for the interior values of u^{n+s+1} by possibly inverting the operator Q_{s+1}, and the numerical boundary conditions

$$u_j^{n+s+1} + \sum_{\sigma=0}^{s+1} B_{j_1,\sigma} u_{1,j'}^{n+\sigma} = g_j^{n+s+1}$$

should determine the boundary values u_j^{n+s+1}, $j_1 = 1-r_1, \ldots, 0$, by possibly giving their expression in terms of finitely many interior values. More precisely, for each $j_1 = 1-r_1, \ldots, 0$, and independently of the tangential variable $j' \in \mathbb{Z}^{d-1}$, u_j^{n+s+1} is assumed to be given by a linear combination of interior values, which amounts to considering a linear combination of $u_{1,j'+\ell'}^{n+\sigma}, \ldots, u_{1+q_1,j'+\ell'}^{n+\sigma}$ with a finite stencil for ℓ' and $\sigma = 0, \ldots, s+1$. Some examples will be given later on. Let us just emphasize here that some more general numerical boundary conditions might probably be considered, including convolution type formula with respect to n, but we restrict in this article to this form of numerical boundary conditions for simplicity.

We wish to deal here simultaneously with explicit and implicit schemes and therefore make the following solvability assumption.

Assumption 1 (Solvability of (1.2)). — The operator Q_{s+1} is an isomorphism on $\ell^2(\mathbb{Z}^d)$. Moreover, for all $(F_j) \in \ell^2(\mathbb{N}^* \times \mathbb{Z}^{d-1})$ and for all $g_{1-r_1,\bullet}, \ldots, g_{0,\bullet} \in \ell^2(\mathbb{Z}^{d-1})$, there exists a unique solution $(u_j)_{j_1 \geqslant 1-r_1} \in \ell^2$ to the system

$$\begin{cases} Q_{s+1} u_j = F_j, & j \in \mathbb{Z}^d, \quad j_1 \geqslant 1, \\ u_j + B_{j_1,s+1} u_{1,j'} = g_j, & j \in \mathbb{Z}^d, \quad j_1 = 1-r_1, \ldots, 0. \end{cases}$$

In particular, Assumption 1 is trivially satisfied in the case of explicit schemes for which Q_{s+1} is the identity ($a_{\ell,s+1} = \delta_{\ell_1,0} \cdots \delta_{\ell_d,0}$ in (1.3), with δ the Kronecker symbol), for in that case, we first determine all interior values by the relation $u_j = F_j$ (F_j is given), and we then use these values to define the boundary values u_j, $j_1 = 1-r_1, \ldots, 0$. The situation is more or less the same as for lower or upper triangular systems. When Q_{s+1} is not the identity, then the values u_j, $j_1 \geqslant 1-r_1$, should be determined all at once.

Using Assumption 1, the first and second equations in (1.2) uniquely determine u_j^{n+s+1} for $j_1 \geqslant 1-r_1$, and one then proceeds to the following time index $n+s+2$. Existence and uniqueness of a solution (u_j^n) to (1.2) thus follows from Assumption 1, so the last requirement for well-posedness is continuous dependence of the solution on the three possible source terms (F_j^n), (g_j^n), (f_j^n). This is a *stability* problem for which several definitions can be chosen according to the functional framework. The following one dates back to [13] in one space dimension and was also considered by Michelson [19] in several space dimensions. It is specifically relevant when the boundary conditions are non-homogeneous ($(g_j^n) \not\equiv 0$).

Definition 1 (Strong stability). — The finite difference approximation (1.2) 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) to (1.2) with $(f_j^0) = \cdots = (f_j^s) = 0$ satisfies the estimate:

$$
(1.4)\quad \frac{\gamma}{\gamma\Delta t+1}\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}|||u^n|||_{1-r_1,+\infty}^2 + \sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j=1-r_1}^{p_1}\|u^n_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})} \leqslant C\Big\{\frac{\gamma\Delta t+1}{\gamma}\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}|||F^n|||_{1,+\infty}^2 + \sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1=1-r_1}^{0}\|g^n_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})}\Big\}.
$$

The main contributions in [13, 19] are to show that strong stability can be characterized by a certain *algebraic condition*, which is usually referred to as the Uniform Kreiss-Lopatinskii Condition, see [6] for an overview of such results. We do not pursue such arguments here but rather assume from the start that (1.2) is strongly stable. We can thus control, with zero initial data, ℓ^2 type norms of the solution to (1.2). Our goal is to understand which kind of stability estimate holds for the solution to (1.2) when one now considers nonzero initial data $(f_j^0), \dots, (f_j^s)$ in ℓ^2. Our main assumption is the following and is related to stability of the recurrence relation:

$$
\sum_{\sigma=0}^{s+1} Q_\sigma u_j^{n+\sigma} = 0,
$$

when applied on all $\mathbb{Z}^d$. Recall that in that case, stability is usually understood in the ℓ^2 sense, and by Fourier analysis, this leads to studying solutions to the latter recurrence relation of the form $z^n \exp(ij\cdot\xi)$ for arbitrary wave vectors $\xi \in \mathbb{R}^d$, hence the so-called dispersion relation (1.5) below, see, e.g., [24, 12].

Assumption 2 (Stability for the discrete Cauchy problem). — For all $\xi \in \mathbb{R}^d$, the dispersion relation

$$
(1.5)\qquad \sum_{\sigma=0}^{s+1}\widehat{Q_\sigma}(\mathrm{e}^{i\xi_1},\dots,\mathrm{e}^{i\xi_d})z^\sigma = 0, \quad \widehat{Q_\sigma}(\kappa) := \sum_{\ell=-r}^{p}\kappa^\ell a_{\ell,\sigma},
$$

has $s+1$ simple roots in $\overline{\mathbb{D}}$. (The von Neumann condition is said to hold when the roots are located in $\overline{\mathbb{D}}$.) In (1.5), we have used the classical notation $\kappa^\ell := \kappa_1^{\ell_1}\cdots\kappa_d^{\ell_d}$ for $\kappa \in (\mathbb{C}\setminus\{0\})^d$ and $\ell \in \mathbb{Z}^d$.

Examples of numerical schemes that satisfy Assumption 2 are given in Section 1.3. At the opposite, Assumption 2 excludes numerical schemes that are based on first performing a space discretization and then using Adams-Bashforth or Adams-Moulton methods of order 3 or higher (such methods have 0 as a root of multiplicity 2 or more at the frequency $\xi = 0$, see [14, Chap. III.3]).

From Assumption 1, we know that Q_{s+1} is an isomorphism on ℓ^2, which implies by Fourier analysis that $\widehat{Q_{s+1}}(\mathrm{e}^{i\xi_1}, \dots, \mathrm{e}^{i\xi_d})$ does not vanish for any $\xi \in \mathbb{R}^d$. In particular, the dispersion relation (1.5) is a polynomial equation of degree $s+1$ in z for any $\xi \in \mathbb{R}^d$. We now make the following assumption, which already appeared in [13, 19] and several other works on the same topic.

Assumption 3 (Noncharacteristic discrete boundary). — For $\ell_1 = -r_1, \dots, p_1$, $z \in \mathbb{C}$ and $\eta \in \mathbb{R}^{d-1}$, let us define

$$a_{\ell_1}(z,\eta) := \sum_{\sigma=0}^{s+1} z^\sigma \sum_{\ell'=-r'}^{p'} a_{\ell,\sigma} \mathrm{e}^{i\ell'\cdot\eta}. \tag{1.6}$$

Then a_{-r_1} and a_{p_1} do not vanish on $\overline{\mathscr{U}} \times \mathbb{R}^{d-1}$, and they have nonzero degree with respect to z for all $\eta \in \mathbb{R}^{d-1}$.

Assumption 3 is used when one performs a Laplace-Fourier transform on (1.2). The Laplace transform refers to the time variable $n \in \mathbb{N}$ and the Fourier transform refers to the tangential space variables $j' \in \mathbb{Z}^{d-1}$. One is then led to a recurrence relation with respect to the space normal variable j_1 which, thanks to Assumption 3 can be either written as an 'evolution' equation for increasing or decreasing j_1. This will be used in Section 3.

Our main result is comparable with [31, Th. 3.3] and [7, Th. 2.4 & 3.5] and shows that strong stability (or GKS stability) is a sufficient condition for incorporating ℓ^2 initial conditions in (1.2) and proving *optimal* semigroup estimates. The main price to pay in Assumption 2 is that the roots of the dispersion relation (1.5), which are nothing but the eigenvalues of the so-called *amplification matrix* for the Cauchy problem, need to be *simple*. This property is satisfied by the leap-frog and modified leap-frog schemes in two space dimensions under an appropriate CFL condition, see Section 1.3 for more examples. Our main result reads as follows.

Theorem 1. — *Let Assumptions* 1, 2 *and* 3 *be satisfied, and assume that the scheme* (1.2) *is strongly stable in the sense of Definition* 1. *Then there exists a constant C such that for all $\gamma > 0$ and all $\Delta t \in (0,1]$, the solution to* (1.2) *satisfies the estimate:*

$$\begin{aligned} \sup_{n\geqslant 0} \mathrm{e}^{-2\gamma n\Delta t} |\!|\!| u^n |\!|\!|^2_{1-r_1,+\infty} + \frac{\gamma}{\gamma\Delta t+1} \sum_{n\geqslant 0} \Delta t \mathrm{e}^{-2\gamma n\Delta t} |\!|\!| u^n |\!|\!|^2_{1-r_1,+\infty} \\ + \sum_{n\geqslant 0} \Delta t \mathrm{e}^{-2\gamma n\Delta t} \sum_{j_1=1-r_1}^{p_1} \| u^n_{j_1,\bullet} \|^2_{\ell^2(\mathbb{Z}^{d-1})} \\ \leqslant C \Big\{ \sum_{\sigma=0}^{s} |\!|\!| f^\sigma |\!|\!|^2_{1-r_1,+\infty} + \frac{\gamma\Delta t+1}{\gamma} \sum_{n\geqslant s+1} \Delta t \mathrm{e}^{-2\gamma n\Delta t} |\!|\!| F^n |\!|\!|^2_{1,+\infty} \\ + \sum_{n\geqslant s+1} \Delta t \mathrm{e}^{-2\gamma n\Delta t} \sum_{j_1=1-r_1}^{0} \| g^n_{j_1,\bullet} \|^2_{\ell^2(\mathbb{Z}^{d-1})} \Big\}. \end{aligned} \tag{1.7}$$

In particular, the scheme (1.2) *is 'semigroup stable' in the sense that there exists a constant C such that for all $\Delta t \in (0,1]$, the solution (u^n_j) to* (1.2) *with $(F^n_j) = (g^n_j) = 0$*

satisfies the estimate

$$\sup_{n \geqslant 0} \| u^n \|_{1-r_1,+\infty}^2 \leqslant C \sum_{\sigma=0}^{s} \| f^\sigma \|_{1-r_1,+\infty}^2 . \tag{1.8}$$

The scheme (1.2) *is also ℓ^2-stable with respect to boundary data, see* [28, Def. 4.5]*, in the sense that there exists a constant C such that for all $\Delta t \in (0,1]$, the solution (u_j^n) to* (1.2) *with $(F_j^n) = (f_j^n) = 0$ satisfies the estimate*

$$\sup_{n \geqslant 0} \| u^n \|_{1-r_1,+\infty}^2 \leqslant C \sum_{n \geqslant s+1} \Delta t \sum_{j_1=1-r_1}^{0} \| g_{j_1,\bullet}^n \|_{\ell^2(\mathbb{Z}^{d-1})}^2 .$$

The semigroup estimate (1.8) as well as ℓ^2-stability with respect to boundary data are indeed trivial consequences of the main estimate (1.7) by letting the parameter γ tend to zero. Our main contribution in this article is to exhibit a suitable multiplier for the multistep recurrence relation in (1.2). With this multiplier, we can readily show that, for zero initial data, the (discrete) derivative of an *energy* can be controlled, as in the work by Rauch [23] on partial differential equations, by the trace estimate of (u_j^n) and this is where strong stability comes into play. This first argument gives Theorem 1 for zero initial data(1). By linearity we can then reduce to the case of zero forcing terms in the interior and on the boundary. The next arguments in [23] use time reversibility, which basically always fails for numerical schemes(2). Hence we must find another argument for dealing with nonzero initial data. Hopefully, the properties of our multiplier enable us to construct an auxiliary problem, where we modify the boundary conditions of (1.2), and for which we can prove optimal semigroup and trace estimates by 'hand-made' calculations. In other words, we exhibit an alternative set of boundary conditions that yields *strict dissipativity.* Using these auxiliary numerical boundary conditions, the proof of Theorem 1 follows from an easy (though lengthy) superposition argument, see, e.g., [3, §4.5] for partial differential equations or [31, 7] for numerical schemes.

Remark 1. — It could seem at first that the general form of (1.2) incorporates not only finite difference approximations of hyperbolic equations but more generally finite difference approximations of any evolutionary constant coefficient partial differential equation. Hence Theorem 1 could apply to more general situations. However, it should be kept in mind that we assume here that each ratio $\Delta t/\Delta x_j$ is constant, and therefore we consider each coefficient in the operators $Q_\sigma, B_{j_1,\sigma}$ as independent of the time and space steps. This point of view has some technical advantages since we may for instance view the Q_σ's as bounded operators with norms that are independent of the time and space steps, and all estimates in Theorem 1 are in fact independent

(1)It would even give the claim of Theorem 1 for nonzero initial data provided that the non-glancing condition of [4] is satisfied, but we do not wish to make such an assumption here.

(2)With the notable exception of the leap-frog scheme that is time reversible! Some schemes based on the Crank-Nicolson integration rule are also time reversible.

of the (only left) small parameter Δt (just divide for instance (1.8) by Δt^d and use the definitions of the norms on either side to simplify all cell volumes). However, our assumption is also a clear limitation when the original system is stiff and one uses implicit schemes in order to get free of CFL constraints (just think of an implicit discretization of the heat equation). In that case, there would be several small parameters involved and the coefficients in $Q_\sigma, B_{j_1,\sigma}$ could not all necessarily be considered as constants (or even bounded). We postpone the extension of this work to parabolic or dispersive equations to a future work.

1.3. Examples

1.3.1. *Examples in one space dimension*. — Our goal is to approximate the outgoing transport equation ($d = 1$ here):

$$\partial_t u + a\partial_x u = 0, \quad u|_{t=0} = u_0, \tag{1.9}$$

with $t, x > 0$ and $a < 0$. The latter transport equation does not require any boundary condition at $x = 0$. However, discretizing (1.9) usually requires prescribing numerical boundary conditions, unless one considers an upwind type scheme with a space stencil 'on the right' (meaning $r_1 = 0$ in (1.2)).

Let us first emphasize that Assumption 3 excludes the case of explicit two level schemes for which $s = 0$ and $Q_1 = I$, for in that case a_{-r_1} and/or a_{p_1} do not depend on z. However, this case has already been dealt with in [31, 7], and we shall see in Section 3 where the assumption that a_{-r_1} and a_{p_1} are *not constant* is involved, and why the proof is actually simpler in the case $s = 0$ and $Q_1 = I$.

We now detail two possible multistep schemes for discretizing (1.9). Both are obtained by the so-called method of lines, which amounts to first discretizing the space derivative $\partial_x u$ and then choosing an integration technique for discretizing the time evolution, see [12].

The leap-frog scheme. — It is obtained by approximating the space derivative $\partial_x u$ by the centered difference $(u_{j+1} - u_{j-1})/(2\Delta x)$, and by then applying the so-called Nyström method of order 2, see [14, Chap. III.1]. The resulting approximation reads

$$u_j^{n+2} + \lambda a(u_{j+1}^{n+1} - u_{j-1}^{n+1}) - u_j^n = 0,$$

which corresponds to $s = p = r = 1$ (here $d = 1$ so we write p instead of p_1 and so on). Recall that $\lambda > 0$ denotes the fixed ratio $\Delta t/\Delta x$. Even though (1.9) does not require any boundary condition at $x = 0$, the leap-frog scheme stencil includes one point to the left, and we therefore need to prescribe some numerical boundary condition at $j = 0$. One possibility is to prescribe the homogeneous or inhomogeneous Dirichlet boundary condition. With all possible source terms, the corresponding scheme reads

$$\begin{cases} u_j^{n+2} + \lambda a(u_{j+1}^{n+1} - u_{j-1}^{n+1}) - u_j^n = \Delta t F_j^{n+2}, & j \geqslant 1, \quad n \geqslant 0, \\ u_0^{n+2} = g_0^{n+2}, & n \geqslant 0, \\ (u_j^0, u_j^1) = (f_j^0, f_j^1), & j \geqslant 0. \end{cases} \tag{1.10}$$

Assumption 1 is trivially satisfied because (1.10) is explicit. More precisely, (1.10) can be written under the form (1.2) by setting:

$$Q_2 := I, \quad Q_1 := \lambda a(\boldsymbol{S} - \boldsymbol{S}^{-1}), \quad Q_0 := -I,$$

and all operators $B_{j,\sigma}$ are zero ($q = 0$ also). The leap-frog scheme satisfies Assumption 2 provided that $\lambda|a| < 1$. In that case, the two roots to the dispersion relation

$$z^2 + 2i\lambda a \sin\xi z - 1 = 0, \tag{1.11}$$

are simple and have modulus 1 for all $\xi \in \mathbb{R}$. Assumption 3 is satisfied as long as the velocity a is nonzero, for in that case $a_1(z) = -a_{-1}(z) = \lambda a z$. The scheme (1.10) is known to be strongly stable, see [11]. In particular, Theorem 1 shows that (1.10) is semigroup stable. More accurate numerical boundary conditions can be considered, and we refer to [12, 20, 25, 28] for some other possible choices which might be more meaningful from a consistency and accuracy point of view.

A scheme based on the backwards differentiation rule. — We still start from the transport equation (1.9), approximate the space derivative $\partial_x u$ by the centered finite difference $(u_{j+1} - u_{j-1})/(2\Delta x)$, and then apply the backwards differentiation formula of order 2, see [14, Chap. III.1]. The resulting scheme reads:

$$\frac{3}{2}u_j^{n+2} + \frac{\lambda a}{2}(u_{j+1}^{n+2} - u_{j-1}^{n+2}) - 2u_j^{n+1} + \frac{1}{2}u_j^n = 0, \tag{1.12}$$

which corresponds to $s = 1$ and

$$Q_2 := \frac{3}{2}I + \frac{\lambda a}{2}(\boldsymbol{S} - \boldsymbol{S}^{-1}), \quad Q_1 := -2I, \quad Q_0 := \frac{1}{2}I.$$

The operator Q_2 is an isomorphism on $\ell^2(\mathbb{Z})$ since Q_2 is an isomorphism for any small λa (as a perturbation of $3/2I$), Q_2 depends continuously on λa, and there holds (uniformly with respect to λa):

$$\frac{3}{2}|\!|\!|u|\!|\!|_{-\infty,+\infty} \leqslant |\!|\!|Q_2 u|\!|\!|_{-\infty,+\infty}.$$

The operator Q_2 is therefore an isomorphism on $\ell^2(\mathbb{Z})$ for any λa (see, e.g., [5, Lem. 4.3]). Let us now study the dispersion relation (1.5), which reads here

$$\Big(\frac{3}{2} + i\lambda a \sin\xi\Big)z^2 - 2z + \frac{1}{2} = 0. \tag{1.13}$$

It is clear that the latter equation has two simple roots in z for any $\xi \in \mathbb{R}$. Moreover, if $\sin\xi = 0$, the roots are 1 and 1/3 which belong to $\overline{\mathbb{D}}$. In the case $\sin\xi \neq 0$, none of the roots belongs to $\mathbb{S}^1$ and examining the case $\lambda a \sin\xi = 1$, we find that for $\sin\xi \neq 0$, both roots belong to $\mathbb{D}$ (which is consistent with the shape of the stability region for the backwards differentiation formula of order 2, see [15, Chap. V.1]). Assumption 2 is therefore satisfied. Assumption 3 is satisfied as long as a is nonzero since there holds $p = r = 1$ and $a_1(z) = -a_{-1}(z) = \lambda a z^2/2$.

Theorem 1 therefore yields semigroup boundedness as long as one uses numerical boundary conditions for which the numerical scheme is well-defined (this is at least the case for λa small enough) and strong stability holds. We shall go back later on to the form of our multiplier for the scheme (1.12) and compare it with another technique that is available in the literature, see, e.g., [8, 9] and references therein.

1.3.2. *Examples in two space dimensions*. — We now wish to approximate the two-dimensional transport equation ($d = 2$):

$$\partial_t u + a_1 \partial_{x_1} u + a_2 \partial_{x_2} u = 0, \quad u|_{t=0} = u_0, \tag{1.14}$$

in the space domain $\{x_1 > 0, x_2 \in \mathbb{R}\}$. When a_1 is negative, the latter problem does not necessitate any boundary condition at $x_1 = 0$. Following [1], we use one of the following two-dimensional versions of the leap-frog scheme, either

$$u_{j,k}^{n+2} + \lambda_1 a_1 (u_{j+1,k}^{n+1} - u_{j-1,k}^{n+1}) + \lambda_2 a_2 (u_{j,k+1}^{n+1} - u_{j,k-1}^{n+1}) - u_{j,k}^n = 0, \tag{1.15}$$

or

$$\begin{aligned} u_{j,k}^{n+2} + \lambda_1 a_1 \Big(\frac{u_{j+1,k+1}^{n+1} + u_{j+1,k-1}^{n+1}}{2} - \frac{u_{j-1,k+1}^{n+1} + u_{j-1,k-1}^{n+1}}{2} \Big) & \\ + \lambda_2 a_2 \Big(\frac{u_{j+1,k+1}^{n+1} + u_{j-1,k+1}^{n+1}}{2} - \frac{u_{j+1,k-1}^{n+1} + u_{j-1,k-1}^{n+1}}{2} \Big) - u_{j,k}^n = 0. & \end{aligned} \tag{1.16}$$

Assumption 1 is trivially satisfied because (1.15) and (1.16) are explicit schemes. The scheme (1.15) satisfies Assumption 2 if and only if $\lambda_1 |a_1| + \lambda_2 |a_2| < 1$, while the scheme (1.16) satisfies Assumption 2 if and only if $\max(\lambda_1 |a_1|, \lambda_2 |a_2|) < 1$. Let us now study when Assumption 3 is valid. For the scheme (1.15), we have $r_1 = p_1 = 1$, and

$$a_1(z, \eta) = \lambda_1 a_1 z, \quad a_{-1}(z, \eta) = -a_1(z, \eta),$$

so Assumption 3 is valid as long as $a_1 \neq 0$. For the scheme (1.16), we have again $r_1 = p_1 = 1$, and

$$a_1(z, \eta) = z(\lambda_1 a_1 \cos \eta + i \lambda_2 a_2 \sin \eta), \quad a_{-1}(z, \eta) = z(-\lambda_1 a_1 \cos \eta + i \lambda_2 a_2 \sin \eta),$$

so Assumption 3 is valid as long as both a_1 and a_2 are nonzero. We refer to [2] for the verification of strong stability depending on the choice of some numerical boundary conditions for (1.15) or (1.16). If strong stability holds, then Theorem 1 yields semigroup boundedness and ℓ^2-stability with respect to boundary data.

Acknowledgments. — This article was completed while the author was visiting the Institute of Mathematical Science at Nanjing University. The author warmly thanks Professor Yin Huicheng and the Institute for their hospitality during this visit.

2. The Leray-Gårding method for fully discrete Cauchy problems

This section is devoted to proving stability estimates for discretized Cauchy problems, which is the first step before considering the discretized initial boundary value problem (1.2). More precisely, we consider the simpler case of the whole space $j \in \mathbb{Z}^d$, and the recurrence relation:

$$\begin{cases} \displaystyle\sum_{\sigma=0}^{s+1} Q_\sigma u_j^{n+\sigma} = 0, & j \in \mathbb{Z}^d, \quad n \geqslant 0, \\ u_j^n = f_j^n, & j \in \mathbb{Z}^d, \quad n = 0, \dots, s, \end{cases} \tag{2.1}$$

where the operators Q_σ are given by (1.3). We recall that in (1.3), the $a_{\ell,\sigma}$ are real numbers and are independent of the small parameter Δt (they may depend on the CFL parameters $\lambda_1, \dots, \lambda_d$), while $\boldsymbol{S}$ denotes the shift operator on the space grid: $(\boldsymbol{S}^\ell v)_j := v_{j+\ell}$ for $j, \ell \in \mathbb{Z}^d$. Stability of (2.1) is defined as follows.

Definition 2 (Stability for the discrete Cauchy problem). — The numerical scheme defined by (2.1) is (ℓ^2-) stable if Q_{s+1} is an isomorphism from $\ell^2(\mathbb{Z}^d)$ onto itself, and if furthermore there exists a constant $C_0 > 0$ such that for all $\Delta t \in (0,1]$, for all initial conditions $(f_j^0)_{j \in \mathbb{Z}^d}, \dots, (f_j^s)_{j \in \mathbb{Z}^d}$ in $\ell^2(\mathbb{Z}^d)$, there holds

$$\sup_{n \in \mathbb{N}} \| u^n \|_{-\infty,+\infty}^2 \leqslant C_0 \sum_{\sigma=0}^{s} \| f^\sigma \|_{-\infty,+\infty}^2 . \tag{2.2}$$

Let us quickly recall, see e.g. [12], that stability in the sense of Definition 2 is in fact independent of $\Delta t \in (0,1]$ (because (2.1) does not involve Δt and (2.2) can be simplified on either side by $\prod_i \Delta x_i$), and can be characterized in terms of the uniform power boundedness of the so-called amplification matrix

$$\mathscr{A}(\kappa) := \begin{pmatrix} -\dfrac{\widehat{Q_s}(\kappa)}{\widehat{Q_{s+1}}(\kappa)} & & \dots & & -\dfrac{\widehat{Q_0}(\kappa)}{\widehat{Q_{s+1}}(\kappa)} \\ 1 & 0 & \dots & & 0 \\ & \ddots & \ddots & & \\ 0 & & \ddots & \ddots & \vdots \\ 0 & 0 & & 1 & 0 \end{pmatrix} \in \mathscr{M}_{s+1}(\mathbb{C}), \tag{2.3}$$

where the $\widehat{Q_\sigma}(\kappa)$'s are defined in (1.5) and where it is understood that $\mathscr{A}$ is defined on the largest open set of $\mathbb{C}^d$ on which $\widehat{Q_{s+1}}$ does not vanish. Let us also recall that if Q_{s+1} is an isomorphism from $\ell^2(\mathbb{Z}^d)$ onto itself, then $\widehat{Q_{s+1}}$ does not vanish on $(\mathbb{S}^1)^d$, and therefore does not vanish on an open neighborhood of $(\mathbb{S}^1)^d$. With the above definition (2.3) for $\mathscr{A}$, the following well-known result holds, see e.g. [12]:

Proposition 1 (Characterization of stability for the fully discrete Cauchy problem)

Assume that Q_{s+1} is an isomorphism from $\ell^2(\mathbb{Z}^d)$ onto itself. Then the scheme (2.1) *is stable in the sense of Definition* 2 *if and only if there exists a constant $C_1 > 0$ such that the amplification matrix $\mathscr{A}$ in* (2.3) *satisfies*

$$\forall\, n \in \mathbb{N}, \quad \forall\, \xi \in \mathbb{R}^d, \quad \big\| \mathscr{A}(\mathrm{e}^{i\xi_1}, \dots, \mathrm{e}^{i\xi_d})^n \big\| \leqslant C_1.$$

In particular, the spectral radius of $\mathscr{A}(\mathrm{e}^{i\xi_1}, \dots, \mathrm{e}^{i\xi_d})$ should not be larger than 1 *(the so-called von Neumann condition).*

The eigenvalues of $\mathscr{A}(\mathrm{e}^{i\xi_1},\dots,\mathrm{e}^{i\xi_d})$ are the roots to the dispersion relation (1.5). When these roots are simple for all $\xi \in \mathbb{R}^d$, the von Neumann condition is both necessary and *sufficient* for stability of (2.1), see, e.g., [6, Prop. 3]. Assumption 2 is therefore a way to assume that (2.1) is stable for the discrete Cauchy problem. Let us also recall that for each eigenvalue of $\mathscr{A}$, the corresponding eigenspace has dimension 1 since $\mathscr{A}$ is a companion matrix. Therefore, assuming that the roots to the dispersion relation (1.5) are simple is equivalent to assuming that $\mathscr{A}$ is diagonalizable.

Our goal is to derive the semigroup estimate (2.2) not by applying Fourier transform to (2.1) and using uniform power boundedness of $\mathscr{A}$, but rather by multiplying the first equation in (2.1) by a suitable *local* multiplier. The analysis relies first on the simpler case where one only considers the time evolution and no additional space variable.

2.1. Stable recurrence relations. — In this section, we consider sequences $(v^n)_{n\in\mathbb{N}}$ with values in $\mathbb{C}$. The index n should be thought of as the discrete time variable, and we therefore introduce the new notation $\boldsymbol{T}$ for the shift operator on the time grid: $(\boldsymbol{T}^m v)^n := v^{n+m}$ for all $m,n\in\mathbb{N}$. We start with the following elementary but crucial Lemma, which is the analogue of [10, Lem. 1.1].

Lemma 1 (The energy-dissipation balance law). — *Let $P\in\mathbb{C}[X]$ be a polynomial of degree $s+1$ whose roots are simple and located in $\overline{\mathbb{D}}$. Then there exists a positive definite Hermitian form q_e on $\mathbb{C}^{s+1}$, and a nonnegative Hermitian form q_d on $\mathbb{C}^{s+1}$, that both depend in a $\mathscr{C}^\infty$ way on P, such that for any sequence $(v^n)_{n\in\mathbb{N}}$ with values in $\mathbb{C}$, there holds*

$$\begin{aligned}\forall\, n\in\mathbb{N},\quad 2\,\mathrm{Re}\left(\overline{\boldsymbol{T}(P'(\boldsymbol{T})v^n)}P(\boldsymbol{T})v^n\right)&\\ =(s+1)|P(\boldsymbol{T})v^n|^2+(\boldsymbol{T}-I)(q_e(v^n,\dots,v^{n+s}))&+q_d(v^n,\dots,v^{n+s}).\end{aligned}$$

In particular, for all sequence $(v^n)_{n\in\mathbb{N}}$ that satisfies the recurrence relation

$$\forall\, n\in\mathbb{N},\quad P(\boldsymbol{T})v^n=0,$$

the sequence $(q_e(v^n,\dots,v^{n+s}))_{n\in\mathbb{N}}$ is non-increasing.

The fact that there exists a Hermitian norm on $\mathbb{C}^{s+1}$ that is non-increasing along solutions to the recurrence relation is not new. In fact, it is easily seen to be a consequence of the Kreiss matrix Theorem, see [26]. However, the important point here is that we can construct a multiplier that yields directly the 'energy boundedness' (or decay). The fact that the coefficients of this multiplier are integer multiples of the coefficients of P will be crucial in the analysis of Section 3, see also Proposition 2 below.

Proof. — We borrow some ideas from [10, Lem. 1.1] and introduce the interpolation polynomials:

$$\forall\, k=1,\dots,s+1,\quad P_k(X):=a\prod_{j\neq k}(X-x_j),$$

where $x_1,\dots,x_{s+1}$ denote the roots of P, and $a\neq 0$ its dominant coefficient. Since the roots of P are pairwise distinct, the P_k's form a basis of $\mathbb{C}_s[X]$ and they depend in a $\mathscr{C}^\infty$ way on the coefficients of P. We have

$$P'=\sum_{k=1}^{s+1} P_k.$$

We then consider a sequence $(v^n)_{n\in\mathbb{N}}$ with values in $\mathbb{C}$ and compute(3)

$$\begin{aligned}
2\,\mathrm{Re}\,&\Big(\overline{\boldsymbol{T}(P'(\boldsymbol{T})v^n)}P(\boldsymbol{T})v^n\Big)-(s+1)|P(\boldsymbol{T})v^n|^2\\
&=\sum_{k=1}^{s+1}\overline{\boldsymbol{T}(P_k(\boldsymbol{T}))v^n}(\boldsymbol{T}-x_k)P_k(\boldsymbol{T})v^n+\boldsymbol{T}(P_k(\boldsymbol{T})v^n)(\boldsymbol{T}-\overline{x_k})\overline{P_k(\boldsymbol{T})v^n}\\
&\qquad\qquad-\sum_{k=1}^{s+1}(\boldsymbol{T}-\overline{x_k})(\overline{P_k(\boldsymbol{T})v^n})(\boldsymbol{T}-x_k)(P_k(\boldsymbol{T})v^n)\\
&=\sum_{k=1}^{s+1}(\boldsymbol{T}-|x_k|^2)|P_k(\boldsymbol{T})v^n|^2.
\end{aligned}$$

The conclusion follows by defining:

$$\forall\,(w^0,\dots,w^s)\in\mathbb{C}^{s+1},\quad q_e(w^0,\dots,w^s):=\sum_{k=1}^{s+1}|P_k(\boldsymbol{T})w^0|^2,\tag{2.4}$$

$$q_d(w^0,\dots,w^s):=\sum_{k=1}^{s+1}(1-|x_k|^2)|P_k(\boldsymbol{T})w^0|^2.\tag{2.5}$$

The form q_e is positive definite because the P_k's form a basis of $\mathbb{C}_s[X]$. The form q_d is nonnegative because the roots of P are located in $\overline{\mathbb{D}}$. Both forms depend in a $\mathscr{C}^\infty$ way on the coefficients of P because the roots of P are simple. □

Lemma 1 shows that the polynomial P' yields the good multiplier $\boldsymbol{T}P'(\boldsymbol{T})v^n$ for the recurrence relation $P(\boldsymbol{T})v^n=0$. Of course, P' is not the only possible choice, though it will be our favorite one in what follows. As in [10, Lem. 1.1], any polynomial of the form(4)

$$Q:=\sum_{k=1}^{s+1}\alpha_k P_k,\quad \alpha_1,\dots,\alpha_{s+1}>0,$$

provides with an energy balance of the form

$$\begin{aligned}
2\,\mathrm{Re}\,&\Big(\overline{\boldsymbol{T}(Q(\boldsymbol{T})v^n)}P(\boldsymbol{T})v^n\Big)\\
&=(\alpha_1+\cdots+\alpha_{s+1})|P(\boldsymbol{T})v^n|^2+(\boldsymbol{T}-I)(q_e(v^n,\dots,v^{n+s}))+q_d(v^n,\dots,v^{n+s}),
\end{aligned}$$

with suitable Hermitian forms q_e,q_d that have the same properties as stated in Lemma 1.

(3) Here we use repeatedly the property $\boldsymbol{T}(\overline{w^n})=\overline{w^{n+1}}=\overline{\boldsymbol{T}w^n}$, as well as $\overline{\boldsymbol{T}w^n}\boldsymbol{T}w^n=|w^{n+1}|^2=\boldsymbol{T}|w^n|^2$.

(4) The sign condition here on the coefficients α_k is the analogue of the separation condition for the roots in [18, 10].

2.2. The energy-dissipation balance for finite difference schemes

In this section, we consider the numerical scheme (2.1). We introduce the following notation:

$$L := \sum_{\sigma=0}^{s+1} \boldsymbol{T}^\sigma Q_\sigma, \quad M := \sum_{\sigma=0}^{s+1} \sigma \boldsymbol{T}^\sigma Q_\sigma, \tag{2.6}$$

so that the discretized Cauchy problem (2.1) reads

$$\begin{cases} L u_j^n = 0, & j \in \mathbb{Z}^d, \quad n \geqslant 0, \\ u_j^n = f_j^n, & j \in \mathbb{Z}^d, \quad n = 0, \dots, s. \end{cases}$$

The operator M will be the 'multiplier' associated with L. Thanks to Fourier analysis, Lemma 1 easily gives the following result:

Proposition 2 (The energy-dissipation balance law). — *Let Assumptions* 1 *and* 2 *be satisfied. Then there exist a continuous coercive quadratic form E_0 and a continuous nonnegative quadratic form D_0 on $\ell^2(\mathbb{Z}^d;\mathbb{R})^{s+1}$ such that for all sequences $(v^n)_{n\in\mathbb{N}}$ with values in $\ell^2(\mathbb{Z}^d;\mathbb{R})$ and for all $n \in \mathbb{N}$, there holds*

$$2\langle M v^n, L v^n\rangle_{-\infty,+\infty} \\ = (s+1) \| L v^n \|^2_{-\infty,+\infty} + (\boldsymbol{T} - I) E_0(v^n, \dots, v^{n+s}) + D_0(v^n, \dots, v^{n+s}).$$

In particular, for all initial data $f^0, \dots, f^s \in \ell^2(\mathbb{Z}^d;\mathbb{R})$, the solution to (2.1) *satisfies*

$$\sup_{n\in\mathbb{N}} E_0(v^n, \dots, v^{n+s}) \leqslant E_0(f^0, \dots, f^s),$$

and (2.1) *is (ℓ^2-)stable.*

Proof. — We use the same notation v^n for the sequence $(v_j^n)_{j\in\mathbb{Z}^d}$ and the corresponding step function on $\mathbb{R}^d$ whose value on the cell $[j_1\Delta x_1, (j_1+1)\Delta x_1) \times \cdots \times [j_d \Delta x_d, (j_d+1)\Delta x_d)$ equals v_j^n. Then Plancherel Theorem gives

$$2\langle M v^n, L v^n\rangle_{-\infty,+\infty} - (s+1) \| L v^n \|^2_{-\infty,+\infty} \\ = \int_{\mathbb{R}^d} 2 \operatorname{Re} \Big(\overline{\boldsymbol{T}(P'_\zeta(\boldsymbol{T}) \widehat{v^n}(\xi))} P_\zeta(\boldsymbol{T}) \widehat{v^n}(\xi) \Big) - (s+1) \big| P_\zeta(\boldsymbol{T}) \widehat{v^n}(\xi) \big|^2 \frac{\mathrm{d}\xi}{(2\pi)^d},$$

where $\widehat{v^n}$ denotes the Fourier transform of v^n, and where we have let

$$P_\zeta(z) := \sum_{\sigma=0}^{s+1} \widehat{Q_\sigma}\big(\mathrm{e}^{i\zeta_1}, \dots, \mathrm{e}^{i\zeta_d}\big) z^\sigma, \quad \zeta_j := \xi_j \Delta x_j,$$

and $P'_\zeta(z)$ denotes the derivative of P_ζ with respect to z.

From Assumption 2, we know that for all $\zeta \in \mathbb{R}^d$, P_ζ has degree $s+1$ and has $s+1$ simple roots in $\overline{\mathbb{D}}$. We can apply Lemma 1 and get

$$2\langle M v^n, L v^n\rangle_{-\infty,+\infty} - (s+1) \| L v^n \|^2_{-\infty,+\infty} \\ = \int_{\mathbb{R}^d} (\boldsymbol{T} - I) q_{e,\zeta}\big(\widehat{v^n}(\xi), \dots, \widehat{v^{n+s}}(\xi)\big) + q_{d,\zeta}\big(\widehat{v^n}(\xi), \dots, \widehat{v^{n+s}}(\xi)\big) \frac{\mathrm{d}\xi}{(2\pi)^d},$$

where $q_{e,\zeta}, q_{d,\zeta}$ depend in a $\mathscr{C}^\infty$ way on $\zeta \in \mathbb{R}^d$ and are 2π-periodic in each ζ_j. Furthermore, $q_{e,\zeta}$ is positive definite and $q_{d,\zeta}$ is nonnegative. We then define

$$\forall\,(w^0,\dots,w^s) \in \ell^2(\mathbb{Z}^d;\mathbb{R})^{s+1}, \quad \begin{aligned} E_0(w^0,\dots,w^s) &:= \int_{\mathbb{R}^d} q_{e,\zeta}\big(\widehat{w^0}(\xi),\dots,\widehat{w^s}(\xi)\big)\,\frac{\mathrm{d}\xi}{(2\pi)^d},\\ D_0(w^0,\dots,w^s) &:= \int_{\mathbb{R}^d} q_{d,\zeta}\big(\widehat{w^0}(\xi),\dots,\widehat{w^s}(\xi)\big)\,\frac{\mathrm{d}\xi}{(2\pi)^d}. \end{aligned}$$

By compactness of $[0,2\pi]^d$, the hermitian forms $q_{e,\zeta}, q_{d,\zeta}$ are uniformly bounded with respect to ζ (because they depend continuously on ζ). Therefore E_0 and D_0 define continuous quadratic forms on $\ell^2(\mathbb{Z}^d;\mathbb{R})^{s+1}$, and D_0 is nonnegative as the integral of a nonnegative quantity. Furthermore, $q_{e,\zeta}$ is a positive Hermitian form for each ζ so by compactness it is uniformly coercive with respect to ζ. Hence E_0 is a coercive quadratic form and the proof of Proposition 2 is complete. □

2.3. Examples. — Let us clarify Proposition 2 in the case of the one-dimensional examples of Section 1.3. For the leap-frog scheme in one space dimension, there holds

$$L = \boldsymbol{T}^2 + \lambda a \boldsymbol{T}(\boldsymbol{S} - \boldsymbol{S}^{-1}) - I,$$

and our multiplier Mu_j^n reads

$$Mu_j^n = 2u_j^{n+2} + \lambda a(u_{j+1}^{n+1} - u_{j-1}^{n+1}) = u_j^{n+2} + u_j^n + \underbrace{Lu_j^n}_{=0},$$

where the equality $Lu_j^n = 0$ is used as long as (u_j^n) corresponds to a solution to the leap-frog scheme. We thus recover the more classical multiplier $u_j^{n+2} + u_j^n$ used in [24], but we emphasize that both multipliers coincide only on solutions to the leap-frog scheme. It will appear more clearly in Section 3 why our choice for Mu_j^n has a major advantage when considering initial boundary value problems because the main energy-dissipation identity of Proposition 2 not only holds for solutions to $Lu_j^n = 0$ but for any sequence (u^n) with values in $\ell^2(\mathbb{Z}^d)$. Let us now look at the energy and dissipation functionals provided by Proposition 2. Since the (two simple) roots z_1, z_2 to (1.11) have modulus 1, if we keep the notation of the proof of Lemma 1 and Proposition 2, we get the expressions

$$\begin{aligned} q_{e,\zeta}(w^0,w^1) &= \big|w^1 - z_1(\zeta)w^0\big|^2 + \big|w^1 - z_2(\zeta)w^0\big|^2\\ &= 2(|w^0|^2 + |w^1|^2) + 4\,\mathrm{Re}\,\big(i\lambda a \sin\zeta\,\overline{w^1}w^0\big),\\ q_{d,\zeta}(w^0,w^1) &= 0, \end{aligned}$$

for all $\zeta \in [0,2\pi]$. After substituting $\zeta = \xi\Delta x$ and integrating with respect to ξ, we get

$$E_0(v^n,v^{n+1}) = 2\sum_{j\in\mathbb{Z}} \Delta x (v_j^n)^2 + \Delta x (v_j^{n+1})^2 + 2\sum_{j\in\mathbb{Z}} \Delta x \lambda a \Big(v_{j+1}^n - v_{j-1}^n\Big) v_j^{n+1},$$

and $D_0 \equiv 0$ (no dissipation). Proposition 2 shows that the energy functional E_0 is preserved for solutions to the leap-frog scheme, a fact that also comes more directly from the relation

$$2\sum_{j\in\mathbb{Z}}\Delta x(u_j^{n+2}+u_j^n)\Big(u_j^{n+2}+\lambda a(u_{j+1}^{n+1}-u_{j-1}^{n+1})-u_j^n\Big)=0,$$

after regrouping

$$\begin{aligned}2\sum_{j\in\mathbb{Z}}\Delta x(u_j^{n+2}+u_j^n)(u_j^{n+2}-u_j^n)\\ =2\sum_{j\in\mathbb{Z}}\Delta x(u_j^{n+1})^2+\Delta x(u_j^{n+2})^2-2\sum_{j\in\mathbb{Z}}\Delta x(u_j^n)^2+\Delta x(u_j^{n+1})^2,\end{aligned}$$

and using summation by parts for the remaining terms. What is important here is that both quantities $|z_1(\zeta)|^2+|z_2(\zeta)|^2$ and $z_1(\zeta)+z_2(\zeta)$ are trigonometric polynomials in ζ, and therefore the expression of E_0 turns out to give a finite sum of 'local' quadratic functionals of one of the forms

$$\sum_{j\in\mathbb{Z}}\Delta x v_{j+\ell_1}^n v_{j+\ell_2}^n,\quad \sum_{j\in\mathbb{Z}}\Delta x v_{j+\ell_1}^n v_{j+\ell_2}^{n+1},\quad \sum_{j\in\mathbb{Z}}\Delta x v_{j+\ell_1}^{n+1} v_{j+\ell_2}^{n+1}.$$

This means that we could also define the energy functional E_0 for sequences (v_j) defined only for $j\geqslant 0$ and not on all $\mathbb{Z}$ by 'localizing' in space (this fact is used in [23] in the context of partial differential equations because the energy functional that arises in [18, 10] turns out to be a local quantity).

Let us now turn to the scheme (1.12) that is based on the backwards differentiation formula of order 2. In that case, we have

$$L=\boldsymbol{T}^2\Big(\frac{3}{2}I+\frac{\lambda a}{2}(\boldsymbol{S}-\boldsymbol{S}^{-1})\Big)-2\boldsymbol{T}+\frac{1}{2}I,\quad M=\boldsymbol{T}^2\left(3I+\lambda a(\boldsymbol{S}-\boldsymbol{S}^{-1})\right)-2\boldsymbol{T},$$

so even if $Lu_j^n=0$, our multiplier Mu_j^n does not coincide with the 'more standard' u_j^{n+2} used in [8, 9]. Here the multiplier Mu_j^n incorporates some terms that take into account the spatial discretization while the multiplier u_j^{n+2} is designed to take advantage of the G-stability of the BDF-2 integration rule, see [15, Chap. V.6]. The energy-dissipation functionals provided by Proposition 2 are not as elegant in this case as what they were in the case of the leap-frog scheme. Namely, we keep the notation of the proofs of Lemma 1 and Proposition 2. Letting z_1, z_2 denote the roots to the dispersion relation (1.13), and introducing the notation

$$\mathscr{Q}(\zeta):=\Big|\frac{3}{2}+i\lambda a\sin\zeta\Big|^2\big(|z_1(\zeta)|^2+|z_2(\zeta)|^2\big),$$

we compute

$$\begin{aligned}q_{e,\zeta}(w^0,w^1)&=\Big|\frac{3}{2}+i\lambda a\sin\zeta\Big|^2\big(\big|w^1-z_1(\zeta)w^0\big|^2+\big|w^1-z_2(\zeta)w^0\big|^2\big)\\&=2\Big|\Big(\frac{3}{2}+i\lambda a\sin\zeta\Big)w^1\Big|^2+\mathscr{Q}(\zeta)|w^0|^2-4\operatorname{Re}\Big(\Big(\frac{3}{2}+i\lambda a\sin\zeta\Big)w^1\overline{w^0}\Big),\end{aligned}$$

$$q_{d,\zeta}(w^0,w^1)=\Big(2\Big|\frac{3}{2}+i\lambda a\sin\zeta\Big|^2-\mathscr{Q}(\zeta)\Big)|w^1|^2+\big(\mathscr{Q}(\zeta)-\frac{1}{2}\big)|w^0|^2$$
$$-4\operatorname{Re}\Big(\overline{\Big(\frac{3}{2}+i\lambda a\sin\zeta\Big)w^1}w^0\Big)-2\Big|\frac{3}{2}+i\lambda a\sin\zeta\Big|^2\operatorname{Re}\big((|z_1|^2z_2+|z_2|^2z_1)\overline{w^1}w^0\big),$$

for all $\zeta\in[0,2\pi]$. The energy functional is then defined as

$$E_0(v^n,v^{n+1})=\int_{\mathbb{R}}q_{e,\xi\Delta x}\big(\widehat{v^n}(\xi),\widehat{v^{n+1}}(\xi)\big)\,\frac{\mathrm{d}\xi}{2\pi},$$

and using the above expression for the Hermitian form $q_{e,\zeta}$, we get

$$E_0(v^n,v^{n+1})=2\sum_{j\in\mathbb{Z}}\Delta x\Big(\frac{3}{2}v_j^{n+1}+\frac{\lambda a}{2}(v_{j+1}^{n+1}-v_{j-1}^{n+1})\Big)^2$$
$$-4\sum_{j\in\mathbb{Z}}\Delta x\Big(\frac{3}{2}v_j^{n+1}+\frac{\lambda a}{2}(v_{j+1}^{n+1}-v_{j-1}^{n+1})\Big)v_j^n+\int_{\mathbb{R}}\mathscr{Q}(\xi\Delta x)|\widehat{v^n}(\xi)|^2\,\frac{\mathrm{d}\xi}{2\pi},$$

and rather similar expression for D_0. The problem at this stage is that there is no obvious reason for $\mathscr{Q}$ to be a trigonometric polynomial in ζ and therefore the last term in the decomposition of E_0 does not obviously decompose as a linear combination of local energy functionals

$$\sum_{j\in\mathbb{Z}}\Delta x v_{j+\ell_1}^n v_{j+\ell_2}^n.$$

Though the energy functional E_0 will be sufficient for our purpose here, its nonlocal feature may prevent from extending the current multiplier technique to obtain stability results on non-Cartesian meshes. We leave this question to further study.

3. Semigroup estimates for fully discrete initial boundary value problems

We now turn to the proof of Theorem 1 for which we shall use the results of Section 2 as a toolbox. By linearity of (1.2), it is sufficient to prove Theorem 1 separately in the case $(f_j^0)=\cdots=(f_j^s)=0$, and in the case $(F_j^n)=0$, $(g_j^n)=0$. The latter case is the most difficult and requires the introduction of an auxiliary set of 'dissipative' boundary conditions. Solutions to (1.2) are always assumed to be real valued, which means that the data are real valued. For complex valued initial data and/or forcing terms, one just uses the linearity of (1.2).

3.1. The case with zero initial data. — We first assume $(f_j^0)=\cdots=(f_j^s)=0$. By strong stability, we already know that (1.4) holds with a constant C that is independent of $\gamma>0$ and $\Delta t\in(0,1]$. Therefore, proving Theorem 1 amounts to showing the existence of a constant C, that is independent of $\gamma>0$ and $\Delta t\in(0,1]$ such that the solution to (1.2) with $(f_j^0)=\cdots=(f_j^s)=0$ satisfies

$$\sup_{n\geqslant 0}\mathrm{e}^{-2\gamma n\Delta t}|||u^n|||_{1-r_1,+\infty}^2\leqslant C\Big\{\frac{\gamma\Delta t+1}{\gamma}\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}|||F^n|||_{1,+\infty}^2+\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1=1-r_1}^{0}\|g_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2\Big\}.\tag{3.1}$$

We thus consider a parameter $\gamma > 0$ and a time step $\Delta t \in (0,1]$, and focus on the numerical scheme (1.2) with zero initial data (that is, $(f_j^0) = \cdots = (f_j^s) = 0$). For all $n \in \mathbb{N}$, we extend the sequence (u_j^n) by zero for $j_1 \leqslant -r_1$:

$$v_j^n := \begin{cases} u_j^n & \text{if } j_1 \geqslant 1 - r_1, \quad j' \in \mathbb{Z}^{d-1}, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that Lv_j^n is not zero for all $j \in \mathbb{Z}^d$ hence the need for the general framework of Proposition 2. We thus use Proposition 2 and compute:

$$(\boldsymbol{T} - I)E_0(v^n, \dots, v^{n+s}) + D_0(v^n, \dots, v^{n+s}) \\ = 2\langle Mv^n, Lv^n\rangle_{-\infty,+\infty} - (s+1)|||Lv^n|||^2_{-\infty,+\infty}.$$

Due to the form of the operator L, see (2.6), and the fact that v_j^n vanishes for $j_1 \leqslant -r_1$, there holds:

$$Lv_j^n = \begin{cases} \Delta t F_j^{n+s+1} & \text{if } j_1 \geqslant 1, \\ 0 & \text{if } j_1 \leqslant -r_1 - p_1, \end{cases}$$

and we thus get

$$(\boldsymbol{T} - I)E_0(v^n, \dots, v^{n+s}) + D_0(v^n, \dots, v^{n+s}) \\ = \Big(\prod_{k=1}^d \Delta x_k\Big) \sum_{j_1 \geqslant 1} \sum_{j' \in \mathbb{Z}^{d-1}} 2\Delta t (Mv_j^n) F_j^{n+s+1} - (s+1)\Delta t^2 (F_j^{n+s+1})^2 \\ + \Big(\prod_{k=1}^d \Delta x_k\Big) \sum_{j_1 = 1 - r_1 - p_1}^{0} \sum_{j' \in \mathbb{Z}^{d-1}} 2(Mv_j^n)Lv_j^n - (s+1)(Lv_j^n)^2.$$

We multiply the latter equality by $\exp(-2\gamma(n+s+1)\Delta t)$, sum with respect to n from 0 to some N and use the fact that D_0 is nonnegative. Recalling that the initial data in (1.2) vanish, we get

$$\text{(3.2)} \quad \mathrm{e}^{-2\gamma(N+s+1)\Delta t} E_0\big(v^{N+1}, \dots, v^{N+s+1}\big) \\ + \underbrace{\big(1 - \mathrm{e}^{-2\gamma\Delta t}\big) \sum_{n=1}^N \mathrm{e}^{-2\gamma(n+s)\Delta t} E_0(v^n, \dots, v^{n+s})}_{\geqslant 0} \leqslant S_{1,N} + S_{2,N},$$

with

$$\text{(3.3)} \quad S_{1,N} \\ := \sum_{n=0}^N \mathrm{e}^{-2\gamma(n+s+1)\Delta t} \Big(2\Delta t \langle Mv^n, F^{n+s+1}\rangle_{1,+\infty} - (s+1)\Delta t^2 |||F^{n+s+1}|||^2_{1,+\infty}\Big),$$

and

$$\text{(3.4)} \quad S_{2,N} \\ := \Big(\prod_{k=1}^d \Delta x_k\Big) \sum_{n=0}^N \mathrm{e}^{-2\gamma(n+s+1)\Delta t} \sum_{j_1 = 1 - r_1 - p_1}^{0} \sum_{j' \in \mathbb{Z}^{d-1}} 2(Mv_j^n)Lv_j^n - (s+1)(Lv_j^n)^2.$$

Let us now estimate the two source terms $S_{1,N}, S_{2,N}$ in (3.2). We begin with the term $S_{2,N}$ defined in (3.4). Let us recall that the ratio $\Delta t/\Delta x_1$ is fixed[(5)]. Furthermore, the form of the operators L and M in (2.6) gives the estimate (recall that v_j^n vanishes for $j_1 \leqslant -r_1$):

$$S_{2,N} \leqslant C\Delta t\bigg(\prod_{k=2}^{d}\Delta x_k\bigg)\sum_{n=0}^{N}\mathrm{e}^{-2\gamma(n+s+1)\Delta t}\sum_{j_1=1-r_1}^{p_1}\sum_{j'\in\mathbb{Z}^{d-1}}(u_j^n)^2+\cdots+(u_j^{n+s+1})^2,$$

for a constant C that does not depend on N, γ nor on Δt. We thus have, uniformly with respect to $N\in\mathbb{N}$, $\gamma>0$ and $\Delta t\in(0,1]$:

$$\begin{aligned}S_{2,N} &\leqslant C\sum_{n=s+1}^{N+s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1=1-r_1}^{p_1}\|u^n_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})}\\ &\leqslant C\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1=1-r_1}^{p_1}\|u^n_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})}\\ &\leqslant C\bigg\{\frac{\gamma\Delta t+1}{\gamma}\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}|||F^n|||^2_{1,+\infty}\\ &\qquad\qquad+\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1=1-r_1}^{0}\|g^n_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})}\bigg\},\end{aligned}$$

where we have used the trace estimate (1.4) that follows from the strong stability assumption.

Let us now focus on the term $S_{1,N}$ in (3.2), see the defining equation (3.3). We use the Cauchy-Schwarz inequality and derive (using now the interior estimate in (1.4) that follows from the strong stability assumption and the fact that the coefficients in the multiplier M are independent of Δt):

$$\begin{aligned}S_{1,N} &\leqslant 2\sum_{n=0}^{N}\Delta t\mathrm{e}^{-2\gamma(n+s+1)\Delta t}|||Mv^n|||_{1,+\infty}|||F^{n+s+1}|||_{1,+\infty}\\ &\leqslant C\sum_{n=0}^{N}\Delta t\mathrm{e}^{-2\gamma(n+s+1)\Delta t}\Big(|||v^{n+1}|||_{1-r_1,+\infty}+\cdots+|||v^{n+s+1}|||_{1-r_1,+\infty}\Big)|||F^{n+s+1}|||_{1,+\infty}\\ &\leqslant C\frac{\gamma}{\gamma\Delta t+1}\sum_{n=s+1}^{N+s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}|||u^n|||^2_{1-r_1,+\infty}+C\frac{\gamma\Delta t+1}{\gamma}\sum_{n=s+1}^{N+s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}|||F^n|||^2_{1,+\infty}\\ &\leqslant C\bigg\{\frac{\gamma\Delta t+1}{\gamma}\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}|||F^n|||^2_{1,+\infty}+\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1=1-r_1}^{0}\|g^n_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})}\bigg\}.\end{aligned}$$

Ignoring the nonnegative term on the left-hand side of (3.2) and using the coercivity of E_0, we have proved that there exists a constant $C>0$ that is uniform with respect

(5) This is one first occurrence where restricting to the 'hyperbolic scaling' for the time and space steps is convenient.

to $N, \gamma, \Delta t$ such that:

$$\mathrm{e}^{-2\gamma(N+s+1)\Delta t} |\!|\!| v^{N+s+1} |\!|\!|_{-\infty,+\infty}^2 \leqslant C \Big\{ \frac{\gamma \Delta t+1}{\gamma} \sum_{n \geqslant s+1} \Delta t \mathrm{e}^{-2\gamma n \Delta t} |\!|\!| F^n |\!|\!|_{1,+\infty}^2 + \sum_{n \geqslant s+1} \Delta t \mathrm{e}^{-2\gamma n \Delta t} \sum_{j_1=1-r_1}^{0} \| g_{j_1,\bullet}^n \|_{\ell^2(\mathbb{Z}^{d-1})}^2 \Big\},$$

which yields (3.1) and therefore the validity of Theorem 1 in the case of zero initial data.

3.2. Construction of dissipative boundary conditions. — In this paragraph, we consider an auxiliary problem for which we shall be able to prove simultaneously an optimal semigroup estimate and a trace estimate for the solution. The argument here is independent of the original numerical scheme (1.2), but the auxiliary scheme introduced in Theorem 2 below will be used later on to decompose the solution to (1.2) into two pieces, each of which being estimated by separate tools. We thus forget temporarily about (1.2) and state the following key result.

Theorem 2. — *Let Assumptions 1, 2 and 3 be satisfied. Then for all $P_1 \in \mathbb{N}$, there exists a constant $C_{P_1} > 0$ such that, for all initial data $(f_j^0), \dots, (f_j^s) \in \ell^2(\mathbb{Z}^d)$ and for all source term $(g_j^n)_{j_1 \leqslant 0, n \geqslant s+1}$ that satisfies*

$$\forall \Gamma > 0, \qquad \sum_{n \geqslant s+1} \mathrm{e}^{-2\Gamma n} \sum_{j_1 \leqslant 0} \| g_{j_1,\bullet}^n \|_{\ell^2(\mathbb{Z}^{d-1})}^2 < +\infty,$$

there exists a unique sequence $(u_j^n)_{j \in \mathbb{Z}^d, n \in \mathbb{N}}$ solution to

$$\begin{cases} L u_j^n = 0, & j_1 \geqslant 1, \quad j' \in \mathbb{Z}^{d-1}, \quad n \geqslant 0, \\ M u_j^n = g_j^{n+s+1}, & j_1 \leqslant 0, \quad j' \in \mathbb{Z}^{d-1}, \quad n \geqslant 0, \\ u_j^n = f_j^n, & j \in \mathbb{Z}^d, \quad n = 0, \dots, s. \end{cases} \tag{3.5}$$

Moreover for all $\gamma > 0$ and $\Delta t \in (0,1]$, this solution satisfies

$$\sup_{n \geqslant 0} \mathrm{e}^{-2\gamma n \Delta t} |\!|\!| u^n |\!|\!|_{-\infty,+\infty}^2 + \frac{\gamma}{\gamma \Delta t + 1} \sum_{n \geqslant 0} \Delta t \mathrm{e}^{-2\gamma n \Delta t} |\!|\!| u^n |\!|\!|_{-\infty,+\infty}^2 + \sum_{n \geqslant 0} \Delta t \mathrm{e}^{-2\gamma n \Delta t} \sum_{j_1=1-r_1}^{P_1} \| u_{j_1,\bullet}^n \|_{\ell^2(\mathbb{Z}^{d-1})}^2 \leqslant C_{P_1} \Big\{ \sum_{\sigma=0}^{s} |\!|\!| f^\sigma |\!|\!|_{-\infty,+\infty}^2 + \sum_{n \geqslant s+1} \Delta t \mathrm{e}^{-2\gamma n \Delta t} \sum_{j_1 \leqslant 0} \| g_{j_1,\bullet}^n \|_{\ell^2(\mathbb{Z}^{d-1})}^2 \Big\}. \tag{3.6}$$

Theorem 2 justifies why we advocate the choice $M u_j^n = 2 u_j^{n+2} + \lambda a (u_{j+1}^{n+1} - u_{j-1}^{n+1})$ rather than the more standard $u_j^{n+2} + u_j^n$ as a multiplier for the leap-frog scheme. Despite repeated efforts, we have not been able to prove the estimate (3.6) when using the numerical boundary condition $u_j^{n+2} + u_j^n$ on $j_1 \leqslant 0$, in conjunction with the leap-frog scheme on $j_1 \geqslant 1$.

Proof. — Let us first quickly observe that the solution to (3.5) is well-defined since, as long as we have determined the solution up to a time index $n+s$, $n \geqslant 0$, then u^{n+s+1} is sought as a solution to an equation of the form

$$Q_{s+1}u^{n+s+1} = F,$$

where F belongs to $\ell^2(\mathbb{Z}^d)$ (this is due to the form of L and M, see (2.6)). Hence u^n is uniquely defined and belongs to $\ell^2(\mathbb{Z}^d)$ for all $n \in \mathbb{N}$.

The proof of Theorem 2 starts again with the application of Proposition 2. Using the non-negativity of the dissipation form D_0, we get[(6)]

$$(\boldsymbol{T} - I)E_0(u^n, \dots, u^{n+s}) + (s+1)|||Lu^n|||^2_{-\infty,+\infty} \leqslant 2\langle Mu^n, Lu^n\rangle_{-\infty,+\infty} = 2\langle g^{n+s+1}, Lu^n\rangle_{-\infty,0}.$$

By the Young inequality

$$2\langle g^{n+s+1}, Lu^n\rangle_{-\infty,0} \leqslant \frac{s+1}{2}|||Lu^n|||^2_{-\infty,0} + \frac{2}{s+1}|||g^{n+s+1}|||^2_{-\infty,0},$$

we get

$$(\boldsymbol{T} - I)E_0(u^n, \dots, u^{n+s}) + \frac{s+1}{2}|||Lu^n|||^2_{-\infty,+\infty} \leqslant \frac{2}{s+1}|||g^{n+s+1}|||^2_{-\infty,0}.$$

We multiply the latter inequality by $\exp(-2\gamma(n+s+1)\Delta t)$, sum from $n=0$ to some arbitrary N and already derive the estimate (here we use again the fact that $\Delta t/\Delta x_1$ is a fixed positive constant):

$$\begin{aligned}\sup_{n\geqslant 1} \mathrm{e}^{-2\gamma(n+s)\Delta t}E_0(u^n, \dots, u^{n+s}) + \big(1-\mathrm{e}^{-2\gamma\Delta t}\big)\sum_{n\geqslant 0}\mathrm{e}^{-2\gamma(n+s)\Delta t}E_0(u^n, \dots, u^{n+s}) \\ + \sum_{n\geqslant 0}\Delta t\mathrm{e}^{-2\gamma(n+s+1)\Delta t}\sum_{j_1\in\mathbb{Z}}\|Lu^n_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})} \\ \leqslant C\Big\{\mathrm{e}^{-2\gamma s\Delta t}E_0(f^0, \dots, f^s) + \sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1\leqslant 0}\|g^n_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})}\Big\}.\end{aligned}$$

Using the coercivity of E_0 and the inequality

$$1-\mathrm{e}^{-2\gamma\Delta t} \geqslant \frac{\gamma\Delta t}{\gamma\Delta t+1},$$

we have therefore derived the estimate

$$\begin{aligned}\sup_{n\geqslant 0} \mathrm{e}^{-2\gamma n\Delta t}|||u^n|||^2_{-\infty,+\infty} + \frac{\gamma}{\gamma\Delta t+1}\sum_{n\geqslant 0}\Delta t\mathrm{e}^{-2\gamma n\Delta t}|||u^n|||^2_{-\infty,+\infty} \\ + \sum_{n\geqslant 0}\Delta t\mathrm{e}^{-2\gamma(n+s+1)\Delta t}\sum_{j_1\in\mathbb{Z}}\|Lu^n_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})} \\ \leqslant C\Big\{\sum_{\sigma=0}^{s}|||f^\sigma|||^2_{-\infty,+\infty} + \sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1\leqslant 0}\|g^n_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})}\Big\},\end{aligned} \tag{3.7}$$

(6)Since $Lu^n_j = 0$ for $j_1 \geqslant 1$, one could also write $|||Lu^n|||^2_{-\infty,0}$ rather than $|||Lu^n|||^2_{-\infty,+\infty}$ on the left hand-side of the inequality.

where the constant C is independent of γ, Δt and on the solution (u_j^n). In order to prove (3.6), the main remaining task is to derive the trace estimate for (u_j^n). This is done by first dealing with the case where $\gamma\Delta t$ is 'large'.

• From the definition of the operator L, see (2.6), there exists a constant $C>0$ and an integer J such that

$$(Lu_j^n)^2 \geqslant \frac{1}{2}(Q_{s+1}u_j^{n+s+1})^2 - C\sum_{\sigma=0}^{s}\sum_{|\ell|\leqslant J}(u_{j+\ell}^{n+\sigma})^2.$$

Since Q_{s+1} is an isomorphism, there exists a constant $c>0$ such that

$$\sum_{j\in\mathbb{Z}^d}(Lu_j^n)^2 \geqslant c\sum_{j\in\mathbb{Z}^d}(u_j^{n+s+1})^2 - \frac{1}{c}\sum_{\sigma=0}^{s}\sum_{j\in\mathbb{Z}^d}(u_j^{n+\sigma})^2.$$

Multiplying by $\exp(-2\gamma(n+s+1)\Delta t)$ and summing with respect to $n\in\mathbb{N}$, we get

$$\begin{aligned}\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1\in\mathbb{Z}}\|u_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2 &\leqslant C\Big\{\sum_{n\geqslant 0}\Delta t\mathrm{e}^{-2\gamma(n+s+1)\Delta t}\sum_{j_1\in\mathbb{Z}}\|Lu_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2\\ &\qquad + \mathrm{e}^{-2\gamma\Delta t}\sum_{n\geqslant 0}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1\in\mathbb{Z}}\|u_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2\Big\}.\end{aligned}\tag{3.8}$$

The second term on the right-hand side is decomposed as

$$\begin{aligned}\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1\in\mathbb{Z}}\|u_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2 + \sum_{\sigma=0}^{s}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1\in\mathbb{Z}}\|f_{j_1,\bullet}^\sigma\|_{\ell^2(\mathbb{Z}^{d-1})}^2\\ \leqslant \sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1\in\mathbb{Z}}\|u_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2 + \lambda_1\sum_{\sigma=0}^{s}|||f^\sigma|||_{-\infty,+\infty}^2.\end{aligned}$$

Choosing $\gamma\Delta t$ large enough, that is $\gamma\Delta t\geqslant \ln R_0$ for some numerical constant $R_0>1$ that depends only on the (fixed) coefficients of the operator L, we can absorb the term

$$\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1\in\mathbb{Z}}\|u_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2,$$

from right to left in (3.8), and we have therefore derived the estimate

$$\begin{aligned}\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1\in\mathbb{Z}}\|u_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2\\ \leqslant C\Big\{\sum_{n\geqslant 0}\Delta t\mathrm{e}^{-2\gamma(n+s+1)\Delta t}\sum_{j_1\in\mathbb{Z}}\|Lu_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2 + \mathrm{e}^{-2\gamma\Delta t}\sum_{\sigma=0}^{s}|||f^\sigma|||_{-\infty,+\infty}^2\Big\}.\end{aligned}$$

It remains to use (3.7) to bound the first term on the right-hand side, and we get an even better estimate than (3.6) which we were originally aiming at:

$$\sum_{n\geqslant s+1} \Delta t \mathrm{e}^{-2\gamma n\Delta t} \sum_{j_1\in\mathbb{Z}} \|u^n_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})} \leqslant C\Big\{\sum_{\sigma=0}^{s} |\!|\!| f^\sigma |\!|\!|^2_{-\infty,+\infty} + \sum_{n\geqslant s+1} \Delta t \mathrm{e}^{-2\gamma n\Delta t} \sum_{j_1\leqslant 0} \|g^n_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})}\Big\}.$$

This gives a control of infinitely many traces and not only finitely many (this restriction to finitely many traces will appear in the regime where $\gamma\Delta t$ can be small).

• From now on, we have fixed a constant $R_0 > 1$ such that (3.6) holds for $\gamma\Delta t \geqslant \ln R_0$ and we thus assume $\gamma\Delta t \in (0, \ln R_0]$. (Getting rid of all large values of $\gamma\Delta t$ will be used to gain 'compactness'.) We also know that the estimate (3.7) holds, independently of the value of $\gamma\Delta t$, and we now wish to estimate the traces of the solution (u^n_j) for finitely many values of j_1.

We first observe from (3.7) that for all $\gamma > 0$ and $\Delta t \in (0,1]$, there exists a constant $C_{\gamma,\Delta t}$ such that

$$\forall\, n\in\mathbb{N},\quad \mathrm{e}^{-2\gamma n\Delta t} \sum_{j\in\mathbb{Z}^d} (u^n_j)^2 \leqslant C_{\gamma,\Delta t}.$$

In particular, for any $j_1 \in \mathbb{Z}$, the Laplace-Fourier transforms $\widehat{u_{j_1}}$ of the step functions

$$u_{j_1} : (t,y) \in \mathbb{R}^+\times\mathbb{R}^{d-1} \longmapsto u^n_j \quad \text{if } (t,y) \in \big[n\Delta t, (n+1)\Delta t\big) \times \prod_{k=2}^{d} \big[j_k\Delta x_k, (j_k+1)\Delta x_k\big),$$

is well-defined on $\{\tau\in\mathbb{C}, \operatorname{Re}\tau > 0\}\times\mathbb{R}^{d-1}$. The dual variables are denoted $\tau = \gamma + i\theta$, $\gamma > 0$, and $\eta = (\eta_2,\dots,\eta_d) \in \mathbb{R}^{d-1}$. It will also be convenient to introduce the notation $\eta_\Delta := (\eta_2\Delta x_2, \dots, \eta_d\Delta x_d)$. Given $\Gamma > 0$, the sequence $(\widehat{u_{j_1}}(\Gamma + i\theta, \eta))_{j_1\in\mathbb{Z}}$ belongs to $\ell^2(\mathbb{Z})$ for almost every $(\theta,\eta) \in \mathbb{R}\times\mathbb{R}^{d-1}$.

We first show the following estimate, which is the Laplace-Fourier analogue of (3.7).

Lemma 2. — *With $R_0 > 1$ fixed as above, there exists a constant $C > 0$ such that for all $\gamma > 0$ and $\Delta t \in (0,1]$ satisfying $\gamma\Delta t \in (0, \ln R_0]$, there holds*

$$\begin{aligned}(3.9)\quad &\sum_{j_1\in\mathbb{Z}} \int_{\mathbb{R}\times\mathbb{R}^{d-1}} \Big| \sum_{\ell_1=-r_1}^{p_1} a_{\ell_1}\big(\mathrm{e}^{(\gamma+i\theta)\Delta t}, \eta_\Delta\big) \widehat{u_{j_1+\ell_1}}(\gamma+i\theta,\eta)\Big|^2 \mathrm{d}\theta\mathrm{d}\eta \\ &+ \sum_{j_1\leqslant 0} \int_{\mathbb{R}\times\mathbb{R}^{d-1}} \Big| \sum_{\ell_1=-r_1}^{p_1} \mathrm{e}^{(\gamma+i\theta)\Delta t} \partial_z a_{\ell_1}\big(\mathrm{e}^{(\gamma+i\theta)\Delta t}, \eta_\Delta\big) \widehat{u_{j_1+\ell_1}}(\gamma+i\theta,\eta)\Big|^2 \mathrm{d}\theta\mathrm{d}\eta \\ &\leqslant C\Big\{\sum_{\sigma=0}^{s} |\!|\!| f^\sigma |\!|\!|^2_{-\infty,+\infty} + \sum_{n\geqslant s+1} \Delta t \mathrm{e}^{-2\gamma n\Delta t} \sum_{j_1\leqslant 0} \|g^n_{j_1,\bullet}\|^2_{\ell^2(\mathbb{Z}^{d-1})}\Big\}.\end{aligned}$$

Proof of Lemma 2. — Given $\tau=\gamma+i\theta$ and η, we compute (here $j_1\in\mathbb{Z}$ is fixed):

$$\sum_{\ell_1=-r_1}^{p_1} a_{\ell_1}\big(\mathrm{e}^{\tau\Delta t},\eta_\Delta\big)\widehat{u_{j_1+\ell_1}}(\tau,\eta) = \widehat{Lu_{j_1,\bullet}}(\tau,\eta)+\frac{1-\mathrm{e}^{-\tau\Delta t}}{\tau}\sum_{\sigma=1}^{s+1}\sum_{\sigma'=0}^{\sigma-1}\mathrm{e}^{(\sigma-\sigma')\tau\Delta t}\mathscr{F}_{j_1}^{\sigma,\sigma'}(\eta), \tag{3.10}$$

$$\sum_{\ell_1=-r_1}^{p_1} \mathrm{e}^{\tau\Delta t}\partial_z a_{\ell_1}\big(\mathrm{e}^{\tau\Delta t},\eta_\Delta\big)\widehat{u_{j_1+\ell_1}}(\tau,\eta) = \widehat{Mu_{j_1,\bullet}}(\tau,\eta)+\frac{1-\mathrm{e}^{-\tau\Delta t}}{\tau}\sum_{\sigma=1}^{s+1}\sum_{\sigma'=0}^{\sigma-1}\sigma\mathrm{e}^{(\sigma-\sigma')\tau\Delta t}\mathscr{F}_{j_1}^{\sigma,\sigma'}(\eta). \tag{3.11}$$

where, in (3.10) and (3.11), we have set

$$\mathscr{F}_{j_1}^{\sigma,\sigma'}(\eta)=\sum_{\ell_1=-r_1}^{p_1}\Big(\sum_{\ell'=-r'}^{p'}a_{\ell,\sigma}\mathrm{e}^{i\ell'\cdot\eta_\Delta}\Big)\widehat{f_{j_1+\ell_1,\bullet}^{\sigma'}}(\eta),$$

which corresponds to the partial Fourier transform with respect to $y=(x_2,\dots,x_d)\in\mathbb{R}^{d-1}$, of the step function associated with the sequence $(Q_\sigma f_j^{\sigma'})$ (no Laplace transform here).

We need to estimate integrals with respect to (θ,η) of the right-hand side of (3.10) and (3.11). The first term on the right of (3.10) and (3.11) are easy. For instance, we have (applying Plancherel Theorem):

$$\begin{aligned}\sum_{j_1\in\mathbb{Z}}\int_{\mathbb{R}\times\mathbb{R}^{d-1}}\big|\widehat{Lu_{j_1,\bullet}}(\tau,\eta)\big|^2\mathrm{d}\theta\mathrm{d}\eta &= (2\pi)^d\sum_{j_1\in\mathbb{Z}}\sum_{n\geqslant0}\int_{n\Delta t}^{(n+1)\Delta t}\mathrm{e}^{-2\gamma s}\|Lu_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2\mathrm{d}s\\ &=(2\pi)^d\frac{1-\mathrm{e}^{-2\gamma\Delta t}}{2\gamma\Delta t}\sum_{n\geqslant0}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1\in\mathbb{Z}}\|Lu_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2.\end{aligned}$$

We now recall that $\gamma\Delta t$ is restricted to the interval $(0,\ln R_0]$, and we use (3.7) to derive

$$\sum_{j_1\in\mathbb{Z}}\int_{\mathbb{R}\times\mathbb{R}^{d-1}}\big|\widehat{Lu_{j_1,\bullet}}(\tau,\eta)\big|^2\mathrm{d}\theta\mathrm{d}\eta \leqslant C\Big\{\sum_{\sigma=0}^{s}|\!|\!|f^\sigma|\!|\!|_{-\infty,+\infty}^2+\sum_{n\geqslant s+1}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1\leqslant0}\|g_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2\Big\}.$$

Similarly, we have

$$\sum_{j_1\leqslant0}\int_{\mathbb{R}\times\mathbb{R}^{d-1}}\big|\widehat{Mu_{j_1,\bullet}}(\tau,\eta)\big|^2\mathrm{d}\theta\mathrm{d}\eta = (2\pi)^d\frac{1-\mathrm{e}^{-2\gamma\Delta t}}{2\gamma\Delta t}\sum_{n\geqslant0}\Delta t\mathrm{e}^{-2\gamma n\Delta t}\sum_{j_1\leqslant0}\|Mu_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2,$$

which we can again uniformly estimate by the right-hand side of (3.9).

Going back to the right-hand side terms in (3.10) and (3.11), we find that there only remains for proving (3.9) to estimate the integral (here there are finitely many values of σ and σ'):

$$\sum_{j_1\in\mathbb{Z}}\int_{\mathbb{R}\times\mathbb{R}^{d-1}}\Big|\frac{1-\mathrm{e}^{-\tau\Delta t}}{\tau}\Big|^2\big|\mathscr{F}_{j_1}^{\sigma,\sigma'}(\eta)\big|^2\mathrm{d}\theta\mathrm{d}\eta \\ =\Big(\int_{\mathbb{R}}\Big|\frac{1-\mathrm{e}^{-\tau\Delta t}}{\tau}\Big|^2\mathrm{d}\theta\Big)\sum_{j_1\in\mathbb{Z}}\int_{\mathbb{R}^{d-1}}\big|\mathscr{F}_{j_1}^{\sigma,\sigma'}(\eta)\big|^2\mathrm{d}\eta,$$

where we have applied Fubini Theorem. Applying first Plancherel Theorem with respect to the $d-1$ last space variables, we get

$$\sum_{j_1\in\mathbb{Z}}\int_{\mathbb{R}^{d-1}}\big|\mathscr{F}_{j_1}^{\sigma,\sigma'}(\eta)\big|^2\mathrm{d}\eta\leqslant C\sum_{j_1\in\mathbb{Z}}\sum_{j'\in\mathbb{Z}^{d-1}}\Big(\prod_{k=2}^{d}\Delta x_k\Big)(f_j^{\sigma'})^2\leqslant\frac{C}{\Delta t}\sum_{\sigma=0}^{s}|||f^\sigma|||^2_{-\infty,+\infty}.$$

The conclusion then follows by computing

$$\int_{\mathbb{R}}\Big|\frac{1-\mathrm{e}^{-\tau\Delta t}}{\tau}\Big|^2\mathrm{d}\theta=2\pi\Delta t\frac{1-\mathrm{e}^{-2\gamma\Delta t}}{2\gamma\Delta t},$$

and by recalling that $\gamma\Delta t$ belongs to $(0,\ln R_0]$. We can eventually bound the integrals on the left-hand side of (3.9) by estimating separately the integrals of each term on the right-hand side of (3.10) and (3.11). □

The conclusion in the proof of Theorem 2 relies on the following crucial result. Here both Assumptions 2 and 3 are heavily used.

Lemma 3 (The trace estimate). — *Let Assumptions* 1, 2 *and* 3 *be satisfied. Let* $R_0>1$ *be fixed as above and let* $P_1\in\mathbb{N}$. *Then there exists a constant* $C_{P_1}>0$ *such that for all* $z\in\mathscr{U}$ *with* $|z|\leqslant R_0$, *for all* $\eta\in\mathbb{R}^{d-1}$ *and for all sequence* $(w_{j_1})_{j_1\in\mathbb{Z}}\in\ell^2(\mathbb{Z};\mathbb{C})$, *there holds*

$$\sum_{j_1=-r_1-p_1}^{P_1}|w_{j_1}|^2 \\ \leqslant C_{P_1}\Big\{\sum_{j_1\in\mathbb{Z}}\Big|\sum_{\ell_1=-r_1}^{p_1}a_{\ell_1}(z,\eta)w_{j_1+\ell_1}\Big|^2+\sum_{j_1\leqslant 0}\Big|\sum_{\ell_1=-r_1}^{p_1}z\partial_z a_{\ell_1}(z,\eta)w_{j_1+\ell_1}\Big|^2\Big\}. \tag{3.12}$$

Recall that the functions a_{ℓ_1}, $\ell_1=-r_1,\dots,p_1$, *are defined in* (1.6).

The proof of Lemma 3 is rather long. Before giving it in full details, we indicate how Lemma 3 yields the result of Theorem 2. We apply Lemma 3 to $z=\exp(\tau\Delta t)$, $\tau=\gamma+i\theta$ with $\gamma\Delta t\in(0,\ln R_0]$, $\eta_\Delta\in\mathbb{R}^{d-1}$ and the sequence $(\widehat{u_{j_1}}(\tau,\eta))_{j_1\in\mathbb{Z}}$. We then integrate (3.12) with respect to (θ,η) and use Lemma 2 to derive

$$\sum_{j_1=-r_1-p_1}^{P_1} \int_{\mathbb{R}\times\mathbb{R}^{d-1}} |\widehat{u_{j_1}}(\gamma+i\theta,\eta)|^2 \,\mathrm{d}\theta\mathrm{d}\eta \\ \leqslant C\Big\{\sum_{\sigma=0}^{s} |\!|\!| f^\sigma |\!|\!|_{-\infty,+\infty}^2 + \sum_{n\geqslant s+1} \Delta t\mathrm{e}^{-2\gamma n\Delta t} \sum_{j_1\leqslant 0} \|g_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2\Big\}.$$

It remains to apply Plancherel Theorem and we get

$$\frac{1-\mathrm{e}^{-2\gamma\Delta t}}{2\gamma\Delta t} \sum_{j_1=-r_1-p_1}^{P_1} \sum_{n\in\mathbb{N}} \Delta t\mathrm{e}^{-2\gamma n\Delta t} \|u_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2 \\ \leqslant C\Big\{\sum_{\sigma=0}^{s} |\!|\!| f^\sigma |\!|\!|_{-\infty,+\infty}^2 + \sum_{n\geqslant s+1} \Delta t\mathrm{e}^{-2\gamma n\Delta t} \sum_{j_1\leqslant 0} \|g_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2\Big\}.$$

Recalling that $\gamma\Delta t$ is restricted to the interval $(0,\ln R_0]$, we have thus derived the trace estimate

$$\sum_{n\in\mathbb{N}} \Delta t\mathrm{e}^{-2\gamma n\Delta t} \sum_{j_1=-r_1-p_1}^{P_1} \|u_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2 \\ \leqslant C\Big\{\sum_{\sigma=0}^{s} |\!|\!| f^\sigma |\!|\!|_{-\infty,+\infty}^2 + \sum_{n\geqslant s+1} \Delta t\mathrm{e}^{-2\gamma n\Delta t} \sum_{j_1\leqslant 0} \|g_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2\Big\}.$$

Combined with the semigroup and interior estimate (3.7), this gives the estimate (3.6) of Theorem 2 for $\gamma\Delta t\in(0,\ln R_0]$. □

Proof of Lemma 3. — Let us recall that the functions a_{ℓ_1} are 2π-periodic with respect to each coordinate of η. We can therefore restrict to $\eta\in[0,2\pi]^{d-1}$ rather than considering $\eta\in\mathbb{R}^{d-1}$. We argue by contradiction and assume that the conclusion to Lemma 3 does not hold. This means the following, up to normalizing and extracting subsequences; there exist three sequences (indexed by $k\in\mathbb{N}$):

– a sequence $(w^k)_{k\in\mathbb{N}}$ with values in $\ell^2(\mathbb{Z};\mathbb{C})$ such that $(w_{-r_1-p_1}^k,\dots,w_{P_1}^k)$ belongs to the unit sphere of $\mathbb{C}^{P_1+r_1+p_1+1}$ for all k, and $(w_{-r_1-p_1}^k,\dots,w_{P_1}^k)$ converges to $(\underline{w}_{-r_1-p_1},\dots,\underline{w}_{P_1})$ as k tends to infinity,

– a sequence $(z^k)_{k\in\mathbb{N}}$ with values in $\mathscr{U}\cap\{\zeta\in\mathbb{C},|\zeta|\leqslant R_0\}$, which converges to $\underline{z}\in\overline{\mathscr{U}}$,

– a sequence $(\eta^k)_{k\in\mathbb{N}}$ with values in $[0,2\pi]^{d-1}$, which converges to $\underline{\eta}\in[0,2\pi]^{d-1}$,

and these sequences satisfy:

(3.13)
$$\lim_{k\to+\infty} \sum_{j_1\in\mathbb{Z}} \Big|\sum_{\ell_1=-r_1}^{p_1} a_{\ell_1}(z^k,\eta^k) w_{j_1+\ell_1}^k\Big|^2 + \sum_{j_1\leqslant 0} \Big|\sum_{\ell_1=-r_1}^{p_1} z^k\partial_z a_{\ell_1}(z^k,\eta^k) w_{j_1+\ell_1}^k\Big|^2 = 0.$$

We are going to show that (3.13) implies that $(\underline{w}_{-r_1-p_1},\dots,\underline{w}_{P_1})$ must be zero, which will yield a contradiction since this vector must have norm 1.

• Let us first show that each component $(w^k_{j_1})_{k\in\mathbb{N}}$, $j_1\in\mathbb{Z}$, has a limit as k tends to infinity. This is already clear for $j_1=-r_1-p_1,\dots,P_1$. For $j_1>P_1$, we argue by induction. From (3.13), we have

$$\lim_{k\to+\infty}\sum_{\ell_1=-r_1}^{p_1} a_{\ell_1}(z^k,\eta^k)\,w^k_{P_1-p_1+1+\ell_1}=0,$$

and by Assumption 3, we know that $a_{p_1}(\underline{z},\underline{\eta})$ is nonzero. Hence $(w^k_{P_1+1})_{k\in\mathbb{N}}$ converges towards

$$-\frac{1}{a_{p_1}(\underline{z},\underline{\eta})}\sum_{\ell_1=-r_1}^{p_1-1} a_{\ell_1}(\underline{z},\underline{\eta})\,\underline{w}_{P_1-p_1+1+\ell_1},$$

which we define as $\underline{w}_{P_1+1}$. We can argue by induction in the same way for all indices $j_1>P_1+1$, but also for indices $j_1<-r_1-p_1$ because the function a_{-r_1} also does not vanish on $\overline{\mathscr{U}}\times\mathbb{R}^{d-1}$.

Using (3.13), we have thus shown that for each $j_1\in\mathbb{Z}$, $(w^k_{j_1})_{k\in\mathbb{N}}$ tends towards some limit $\underline{w}_{j_1}$ as k tends to infinity, and the sequence $\underline{w}$, which does not necessarily belong to $\ell^2(\mathbb{Z};\mathbb{C})$, satisfies the induction relations:

$$\forall\, j_1\in\mathbb{Z},\quad \sum_{\ell_1=-r_1}^{p_1} a_{\ell_1}(\underline{z},\underline{\eta})\,\underline{w}_{j_1+\ell_1}=0, \tag{3.14}$$

$$\forall\, j_1\leqslant 0,\quad \sum_{\ell_1=-r_1}^{p_1} \underline{z}\,\partial_z a_{\ell_1}(\underline{z},\underline{\eta})\,\underline{w}_{j_1+\ell_1}=0. \tag{3.15}$$

• The induction relation (3.14) is the one that arises in [13, 19] and all the works that deal with strong stability. The main novelty here is to use simultaneously (3.14) for controlling the unstable components of $(\underline{w}_{-r_1-p_1},\dots,\underline{w}_{-1})$ and (3.15) for controlling the stable components of $(\underline{w}_{-r_1-p_1},\dots,\underline{w}_{-1})$. The fact that $\underline{w}$ satisfies simultaneously (3.14) and (3.15) for $j_1\leqslant 0$ automatically annihilates the central components. This sketch of proof is made precise below.

We define the source terms:

$$F^k_{j_1}:=\sum_{\ell_1=-r_1}^{p_1} a_{\ell_1}(z^k,\eta^k)\,w^k_{j_1+\ell_1},\quad G^k_{j_1}:=\sum_{\ell_1=-r_1}^{p_1} z^k\,\partial_z a_{\ell_1}(z^k,\eta^k)\,w^k_{j_1+\ell_1},$$

which, according to (3.13), satisfy

$$\lim_{k\to 0}\sum_{j_1\in\mathbb{Z}}|F^k_{j_1}|^2=0,\quad \lim_{k\to 0}\sum_{j_1\leqslant 0}|G^k_{j_1}|^2=0. \tag{3.16}$$

We also introduce the vectors (here T denotes transposition)

$$W^k_{j_1}:=\big(w^k_{j_1+p_1},\dots,w^k_{j_1+1-r_1}\big)^T,\quad \underline{W}_{j_1}:=\big(\underline{w}_{j_1+p_1},\dots,\underline{w}_{j_1+1-r_1}\big)^T,$$

and the matrices:

$$(3.17)\qquad \mathbb{L}(z,\eta) := \begin{pmatrix} -\dfrac{a_{p_1-1}(z,\eta)}{a_{p_1}(z,\eta)} & & \dots & & -\dfrac{a_{-r_1}(z,\eta)}{a_{p_1}(z,\eta)} \\ 1 & 0 & \dots & & 0 \\ & \ddots & \ddots & & \\ 0 & & \ddots & \ddots & \vdots \\ 0 & 0 & & 1 & 0 \end{pmatrix} \in \mathscr{M}_{p_1+r_1}(\mathbb{C}),$$

$$(3.18)\qquad \mathbb{M}(z,\eta) := \begin{pmatrix} -\dfrac{\partial_z a_{p_1-1}(z,\eta)}{\partial_z a_{p_1}(z,\eta)} & & \dots & & -\dfrac{\partial_z a_{-r_1}(z,\eta)}{\partial_z a_{p_1}(z,\eta)} \\ 1 & 0 & \dots & & 0 \\ & \ddots & \ddots & & \\ 0 & & \ddots & \ddots & \vdots \\ 0 & 0 & & 1 & 0 \end{pmatrix} \in \mathscr{M}_{p_1+r_1}(\mathbb{C}).$$

The matrix $\mathbb{L}$ is well-defined on $\overline{\mathscr{U}} \times \mathbb{R}^{d-1}$ according to Assumption 3. The matrix $\mathbb{M}$ is also well-defined on $\overline{\mathscr{U}} \times \mathbb{R}^{d-1}$ because for any $\eta \in \mathbb{R}^{d-1}$, Assumption 3 asserts that $a_{p_1}(\cdot,\eta)$ is a non-constant polynomial whose roots lie in $\mathbb{D}$. From the Gauss-Lucas Theorem, the roots of $\partial_z a_{p_1}(\cdot,\eta)$ lie in the convex hull of those of $a_{p_1}(\cdot,\eta)$. Therefore $\partial_z a_{p_1}(\cdot,\eta)$ does not vanish on $\overline{\mathscr{U}}$. In the same way, $\partial_z a_{-r_1}(\cdot,\eta)$ does not vanish on $\overline{\mathscr{U}}$.

With our above notation, the vectors $W^k_{j_1}$, $\underline{W}_{j_1}$, satisfy the one step induction relations:

$$(3.19)\quad \forall\, j_1 \in \mathbb{Z},\qquad \begin{aligned} W^k_{j_1+1} &= \mathbb{L}(z^k,\eta^k)W^k_{j_1} + \big(F^k_{j_1+1}/a_{p_1}(z^k,\eta^k),0,\dots,0\big)^T, \\ \underline{W}_{j_1+1} &= \mathbb{L}(\underline{z},\underline{\eta})\underline{W}_{j_1}, \end{aligned}$$

$$(3.20)\quad \forall\, j_1 \leqslant -1,\qquad \begin{aligned} W^k_{j_1+1} &= \mathbb{M}(z^k,\eta^k)W^k_{j_1} + \big(G^k_{j_1+1}/(z^k\partial_z a_{p_1}(z^k,\eta^k)),0,\dots,0\big)^T, \\ \underline{W}_{j_1+1} &= \mathbb{M}(\underline{z},\underline{\eta})\underline{W}_{j_1}. \end{aligned}$$

• From Assumption 3 and the above application of the Gauss-Lucas Theorem, we already know that both matrices $\mathbb{L}(z,\eta)$ and $\mathbb{M}(z,\eta)$ are invertible for $(z,\eta) \in \overline{\mathscr{U}} \times \mathbb{R}^{d-1}$. Furthermore, Assumption 2 shows that $\mathbb{L}(z,\eta)$ has no eigenvalue on $\mathbb{S}^1$ for $(z,\eta) \in \mathscr{U} \times \mathbb{R}^{d-1}$. This property dates back at least to [16]. However, central eigenvalues on $\mathbb{S}^1$ may occur for $\mathbb{L}$ when z belongs to $\mathbb{S}^1$. The crucial point for proving Lemma 3 is that Assumption 2 precludes central eigenvalues of $\mathbb{M}$ for all $z \in \overline{\mathscr{U}}$. Namely, for all $z \in \overline{\mathscr{U}}$ and all $\eta \in \mathbb{R}^{d-1}$, $\mathbb{M}(z,\eta)$ has no eigenvalue on $\mathbb{S}^1$. This property holds because otherwise, for some $(z,\eta) \in \overline{\mathscr{U}} \times \mathbb{R}^{d-1}$, there would exist a solution $\kappa_1 \in \mathbb{S}^1$ to the dispersion relation

$$\sum_{\ell_1=-r_1}^{p_1} z\partial_z a_{\ell_1}(z,\eta)\kappa_1^{\ell_1} = 0.$$

For convenience, the coordinates of η are denoted $(\eta_2,\dots,\eta_d)$. Using the definition (1.6) of a_{ℓ_1}, and defining $\kappa := (\kappa_1, \mathrm{e}^{i\eta_2},\dots,\mathrm{e}^{i\eta_d})$, we have found a root $z \in \overline{\mathscr{U}}$ to the relation

$$\sum_{\sigma=1}^{s+1} \sigma \widehat{Q_\sigma}(\kappa) z^{\sigma-1} = 0, \tag{3.21}$$

but this is not possible because the $s+1$ roots (in z) to the dispersion relation (1.5) are simple and belong to $\overline{\mathbb{D}}$. The Gauss-Lucas Theorem thus shows that the roots to the relation (3.21) belong to $\mathbb{D}$ (and therefore not to $\overline{\mathscr{U}}$).

At this stage, we know that the eigenvalues of $\mathbb{M}(z,\eta)$, $(z,\eta) \in \overline{\mathscr{U}} \times \mathbb{R}^{d-1}$, split into two groups: those in $\mathscr{U}$, which we call the unstable ones, and those in $\mathbb{D}$, which we call the stable ones. For $(z,\eta) \in \overline{\mathscr{U}} \times \mathbb{R}^{d-1}$, we then introduce the spectral projector $\Pi^s_{\mathbb{M}}(z,\eta)$, resp. $\Pi^u_{\mathbb{M}}(z,\eta)$, of $\mathbb{M}(z,\eta)$ on the generalized eigenspace associated with eigenvalues in $\mathbb{D}$, resp. $\mathscr{U}$. We can then integrate the first induction relation in (3.20) and get

$$\Pi^s_{\mathbb{M}}(z^k,\eta^k) W_0^k = \frac{1}{z^k \partial_z a_{p_1}(z^k,\eta^k)} \sum_{j_1 \leqslant 0} \mathbb{M}(z^k,\eta^k)^{|j_1|} \Pi^s_{\mathbb{M}}(z^k,\eta^k) \big(G^k_{j_1},0,\dots,0\big)^T.$$

The projector $\Pi^s_{\mathbb{M}}$ depends continuously on $(z,\eta) \in \overline{\mathscr{U}} \times \mathbb{R}^{d-1}$. Furthermore, since the spectrum of $\mathbb{M}$ does not meet $\mathbb{S}^1$ even for $z \in \mathbb{S}^1$, there exists a constant $C>0$ and a $\delta \in (0,1)$ that are independent of $k \in \mathbb{N}$ and such that

$$\forall\, j_1 \leqslant 0, \quad \big\|\mathbb{M}(z^k,\eta^k)^{|j_1|} \Pi^s_{\mathbb{M}}(z^k,\eta^k)\big\| \leqslant C \delta^{|j_1|}.$$

We thus get a uniform estimate with respect to k:

$$\big|\Pi^s_{\mathbb{M}}(z^k,\eta^k) W_0^k\big|^2 \leqslant C \sum_{j_1 \leqslant 0} |G^k_{j_1}|^2.$$

Passing to the limit and using (3.16), we get $\Pi^s_{\mathbb{M}}(\underline{z},\underline{\eta}) \underline{W}_0 = 0$, or in other words $\underline{W}_0 = \Pi^u_{\mathbb{M}}(\underline{z},\underline{\eta}) \underline{W}_0$.

• The sequence $(\underline{W}_{j_1})_{j_1 \leqslant 0}$ satisfies both induction relations (3.19) and (3.20). Due to the form of the companion matrices $\mathbb{L}$ and $\mathbb{M}$, see (3.17)-(3.18), we can conclude that the vector $\underline{W}_0$ belongs to the generalized eigenspace (of either $\mathbb{L}$ or $\mathbb{M}$) associated with the common eigenvalues of $\mathbb{M}(\underline{z},\underline{\eta})$ and $\mathbb{L}(\underline{z},\underline{\eta})$. We have already seen that $\mathbb{M}(\underline{z},\underline{\eta})$ has no eigenvalue on $\mathbb{S}^1$ and $\underline{W}_0 = \Pi^u_{\mathbb{M}}(\underline{z},\underline{\eta}) \underline{W}_0$, so we can conclude that $\underline{W}_0$ belongs to the generalized eigenspace of $\mathbb{L}$ associated with those common eigenvalues of $\mathbb{M}(\underline{z},\underline{\eta})$ and $\mathbb{L}(\underline{z},\underline{\eta})$ in $\mathscr{U}$.

The matrix $\mathbb{L}(\underline{z},\underline{\eta})$ has N^u eigenvalues in $\mathscr{U}$, N^s in $\mathbb{D}$ and N^c on $\mathbb{S}^1$. (Since $\underline{z}$ may belong to $\mathbb{S}^1$, N^c is not necessarily zero.) With obvious notations, we let $\Pi^{u,s,c}_{\mathbb{L}}(z,\eta)$ denote the corresponding spectral projectors of $\mathbb{L}$ for (z,η) sufficiently close to $(\underline{z},\underline{\eta})$. In particular, the eigenvalues corresponding to $\Pi^u_{\mathbb{L}}(z,\eta)$ lie in $\mathscr{U}$ uniformly away from $\mathbb{S}^1$ for (z,η) sufficiently close to $(\underline{z},\underline{\eta})$. We can then integrate the first induction relation in (3.19) and derive (for k sufficiently large):

$$\Pi^u_{\mathbb{L}}(z^k,\eta^k)W_0^k = -\frac{1}{a_{p_1}(z^k,\eta^k)}\sum_{j_1\geqslant 0}\mathbb{L}(z^k,\eta^k)^{-j_1-1}\Pi^u_{\mathbb{L}}(z^k,\eta^k)\big(F^k_{j_1},0,\dots,0\big)^T.$$

Using the uniform exponential decay of $\mathbb{L}(z^k,\eta^k)^{-j_1-1}\Pi^u_{\mathbb{L}}(z^k,\eta^k)$ and (3.16), we finally end up with

$$\Pi^u_{\mathbb{L}}(\underline{z},\underline{\eta})\underline{W}_0 = 0.$$

Since $\underline{W}_0$ belongs to the generalized eigenspace of $\mathbb{L}$ associated with those common eigenvalues of $\mathbb{M}(\underline{z},\underline{\eta})$ and $\mathbb{L}(\underline{z},\underline{\eta})$ in $\mathscr{U}$, we can conclude that $\underline{W}_0$ equals zero. Applying the induction relation (3.19), the whole sequence $(\underline{W}_{j_1})_{j_1\in\mathbb{Z}}$ is zero which yields the expected contradiction. □

The crucial property that we use in the proof of Lemma 3 is the fact that up to $z\in\mathbb{S}^1$, the eigenvalues of $\mathbb{M}(z,\eta)$ lie either in $\mathbb{D}$ or $\mathscr{U}$. For the leap-frog scheme, this property would not be true if we had imposed the auxiliary numerical boundary condition $u_j^{n+2}+u_j^n$ rather than $2u_j^{n+2}+\lambda a(u_{j+1}^{n+1}-u_{j-1}^{n+1})$.

Let us also observe that we have used the fact that a_{p_1} and a_{-r_1} are non-constant in order to study the induction relation (3.15). There might be some schemes for which a_{p_1} and/or a_{-r_1} are constant but for which one can still apply similar arguments as in the previous proof, even though (3.15) is an induction relation with fewer steps than (3.14). In this respect, Assumption 3 might be relaxed in specific applications.

Remark 2. — The auxiliary problem (3.5) is in general not of the same form as (1.2) because in (3.5) one has to impose infinitely many numerical boundary conditions. This is due to the fact that the stencil of M incorporates points 'on the left' with respect to the first space variable. A remarkable exception occurs for explicit schemes with $s=0$, for in that case the multiplier Mv_j^n reads v_j^{n+1} and (3.5) is exactly the auxiliary problem considered in [7] (and labeled (2.7) there) where one imposes Dirichlet boundary conditions on finitely many boundary meshes (just use $g_j^n=0$ for $j_1\leqslant -r_1$). In full generality, there still remains an open problem of constructing a set of dissipative numerical boundary conditions of the same form as (1.2) with $s\geqslant 1$, that is with finitely many numerical boundary conditions, and for which one can prove by hand both a semigroup and a trace estimate as in Theorem 2.

3.3. End of the proof.

— As explained in the introduction of Section 3, the linearity of (1.2) reduces the proof of Theorem 1 to the case $(F_j^n)=0$, $(g_j^n)=0$, since we have already dealt with the case of zero initial data. We thus focus on (1.2) with $(F_j^n)=0$ and $(g_j^n)=0$, and write the corresponding solution (u_j^n) as $u_j^n=v_j^n+w_j^n$, where the sequence (v_j^n) solves:

$$\begin{cases} Lv_j^n=0, & j_1\geqslant 1,\quad j'\in\mathbb{Z}^{d-1},\quad n\geqslant 0,\\ Mv_j^n=0, & j_1\leqslant 0,\quad j'\in\mathbb{Z}^{d-1},\quad n\geqslant 0,\\ v_j^n=f_j^n, & j\in\mathbb{Z}^d,\quad n=0,\dots,s,\end{cases}\tag{3.22}$$

and (w_j^n) solves:

$$
\begin{cases}
Lw_j^n = 0, & j \in \mathbb{Z}^d,\ j_1 \geqslant 1,\ n \geqslant 0, \\
w_j^{n+s+1} + \displaystyle\sum_{\sigma=0}^{s+1} B_{j_1,\sigma} w_{1,j'}^{n+\sigma} = \widetilde{g}_j^{n+s+1}, & j \in \mathbb{Z}^d,\ j_1 = 1-r_1, \dots, 0,\ n \geqslant 0, \\
w_j^n = 0, & j \in \mathbb{Z}^d,\ n = 0, \dots, s.
\end{cases}
\tag{3.23}
$$

For $(v_j^n + w_j^n)_{j_1 \geqslant 1-r_1}$ to coincide with the solution (u_j^n) to (1.2), it is sufficient to extend the initial data $f_j^0, \dots, f_j^s$ by zero for $j_1 \leqslant -r_1$, which provides with the initial data in (3.22) on all $\mathbb{Z}^d$, and to define the boundary source term in (3.23) by:

$$
\widetilde{g}_j^{n+s+1} := -v_j^{n+s+1} - \sum_{\sigma=0}^{s+1} B_{j_1,\sigma} v_{1,j'}^{n+\sigma}.
\tag{3.24}
$$

We can estimate the solution (v_j^n) to (3.22) by applying Theorem 2. In particular, the trace estimate:

$$
\sum_{n \geqslant 0} \Delta t \mathrm{e}^{-2\gamma n \Delta t} \sum_{j_1=1-r_1}^{P_1} \|v_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2 \leqslant C \sum_{\sigma=0}^{s} |\!|\!| f^\sigma |\!|\!|_{1-r_1,+\infty}^2,
$$

for $P_1 = \max(p_1, q_1+1)$ gives (recall the definition (3.24) of $\widetilde{g}_j^{n+s+1}$):

$$
\begin{aligned}
\sum_{n \geqslant s+1} \Delta t \mathrm{e}^{-2\gamma n \Delta t} \sum_{j_1=1-r_1}^{0} \|\widetilde{g}_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2 &\leqslant C \sum_{n \geqslant 0} \Delta t \mathrm{e}^{-2\gamma n \Delta t} \sum_{j_1=1-r_1}^{\max(p_1,q_1+1)} \|v_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2 \\
&\leqslant C \sum_{\sigma=0}^{s} |\!|\!| f^\sigma |\!|\!|_{1-r_1,+\infty}^2.
\end{aligned}
$$

We can apply Theorem 1 to the solution (w_j^n) to (3.23) because the initial data in (3.23) vanish. We get:

$$
\begin{aligned}
\sup_{n \geqslant 0} \mathrm{e}^{-2\gamma n \Delta t} |\!|\!| w^n |\!|\!|_{1-r_1,+\infty}^2 + \frac{\gamma}{\gamma \Delta t + 1} \sum_{n \geqslant 0} \Delta t \mathrm{e}^{-2\gamma n \Delta t} |\!|\!| w^n |\!|\!|_{1-r_1,+\infty}^2 \\
+ \sum_{n \geqslant 0} \Delta t \mathrm{e}^{-2\gamma n \Delta t} \sum_{j_1=1-r_1}^{p_1} \|w_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2 \\
\leqslant C \sum_{n \geqslant s+1} \Delta t \mathrm{e}^{-2\gamma n \Delta t} \sum_{j_1=1-r_1}^{0} \|\widetilde{g}_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2 \leqslant C \sum_{\sigma=0}^{s} |\!|\!| f^\sigma |\!|\!|_{1-r_1,+\infty}^2.
\end{aligned}
$$

Combining with the similar estimate provided by Theorem 2 for (v_j^n), we end up with the expected estimate:

$$
\begin{aligned}
\sup_{n \geqslant 0} \mathrm{e}^{-2\gamma n \Delta t} |\!|\!| u^n |\!|\!|_{1-r_1,+\infty}^2 + \frac{\gamma}{\gamma \Delta t + 1} \sum_{n \geqslant 0} \Delta t \mathrm{e}^{-2\gamma n \Delta t} |\!|\!| u^n |\!|\!|_{1-r_1,+\infty}^2 \\
+ \sum_{n \geqslant 0} \Delta t \mathrm{e}^{-2\gamma n \Delta t} \sum_{j_1=1-r_1}^{p_1} \|u_{j_1,\bullet}^n\|_{\ell^2(\mathbb{Z}^{d-1})}^2 \leqslant C \sum_{\sigma=0}^{s} |\!|\!| f^\sigma |\!|\!|_{1-r_1,+\infty}^2,
\end{aligned}
$$

which completes the proof of Theorem 1.

4. Conclusion and perspectives

Let us first observe that in [30], Wade has constructed symmetrizers for deriving stability estimates for multistep schemes, even in the case of variable coefficients. His conditions for constructing a symmetrizer are less restrictive than Assumption 2. However, the symmetrizer in [30] is genuinely nonlocal and it is therefore not clear that it may be useful for boundary value problems. The main novelty here is to construct a *local* multiplier whose properties allow for the design of an auxiliary *dissipative* boundary value problem. This is the key to Theorem 1, despite the nonlocal feature of our energy functional.

The main possible improvement of Theorem 1 would consist of assuming that only the roots to (1.5) that lie on $\mathbb{S}^1$ are simple. Here we have assumed that all the roots, including those in $\mathbb{D}$ are simple. If we could manage to deal with multiple roots in $\mathbb{D}$, then Theorem 1 would be applicable to any stable numerical approximation of the transport equation (1.9) (recall that uniform power boundedness for the amplification matrix $\mathscr{A}$ given in (2.3) requires only that eigenvalues of modulus 1 be simple).

The results in this paper achieve the proof of a 'weak form' of the conjecture in [17] that strong stability, in the sense of Definition 1, implies semigroup stability. However, an even stronger assumption was made in [17], namely that the sole fulfillment of the interior estimate

$$\frac{\gamma}{\gamma\Delta t+1}\sum_{n\geqslant s+1}\Delta t\,\mathrm{e}^{-2\gamma n\Delta t}|||u^n|||^2_{1-r_1,+\infty}\leqslant C\,\frac{\gamma\Delta t+1}{\gamma}\sum_{n\geqslant s+1}\Delta t\,\mathrm{e}^{-2\gamma n\Delta t}|||F^n|||^2_{1,+\infty},$$

when *both initial and boundary data* for (1.2) vanish, does imply semigroup stability. The analogous conjecture for partial differential equations seems to be still open so far, but we do hope that our multiplier technique may yield some insight for dealing with the strong form of the conjecture in [17]. We also hope to extend our multiplier technique to prove some stability estimates for some multistep finite volume schemes on non-Cartesian meshes.

References

[1] S. Abarbanel & D. Gottlieb – "A note on the leap-frog scheme in two and three space dimensions", *J. Comput. Phys.* **21** (1976), no. 3, p. 351–355.

[2] ———, "Stability of two-dimensional initial boundary value problems using leap-frog type schemes", *Math. Comp.* **33** (1979), no. 148, p. 1145–1155.

[3] S. Benzoni-Gavage & D. Serre – *Multidimensional hyperbolic partial differential equations. first-order systems and applications*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, Oxford, 2007.

[4] J.-F. Coulombel – "Fully discrete hyperbolic initial boundary value problems with nonzero initial data", to appear in Confluentes Math.

[5] ———, "Stability of finite difference schemes for hyperbolic initial boundary value problems", *SIAM J. Numer. Anal.* **47** (2009), no. 4, p. 2844–2871.

[6] ———, "Stability of finite difference schemes for hyperbolic initial boundary value problems", in *HCDTE lecture notes. Part I. Nonlinear hyperbolic PDEs, dispersive and transport equations*, AIMS Ser. Appl. Math., vol. 6, Am. Inst. Math. Sci. (AIMS), Springfield, MO, 2013, p. 146.

[7] J.-F. Coulombel & A. Gloria – "Semigroup stability of finite difference schemes for multidimensional hyperbolic initial boundary value problems", *Math. Comp.* **80** (2011), no. 273, p. 165–203.

[8] E. Emmrich – "Convergence of the variable two-step BDF time discretisation of nonlinear evolution problems governed by a monotone potential operator", *BIT* **49** (2009), no. 2, p. 297–323.

[9] ———, "Two-step BDF time discretisation of nonlinear evolution problems governed by monotone operators with strongly continuous perturbations", *Comput. Methods Math.* **9** (2009), no. 1, p. 37–62.

[10] L. Gårding – "Solution directe du problème de Cauchy pour les équations hyperboliques", in *La théorie des équations aux dérivées partielles*, Colloques Internationaux du C.N.R.S., C.N.R.S., Paris, 1956, p. 71–90.

[11] M. Goldberg & E. Tadmor – "Scheme-independent stability criteria for difference approximations of hyperbolic initial-boundary value problems. II", *Math. Comp.* **36** (1981), no. 154, p. 603–626.

[12] B. Gustafsson, H.-O. Kreiss & J. Oliger – *Time dependent problems and difference methods*, Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, 1995.

[13] B. Gustafsson, H.-O. Kreiss & A. Sundström – "Stability theory of difference approximations for mixed initial boundary value problems. II", *Math. Comp.* **26** (1972), no. 119, p. 649–686.

[14] E. Hairer, S. P. Nørsett & G. Wanner – *Solving ordinary differential equations I. Nonstiff problems*, second ed., Springer Series in Computational Mathematics, vol. 8, Springer-Verlag, Berlin, 1993.

[15] E. Hairer & G. Wanner – *Solving ordinary differential equations II. Stiff and differential-algebraic problems*, second ed., Springer Series in Computational Mathematics, vol. 14, Springer-Verlag, Berlin, 1996.

[16] H.-O. Kreiss – "Stability theory for difference approximations of mixed initial boundary value problems. I", *Math. Comp.* **22** (1968), p. 703–714.

[17] H.-O. Kreiss & L. Wu – "On the stability definition of difference approximations for the initial-boundary value problem", *Appl. Numer. Math.* **12** (1993), no. 1-3, p. 213–227.

[18] J. Leray – *Hyperbolic differential equations*, The Institute for Advanced Study, Princeton, N.J., 1953.

[19] D. Michelson – "Stability theory of difference approximations for multidimensional initial-boundary value problems", *Math. Comp.* **40** (1983), no. 161, p. 1–45.

[20] J. Oliger – "Fourth order difference methods for the initial boundary-value problem for hyperbolic equations", *Math. Comp.* **28** (1974), p. 15–25.

[21] S. Osher – "Stability of difference approximations of dissipative type for mixed initial boundary value problems. I", *Math. Comp.* **23** (1969), p. 335–340.

[22] ———, "Systems of difference equations with general homogeneous boundary conditions", *Trans. Amer. Math. Soc.* **137** (1969), p. 177–201.

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

[24] R. D. Richtmyer & K. W. Morton – *Difference methods for initial-value problems. Theory and applications*, second ed., Interscience Tracts in Pure and Applied Mathematics, vol. 4, Interscience Publishers John Wiley & Sons, Inc., New York-London-Sydney, 1967.

[25] D. M. Sloan – "Boundary conditions for a fourth order hyperbolic difference scheme", *Math. Comp.* **41** (1983), p. 1–11.

[26] J. C. Strikwerda & B. A. Wade – "A survey of the Kreiss matrix theorem for power bounded families of matrices and its extensions", in *Linear operators (Warsaw, 1994)*, Banach Center Publ., vol. 38, Polish Acad. Sci., Warsaw, 1997, p. 339–360.

[27] J. M. Thomas – "Discrétisation des conditions aux limites dans les schémas saute-mouton", *ESAIM Math. Model. Numer. Anal.* **6** (1972), no. R-2, p. 31–44.

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

[29] L. N. Trefethen & M. Embree – *Spectra and pseudospectra. the behavior of nonnormal matrices and operators*, Princeton University Press, Princeton, N.J., 2005.

[30] B. A. Wade – "Symmetrizable finite difference operators", *Math. Comp.* **54** (1990), no. 190, p. 525–543.
[31] L. Wu – "The semigroup stability of the difference approximations for initial-boundary value problems", *Math. Comp.* **64** (1995), no. 209, p. 71–88.

Manuscript received June 12, 2015
accepted November 6, 2015

Jean-François Coulombel, CNRS & Université de Nantes, Laboratoire de Mathématiques Jean Leray (UMR CNRS 6629)
2 rue de la Houssinière, BP 92208, 44322 Nantes Cedex 3, France
E-mail : jean-francois.coulombel@univ-nantes.fr
Url : http://www.math.sciences.univ-nantes.fr/~coulombel/