# On some stable boundary closures of finite difference schemes for the transport equation

Jean-François Coulombel, Tomas Lundquist

# On some stable boundary closures of finite difference schemes for the transport equation

Jean-François Coulombel* & Tomas Lundquist†

January 6, 2020

**Abstract**

We explore in this article the possibilities and limitations of the so-called energy method for analyzing the stability of finite difference approximations to the transport equation with extrapolation numerical boundary conditions at the outflow boundary. We first show that for the most simple schemes, namely the explicit schemes with a three point stencil, the energy method can be applied for proving stability estimates when the scheme is implemented with either the first or second order extrapolation boundary condition. We then examine the case of five point stencils and give several examples of schemes and second order extrapolation numerical boundary conditions for which the energy method produces stability estimates. However, we also show that for the standard first or second order translatory extrapolation boundary conditions, the energy method cannot be applied for proving stability of the classical fourth order scheme originally proposed by Strang. This gives a clear limitation of the energy method with respect to the more general approach based on the normal mode decomposition.

**AMS classification:** 65M12, 65M06, 65M20.
**Keywords:** transport equation, numerical schemes, extrapolation boundary condition, energy, stability.

## 1 Introduction

A general approach for studying the stability of numerical boundary conditions for discretized hyperbolic equations has been initiated in the fundamental contribution [GKS72]. However, for technical reasons, the stability estimates in [GKS72] are restricted to zero initial data and can be obtained only after verifying the fulfillment of some "algebraic" condition which is commonly referred to as the Uniform Kreiss-Lopatinskii Condition. The theory in [GKS72] has been successfully applied to some well-known extrapolation procedures at outflow boundaries, see a preliminary announcement in [Kre66] and the complete proof in [Gol77]. In the recent work [CL20], one of the authors has revisited the stability estimates for the outflow extrapolation procedures in [Kre66, Gol77] and shown that a suitably devised energy argument could bypass the (technical and lenghty) arguments of [GKS72]. In this article, we examine in a systematic way the possible applications but also the limitations of the energy method

*Institut de Mathématiques de Toulouse ; UMR5219, Université de Toulouse ; CNRS, Université Paul Sabatier, F-31062 Toulouse Cedex 9, France. Email: jean-francois.coulombel@math.univ-toulouse.fr. Research of the author was supported by ANR project NABUCO, ANR-17-CE40-0025.

†Institut de Mathématiques de Toulouse ; UMR5219, Université de Toulouse ; CNRS, Université Paul Sabatier, F-31062 Toulouse Cedex 9, France. Email: tomas.lundquist@math.univ-toulouse.fr. Research of the author was supported by ANR project NABUCO, ANR-17-CE40-0025.

for analyzing the stability of extrapolation procedures at an outflow boundary for the most simple one-dimensional transport equation. We begin by reviewing some central concepts and notation that will be extensively used throughout the remainder of this paper.

### 1.1 Continuous problem

We consider the scalar advection equation in $1D$ on a semi-infinite interval with an outflow boundary at $x = 0$,

$$u_t + u_x = 0, \quad -\infty \le x \le 0\,, \quad t \ge 0\,. \tag{1}$$

The differential operator $\partial_t + \partial_x$ is *semi-bounded* on this domain, i.e. for all $0 \le t_1 \le t_2$ and integrable smooth functions $\Phi(x,t)$ where $\Phi(-\infty,t) = 0$, we have,

$$\Phi_t + \Phi_x = 0 \quad \Rightarrow \quad \|\Phi\|^2(t_2) \;=\; \|\Phi\|^2(t_1) - \int_{t_1}^{t_2} \Phi(0,\tau)^2 \,\mathrm{d}\tau \;\le\; \|\Phi\|^2(t_1)\,, \tag{2}$$

where we have used the standard definition of the $L_2$ inner product and norm (with respect to the space variable $x$) for real valued functions,

$$(\Phi,\Psi) \;:=\; \int_{-\infty}^{0} \Phi(x)\,\Psi(x)\,\mathrm{d}x\,, \quad \|\Phi\|^2 \;:=\; (\Phi,\Phi)\,.$$

When later discretizing (1) in time and space, we shall seek to mimic the energy balance (2).

### 1.2 Discretizing on the whole real line

We seek to discretize the problem (1) in time and space using a finite difference method. In the interior of the spatial domain we use a repeated interior stencil, which we analyze for the Cauchy version of (1). We introduce an equispaced spatial grid over the whole real line,

$$x_j \;:=\; j\,\Delta x, \quad -\infty \le j \le \infty\,,$$

where $\Delta x > 0$ is the grid spacing. Unless otherwise stated, sequences $(\Phi_j)$ on the grid are assumed to be real valued, and the norm on $\ell^2(\mathbb{Z})$ is defined by,

$$\|\Phi\|^2_{\ell^2(\mathbb{Z})} \;:=\; \sum_{j\in\mathbb{Z}} \Phi_j^2\,.$$

Next, we introduce a time step $\Delta t > 0$ (and accordingly, $t_n = n\Delta t$) and consider the ratio $z := \Delta t/\Delta x$ (the so-called Courant-Friedrichs-Lewy parameter [CFL28]) as a constant. In what follows, $u_j^n$ stands for an approximation of the solution $u$ to the Cauchy version of (1) in the neighborhood of $(x_j, t_n)$. We consider explicit finite difference discretizations to (1) of the form:

$$u_j^{n+1} \;=\; A(z)\,u_j^n\,, \quad -\infty \le j \le \infty\,, \quad n \in \mathbb{N}\,, \tag{3}$$

where $A(z)$ is a polynomial with respect to $z = \Delta t/\Delta x$ of difference operators. E.g, decomposing $A(z)$ in terms of simple shift operators, $A(z)$ then acts on sequences $(\Phi_j)$ indexed by $j$ according to,

$$A(z)\,\Phi_j \;:=\; \sum_{\ell=-\ell_-}^{\ell_+} a_\ell(z)\,\Phi_{j+\ell}\,,$$

where the integers $\ell_\pm$ mark the extent (or bandwidth) of the difference stencil, and where the coefficients $a_\ell(z)$ are polynomial expressions with respect to $z$. Most of the examples we will consider below fall into the classes of three point stencils ($\ell_- = \ell_+ = 1$) and five point stencils ($\ell_- = \ell_+ = 2$). It is clear that $A(z)$ acts boundedly on $\ell^2(\mathbb{Z})$, and we shall say that $A(z)$ is a contraction if the norm of $A(z)$ as an operator on $\ell^2(\mathbb{Z})$ is not larger than 1.

In what follows, we only consider finite difference schemes that are consistent with the transport equation (1). In other words, see for instance [RM94, GKO95], we only consider operators $A(z)$ in (3) that satisfy at least the first order accuracy relations,

$$\sum_{\ell=-\ell_-}^{\ell_+} a_\ell(z) \,=\, 1\,, \quad \text{and} \quad \sum_{\ell=-\ell_-}^{\ell_+} \ell\, a_\ell(z) \,=\, -z\,.$$

In order to prove stability estimates, be they on the whole space $\mathbb{Z}$ or on the half-space $\mathbb{Z}^-$, we aim to use the energy method in order to prove that the scheme is a contraction with respect to a certain norm. For the Cauchy problem, we thus need for $A(z)$ to be a contraction on $\ell^2(\mathbb{Z})$. Since the coefficients of the operator $A(z)$ do not depend on $j$, we can achieve this by splitting each local energy balance term into a telescopic and a dissipative part. In other words, the goal will be to show that, for an arbitrary sequence $(\Phi_j)_{j\in\mathbb{Z}}$,

$$\Psi_j \,=\, A(z)\,\Phi_j \quad \Rightarrow \quad \Psi_j^2 \,=\, \Phi_j^2 \,+\, T_j - T_{j-1} + S_j\,, \quad S_j \le 0\,, \tag{4}$$

for all values of $z$ within a certain specified interval. In (4), $T_j$, $T_{j-1}$ and $S_j$ are quadratic quantities over the $\Phi_k$ entries for $j-\ell_- \le k \le j+\ell_+$. When summing over all whole numbers, the telescopic part cancels and we are left with

$$\|\Psi\|_{\ell^2(\mathbb{Z})}^2 \,=\, \|\Phi\|_{\ell^2(\mathbb{Z})}^2 + \sum_{j\in\mathbb{Z}} S_j \,\le\, \|\Phi\|_{\ell^2(\mathbb{Z})}^2\,,$$

showing that $A(z)$ is a contraction, just as the continuous problem (2).

The main goal of analyzing schemes of the form (3) will be to derive *integration by parts decompositions* of the form (4) with $S_j \le 0$ for any $j$. There is a lot of freedom involved in splitting the difference $\Psi_j^2 - \Phi_j^2$ into telescopic and non-telescopic parts as in (4), see for instance [CL20, Lemma A.1]. In this paper we will take advantage of this in order to expand and refine the energy method arguments used in [CL20].

### 1.2.1 Finite difference operators

The basic unit for constructing all finite difference stencils will be the backward difference operator,

$$D\,\Phi_j \,:=\, \Phi_j - \Phi_{j-1}\,. \tag{5}$$

In order to define a parametric family of schemes with a three point stencil, we will use the standard first and second order derivative approximations (the "centered finite difference" and the "discrete Laplacian", respectively). On normalized form with respect to the spacing $\Delta x$, we thus consider,

$$D_0\,\Phi_j \,=\frac{1}{2}\,(D\,\Phi_j + D\,\Phi_{j+1}) \,=\, \frac{1}{2}\,(\Phi_{j+1} - \Phi_{j-1})\,, \tag{6}$$

$$\Delta\,\Phi_j \,=D\,\Phi_{j+1} - D\,\Phi_j \,=\, \Phi_{j-1} - 2\,\Phi_j + \Phi_{j+1}\,. \tag{7}$$

Note that both of these operators can easily be shown to be of accuracy order 2. To define a family of five point stencils, we also need to approximate the third and fourth derivatives, which we define using,

$$D_0 \, \Delta \, \Phi_j \, = \frac{1}{2} \, \left( -\Phi_{j-2} + 2 \, \Phi_{j-1} - 2 \, \Phi_{j+1} + \Phi_{j+2} \right) \, , \tag{8}$$

$$\Delta^2 \, \Phi_j \, = \Phi_{j-2} - 4 \, \Phi_{j-1} + 6 \, \Phi_j - 4 \, \Phi_{j+1} + \Phi_{j+2} \, . \tag{9}$$

Since we will not consider schemes with more than a five point stencil, in this paper we will limit our attention to only the finite difference operators introduced above.

### 1.3 Outflow boundary conditions

We close the discrete domain at the outflow boundary $x = 0$, and define a general vector on the grid as,

$$\mathbf{\Phi} \; = \; \left( \ldots, \Phi_{-2}, \Phi_{-1}, \Phi_0 \right) \, .$$

At the points nearest to the outflow boundary at $x = 0$, extrapolation to ghost points outside the computational domain are needed in order to close the finite difference stencils and thus the numerical scheme (3) itself. Namely, we need a method to define the $\Phi_1, \ldots, \Phi_{\ell_+}$ values in order to compute $\Psi_j := A(z) \, \Phi_j$ for all $j \in \mathbb{Z}^-$. Note that this extrapolation step may or may not reduce the local accuracy order from that of the interior stencil itself.

Next, we define an inner product on $\mathbb{Z}^-$ in the form of a quadrature rule, such that in general only the last $r$ quadrature weights $h_j > 0$ are non-unity,

$$\| \mathbf{\Phi} \|_H^2 \; := \; (\mathbf{\Phi}, \mathbf{\Phi})_H \, , \quad (\mathbf{\Phi}, \mathbf{\Psi})_H \; := \; \sum_{j \le -r} \Phi_j \, \Psi_j + \sum_{j=-r+1}^{0} h_j \, \Phi_j \, \Psi_j \, . \tag{10}$$

Now let's assume that contractivity (4) holds for all sequences $(\Phi_j) \in \ell^2(\mathbb{Z})$. Then, using an arbitrary set of extrapolation conditions to define the $\Phi_j$ values with positive indices $j = 1, \ldots, \ell_+$, we have,

$$\forall \, j \in \mathbb{Z}^- \, , \quad \Psi_j \; = \; A(z) \, \Phi_j \quad \Rightarrow \quad \| \mathbf{\Psi} \|_H^2 - \| \mathbf{\Phi} \|_H^2 - \sum_{j \le -r} S_j \; = \; \sum_{j=-r}^{0} (h_j - h_{j+1}) \, T_j + \sum_{j=-r+1}^{0} h_j \, S_j \; =: \; E \, , \tag{11}$$

where we have defined $h_{-r} := 1$ and $h_1 := 0$. Since the local dissipation $S_j$ is nonpositive for any $j$, we get from (11):

$$\forall \, j \in \mathbb{Z}^- \, , \quad \Psi_j \; = \; A(z) \, \Phi_j \quad \Rightarrow \quad \| \mathbf{\Psi} \|_H^2 \; \le \; \| \mathbf{\Phi} \|_H^2 + E \, .$$

The goal at this point is to determine the extrapolation procedure and the quadrature weights such that $E$ is, at least, non-positive. In that case, the combination of the operator $A(z)$ with the extrapolation procedure yields a contraction on $\ell^2(\mathbb{Z}^-)$ provided that the latter space is equipped with the norm (10). We formalize this notion below.

**Definition 1.** *Together with suitable extrapolation conditions at an outflow boundary, we say that the operator $A(z)$ in (3) is semi-bounded with respect to the inner product defined in (10) if, for a given $z \ge 0$, the following two conditions are satified.*

1. *The scheme is contractive, that is $A(z)$ admits a decomposition (4) with $S_j \le 0$.*

2. *The boundary contribution is non-positive, i.e.*

$$E := \sum_{j=-r}^{0} (h_j - h_{j+1})\, T_j + \sum_{j=-r+1}^{0} h_j\, S_j \leq 0\,, \tag{12}$$

*where we use the convention* $h_{-r} := 1$ *and* $h_1 := 0$.

As a matter of fact, it is useful for the purpose of convergence proofs (see, e.g., [Gus75, GKO95, CL20]) to infer from (11) a trace estimate on the sequence $\mathbf{\Phi}$. This will be the case provided that $E$ is a negative definite quadratic form of its arguments, as evidenced below on several examples. This energy argument bypasses, as in [Wu95, CG11, CL20], the lengthy verification of the fulfillment of the so-called Uniform Lopatinskii Condition. We refer to [GKS72, Cou13] for a general presentation of this topic.

The rest of this article is organized as follows. In Section 2, we examine the application of the energy method described above in the case of three point schemes ($\ell_- = \ell_+ = 1$). Since such schemes have at most second order accuracy [GKO95], we restrict our analysis to the case of first and second order extrapolation at the outflow boundary. We then examine the case of five point schemes ($\ell_- = \ell_+ = 2$) in Section 3. Since the involved algebra becomes heavier, we restrict sometimes the analysis to three particular cases which are analogues of some well-known three point schemes. We give examples of extrapolation procedures for which we can construct a suitable energy that yields semi-boundedness. In contrast to what is more commonly used, these energy stable extrapolation procedures are not translation invariant. Finally, in section 5 we demonstrate the limitation of more standard extrapolation procedures, proving that for a fourth order five point scheme combined with the most common first and second order extrapolation procedures, the energy method actually fails to work, though the corresponding closure of the operator $A(z)$ is known to be power-bounded. This shows that the normal mode analysis developed in [GKS72] and subsequent works is sometimes necessary to capture stability in general.

## 2 Three point schemes

The family of three point stencils of at least first order accuracy can be defined using a single free parameter $\nu$ as follows (see [RM94, GKO95]),

$$A(z) \;=\; I - z\, D_0 + \frac{\nu}{2}\,\Delta\,. \tag{13}$$

Famous examples include the Euler forward ($\nu = 0$), Lax-Friedrichs ($\nu = 1$), the two-point upwind ($\nu = z$) as well as the second order accurate Lax-Wendroff ($\nu = z^2$) schemes.

### 2.1 Contractivity on the whole real line

For the family of schemes considered in (13), our first result is the following.

**Lemma 1.** *Let* $z, \nu \in \mathbb{R}$, *and consider the scheme (3) with* $A(z)$ *as in* (13). *Then (4) holds with*

$$T_j := \begin{pmatrix} \Phi_j \\ D\,\Phi_{j+1} \end{pmatrix}^T Q \begin{pmatrix} \Phi_j \\ D\,\Phi_{j+1} \end{pmatrix}, \quad Q := \begin{pmatrix} -z & \dfrac{\nu - z}{2} \\ \dfrac{\nu - z}{2} & \dfrac{\nu\,(1-z)}{2} \end{pmatrix}, \tag{14}$$

*and*

$$S_j \;=\; \begin{pmatrix} D\,\Phi_j \\ D\,\Phi_{j+1} \end{pmatrix}^T M \begin{pmatrix} D\,\Phi_j \\ D\,\Phi_{j+1} \end{pmatrix}, \quad M \;:=\; a_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + a_2 \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \tag{15}$$

*where*

$$a_1 \;:=\; -\frac{\nu - z^2}{2}, \quad a_2 \;:=\; \frac{\nu^2 - z^2}{4}.$$

*In particular, $A(z)$ is a contraction on $\ell^2(\mathbb{Z})$, i.e. $S_j \leq 0$, if $z$ and $\nu$ satisfy the relation $z^2 \leq \nu \leq 1$.*

*Proof.* With $A(z)$ as in (13), we have in (4),

$$\Psi_j^2 - \Phi_j^2 \;=\; -2\,z\,\Phi_j\,(D_0\,\Phi_j) + z^2\,(D_0\,\Phi_j)^2 + \nu\,\Phi_j\,(\Delta\,\Phi_j) - z\,\nu\,(D_0\,\Phi_j)\,(\Delta\,\Phi_j) + \frac{\nu^2}{4}\,(\Delta\,\Phi_j)^2. \tag{16}$$

In order to split this expression into telescopic and non-telescopic parts, we first note that the $(D_0\,\Phi_j)^2$ term can be expressed as the sum of squares (this is nothing but the parallelogram identity),

$$(D_0\,\Phi_j)^2 \;=\; \frac{1}{4}\left[2\,(D\,\Phi_j)^2 + 2\,(D\,\Phi_{j+1})^2 - (\Delta\,\Phi_j)^2\right]. \tag{17}$$

Moreover, we will use the purely telescopic formulas,

$$2\,\Phi_j\,D_0\,\Phi_j \;=\; \Phi_j\,\Phi_{j+1} - \Phi_{j-1}\,\Phi_j \;=\; \left(\Phi_j^2 + \Phi_j\,D\,\Phi_{j+1}\right) - \left(\Phi_{j-1}^2 + \Phi_{j-1}\,D\,\Phi_j\right), \tag{18}$$

$$2\,D_0\,\Phi_j\,\Delta\,\Phi_j \;=\; (D\,\Phi_{j+1})^2 - (D\,\Phi_j)^2. \tag{19}$$

Finally, the $\Phi_j\,\Delta\,\Phi_j$ term can be written as the combination of non-telescopic (sum of squares) and telescopic parts[1],

$$\begin{aligned} 2\,\Phi_j\,\Delta\,\Phi_j \;&=\; -(D\,\Phi_j)^2 - (D\,\Phi_{j+1})^2 + \Phi_{j+1}^2 - 2\,\Phi_j^2 + \Phi_{j-1}^2 \\ &=\; -(D\,\Phi_j)^2 - (D\,\Phi_{j+1})^2 + \left[2\,\Phi_j\,D\,\Phi_{j+1} + (D\,\Phi_{j+1})^2\right] - \left[2\,\Phi_{j-1}\,D\,\Phi_j + (D\,\Phi_j)^2\right]. \end{aligned} \tag{20}$$

Inserting the formulas (17) through (20) into (16), we get the desired form $\Psi_j^2 - \Phi_j^2 = T_j - T_{j-1} + S_j$ with $T_j$ directly as in (14) and $S_j$ given by

$$S_j \;=\; a_1\left[(D\,\Phi_j)^2 + (D\,\Phi_{j+1})^2\right] + a_2\,(\Delta\,\Phi_j)^2.$$

From the definition of the discrete Laplacian $\Delta$ in (7), we can further rewrite this into (15). This proves the first part of the Lemma.

Next, an orthogonal set of eigenvectors to $M$ is given by $(1\ 1)^T$ and $(1\ -1)^T$, and the associated eigenvalues are,

$$\lambda_1 \;=\; a_1 \;=\; -\frac{\nu - z^2}{2}, \quad \lambda_2 \;=\; a_1 + 2\,a_2 \;=\; -\frac{\nu\,(1-\nu)}{2}.$$

We have,

$$\begin{aligned} \lambda_1 \leq 0 \quad &\Leftrightarrow \quad z^2 \leq \nu, \\ \lambda_2 \leq 0 \quad &\Leftrightarrow \quad \nu \in [0,1]. \end{aligned}$$

and we have thus shown that the scheme defined by $A(z)$ is a contraction if $z^2 \leq \nu \leq 1$. □

[1] The decomposition mimics the equality $2\,u\,u'' = -2\,(u')^2 + (u^2)''$.

It is worth to mention here that the same exact bounds on $z, \nu$ for $\ell^2$ stability given in Lemma 1 above can also be derived using Fourier analysis (they are both necessary and sufficient for $A(z)$ to be a contraction). However, for a wider stencil we can not in general guarantee that the energy method will yield the same bounds as obtained with the Fourier method. Since we will use the energy method to derive extrapolation boundary conditions leading to semi-boundedness, we shall in this paper only consider stencils for which contractivity can be shown using the energy method in a similar fashion to the one above.

To conclude this section, we write down the eigenvalues associated with the quadratic form $S_j$ for the important special cases mentioned at the beginning of this section. The Euler forward method ($\nu = 0$) yields,

$$\lambda_1 \,=\, \frac{z^2}{2}\,, \quad \lambda_2 \,=\, 0\,,$$

showing that this scheme is not contractive for any $z$. Moreover, the Lax-Friedrichs scheme ($\nu = 1$) yields,

$$\lambda_1 \,=\, -\frac{1-z^2}{2}\,, \quad \lambda_2 \,=\, 0\,,$$

the upwind scheme ($\nu = z$) yields,

$$\lambda_1 \,=\, \lambda_2 \,=\, -\frac{z\,(1-z)}{2}\,,$$

and finally the Lax-Wendroff ($\nu = z^2$) scheme,

$$\lambda_1 \,=\, 0\,, \quad \lambda_2 \,=\, -\frac{z^2\,(1-z^2)}{4}\,.$$

This shows that the Lax-Friedrichs and Lax-Wendroff schemes are both contractive for all $|z| \le 1$ (meaning that we can freely change the sign in front of $u_x$ in (1) without modifying the stability property of the scheme), while the upwind scheme is contractive for $0 \le z \le 1$.

### 2.2 Outflow boundary conditions

For simplicity, in (10) we let $r = 1$ from the start, i.e. the norm of $\mathbf{\Phi} \in \ell^2(\mathbb{Z}^-)$ in (10) is defined by

$$\|\mathbf{\Phi}\|_H^2 \,=\, \sum_{j \le -1} \Phi_j^2 + h_0\,\Phi_0^2\,,$$

and for now we consider the last quadrature weight $h_0$ as a free (positive) parameter. In (12) we then have, using Lemma 1 from the previous paragraph,

$$\begin{aligned} E \,&=(1-h_0)\,T_{-1} + h_0\,T_0 + h_0\,S_0 \\ &=(1-h_0)\begin{pmatrix}\Phi_{-1}\\ D\,\Phi_0\end{pmatrix}^T Q \begin{pmatrix}\Phi_{-1}\\ D\,\Phi_0\end{pmatrix} + h_0 \begin{pmatrix}\Phi_0\\ D\,\Phi_1\end{pmatrix}^T Q \begin{pmatrix}\Phi_0\\ D\,\Phi_1\end{pmatrix} + h_0 \begin{pmatrix}D\,\Phi_0\\ D\,\Phi_1\end{pmatrix}^T M \begin{pmatrix}D\,\Phi_0\\ D\,\Phi_1\end{pmatrix}. \end{aligned} \tag{21}$$

Inserting (14) and (15) into the above expression (21) for $E$ gives us right away a stability estimate for the most popular first order extrapolation boundary closure.

**Corollary 1.** *Let $z \in (0,1]$ and let $\nu$ satisfy $z^2 \le \nu \le 1$. Then the scheme (13) together with the first order extrapolation condition $\Phi_1 = \Phi_0$ is semi-bounded. In particular, the energy balance is given by,*

$$\sum_{j\le 0} \Psi_j^2 \le \sum_{j\le 0} \Phi_j^2 - z\,\Phi_0^2 \le \sum_{j\le 0} \Phi_j^2\,, \tag{22}$$

*which exactly mimics (2) with added dissipation.*

*Proof.* We choose $h_0 = 1$ for the last quadrature weight. From the extrapolation condition we can use $D\,\Phi_1 = 0$ in (21) to immediately get,

$$E \;=\; \begin{pmatrix}\Phi_0\\ 0\end{pmatrix}^T Q \begin{pmatrix}\Phi_0\\ 0\end{pmatrix} + \begin{pmatrix}D\,\Phi_0\\ 0\end{pmatrix}^T M \begin{pmatrix}D\,\Phi_0\\ 0\end{pmatrix} \;=\; -z\,\Phi_0^2 + S_0\,.$$

The energy balance (22) then follows from (11) and the fact that $S_j \le 0$ for any $j \le 0$ if $z^2 \le \nu \le 1$. □

We now examine the case of a second order extrapolation condition at the outflow boundary, that is we impose $\Delta\,\Phi_0 = 0$. Our result is the following.

**Corollary 2.** *Let $z \in (0,1]$ and let $\nu$ satisfy $z^2 \le \nu \le 1$, and consider the scheme (13) together with the second order extrapolation condition $\Delta\,\Phi_0 = 0$. Then $E$ in (12) is given by*

$$E \;=\; \begin{pmatrix}\Phi_0\\ D\,\Phi_0\end{pmatrix}^T B \begin{pmatrix}\Phi_0\\ D\,\Phi_0\end{pmatrix}\,, \quad B \;=\; \begin{pmatrix} -z & z(\frac{1}{2} - h_0) + \frac{\nu}{2} \\ z(\frac{1}{2} - h_0) + \frac{\nu}{2} & h_0 z^2 - \frac{\nu(1+z)}{2}\end{pmatrix}\,. \tag{23}$$

*Moreover, $B$ is negative definite (and thus the scheme is semi-bounded) for the choice $h_0 = (1 - z + \nu/z)/2$.*

*Proof.* From the second order extrapolation condition, we have,

$$\Delta\,\Phi_0 \;=\; 0 \quad \Rightarrow \quad D\,\Phi_1 \;=\; D\,\Phi_0\,.$$

We also have, from the definition of $D$ in (5),

$$\Phi_{-1} \;=\; \Phi_0 - D\,\Phi_0\,.$$

Going back to (21), we now write $E$ as,

$$E \;=\; (1-h_0)\begin{pmatrix}\Phi_0 - D\,\Phi_0\\ D\,\Phi_0\end{pmatrix}^T Q \begin{pmatrix}\Phi_0 - D\,\Phi_0\\ D\,\Phi_0\end{pmatrix} + h_0 \begin{pmatrix}\Phi_0\\ D\,\Phi_0\end{pmatrix}^T Q \begin{pmatrix}\Phi_0\\ D\,\Phi_0\end{pmatrix} + h_0 \begin{pmatrix}D\,\Phi_0\\ D\,\Phi_0\end{pmatrix}^T M \begin{pmatrix}D\,\Phi_0\\ D\,\Phi_0\end{pmatrix}\,,$$

which, using (14) and (15), can be further simplified into (23).

In order to see whether $B$ in (23) is negative definite, we rotate to diagonal form using the similarity transformation,

$$R \;:=\; \begin{pmatrix} 1 & \left(\dfrac{\nu+z}{2} - z\,h_0\right)/z \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad R^T B\,R \;=\; \begin{pmatrix} -z & 0 \\ 0 & \gamma\end{pmatrix}\,,$$

where

$$\gamma \;:=\; \Big[(\nu + z - 2\,z\,h_0)^2 + 4\,h_0\,z^3 - 2\,z\,(1+z)\,\nu\Big]\,/(4\,z)\,.$$

The minimum of $\gamma$ with respect to $h_0$ is obtained for

$$\frac{\partial \gamma}{\partial h_0} \;=\; z^2 - z - \nu + 2\,h_0\,z \;=\; 0 \quad \Leftrightarrow \quad h_0 \;=\; \frac{\frac{\nu}{z} - z + 1}{2} \geq \frac{1}{2}\,,$$

which in all cases leads to the same negative value of $\gamma$,

$$\gamma \;=\; -z^3/4\,.$$

With this choice of $h_0$, the quantity (10) (with $r = 1$) defines a norm on $\ell^2(-\infty, 0)$ and the second condition of Definition 1 is always satisfied for $z \in (0, 1]$ and $\nu \in [z^2, 1]$. Semi-boundedness is thus obtained as long as the interior scheme is contractive. □

In particular, the second order accurate Lax-Wendroff scheme ($\nu = z^2$) is associated with the likewise second order accurate Trapezoidal rule with the constant value $h_0 = 1/2$ (independently of $z \in (0, 1]$). For all other three-point schemes, the above value for $h_0$ as obtained with this method depends on $z$ and grows to infinity as $z$ approaches 0 ($\nu$ being fixed).

## 3 Five point schemes

In this preliminary work we do not aim for a complete theory encompassing finite difference schemes of any order. Instead we go on to consider a slightly more involved example than in the previous section, namely the family of schemes with a five point stencil and at least second order accuracy. Such schemes can be parametrized with two free parameters $\sigma, \tau$ as follows,

$$A(z) \;=\; I - z\,D_0 + \frac{z^2}{2}\,\Delta + \sigma\,D_0\,\Delta + \tau\,\Delta^2\,. \tag{24}$$

We first derive a general integration by parts decomposition formula for $A(z)$ and thereby obtain three interesting examples for which $A(z)$ is contractive.

### 3.1 Contractivity on the whole real line

Generalizing Lemma 1 to the family of operators in (24), our result is the following.

**Lemma 2.** *Let $z, \sigma, \tau \in \mathbb{R}$, and consider the scheme (3) with $A(z)$ as in (24). Then (4) holds with*

$$T_j \;:=\; \begin{pmatrix} \Phi_j \\ D\,\Phi_{j+1} \\ \Delta\,\Phi_j \\ D\,\Delta\,\Phi_{j+1} \end{pmatrix}^T Q \begin{pmatrix} \Phi_j \\ D\,\Phi_{j+1} \\ \Delta\,\Phi_j \\ D\,\Delta\,\Phi_{j+1} \end{pmatrix},$$

$$Q \;:=\; \begin{pmatrix} -z & -\frac{z\,(1-z)}{2} & \sigma & \frac{\sigma}{2} + \tau \\ -\frac{z\,(1-z)}{2} & \frac{z^2\,(1-z)}{2} - \sigma & \frac{\sigma}{2} - z\,\sigma - \tau & -z\,\left(\frac{\sigma}{2} + \tau\right) \\ \sigma & \frac{\sigma}{2} - z\,\sigma - \tau & \frac{z^2\,\sigma}{2} + z\,\tau & \frac{z\,(1+z)}{2}\,\left(\frac{\sigma}{2} + \tau\right) - \frac{a_1}{3} \\ \frac{\sigma}{2} + \tau & -z\,\left(\frac{\sigma}{2} + \tau\right) & \frac{z\,(1+z)}{2}\,\left(\frac{\sigma}{2} + \tau\right) - \frac{a_1}{3} & \frac{z\,\sigma}{4} + \frac{z^2\,\tau}{2} + \sigma\,\tau - \frac{a_1}{3} \end{pmatrix}, \tag{25}$$

*and*

$$S_j := \begin{pmatrix} \Delta\,\Phi_{j-1} \\ \Delta\,\Phi_j \\ \Delta\,\Phi_{j+1} \end{pmatrix}^T M \begin{pmatrix} \Delta\,\Phi_{j-1} \\ \Delta\,\Phi_j \\ \Delta\,\Phi_{j+1} \end{pmatrix}, \; M := \frac{a_1}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + a_2 \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} + a_3 \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}, \tag{26}$$

*where*

$$a_1 := \frac{z^2\,(z^2-1)}{4} + 2\,z\,\sigma + 2\,\tau\,, \quad a_2 := -\frac{z\,\sigma}{4} + \frac{\sigma^2}{2} - \frac{z^2\,\tau}{2}\,, \quad a_3 := -\frac{\sigma^2}{4} + \tau^2\,.$$

*In particular, $A(z)$ is a contraction on $\ell^2(\mathbb{Z})$ in each of the three following cases:*

- $\sigma = 0$, $\tau = -(1-z^2)/12$ *and* $z \in [-1,1]$,
- $\sigma = z\,(1-z^2)/6$, $\tau = -\sigma/2$ *and* $z \in [0,1]$,
- $\sigma = z\,(1-z^2)/6$, $\tau = -z^2\,(1-z^2)/24$ *and* $z \in [-1,1]$.

*Proof.* Using (24) in (4), we now have,

$$\begin{aligned} \Psi_j^2 - \Phi_j^2 \;=\; & -2\,z\,\Phi_j\,D_0\,\Phi_j + z^2\,(D_0\,\Phi_j)^2 + z^2\Phi_j\,\Delta\,\Phi_j - z^3\,(D_0\,\Phi_j)\,(\Delta\,\Phi_j) + \frac{z^4}{4}\,(\Delta\,\Phi_j)^2 \\ & + 2\,\sigma\,\Phi_j\,D_0\,\Delta\,\Phi_j - 2\,z\,\sigma\,(D_0\Phi_j)\,(D_0\,\Delta\,\Phi_j) + z^2\,\sigma\,(\Delta\,\Phi_j)\,(D_0\,\Delta\,\Phi_j) + \sigma^2\,(D_0\,\Delta\,\Phi_j)^2 \\ & + 2\,\tau\,\Phi_j\,(\Delta^2\,\Phi_j) - 2\,z\,\tau\,(D_0\,\Phi_j)\,(\Delta^2\,\Phi_j) + z^2\,\tau\,(\Delta\,\Phi_j)\,(\Delta^2\,\Phi_j) + 2\,\sigma\,\tau\,(D_0\,\Delta\,\Phi_j)\,(\Delta^2\,\Phi_j) \\ & + \tau^2\,(\Delta^2\,\Phi_j)^2\,. \end{aligned} \tag{27}$$

The terms in the first line on the right hand side above are covered by the previous formulas (17), (18), (19) and (20). Further, by substituting $\Phi_j$ for $\Delta\,\Phi_j$ in the same formulas, we get,

$$(D_0\,\Delta\,\Phi_j)^2 \;=\; \frac{1}{4}\,\left[2\,(D\,\Delta\,\Phi_j)^2 + 2\,(D\,\Delta\,\Phi_{j+1})^2 - (\Delta^2\,\Phi_j)^2\right], \tag{28}$$

as well as

$$2\,(\Delta\,\Phi_j)\,(D_0\,\Delta\,\Phi_j) \;=\; \left[(\Delta\,\Phi_j)^2 + \Delta\,\Phi_j\,D\,\Delta\,\Phi_{j+1}\right] - \left[(\Delta\,\Phi_{j-1})^2 + \Delta\,\Phi_{j-1}\,D\,\Delta\,\Phi_j\right], \tag{29}$$

$$2\,(D_0\,\Delta\,\Phi_j)\,(\Delta^2\,\Phi_j) \;=\; (D\,\Delta\,\Phi_{j+1})^2 - (D\,\Delta\,\Phi_j)^2\,, \tag{30}$$

and

$$\begin{aligned} 2\,\Delta\,\Phi_j\,\Delta^2\,\Phi_j \;=\; & -(D\,\Delta\,\Phi_j)^2 - (D\,\Delta\,\Phi_{j+1})^2 \\ & + \left[2\,\Delta\,\Phi_j\,D\,\Delta\,\Phi_{j+1} + (D\,\Delta\,\Phi_{j+1})^2\right] - \left[2\,\Delta\,\Phi_{j-1}\,D\,\Delta\,\Phi_j + (D\,\Delta\,\Phi_j)^2\right]. \end{aligned} \tag{31}$$

For the remaining terms on the right hand side of (27), we need to introduce the following four new high order formulas, the proof of which can be found in Appendix,

$$\begin{aligned} 2\,D_0\,\Phi_j\,D_0\,\Delta\,\Phi_j \;=\; & -2\,(\Delta\,\Phi_j)^2 + \frac{1}{4}\,\left[(D\,\Delta\,\Phi_j)^2 + (D\,\Delta\,\Phi_{j+1})^2\right] \\ & + \left[(D\,\Phi_{j+1} - \frac{1}{4}\,D\,\Delta\,\Phi_{j+1})\,(D\,\Delta\,\Phi_{j+1} + 2\,\Delta\,\Phi_j)\right] - \left[(D\,\Phi_j - \frac{1}{4}\,D\,\Delta\,\Phi_j)\,(D\,\Delta\,\Phi_j + 2\,\Delta\,\Phi_{j-1})\right], \end{aligned} \tag{32}$$

as well as,

$$2\,\Phi_j\,D_0\,\Delta\,\Phi_j = [\Phi_j\,(D\,\Delta\,\Phi_{j+1} + 2\,\Delta\,\Phi_j) + (\Delta\,\Phi_j - D\,\Phi_{j+1})\,D\,\Phi_{j+1}] \\ - [\Phi_{j-1}\,(D\,\Delta\,\Phi_j + 2\,\Delta\,\Phi_{j-1}) + (\Delta\,\Phi_{j-1} - D\,\Phi_j)\,D\,\Phi_j]\,, \tag{33}$$

$$2\,D_0\,\Phi_j\,\Delta^2\,\Phi_j = [-(D\,\Delta\,\Phi_{j+1} + \Delta\,\Phi_j)\,\Delta\,\Phi_j + 2\,D\,\Phi_{j+1}\,D\,\Delta\,\Phi_{j+1}] \\ - [-(D\,\Delta\,\Phi_j + \Delta\,\Phi_{j-1})\,\Delta\,\Phi_{j-1} + 2\,D\,\Phi_j\,D\,\Delta\,\Phi_j]\,, \tag{34}$$

and

$$2\,\Phi_j\,\Delta^2\,\Phi_j \;=\; 2\,(\Delta\,\Phi_j)^2 + [2\,\Phi_j\,D\,\Delta\,\Phi_{j+1} - 2\,D\,\Phi_{j+1}\,\Delta\,\Phi_j] - [2\,\Phi_{j-1}\,D\,\Delta\,\Phi_j - 2\,D\,\Phi_j\,\Delta\,\Phi_{j-1}]\,. \tag{35}$$

Note that all the non-telescopic parts above are expressed in terms of the three variables $(\Delta\,\Phi_j)^2$, $(\Delta^2\,\Phi_j)^2$ and $(D\,\Delta\,\Phi_j)^2 + (D\,\Delta\,\Phi_{j+1})^2$. The last two of these can be written as,

$$(\Delta^2\,\Phi_j)^2 \;=\; \begin{pmatrix}\Delta\,\Phi_{j-1}\\ \Delta\,\Phi_j\\ \Delta\,\Phi_{j+1}\end{pmatrix}^T \begin{pmatrix}1\\-2\\1\end{pmatrix} \begin{pmatrix}1 & -2 & 1\end{pmatrix} \begin{pmatrix}\Delta\,\Phi_{j-1}\\ \Delta\,\Phi_j\\ \Delta\,\Phi_{j+1}\end{pmatrix} = \begin{pmatrix}\Delta\,\Phi_{j-1}\\ \Delta\,\Phi_j\\ \Delta\,\Phi_{j+1}\end{pmatrix}^T \begin{pmatrix}1 & -2 & 1\\ -2 & 4 & -2\\ 1 & -2 & 1\end{pmatrix} \begin{pmatrix}\Delta\,\Phi_{j-1}\\ \Delta\,\Phi_j\\ \Delta\,\Phi_{j+1}\end{pmatrix}\,,$$

and

$$(D\,\Delta\,\Phi_j)^2 + (D\,\Delta\,\Phi_{j+1})^2 \;=\; \begin{pmatrix}\Delta\,\Phi_{j-1}\\ \Delta\,\Phi_j\\ \Delta\,\Phi_{j+1}\end{pmatrix}^T \begin{pmatrix}1 & -1 & 0\\ -1 & 2 & -1\\ 0 & -1 & 1\end{pmatrix} \begin{pmatrix}\Delta\,\Phi_{j-1}\\ \Delta\,\Phi_j\\ \Delta\,\Phi_{j+1}\end{pmatrix}\,.$$

Thus, in order to analyze the sign of the complete non-telescopic part, it is convenient to expand $(\Delta\,\Phi_j)^2$ into a similar matrix form with a telescopic correction,

$$(\Delta\,\Phi_j)^2 \;=\; \frac{1}{3}\begin{pmatrix}\Delta\,\Phi_{j-1}\\ \Delta\,\Phi_j\\ \Delta\,\Phi_{j+1}\end{pmatrix}^T \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} \begin{pmatrix}\Delta\,\Phi_{j-1}\\ \Delta\,\Phi_j\\ \Delta\,\Phi_{j+1}\end{pmatrix} \\ - \frac{1}{3}\,\left[(D\Delta\,\Phi_j + \Delta\,\Phi_j)^2 + (\Delta\,\Phi_j)^2\right] + \frac{1}{3}\,\left[(D\Delta\,\Phi_{j-1} + \Delta\,\Phi_{j-1})^2 + (\Delta\,\Phi_{j-1})^2\right]\,.$$

Inserting (28) through (35) into (27) and then rewriting $(\Delta\,\Phi_j)^2$, $(\Delta^2\,\Phi_j)^2$ and $(D\,\Delta\,\Phi_j)^2 + (D\,\Delta\,\Phi_{j+1})^2$ according to the above, the first part of the Lemma follows. Note that the term $a_1\,(\Delta\,\Phi_j)^2$ has been split between between the telescopic and dissipative parts.

The same orthogonal set of eigenvectors to each of the three terms in (26) is given by $(1,1,1)^T$, $(-1,0,1)^T$ and $(1,-2,1)^T$, with the associated eigenvalues of $M$,

$$\lambda_1 \;:=\; \frac{a_1}{3}\,, \quad \lambda_2 \;:=\; \frac{a_1}{3} + a_2\,, \quad \lambda_3 \;:=\; \frac{a_1}{3} + 3\,a_2 + 6\,a_3\,.$$

The operator $A(z)$ is a contraction on $\ell^2(\mathbb{Z})$ if all three eigenvalues $\lambda_{1,2,3}$ of $M$ are non-positive. In contrast to the three point stencil case, we can not easily rewrite the conditions for contractivity into explicit relations for $\sigma$ and $\tau$. However, three special cases seem to be worth to mention. First, if we let $\sigma = 0$ and $\tau = -(1-z^2)/12$, then the eigenvalues of $M$ simplify into,

$$\lambda_1 \;=\; -\frac{(1-z^2)\,(2+3\,z^2)}{36}\,, \quad \lambda_2 \;=\; -\frac{(1-z^2)\,(4+3\,z^2)}{72}\,, \quad \lambda_3 \;=\; -\frac{1-z^2}{72}\,.$$

Hence $A(z)$ is a contraction for any $z \in [-1,1]$ in that case. Recall that for the second order Lax-Wendroff scheme, the dissipation approaches 0 as $z$ approaches 0, which is not the case here. In terms of dissipation properties, this new scheme can instead be seen in some sense to be analogous to the three-point Lax-Friedrichs scheme (though one eigenvalue for the Lax-Friedrichs scheme is zero, which is not the case here).

A skewed $4-$point stencil of accuracy order 3, which can be seen as analogous to the two-point upwind scheme, is given by $\sigma := z\,(1-z^2)/6$ and $\tau := -\sigma/2 = -z(1-z^2)/12$, yielding,

$$\lambda_1 \;=\; -\frac{z\,(1-z^2)\,(2-z)}{36}\,, \quad \lambda_2 \;=\; -\,\frac{z\,(1-z^2)\,(1+z)\,(2-z)^2}{72}\,,$$
$$\lambda_3 \;=\; -\,\frac{z\,(1-z^2)\,(2-z)\,(2+3\,z-3\,z^2)}{72}\,.$$

Again, $A(z)$ is a contraction on $\ell^2(\mathbb{Z})$ for any $z \in [0,1]$. Finally, the so-called Strang scheme of fourth order accuracy [Str62] is given by $\sigma := z\,(1-z^2)/6$ and $\tau := -z^2\,(1-z^2)/24$, with the corresponding eigenvalues,

$$\lambda_1 \;=\; 0\,, \quad \lambda_2 \;=\; -\frac{z^2\,(1-z^2)\,(4-z^2)}{144}\,, \quad \lambda_3 \;=\; -\frac{z^2\,(1-z^2)\,(3-z^2)\,(4-z^2)}{96}\,.$$

The operator $A(z)$ is again a contraction on $\ell^2(\mathbb{Z})$ for any $z \in [-1,1]$. This scheme is a high (i.e., fourth) order analogue to the second order accurate Lax-Wendroff scheme. □

We note that the corresponding stability analysis for the fourth order Strang scheme in [Str62] is performed by means of the Fourier transform and is therefore not applicable for the outflow problem, which will be considered next.

### 3.2 Outflow boundary conditions

We examine in this paragraph two sets of outflow boundary conditions based on second order accurate extrapolation, and thus leave higher order boundary conditions as a topic for future work. As opposed to the three point stencil case, we have not found a single set of second order boundary conditions which is stable for all of the three example schemes listed in Lemma 2, which is why we consider two different alternatives below.

In what follows, we shall use the inner product (10) with $r = 1$ and $h_0 = 1/2$. In (12), we thus have,

$$E \;=\; \frac{1}{2}\begin{pmatrix} \Phi_{-1} \\ D\,\Phi_0 \\ \Delta\,\Phi_{-1} \\ D\,\Delta\,\Phi_0 \end{pmatrix}^T Q \begin{pmatrix} \Phi_{-1} \\ D\,\Phi_0 \\ \Delta\,\Phi_{-1} \\ D\,\Delta\,\Phi_0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \Phi_0 \\ D\,\Phi_1 \\ \Delta\,\Phi_0 \\ D\,\Delta\,\Phi_1 \end{pmatrix}^T Q \begin{pmatrix} \Phi_0 \\ D\,\Phi_1 \\ \Delta\,\Phi_0 \\ D\,\Delta\,\Phi_1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \Delta\,\Phi_{-1} \\ \Delta\,\Phi_0 \\ \Delta\,\Phi_1 \end{pmatrix}^T M \begin{pmatrix} \Delta\,\Phi_{-1} \\ \Delta\,\Phi_0 \\ \Delta\,\Phi_1 \end{pmatrix}, \tag{36}$$

with $Q$ given in (25) and $M$ given in (26).

#### 3.2.1 Second order extrapolation of type 1

We can prove

**Proposition 1.** *Let $r = 1$, $h_0 = 1/2$, and consider the set of second order extrapolation conditions $\Delta\,\Phi_0 = D_0\,\Delta\,\Phi_0 = 0$. In (36), we then have,*

$$E = \begin{pmatrix} \Phi_0 \\ D\,\Phi_0 \\ D^2\,\Phi_0 \end{pmatrix}^T B \begin{pmatrix} \Phi_0 \\ D\,\Phi_0 \\ D^2\,\Phi_0 \end{pmatrix}, \; B := \begin{pmatrix} -z & \frac{z^2}{2} & \frac{\sigma}{2} \\ \frac{z^2}{2} & -\frac{z^3}{2} - \sigma & \frac{-z\sigma}{2} \\ \frac{\sigma}{2} & \frac{-z\sigma}{2} & \frac{2\,\tau}{3} + \sigma\,\tau + \frac{5\,z\,\sigma}{12} - \frac{z^2\,\tau}{2} + 2\,\tau^2 - \frac{z^2\,(1-z^2)}{12} \end{pmatrix}. \tag{37}$$

*In particular, the scheme is semi-bounded (i.e. $B$ is negative semi-definite) for all $z \in (0,1)$ at least in the following two special cases,*

- *$\sigma = 0$ and $\tau = -(1-z^2)/12$,*
- *$\sigma = z\,(1-z^2)/6$ and $\tau = -\sigma/2$.*

*Proof.* The two extrapolation conditions readily yield,

$$D\,\Phi_1 \;=\; D\,\Phi_0 \qquad \Delta\,\Phi_1 \;=\; \Delta\,\Phi_{-1}\,.$$

By combining the two we also get $D\,\Delta\,\Phi_1 = \Delta\,\Phi_{-1}$, and by definition we also have $\Phi_{-1} = \Phi_0 - D\,\Phi_0$ as well as $\Delta\,\Phi_{-1} = D^2\,\Phi_0$. With these formulas, we can now simplify $E$ in (36) into,

$$E = \frac{1}{2} \begin{pmatrix} \Phi_0 - D\,\Phi_0 \\ D\,\Phi_0 \\ D^2\,\Phi_0 \\ -D^2\,\Phi_0 \end{pmatrix}^T Q \begin{pmatrix} \Phi_0 - D\,\Phi_0 \\ D\,\Phi_0 \\ D^2\,\Phi_0 \\ -D^2\,\Phi_0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \Phi_0 \\ D\,\Phi_0 \\ 0 \\ D^2\,\Phi_0 \end{pmatrix}^T Q \begin{pmatrix} \Phi_0 \\ D\,\Phi_0 \\ 0 \\ D^2\,\Phi_0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} D^2\,\Phi_0 \\ 0 \\ D^2\,\Phi_0 \end{pmatrix}^T M \begin{pmatrix} D^2\,\Phi_0 \\ 0 \\ D^2\,\Phi_0 \end{pmatrix},$$

which in turn, after some straightforward algebra, leads to (37).

For the semi-boundedness part of the Proposition, let us first consider the case $\sigma = 0$ and $\tau = -(1-z^2)/12$. The general expression for the matrix $B$ in (37) reduces to,

$$B := \begin{pmatrix} -z & z^2/2 & 0 \\ z^2/2 & -\frac{z^3}{2} & 0 \\ 0 & 0 & -\frac{(1-z^2)\,(3+4\,z^2)}{72} \end{pmatrix}.$$

It is a simple exercise to verify that the upper left $2 \times 2$ block is negative definite for $z \in (0,1)$, hence $B$ is negative definite for $z \in (0,1)$. We get in that case the conclusion of Proposition 1 as an immediate consequence. We now focus on the case $\sigma = z\,(1-z^2)/6$ and $\tau = -\sigma/2$, for which the matrix $B$ reads:

$$B = -z \begin{pmatrix} 1 & -z/2 & -(1-z^2)/12 \\ -z/2 & \frac{1+2\,z^2}{6} & z\,(1-z^2)/12 \\ -(1-z^2)/12 & z\,(1-z^2)/12 & \frac{(1-z^2)\,(1+z)\,(4-3\,z)}{72} \end{pmatrix}.$$

We then compute:

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix}^T B \begin{pmatrix} U \\ V \\ W \end{pmatrix} = -z \left( \left( U - \frac{z}{2} V - \frac{1-z^2}{12} W \right)^2 + \frac{z^2}{4} V^2 + \frac{(1-z^2)}{6} \left( \left( V + \frac{z}{4} W \right)^2 + \frac{14 + 4\,z - 8\,z^2}{48} W^2 \right) \right),$$

thereby showing that $B$ is negative semi-definite for $z \in (0,1)$. This completes the proof of Proposition 1. □

### 3.2.2 Outflow boundary: second order extrapolation of type 2

We examine in this paragraph a second set of outflow boundary conditions with second order accuracy.

**Proposition 2.** *Let $r = 1$, $h_0 = 1/2$, and consider the set of second order extrapolation conditions $\Delta\,\Phi_0 = \Delta^2\,\Phi_0 = 0$. Then we have*

$$E = \begin{pmatrix} \Phi_0 \\ D\,\Phi_0 \\ D^2\,\Phi_0 \end{pmatrix}^T B \begin{pmatrix} \Phi_0 \\ D\,\Phi_0 \\ D^2\,\Phi_0 \end{pmatrix}, \quad B := \begin{pmatrix} -z & z^2/2 & -\tau \\ z^2/2 & -\frac{z^3}{2} - \sigma & z\,\tau \\ -\tau & z\,\tau & \frac{2\,\tau}{3} + \sigma\,\tau + \frac{5\,z\,\sigma}{12} - \frac{z^2\,\tau}{2} + \frac{\sigma^2}{2} - \frac{z^2\,(1-z^2)}{12} \end{pmatrix}. \tag{38}$$

*In particular, the scheme is semi-bounded (i.e. $B$ is negative semi-definite) for all $z \in (0,1)$ at least in the following two special cases,*

- *$\sigma = z\,(1-z^2)/6$ and $\tau = -\sigma/2$,*
- *$\sigma = z\,(1-z^2)/6$ and $\tau = -z^2\,(1-z^2)/24$.*

*Proof.* We use (36) and insert the second order extrapolation conditions $\Delta\,\Phi_0 = \Delta^2\,\Phi_0 = 0$, from which we deduce the relations $D\,\Phi_1 = D\,\Phi_0$, and $D\,\Delta\,\Phi_1 = \Delta\,\Phi_1 = -\Delta\,\Phi_{-1}$. The boundary contribution $E$ in the energy balance (11) reads

$$E = \frac{1}{2} \begin{pmatrix} \Phi_0 - D\,\Phi_0 \\ D\,\Phi_0 \\ D^2\,\Phi_0 \\ -D^2\,\Phi_0 \end{pmatrix}^T Q \begin{pmatrix} \Phi_0 - D\,\Phi_0 \\ D\,\Phi_0 \\ D^2\,\Phi_0 \\ -D^2\,\Phi_0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \Phi_0 \\ D\,\Phi_0 \\ 0 \\ -D^2\,\Phi_0 \end{pmatrix}^T Q \begin{pmatrix} \Phi_0 \\ D\,\Phi_0 \\ 0 \\ -D^2\,\Phi_0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} D^2\,\Phi_0 \\ 0 \\ -D^2\,\Phi_0 \end{pmatrix}^T M \begin{pmatrix} D^2\,\Phi_0 \\ 0 \\ -D^2\,\Phi_0 \end{pmatrix},$$

which leads to (38).

We first focus on the case $\sigma = z\,(1-z^2)/6$ and $\tau = -\sigma/2$, for which the matrix $B$ reads:

$$B = -z \begin{pmatrix} 1 & -z/2 & -(1-z^2)/12 \\ -z/2 & \frac{1+2\,z^2}{6} & z\,(1-z^2)/12 \\ -(1-z^2)/12 & z\,(1-z^2)/12 & \frac{(1-z^2)\,(1+z)\,(4-3\,z)}{72} \end{pmatrix},$$

which we have already analyzed in the proof of Proposition 1. In particular, we have already shown that $B$ is negative semi-definite for any $z \in (0,1)$ and the result of Proposition 2 follows in that case. Finally, and most importantly, the fourth order Strang scheme $\sigma = z\,(1-z^2)/6$, $\tau = -z^2\,(1-z^2)/24$ yields the expression,

$$B \;=\; -z \begin{pmatrix} 1 & -z/2 & -z\,(1-z^2)/24 \\ -z/2 & \dfrac{1+2\,z^2}{6} & z^2\,(1-z^2)/24 \\ -z\,(1-z^2)/24 & z^2\,(1-z^2)/24 & \frac{z\,(1-z^2)\,(4-z^2+z\,(1-z^2))}{144} \end{pmatrix}.$$

We thus compute:

$$\begin{aligned} &\begin{pmatrix} U \\ V \\ W \end{pmatrix}^T B \begin{pmatrix} U \\ V \\ W \end{pmatrix} \\ &\quad = -z \left( \left( U - \frac{z}{2}\,V - \frac{z\,(1-z^2)}{24}\,W \right)^2 + \frac{1}{6}\,V^2 + \frac{z^2}{12} \left( V + \frac{1-z^2}{4}\,W \right)^2 + \frac{z\,(1-z^2)\,(4-z^2)}{144}\,W^2 \right), \end{aligned}$$

showing that $B$ is negative semi-definite for any $z \in (0,1)$. Hence the result of Proposition 2 follows in the case of the fourth order Strang scheme. □

In view of all above results, the energy method seems to be a rather efficient tool to prove stability estimates for numerical schemes that are contractive in the whole space $\mathbb{Z}$ combined with some carefully selected extrapolation conditions at an outflow boundary. It is the purpose of the next section to illustrate the limitations of this energy approach, namely that it can not be applied to the more often considered translation invariant extrapolation conditions. For such boundary conditions and more general schemes (based for instance on multistep quadrature methods in time), the general theory for proving stability of numerical boundary conditions initiated in [GKS72] remains the only available one.

## 4 On the limitation of the energy method for analyzing numerical boundary conditions

In this paragraph, we consider the fourth order Strang scheme [Str62]:

$$A(z) \;:=\; I - z\,D_0 + \frac{z^2}{2}\,\Delta + \frac{z\,(1-z^2)}{6}\,D_0\,\Delta - \frac{z^2\,(1-z^2)}{24}\,\Delta^2. \tag{39}$$

With a prescribed five point stencil, the finite difference scheme (39) is the only one that achieves fourth order accuracy (with respect to both space and time), just like the Lax-Wendroff scheme is the only three point scheme that achieves second order accuracy. Our goal is to study the semi-boundedness of (39) when implemented on a half line $\mathbb{Z}^-$ with extrapolation numerical boundary conditions, be they for instance of order 1 or 2. The extrapolation boundary conditions considered in [Kre66, Gol77, CL20] are, in the terminology of [GT78, GT81], *translatory*, meaning that they are of the exact same form in each ghost cell (as opposed to the extrapolation conditions considered in Propositions 1 and 2). The main results in [Kre66, Kre68, Gol77, CL20] show that, whatever the extrapolation order at the outflow boundary, the corresponding operator on $\ell^2(-\infty,0)$ is power-bounded. The proof of this result in [Kre66, Kre68, Gol77] relies on the normal mode decomposition and power-boundedness follows from the general result in [Wu95], while the more direct proof of the same result in [CL20] relies first on the energy method for the Dirichlet

boundary condition and on an induction argument with respect to the extrapolation order at the outflow boundary. Our goal below is to determine whether, as in Corollary 2, such stability estimates can be achieved by means of a straightforward energy method without using "auxiliary problems". The answer is *negative*, see Theorems 1 and 2 below, which seems to indicate that the induction argument in [CL20] is more or less the shortest way to derive stability estimates for translatory extrapolation numerical boundary conditions. Examples of uniformly stable, though non dissipative, boundary conditions have been known for quite some time in the context of hyperbolic partial differential equations, see for instance the examples provided in [BGRSZ02, Ben14]. As far as we are aware of, the example given by Theorem 1 seems to be the first in the fully discrete setting.

Our first main result for the Strang scheme (39) is the following.

**Theorem 1.** *There does not exist a parameter $z_0 > 0$, an integer $r \in \mathbb{N}$, and a continuous map $H$ from $[0, z_0]$ with values in $\mathscr{M}_r(\mathbb{R})$ such that for all $z \in [0, z_0]$, $H(z)$ is a symmetric matrix satisfying the following property: for any sequence $\mathbf{\Phi} \in \ell^2(-\infty, 2)$ satisfying the first order extrapolation boundary condition $\Phi_2 = \Phi_1 = \Phi_0$, the following energy inequality holds:*

$$\sum_{j \leq -r} \Psi_j^2 + \begin{pmatrix} \Psi_{-r+1} \\ \vdots \\ \Psi_0 \end{pmatrix}^T H(z) \begin{pmatrix} \Psi_{-r+1} \\ \vdots \\ \Psi_0 \end{pmatrix} - \sum_{j \leq -r} \Phi_j^2 - \begin{pmatrix} \Phi_{-r+1} \\ \vdots \\ \Phi_0 \end{pmatrix}^T H(z) \begin{pmatrix} \Phi_{-r+1} \\ \vdots \\ \Phi_0 \end{pmatrix} \leq 0 \,, \tag{40}$$

*with the sequence $\mathbf{\Psi} \in \ell^2(\mathbb{Z}^-)$ being defined by:*

$$\forall\, j \leq 0 \,, \quad \Psi_j := A(z)\, \Phi_j = \Phi_j - z\, D_0\, \Phi_j + \frac{z^2}{2}\, \Delta\, \Phi_j + \frac{z\,(1 - z^2)}{6}\, D_0\, \Delta\, \Phi_j - \frac{z^2\,(1 - z^2)}{24}\, \Delta^2\, \Phi_j \,.$$

Let us observe that we do not even assume the matrix $H(z)$ to be positive definite, which would be necessary to make the quantity:

$$\sum_{j \leq -r} \Phi_j^2 + \begin{pmatrix} \Phi_{-r+1} \\ \vdots \\ \Phi_0 \end{pmatrix}^T H(z) \begin{pmatrix} \Phi_{-r+1} \\ \vdots \\ \Phi_0 \end{pmatrix} ,$$

the square of a norm on the space of sequences $\mathbf{\Phi} \in \ell^2(-\infty, 2)$ with the prescribed boundary conditions. The obstacle for the existence of $H(z)$ in Theorem 1 comes from the small values of $z$. It could very well be that for some $z \in (0, 1)$, one can construct a real symmetric positive definite matrix $H(z)$ of size $r \in \mathbb{N}$ such that one has the optimal energy balance:

$$\sum_{j \leq -r} \Psi_j^2 + \begin{pmatrix} \Psi_{-r+1} \\ \vdots \\ \Psi_0 \end{pmatrix}^T H(z) \begin{pmatrix} \Psi_{-r+1} \\ \vdots \\ \Psi_0 \end{pmatrix} - \sum_{j \leq -r} \Phi_j^2 - \begin{pmatrix} \Phi_{-r+1} \\ \vdots \\ \Phi_0 \end{pmatrix}^T H(z) \begin{pmatrix} \Phi_{-r+1} \\ \vdots \\ \Phi_0 \end{pmatrix} + c\,(\Phi_{-1}^2 + \Phi_0^2) \leq 0 \,,$$

with $c$ a positive constant. Theorem 1 shows, however, that such a construction will not be possible for all values of $z \in (0, 1)$ with the additional requirement that $H(z)$ extends continuously to $z = 0$ (as was the case, for instance, for the Lax-Wendroff scheme with second order extrapolation by Corollary 2).

*Proof of Theorem 1.* The proof of Theorem 1 is based on an induction argument with respect to the integer $r$. As will follow from the argument below, it appears that the initial step of the induction

argument corresponds to $r = 2$. The first two cases $r = 0$ and $r = 1$ are dealt with separately, though the argument is similar to the one for $r = 2$. Let us start with a general argument from which the conclusion of Theorem 1 will follow.

• The derivative of the energy balance with respect to $z$ at 0. In all what follows, we assume that the opposite statement to Theorem 1 holds, meaning that we assume that there exist a parameter $z_0 > 0$, an integer $r \in \mathbb{N}$, and a continuous map $H$ from $[0, z_0]$ with values in $\mathscr{M}_r(\mathbb{R})$ such that for any $z \in [0, z_0]$, $H(z)$ is a real symmetric matrix with the previously stated property. Let us consider a sequence $\mathbf{\Phi} \in \ell^2(-\infty, 2)$ that satisfies $\Phi_2 = \Phi_1 = \Phi_0$, and let us then denote with $f(z)$ the quantity on the left hand side of (40). Since the sequence $\mathbf{\Psi}(z)$ depends in a $\mathscr{C}^1$ fashion on $z$ in $\ell^2$, and since $H$ is continuous at 0, the function $f$ is differentiable at 0 and it holds that,

$$f'(0) \;=\; \sum_{j \leq -r} 2\,\Psi_j(0)\,\Psi'_j(0) + 2 \begin{pmatrix} \Psi'_{-r+1}(0) \\ \vdots \\ \Psi'_0(0) \end{pmatrix}^T H(0) \begin{pmatrix} \Phi_{-r+1} \\ \vdots \\ \Phi_0 \end{pmatrix}. \tag{41}$$

Substituting the value of $\Psi'_j(0)$, (41) can be expressed in terms of the sequence $\mathbf{\Phi}$ only, i.e.

$$\begin{aligned} f'(0) \;=\; & \sum_{j \leq -r} -2\,\Phi_j\,D_0\,\Phi_j + \frac{1}{3}\,\Phi_j\,D_0\,\Delta\,\Phi_j \\ & + \begin{pmatrix} -(\Phi_{-r+2} - \Phi_{-r}) + (\Phi_{-r+3} - 2\,\Phi_{-r+2} + 2\,\Phi_{-r} - \Phi_{-r-1})/6 \\ \vdots \\ -(\Phi_1 - \Phi_{-1}) + (\Phi_2 - 2\,\Phi_1 + 2\,\Phi_{-1} - \Phi_{-2})/6 \end{pmatrix}^T H(0) \begin{pmatrix} \Phi_{-r+1} \\ \vdots \\ \Phi_0 \end{pmatrix} \\ =\; & -\frac{4}{3}\,\Phi_{-r}\,\Phi_{-r+1} + \frac{1}{6}\,(\Phi_{-r-1}\,\Phi_{-r+1} + \Phi_{-r}\,\Phi_{-r+2}) \\ & + \begin{pmatrix} -(\Phi_{-r+2} - \Phi_{-r}) + (\Phi_{-r+3} - 2\,\Phi_{-r+2} + 2\,\Phi_{-r} - \Phi_{-r-1})/6 \\ \vdots \\ -(\Phi_1 - \Phi_{-1}) + (\Phi_2 - 2\,\Phi_1 + 2\,\Phi_{-1} - \Phi_{-2})/6 \end{pmatrix}^T H(0) \begin{pmatrix} \Phi_{-r+1} \\ \vdots \\ \Phi_0 \end{pmatrix}. \end{aligned}$$

We observe that $f$ vanishes at $z = 0$ because $\Psi_j(0) = \Phi_j$ for all $j$, and $f$ takes nonpositive values on $[0, z_0]$ by (40). Hence 0 is a maximum of $f$ on $[0, z_0]$ and the derivative $f'(0)$ is nonpositive. In other words, we have just obtained that there exists a real symmetric matrix $\mathscr{H}$ of size $r$ such that for any sequence $\mathbf{\Phi} \in \ell^2(-\infty, 2)$ satisfying $\Phi_2 = \Phi_1 = \Phi_0$, the following inequality holds:

$$\begin{aligned} & -\frac{4}{3}\,\Phi_{-r}\,\Phi_{-r+1} + \frac{1}{6}\,(\Phi_{-r-1}\,\Phi_{-r+1} + \Phi_{-r}\,\Phi_{-r+2}) \\ & \qquad + \begin{pmatrix} -(\Phi_{-r+2} - \Phi_{-r}) + (\Phi_{-r+3} - 2\,\Phi_{-r+2} + 2\,\Phi_{-r} - \Phi_{-r-1})/6 \\ \vdots \\ -(\Phi_1 - \Phi_{-1}) + (\Phi_2 - 2\,\Phi_1 + 2\,\Phi_{-1} - \Phi_{-2})/6 \end{pmatrix}^T \mathscr{H} \begin{pmatrix} \Phi_{-r+1} \\ \vdots \\ \Phi_0 \end{pmatrix} \leq 0\,. \end{aligned} \tag{42}$$

It remains to examine the consequences of (42), which is where the specific value of $r$ comes into play because the values $\Phi_1$ and $\Phi_2$ are not arbitrary due to the extrapolation boundary conditions.

• The case $r = 0$. In that case, the inequality (42) reduces to:

$$-\frac{4}{3}\,\Phi_0\,\Phi_1 + \frac{1}{6}\,(\Phi_{-1}\,\Phi_1 + \Phi_0\,\Phi_2) \;\leq\; 0\,.$$

Using the extrapolation boundary conditions $\Phi_2 = \Phi_1 = \Phi_0$, we end up with:

$$-\frac{7}{6}\,\Phi_0^2 + \frac{1}{6}\,\Phi_0\,\Phi_{-1} \,\leq\, 0\,,$$

which is obviously impossible since the values $\Phi_0, \Phi_{-1}$ are arbitrary. This means that we cannot use the standard $\ell^2$ norm on $\mathbb{Z}^-$ for proving the stability of (39) with first order extrapolation condition. Let us now deal with the next case ($r = 1$) in the induction argument.

• The case $r = 1$. In that case, the inequality (42) reduces to:

$$-\frac{4}{3}\,\Phi_{-1}\,\Phi_0 + \frac{1}{6}\,(\Phi_{-2}\,\Phi_0 + \Phi_{-1}\,\Phi_1) - \mathscr{H}\,\Phi_0\,(\Phi_1 - \Phi_{-1}) + \frac{\mathscr{H}}{6}\,\Phi_0\,(\Phi_2 - 2\,\Phi_1 + 2\,\Phi_{-1} - \Phi_{-2}) \,\leq\, 0\,,$$

where $\mathscr{H}$ is a real number. After using the boundary conditions $\Phi_2 = \Phi_1 = \Phi_0$, we get:

$$-\frac{7}{6}\,\Phi_{-1}\,\Phi_0 + \frac{1}{6}\,\Phi_{-2}\,\Phi_0 - \mathscr{H}\,\Phi_0\,(\Phi_0 - \Phi_{-1}) - \frac{\mathscr{H}}{6}\,\Phi_0\,(\Phi_0 - 2\,\Phi_{-1} + \Phi_{-2}) \,\leq\, 0\,, \tag{43}$$

where now the three values $\Phi_0, \Phi_{-1}, \Phi_{-2}$ are arbitrary. It is useful to introduce the new variables:

$$y_3 \,:=\, \Phi_0\,, \quad y_2 \,:=\, \Phi_0 - \Phi_{-1}\,, \quad y_1 \,:=\, \Phi_0 - 2\,\Phi_{-1} + \Phi_{-2}\,,$$

with which (43) is rewritten as follows:

$$\forall\, y \in \mathbb{R}^3\,, \quad -y_3^2 + \left(\frac{5}{6} - \mathscr{H}\right) y_3\,y_2 + \frac{1 - \mathscr{H}}{6}\,y_3\,y_1 \,\leq\, 0\,, \tag{44}$$

The latter inequality is obviously impossible since the only available parameter $\mathscr{H}$ should equal both 1 and 5/6 to cancel the off-diagonal terms $y_3\,y_2$ and $y_3\,y_1$. This completes the proof of Theorem 1 in the case $r = 1$. Let us now deal with the case $r = 2$.

• The case $r = 2$. This is really the starting point of the induction argument, and we shall borrow the methodology introduced for the case $r = 1$. Namely, when $r$ equals 2, the inequality (42) reduces to:

$$-\frac{4}{3}\,\Phi_{-2}\,\Phi_{-1} + \frac{1}{6}\,(\Phi_{-3}\,\Phi_{-1} + \Phi_{-2}\,\Phi_0) + \begin{pmatrix} -(\Phi_0 - \Phi_{-2}) + (\Phi_1 - 2\,\Phi_0 + 2\,\Phi_{-2} - \Phi_{-3})/6 \\ -(\Phi_1 - \Phi_{-1}) + (\Phi_2 - 2\,\Phi_1 + 2\,\Phi_{-1} - \Phi_{-2})/6 \end{pmatrix}^T \mathscr{H} \begin{pmatrix} \Phi_{-1} \\ \Phi_0 \end{pmatrix} \,\leq\, 0\,. \tag{45}$$

After using the boundary conditions $\Phi_2 = \Phi_1 = \Phi_0$, (45) reduces to:

$$-\frac{4}{3}\,\Phi_{-2}\,\Phi_{-1} + \frac{1}{6}\,(\Phi_{-3}\,\Phi_{-1} + \Phi_{-2}\,\Phi_0) - \begin{pmatrix} (\Phi_0 - \Phi_{-2}) + (\Phi_0 - 2\,\Phi_{-2} + \Phi_{-3})/6 \\ (\Phi_0 - \Phi_{-1}) + (\Phi_0 - 2\,\Phi_{-1} + \Phi_{-2})/6 \end{pmatrix}^T \mathscr{H} \begin{pmatrix} \Phi_{-1} \\ \Phi_0 \end{pmatrix} \,\leq\, 0\,. \tag{46}$$

Let us extend the strategy used in the case $r = 1$, and introduce the new variables:

$$y_4 \,:=\, \Phi_0\,, \quad y_3 \,:=\, D\,\Phi_0\,, \quad y_2 \,:=\, D^2\,\Phi_0\,, \quad y_1 \,:=\, D^3\,\Phi_0\,.$$

The inequality (46) can be equivalently rewritten as:

$$-y_4^2 + 3\,y_4\,y_3 - \frac{2}{3}\,y_4\,y_2 - \frac{1}{6}\,y_4\,y_1 - \frac{13}{6}\,y_3^2 + \frac{5}{6}\,y_3\,y_2 + \frac{1}{6}\,y_3\,y_1 - \frac{1}{6}\begin{pmatrix} 13\,y_3 - 5\,y_2 - y_1 \\ 6\,y_3 + y_2 \end{pmatrix}^T \mathscr{H} \begin{pmatrix} y_4 - y_3 \\ y_4 \end{pmatrix} \,\leq\, 0\,, \tag{47}$$

where (47) holds for all $y \in \mathbb{R}^4$, since $\Phi_0, \Phi_{-1}, \Phi_{-2}, \Phi_{-3}$ in (46) are arbitrary. It is useful at this stage to introduce the coefficients of the symmetric matrix $\mathscr{H}$, and we thus write:

$$\mathscr{H} \,=\, \begin{pmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{pmatrix} .$$

The quadratic form in $y \in \mathbb{R}^4$ on the left hand side of (47) is nonpositive, and furthermore it has no $y_1^2$ term. This implies that the coefficients of the cross products $y_4\, y_1$ and $y_3\, y_1$ must vanish. Computing those coefficients, we get:

$$h_{11} \,=\, 1\,, \quad h_{12} \,=\, 0\,,$$

which means that the matrix $\mathscr{H}$ reads:

$$\mathscr{H} \,=\, \begin{pmatrix} 1 & 0 \\ 0 & h_{22} \end{pmatrix} ,$$

and then (47) reduces to:

$$-y_4^2 + \left( \frac{5}{6} - h_{22} \right) y_4\, y_3 + \frac{1 - h_{22}}{6}\, y_4\, y_2 \,\leq\, 0\,,$$

which is nothing else but the inequality (44) we had obtained in the analysis of the case $r = 1$ except for the shift in the indeces (one should only substitute $(y_4, y_3, y_2)$ in place of $(y_3, y_2, y_1)$ in (44), and $h_{22}$ in place of $\mathscr{H}$). As already observed in the analysis of the case $r = 1$, we are led to a contradiction, which completes the proof of Theorem 1 in the case $r = 2$.

• The general case $r \geq 3$. We go back to (42) and assume $r \geq 3$. In particular, the first line on the left hand side of (42) does not involve the ghost cell values $\Phi_1, \Phi_2$. Substituting the first order extrapolation boundary conditions $\Phi_2 = \Phi_1 = \Phi_0$ in (42), we get the inequality:

$$\begin{aligned} &-\frac{4}{3}\,\Phi_{-r}\,\Phi_{-r+1} + \frac{1}{6}\,(\Phi_{-r-1}\,\Phi_{-r+1} + \Phi_{-r}\,\Phi_{-r+2}) \\ &\qquad + \begin{pmatrix} -(\Phi_{-r+2} - \Phi_{-r}) + (\Phi_{-r+3} - 2\,\Phi_{-r+2} + 2\,\Phi_{-r} - \Phi_{-r-1})/6 \\ \vdots \\ -(\Phi_{-1} - \Phi_{-3}) + (\Phi_0 - 2\,\Phi_{-1} + 2\,\Phi_{-3} - \Phi_{-4})/6 \\ -(\Phi_0 - \Phi_{-2}) - (\Phi_0 - 2\,\Phi_{-2} + \Phi_{-3})/6 \\ -(\Phi_0 - \Phi_{-1}) - (\Phi_0 - 2\,\Phi_{-1} + \Phi_{-2})/6 \end{pmatrix}^T \mathscr{H} \begin{pmatrix} \Phi_{-r+1} \\ \vdots \\ \Phi_{-2} \\ \Phi_{-1} \\ \Phi_0 \end{pmatrix} \leq\, 0\,. \end{aligned} \tag{48}$$

We introduce the variable $y \in \mathbb{R}^{r+2}$ defined by:

$$\forall\, \ell \,=\, 1, \ldots, r+2\,, \quad y_\ell \,:=\, D^{r+2-\ell}\,\Phi_0\,,$$

which, conversely, corresponds to:

$$\forall\, \ell \,=\, 0, \ldots, r+1\,, \quad \Phi_{-\ell} \,=\, D^\ell\, y_{r+2}\,.$$

Using from now on the coordinates of $y \in \mathbb{R}^{r+2}$ as free parameters, (48) reads:

$$-\frac{4}{3}\,D^r\,y_{r+2}\,D^{r-1}\,y_{r+2}+\frac{1}{6}\,(D^{r+1}\,y_{r+2}\,D^{r-1}\,y_{r+2}+D^r\,y_{r+2}\,D^{r-2}\,y_{r+2})$$
$$+\begin{pmatrix}-2\,D^{r-2}\,y_{r+1}+D^{r-2}\,y_r+(2\,D^{r-3}\,y_{r-1}-D^{r-3}\,y_{r-2})/6\\ \vdots\\ -2\,D\,y_{r+1}+D\,y_r+(2\,y_{r-1}-y_{r-2})/6\\ -(13\,y_{r+1}-5\,y_r-y_{r-1})/6\\ -y_{r+1}-y_r/6\end{pmatrix}^T\mathscr{H}\begin{pmatrix}D^{r-1}\,y_{r+2}\\ \vdots\\ D^2\,y_{r+2}\\ D\,y_{r+2}\\ y_{r+2}\end{pmatrix}\le 0\,. \quad (49)$$

We are not going to compute all the coefficients of the quadratic form (in $y$) arising on the left hand side of (49). It is useful however to observe that the first two coordinates $y_1$ and $y_2$ of $y$ do not appear in the expressions of $D^{r-1}\,y_{r+2},\,\ldots,\,D\,y_{r+2},\,y_{r+2}$. Therefore, if we rewrite the quadratic form (in $y$) arising on the left hand side of (49) as $y^T\,S\,y$, with $S$ a real symmetric matrix of size $r+2$, then not only $S$ is nonpositive because of (49), but $S$ also reads:

$$S\,=\,\begin{pmatrix}\widetilde{S} & \Upsilon_2 & \Upsilon_1\\ \Upsilon_2^T & 0 & 0\\ \Upsilon_1^T & 0 & 0\end{pmatrix}\,,$$

with $\Upsilon_1, \Upsilon_1 \in \mathbb{R}^r$, and $\widetilde{S}$ a real symmetric matrix of size $r$. Since $S$ is nonnegative, we must necessarily have $\Upsilon_1 = \Upsilon_2 = 0$. In other words, this means that no cross product of the form $y_1\,y_3,\ldots,y_1\,y_{r+2}$ or $y_2\,y_3,\ldots,y_2\,y_{r+2}$ arises on the left hand side of (49), or, equivalently, that the quantity on the left hand side of (49) does not depend on $(y_1, y_2)$. Computing the partial derivative with respect to $y_1$, we get the relation:

$$\frac{(-1)^{r+1}}{6}\left((1-\mathscr{H}_{11})\,D^{r-1}\,y_{r+2}+\sum_{\ell=2}^{r}\mathscr{H}_{1\ell}\,D^{r-\ell}\,y_{r+2}\right)\,=\,0\,,$$

from which we deduce that the first line of $\mathscr{H}$ should read:

$$\begin{pmatrix}1 & 0 & \cdots & 0\end{pmatrix}\,.$$

Since $\mathscr{H}$ is symmetric, (49) reduces to:

$$-\frac{4}{3}\,D^r\,y_{r+2}\,D^{r-1}\,y_{r+2}+\frac{1}{6}\,(D^{r+1}\,y_{r+2}\,D^{r-1}\,y_{r+2}+D^r\,y_{r+2}\,D^{r-2}\,y_{r+2})$$
$$+\Big(-2\,D^{r-2}\,y_{r+1}+D^{r-2}\,y_r+\frac{1}{3}\,D^{r-3}\,y_{r-1}-\frac{1}{6}\,D^{r-3}\,y_{r-2}\Big)\,D^{r-1}\,y_{r+2} \quad (50)$$
$$+\begin{pmatrix}-2\,D^{r-3}\,y_{r+1}+D^{r-3}\,y_r+(2\,D^{r-4}\,y_{r-1}-D^{r-4}\,y_{r-2})/6\\ \vdots\\ -2\,D\,y_{r+1}+D\,y_r+(2\,y_{r-1}-y_{r-2})/6\\ -(13\,y_{r+1}-5\,y_r-y_{r-1})/6\\ -y_{r+1}-y_r/6\end{pmatrix}^T\mathscr{H}_\sharp\begin{pmatrix}D^{r-2}\,y_{r+2}\\ \vdots\\ D^2\,y_{r+2}\\ D\,y_{r+2}\\ y_{r+2}\end{pmatrix}\le 0\,,$$

where the real symmetric matrix $\mathscr{H}_\sharp$ of size $r-1$ corresponds to the block decomposition of $\mathscr{H}$:

$$\mathscr{H}\,=\,\begin{pmatrix}1 & 0\\ 0 & \mathscr{H}_\sharp\end{pmatrix}\,.$$

We can simplify the first two lines of (50) by using the relation:

$$
\begin{aligned}
&- 2\,D^{r-2}\,y_{r+1} + D^{r-2}\,y_r + \frac{1}{3}\,D^{r-3}\,y_{r-1} - \frac{1}{6}\,D^{r-3}\,y_{r-2} \\
&\qquad = -\big(D^{r-2}\,y_{r+2} - D^r\,y_{r+2}\big) + \frac{1}{6}\,\big(D^{r-3}\,y_{r+2} - 2\,D^{r-2}\,y_{r+2} + 2\,D^r\,y_{r+2} - D^{r+1}\,y_{r+2}\big)\,,
\end{aligned}
$$

and (50) can be rewritten as:

$$
\begin{aligned}
&- \frac{4}{3}\,D^{r-1}\,y_{r+2}\,D^{r-2}\,y_{r+2} + \frac{1}{6}\,(D^r\,y_{r+2}\,D^{r-2}\,y_{r+2} + D^{r-1}\,y_{r+2}\,D^{r-3}\,y_{r+2}) \\
&\qquad + \begin{pmatrix} -2\,D^{r-3}\,y_{r+1} + D^{r-3}\,y_r + (2\,D^{r-4}\,y_{r-1} - D^{r-4}\,y_{r-2})/6 \\ \vdots \\ -2\,D\,y_{r+1} + D\,y_r + (2\,y_{r-1} - y_{r-2})/6 \\ -(13\,y_{r+1} - 5\,y_r - y_{r-1})/6 \\ -y_{r+1} - y_r/6 \end{pmatrix}^T \mathscr{H}_\sharp \begin{pmatrix} D^{r-2}\,y_{r+2} \\ \vdots \\ D^2\,y_{r+2} \\ D\,y_{r+2} \\ y_{r+2} \end{pmatrix} \le 0\,. \qquad (51)
\end{aligned}
$$

Shfiting the indeces in the variables, that is introducing the vector:

$$
(\tilde{y}_{r+1}, \dots, \tilde{y}_1) := (y_{r+2}, \dots, y_2)\,,
$$

and forgetting about the tilde, we see that (51) is exactly the same as (49) with the integer $r-1$ in place of $r$. By a finite induction process, we can therefore show that the validity of (49) for some real symmetric matrix $\mathscr{H}$ of size $r$ implies the validity of (47) (which is exactly (49) in the particular case $r = 2$), and we have already seen that this leads to a contradiction. The proof of Theorem 1 is now complete. □

The exact same argument of proof can be used to deal with the case of second order extrapolation at the boundary. We shall not reproduce the proof here and leave the (minor) modifications to the interested reader. We thus only state the final result, which is entirely similar to Theorem 1 above except for the extrapolation conditions.

**Theorem 2.** *There does not exist a parameter $z_0 > 0$, an integer $r \in \mathbb{N}$, and a continuous map $H$ from $[0, z_0]$ with values in $\mathscr{M}_r(\mathbb{R})$ such that for any $z \in [0, z_0]$, $H(z)$ is a symmetric matrix satisfying the following property: for any sequence $\mathbf{\Phi} \in \ell^2(-\infty, 2)$ verifying the second order extrapolation boundary conditions $\Delta\,\Phi_1 = \Delta\,\Phi_0 = 0$, with the sequence $\mathbf{\Psi}$ being defined by:*

$$
\forall\, j \le 0\,, \quad \Psi_j := A(z)\,\Phi_j = \Phi_j - z\,D_0\,\Phi_j + \frac{z^2}{2}\,\Delta\,\Phi_j + \frac{z\,(1-z^2)}{6}\,D_0\,\Delta\,\Phi_j - \frac{z^2\,(1-z^2)}{24}\,\Delta^2\,\Phi_j\,,
$$

*then the following energy inequality holds:*

$$
\sum_{j \le -r} \Psi_j^2 + \begin{pmatrix} \Psi_{-r+1} \\ \vdots \\ \Psi_0 \end{pmatrix}^T H(z) \begin{pmatrix} \Psi_{-r+1} \\ \vdots \\ \Psi_0 \end{pmatrix} - \sum_{j \le -r} \Phi_j^2 - \begin{pmatrix} \Phi_{-r+1} \\ \vdots \\ \Phi_0 \end{pmatrix}^T H(z) \begin{pmatrix} \Phi_{-r+1} \\ \vdots \\ \Phi_0 \end{pmatrix} \le 0\,.
$$

Theorems 1 and 2 show that for the fourth order Strang scheme (39), stability for the *translatory* first or second order extrapolation conditions:

$$
\begin{aligned}
\text{(first order)} \quad & \Phi_2 = \Phi_1 = \Phi_0\,, \\
\text{(second order)} \quad & \Delta\,\Phi_1 = \Delta\,\Phi_0 = 0\,,
\end{aligned}
$$

cannot be obtained by a 'straightforward' energy argument (at least for all relevant values of the CFL parameter $z$), meaning by the construction of an energy that is a finite rank perturbation of the identity that is non-increasing for the associated evolution operator. It is known nevertheless that these numerical boundary conditions satisfy the strong stability condition of [GKS72], see [Gol77]. These two above examples clearly indicate that the theory initiated in [GKS72] is the only one able to capture stability for numerical boundary conditions in general.

# A Higher order integration by parts decompositions

This appendix is devoted to the proof of the relations (32)-(35) which we have used in the proof of Lemma 2. Let us start with the proof of formula (32), which we rewrite here for the reader's convenience:

$$\begin{aligned} 2\, D_0\, \Phi_j\, D_0\, \Delta\, \Phi_j \;=\; & -2\, (\Delta\, \Phi_j)^2 + \frac{1}{4}\, \left[ (D\, \Delta\, \Phi_j)^2 + (D\, \Delta\, \Phi_{j+1})^2 \right] \\ & + \left[ (D\, \Phi_{j+1} - \frac{1}{4}\, D\, \Delta\, \Phi_{j+1})\, (D\, \Delta\, \Phi_{j+1} + 2\, \Delta\, \Phi_j) \right] - \left[ (D\, \Phi_j - \frac{1}{4}\, D\, \Delta\, \Phi_j)\, (D\, \Delta\, \Phi_j + 2\, \Delta\, \Phi_{j-1}) \right] . \end{aligned} \tag{52}$$

We first apply the formula (20) to get:

$$\begin{aligned} 2\, D_0\, \Phi_j\, \Delta\, D_0\, \Phi_j \;=\; & -(D\, D_0\, \Phi_j)^2 - (D\, D_0\, \Phi_{j+1})^2 \\ & + \left[ 2\, D_0\, \Phi_j\, D\, D_0\, \Phi_{j+1} + (D\, D_0\, \Phi_{j+1})^2 \right] - \left[ 2\, D_0\, \Phi_{j-1}\, D\, D_0\, \Phi_j + (D\, D_0\, \Phi_j)^2 \right] . \end{aligned} \tag{53}$$

and then rewrite the first line on the right hand side of (53) (that is, the symmetric terms) as:

$$\begin{aligned} (D\, D_0\, \Phi_j)^2 + (D\, D_0\, \Phi_{j+1})^2 \;=\; & \frac{1}{4}\, (\Delta\, \Phi_{j-1} + \Delta\, \Phi_j)^2 + \frac{1}{4}\, (\Delta\, \Phi_j + \Delta\, \Phi_{j+1})^2 \\ = \; & \frac{1}{4}\, (-D\, \Delta\, \Phi_j + 2\, \Delta\, \Phi_j)^2 + \frac{1}{4}\, (2\, \Delta\, \Phi_j + D\, \Delta\, \Phi_{j+1})^2 \\ = \; & 2\, (\Delta\, \Phi_j)^2 + \Delta\, \Phi_j\, \Delta^2\, \Phi_j + \frac{1}{4}\, \left( (D\, \Delta\, \Phi_j)^2 + (D\, \Delta\, \Phi_{j+1})^2 \right) \end{aligned}$$

We use again (20) for the term $\Delta\, \Phi_j\, \Delta^2\, \Phi_j$ and get:

$$\begin{aligned} (D\, D_0\, \Phi_j)^2 + (D\, D_0\, \Phi_{j+1})^2 \;=\; & 2\, (\Delta\, \Phi_j)^2 - \frac{1}{4}\, \left( (D\, \Delta\, \Phi_j)^2 + (D\, \Delta\, \Phi_{j+1})^2 \right) \\ & + \left[ \frac{1}{2} (\Delta\, \Phi_{j+1})^2 - \frac{1}{2} (\Delta\, \Phi_j)^2 \right] - \left[ \frac{1}{2} (\Delta\, \Phi_j)^2 - \frac{1}{2} (\Delta\, \Phi_{j-1})^2 \right] . \end{aligned}$$

Substituting this relation in the first line on the right hand side of (53), we have thus obtained the decomposition:

$$2\, D_0\, \Phi_j\, \Delta\, D_0\, \Phi_j \;=\; -2\, (\Delta\, \Phi_j)^2 + \frac{1}{4}\, \left( (D\, \Delta\, \Phi_j)^2 + (D\, \Delta\, \Phi_{j+1})^2 \right) + T_j - T_{j-1}\, ,$$

with

$$\begin{aligned} T_j \;:=\; & 2\, D_0\, \Phi_j\, D\, D_0\, \Phi_{j+1} + (D\, D_0\, \Phi_{j+1})^2 - \frac{1}{2} (\Delta\, \Phi_{j+1})^2 + \frac{1}{2} (\Delta\, \Phi_j)^2 \\ = \; & (D_0\, \Phi_{j+1})^2 - (D_0\, \Phi_j)^2 - \frac{1}{2} (\Delta\, \Phi_{j+1})^2 + \frac{1}{2} (\Delta\, \Phi_j)^2 . \end{aligned}$$

The decomposition (52) then follows by just rewriting the quantity $T_j$ in terms of $D\,\Phi_{j+1}$, $\Delta\,\Phi_j$ and $D\,\Delta\,\Phi_{j+1}$. The (small) details are left to the reader.

We now wish to justify the purely telescopic formula (33), which we also rewrite here for convenience:

$$\begin{aligned} 2\,\Phi_j\,D_0\,\Delta\,\Phi_j \;=\; & [\Phi_j\,(D\,\Delta\,\Phi_{j+1}+2\,\Delta\,\Phi_j)+(\Delta\,\Phi_j-D\,\Phi_{j+1})\,D\,\Phi_{j+1}] \\ & -[\Phi_{j-1}\,(D\,\Delta\,\Phi_j+2\,\Delta\,\Phi_{j-1})+(\Delta\,\Phi_{j-1}-D\,\Phi_j)\,D\,\Phi_j]\;. \end{aligned} \tag{54}$$

Let us write the left hand side of (54) as:

$$2\,\Phi_j\,D_0\,\Delta\,\Phi_j \;=\; \begin{pmatrix}\Phi_{j+2}\\ \Phi_{j+1}\\ \Phi_j\\ \Phi_{j-1}\\ \Phi_{j-2}\end{pmatrix}^T \begin{pmatrix} 0 & 0 & 1/2 & 0 & 0\\ 0 & 0 & -1 & 0 & 0\\ 1/2 & -1 & 0 & 1 & -1/2\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & -1/2 & 0 & 0\end{pmatrix} \begin{pmatrix}\Phi_{j+2}\\ \Phi_{j+1}\\ \Phi_j\\ \Phi_{j-1}\\ \Phi_{j-2}\end{pmatrix}\;,$$

and decompose the corresponding symmetric matrix in a telescopic way:

$$\begin{pmatrix} 0 & 0 & 1/2 & 0 & 0\\ 0 & 0 & -1 & 0 & 0\\ 1/2 & -1 & 0 & 1 & -1/2\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & -1/2 & 0 & 0\end{pmatrix} = \begin{pmatrix} 0 & 0 & 1/2 & 0 & 0\\ 0 & 0 & -1 & 1/2 & 0\\ 1/2 & -1 & 0 & 0 & 0\\ 0 & 1/2 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1/2 & 0\\ 0 & 0 & 0 & -1 & 1/2\\ 0 & 1/2 & -1 & 0 & 0\\ 0 & 0 & 1/2 & 0 & 0\end{pmatrix}\;.$$

At this stage, we have obtained the telescopic decomposition:

$$2\,\Phi_j\,D_0\,\Delta\,\Phi_j \;=\; [\Phi_j\,\Phi_{j+2}-2\,\Phi_j\,\Phi_{j+1}+\Phi_{j-1}\,\Phi_{j+1}]-[\Phi_{j-1}\,\Phi_{j+1}-2\,\Phi_{j-1}\,\Phi_j+\Phi_{j-2}\,\Phi_j]\;,$$

and the proof of (54) follows by rewriting the telescopic term as:

$$\begin{aligned} \Phi_j\,\Phi_{j+2}-2\,\Phi_j\,\Phi_{j+1}+\Phi_{j-1}\,\Phi_{j+1} &= \Phi_j\,\Delta\,\Phi_{j+1}+\Phi_j\,\Delta\,\Phi_j-D\,\Phi_j\,D\,\Phi_{j+1} \\ &= \Phi_j\,(D\,\Delta\,\Phi_{j+1}+2\,\Delta\,\Phi_j)+(\Delta\,\Phi_j-D\,\Phi_{j+1})\,D\,\Phi_{j+1}\,. \end{aligned}$$

This completes the proof of (54).

We now turn to the proof of the telescopic formula (34), which we also rewrite here for convenience:

$$\begin{aligned} 2\,D_0\,\Phi_j\,\Delta^2\,\Phi_j \;=\; & [-(D\,\Delta\,\Phi_{j+1}+\Delta\,\Phi_j)\,\Delta\,\Phi_j+2\,D\,\Phi_{j+1}\,D\,\Delta\,\Phi_{j+1}] \\ & -[-(D\,\Delta\,\Phi_j+\Delta\,\Phi_{j-1})\,\Delta\,\Phi_{j-1}+2\,D\,\Phi_j\,D\,\Delta\,\Phi_j]\;. \end{aligned} \tag{55}$$

We expand the left hand side of (55) as follows:

$$\begin{aligned} 2\,D_0\,\Phi_j\,\Delta^2\,\Phi_j &= (\Phi_{j+1}-\Phi_{j-1})\,(\Phi_{j+2}-4\,\Phi_{j+1}+6\,\Phi_j-4\,\Phi_{j-1}+\Phi_{j-2}) \\ &= 6\,\Phi_j\,(\Phi_{j+1}-\Phi_{j-1})-4\,(\Phi_{j+1}^2-\Phi_{j-1}^2)+(\Phi_{j+1}-\Phi_{j-1})\,(\Phi_{j+2}+\Phi_{j-2}) \\ &= \big[6\,\Phi_j\,\Phi_{j+1}-4\,\Phi_j^2-4\,\Phi_{j+1}^2\big]-\big[6\,\Phi_{j-1}\,\Phi_j-4\,\Phi_{j-1}^2-4\,\Phi_j^2\big] \\ &\quad +(\Phi_{j+1}-\Phi_{j-1})\,(\Phi_{j+2}+\Phi_{j-2}) \\ &= T_j-T_{j-1}\,, \end{aligned}$$

with

$$\begin{aligned} T_j &:= 7\,\Phi_j\,\Phi_{j+1} - 4\,\Phi_j^2 - 4\,\Phi_{j+1}^2 + \Phi_{j+1}\,\Phi_{j+2} - \Phi_{j-1}\,\Phi_{j+2} + \Phi_{j-1}\,\Phi_j \\ &= -\,4\,(D\,\Phi_{j+1})^2 + (\Phi_{j+2} - \Phi_j)\,(\Phi_{j+1} - \Phi_{j-1}) \\ &= -\,4\,(D\,\Phi_{j+1})^2 + (\Delta\,\Phi_{j+1} + 2\,D\,\Phi_{j+1})\,(2\,D\,\Phi_{j+1} - \Delta\,\Phi_j) \\ &= -\,\Delta\,\Phi_j\,\Delta\,\Phi_{j+1} + 2\,D\,\Phi_{j+1}\,D\,\Delta\,\Phi_{j+1} \;=\; -(D\,\Delta\,\Phi_{j+1} + \Delta\,\Phi_j)\,\Delta\,\Phi_j + 2\,D\,\Phi_{j+1}\,D\,\Delta\,\Phi_{j+1}\,. \end{aligned}$$

This completes the proof of (55).

It only remains to prove the formula (35), that is:

$$\Phi_j\,\Delta^2\,\Phi_j \;=\; (\Delta\,\Phi_j)^2 + [\Phi_j\,D\,\Delta\,\Phi_{j+1} - D\,\Phi_{j+1}\,\Delta\,\Phi_j] - [\Phi_{j-1}\,D\,\Delta\,\Phi_j - D\,\Phi_j\,\Delta\,\Phi_{j-1}]\;, \tag{56}$$

which is the discrete counterpart of the relation $u\,u'''' = (u'')^2 + (u\,u''' - u'\,u'')'$. We compute:

$$\Phi_j\,\Delta^2\,\Phi_j - (\Delta\,\Phi_j)^2 \;=\; \begin{pmatrix} \Phi_{j+2} \\ \Phi_{j+1} \\ \Phi_j \\ \Phi_{j-1} \\ \Phi_{j-2} \end{pmatrix}^T \begin{pmatrix} 0 & 0 & 1/2 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 \\ 1/2 & 0 & 2 & 0 & 1/2 \\ 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \end{pmatrix} \begin{pmatrix} \Phi_{j+2} \\ \Phi_{j+1} \\ \Phi_j \\ \Phi_{j-1} \\ \Phi_{j-2} \end{pmatrix}\;,$$

and decompose the corresponding symetric matrix in a telescopic way:

$$\begin{pmatrix} 0 & 0 & 1/2 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 \\ 1/2 & 0 & 2 & 0 & 1/2 \\ 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1/2 & 0 & 0 \\ 0 & -1 & 0 & -1/2 & 0 \\ 1/2 & 0 & 1 & 0 & 0 \\ 0 & -1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & -1 & 0 & -1/2 \\ 0 & 1/2 & 0 & 1 & 0 \\ 0 & 0 & -1/2 & 0 & 0 \end{pmatrix}\;.$$

We have thus obtained the telescopic decomposition:

$$\begin{aligned} \Phi_j\,\Delta^2\,\Phi_j - (\Delta\,\Phi_j)^2 \;=\; &\big[\Phi_j\,\Phi_{j+2} - (\Phi_{j+1})^2 + (\Phi_j)^2 - \Phi_{j-1}\,\Phi_{j+1}\big] \\ &- \big[\Phi_{j-1}\,\Phi_{j+1} - (\Phi_j)^2 + (\Phi_{j-1})^2 - \Phi_{j-2}\,\Phi_j\big]\;, \end{aligned}$$

and (56) eventually follows from rewriting the telescopic quantity in terms of $\Phi_j$, $D\,\Phi_{j+1}$, $\Delta\,\Phi_j$ and $D\,\Delta\,\Phi_{j+1}$:

$$\Phi_j\,\Phi_{j+2} - (\Phi_{j+1})^2 + (\Phi_j)^2 - \Phi_{j-1}\,\Phi_{j+1} \;=\; \Phi_j\,(\Phi_{j+2} - 3\,\Phi_{j+1} + 3\,\Phi_j - \Phi_{j-1}) - (\Phi_{j+1} - \Phi_j)\,(\Phi_{j+1} - 2\,\Phi_j + \Phi_{j-1})\,.$$

## References


[Ben14] A. Benoit. Geometric optics expansions for linear hyperbolic boundary value problems and optimality of energy estimates for surface waves. *Differential Integral Equations*, 27(5-6):531–562, 2014.

[BGRSZ02] S. Benzoni-Gavage, F. Rousset, D. Serre, and K. Zumbrun. Generic types and transitions in hyperbolic initial-boundary-value problems. *Proc. Roy. Soc. Edinburgh Sect. A*, 132(5):1073–1104, 2002.

[CFL28] R. Courant, K. Friedrichs, and H. Lewy. Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Ann.*, 100(1):32–74, 1928.

[CG11] J.-F. Coulombel and A. Gloria. Semigroup stability of finite difference schemes for multidimensional hyperbolic initial boundary value problems. *Math. Comp.*, 80(273):165–203, 2011.

[CL20] J.-F. Coulombel and F. Lagoutière. The neumann numerical boundary condition for transport equations. *Kinet. Relat. Models*, 13(1):1–32, 2020.

[Cou13] J.-F. Coulombel. Stability of finite difference schemes for hyperbolic initial boundary value problems. In *HCDTE Lecture Notes. Part I. Nonlinear Hyperbolic PDEs, Dispersive and Transport Equations*, pages 97–225. American Institute of Mathematical Sciences, 2013.

[GKO95] B. Gustafsson, H.-O. Kreiss, and J. Oliger. *Time dependent problems and difference methods*. John Wiley & Sons, 1995.

[GKS72] B. Gustafsson, H.-O. Kreiss, and A. Sundström. Stability theory of difference approximations for mixed initial boundary value problems. II. *Math. Comp.*, 26(119):649–686, 1972.

[Gol77] M. Goldberg. On a boundary extrapolation theorem by Kreiss. *Math. Comp.*, 31(138):469–477, 1977.

[GT78] M. Goldberg and E. Tadmor. Scheme-independent stability criteria for difference approximations of hyperbolic initial-boundary value problems. I. *Math. Comp.*, 32(144):1097–1107, 1978.

[GT81] M. Goldberg and E. Tadmor. Scheme-independent stability criteria for difference approximations of hyperbolic initial-boundary value problems. II. *Math. Comp.*, 36(154):603–626, 1981.

[Gus75] B. Gustafsson. The convergence rate for difference approximations to mixed initial boundary value problems. *Math. Comp.*, 29(130):396–406, 1975.

[Kre66] H.-O. Kreiss. Difference approximations for hyperbolic differential equations. In *Numerical Solution of Partial Differential Equations (Proc. Sympos. Univ. Maryland, 1965)*, pages 51–58. Academic Press, 1966.

[Kre68] H.-O. Kreiss. Stability theory for difference approximations of mixed initial boundary value problems. I. *Math. Comp.*, 22:703–714, 1968.

[RM94] R. D. Richtmyer and K. W. Morton. *Difference methods for initial-value problems*. Robert E. Krieger Publishing Co., second edition, 1994.

[Str62] G. Strang. Trigonometric polynomials and difference methods of maximum accuracy. *J. Math. Phys.*, 41:147–154, 1962.

[Wu95] L. Wu. The semigroup stability of the difference approximations for initial-boundary value problems. *Math. Comp.*, 64(209):71–88, 1995.