High order numerical schemes for transport equations on bounded domains

Benjamin Boutin, T.H.T. Nguyen, Abraham Sylla, Sébastien Tran-Tien, Jean-François Coulombel

High order numerical schemes for transport equations on bounded domains

Benjamin Boutin* & Thi Hoai Thuong Nguyen† & Abraham Sylla‡
& Sébastien Tran-Tien§ & Jean-François Coulombel¶

December 5, 2019

Abstract

This article is an account of the NABUCO project achieved during the summer camp CEMRACS 2019 devoted to geophysical fluids and gravity flows. The goal is to construct finite difference approximations of the transport equation with nonzero incoming boundary data that achieve the best possible convergence rate in the maximum norm. We construct, implement and analyze the so-called inverse Lax-Wendroff procedure at the incoming boundary. Optimal convergence rates are obtained by combining sharp stability estimates for extrapolation boundary conditions with numerical boundary layer expansions. We illustrate the results with the Lax-Wendroff and $O3$ schemes.

1 Introduction

1.1 Context

The goal of this article is to propose a high order numerical treatment of nonzero incoming boundary data for the advection equation. The methodology is developed here for the one-dimensional problem but it is our hope that the tools used below will be useful for higher dimensional problems. We are thus given a fixed constant velocity $a > 0$, an interval length $L > 0$ and we consider the (continuous) problem:

$$\begin{cases} \partial_t u + a\,\partial_x u \,=\, 0\,, & t \geq 0\,, \quad x \in (0,L)\,, \\ u(0,x) \,=\, f(x)\,, & x \in (0,L)\,, \\ u(t,0) \,=\, g(t)\,, & t \geq 0\,. \end{cases} \tag{1.1}$$

*IRMAR (UMR CNRS 6625), Université de Rennes, Campus de Beaulieu, 35042 Rennes Cedex, France ; benjamin.boutin@univ-rennes1.fr

†IRMAR (UMR CNRS 6625), Université de Rennes, Campus de Beaulieu, 35042 Rennes Cedex, France ; thi-hoai-thuong.nguyen@univ-rennes1.fr

‡Institut Denis Poisson CNRS UMR 7013, Université de Tours, Faculté des Sciences et Techniques, Bâtiment E2, Parc de Grandmont, 37200 Tours, France ; Abraham.Sylla@lmpt.univ-tours.fr

§École Normale Supérieure de Lyon, 15 parvis René Descartes, BP 7000, 69342 Lyon Cedex 07, France ; sebastien.tran-tien@ens-lyon.fr

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

The requirements on the initial and boundary data, namely f and g, will be made precise below. The solution to (1.1) is given by the method of characteristics, which yields the formula:

$$\forall\,(t,x)\in\mathbb{R}^+\times(0,L)\,,\quad u(t,x)\,=\,\begin{cases} f\,(x-a\,t)\,, & \text{if } x\geq a\,t\,,\\ g\left(t-\dfrac{x}{a}\right), & \text{if } x\leq a\,t\,. \end{cases} \tag{1.2}$$

The question we address is how to construct high order numerical approximations of the solution (1.2) to (1.1) by means of (explicit) finite difference approximations. This problem has been addressed in [CL20] in the case of *zero* incoming boundary data (that is, $g=0$ in (1.1)). The focus in [CL20] is on the outflow boundary ($x=L$ here since a is positive), for which *extrapolation* numerical boundary conditions are analyzed. Fortunately for us, a large part of the analysis in [CL20] can be used here as a black box and we therefore focus on the incoming boundary. To motivate the analysis of this paper, let us present a very simple -though illuminating- example for which we just need to introduce the basic notations that will be used throughout this article. In all what follows, we consider a positive integer J, that is meant to be large, and define the space step Δx and the grid points $(x_j)_{j\in\mathbb{Z}}$ by

$$\Delta x\,:=\,\frac{L}{J}\,,\quad x_j\,:=\,j\,\Delta x\quad(j\in\mathbb{Z})\,.$$

The interval $(0,L)$ corresponds to the cells (x_{j-1},x_j) with $j=1,\dots,J$, but considering the whole real line $\{j\in\mathbb{Z}\}$ will be useful in some parts of the analysis. The time step Δt is then defined as $\Delta t:=\lambda\,\Delta x$, where $\lambda>0$ is a constant that is fixed so that assumption 1.1 below is satisfied. We use from now on the notation $t^n:=n\,\Delta t$, $n\in\mathbb{N}$; the quantity u_j^n will play the role of an approximation for the solution u to (1.1) at the time t^n on the cell (x_{j-1},x_j).

We now examine an example where the exact solution to (1.1) is approximated by means of the Lax-Wendroff scheme. The approximation reads:

$$u_j^{n+1}\,=\,u_j^n-\frac{\lambda\,a}{2}\,(u_{j+1}^n-u_{j-1}^n)+\frac{(\lambda\,a)^2}{2}\,(u_{j+1}^n-2\,u_j^n+u_{j-1}^n)\,,\quad n\in\mathbb{N}\,,\quad j=1,\dots,J\,, \tag{1.3}$$

where we recall that $\lambda=\Delta t/\Delta x$ is a fixed constant and $a>0$ is the transport velocity in (1.1). The initial condition for (1.3) is defined, for instance, by computing the cell averages of the initial condition f in (1.1), namely:

$$\forall\,j=1,\dots,J\,,\quad u_j^0\,:=\,\frac{1}{\Delta x}\int_{x_{j-1}}^{x_j}f(x)\,\mathrm{d}x\,. \tag{1.4}$$

Without any boundary, the Lax-Wendroff scheme is a second order approximation to the transport equation [GKO95]. We would like, of course, to maintain the second order accuracy property when implementing (1.3) on an interval. This implementation, however, requires, at each time iteration n, the definition of the boundary (or *ghost cell*) values u_0^n and u_{J+1}^n. At the outflow boundary, we prescribe an extrapolation condition [Kre66, Gol77], the significance of which will be thoroughly justified in the next sections:

$$u_{J+1}^n\,=\,2\,u_J^n-u_{J-1}^n\,,\quad n\in\mathbb{N}\,. \tag{1.5}$$

Combining (1.3) with (1.5), the last interior cell value u_J^n obeys the induction formula:

$$u_J^{n+1}\,=\,u_J^n-\lambda\,a\,(u_J^n-u_{J-1}^n)\,,\quad n\in\mathbb{N}\,,$$

which is nothing but the upwind scheme. It then only remains to determine the inflow numerical boundary condition u_0^n. Since we wish to approximate the exact solution to (1.1) and u_0^n is meant, at least, to

approximate the trace $u(t^n, 0)$, it seems reasonable at first sight to prescribe the Dirichlet boundary condition:

$$u_0^n \,=\, g(t^n)\,, \quad n \in \mathbb{N}\,. \tag{1.6}$$

In the case of *zero* incoming boundary data ($g = 0$), and for any sufficiently smooth initial condition f that is "flat" at the incoming boundary, the main result of [CL20] shows that the above numerical scheme (1.3), (1.4), (1.5), (1.6) converges towards the exact solution to (1.1) with a rate of convergence 3/2 in the maximum norm. Numerical simulations even predict that the rate of convergence should be 2, or at least close to 2, for smooth initial data. However, implementing the above numerical scheme[1] quickly shows that the rate of convergence falls down to 1 when g is nonzero and satisfies the compatibility conditions[2] described hereafter with the initial condition f.

Our goal is to provide with a thorough treatment of nonzero incoming boundary data and to design numerical boundary conditions that recover the optimal rate of convergence in the maximum norm (at least, the same rate of convergence as the one in [CL20] for zero boundary data). The strategy is not new and is now referred to as the *inverse Lax-Wendroff method*. It consists, as detailed below, in writing Taylor expansions with respect to the space variable x close to the incoming boundary and then using the advection equation (1.1) to substitute the normal derivatives $\partial_x^m u(t, 0)$ for tangential derivatives $\partial_t^m u(t, 0)$, the latter being computed thanks to the boundary condition in (1.1). This strategy is available when the boundary is non-characteristic [BGS07].

The inverse Lax-Wendroff method is a general strategy that has been followed in various directions. We refer for instance to [TS10, ST17, FY13, VS15, DDJ18] for various implementations related to either hyperbolic or kinetic partial differential equations. In these works, most of the time, the incoming numerical boundary condition prescribes the ghost cell value u_0^n in terms of the boundary datum g but also of *interior cell* values u_j^n with $j \geq 1$. This is the reason why *stability* is a real issue in these works, see for instance the discussion in [VS15, Section 4], and many rigorous justifications are still open. We develop here a simplified version of some of those previously proposed boundary treatments, but we rigorously justify the convergence with an (almost) optimal rate of convergence. As in [CL20], the key ingredient in our analysis is an *unconditional* stability result for the Dirichlet boundary conditions which dates back to [GT78, GT81], see an alternative proof in [CG11].

1.2 The inverse Lax-Wendroff method

We first fix from now on some notations. In all this article, we are given some fixed integers $p, r \in \mathbb{N}$ and consider an explicit two time step approximation for the solution to (1.1):

$$u_j^{n+1} \,=\, \sum_{\ell=-r}^{p} a_\ell \, u_{j+\ell}^n\,, \quad n \in \mathbb{N}\,, \quad j = 1, \dots, J\,. \tag{1.7}$$

In (1.7), the numbers $a_{-r}, \dots, a_p$ are defined in terms of the parameter λ and of the velocity a (see, for instance, (1.3) for which $p = r = 1$). These numbers are fixed, which means that (1.7) is linear with respect to (u_j^n). For simplicity, we follow [CL20] and choose as initial data for (1.7) the cell averages of the initial condition f in (1.1). This means that the vector $(u_1^0, \dots, u_J^0)$ is defined by (1.4). For (1.7) to define inductively (with respect to n) the vector $(u_1^n, \dots, u_J^n)$, we need to prescribe the ghost cell values $u_{1-r}^n, \dots, u_0^n$ and $u_{J+1}^n, \dots, u_{J+p}^n$. As explained above, we focus here on the inflow boundary and we

[1] One can choose for instance $a = 1$, $\lambda = 5/6$, $L = 6$, $f(x) = \sin(x)$, $g(t) = -\sin(t)$ and increase the integer J geometrically.

[2] The rate of convergence could be even smaller than 1 when the compatibility conditions are not satisfied but that would just reflect the fact that the exact solution (1.2) is not smooth (for instance, not even continuous if $f(0) \neq g(0)$).

therefore follow the extrapolation boundary treatment of [CL20] for the outflow boundary. Namely, if we define the finite difference operator D_- as:

$$(D_- v)_j \,:=\, v_j - v_{j-1}\,,$$

and its iterates D_-^m accordingly, we choose from now on an extrapolation order $k_b \in \mathbb{N}$ for the outflow boundary and prescribe:

$$(D_-^{k_b} u^n)_{J+\ell} \,=\, 0\,, \quad n \in \mathbb{N}\,, \quad \ell = 1, \dots, p\,. \tag{1.8}$$

The example (1.5) corresponds to $k_b = 2$ (recall $p = 1$ for the Lax-Wendroff scheme so there is only one ghost cell). It now remains to prescribe the inflow values $u^n_{1-r}, \dots, u^n_0$. Unlike some previous works, we are going to prescribe Dirichlet boundary conditions, meaning for instance that the value u^n_0 will be determined in terms of the boundary datum g only. Let us assume for a while that u^n_j is a second order approximation of $u(t^n, (x_{j-1} + x_j)/2)$ where u is the exact solution (1.2) of the continuous problem (1.1). Then we *formally* have:

$$u^n_0 \approx u\left(t^n, -\frac{\Delta x}{2}\right) \approx u(t^n, 0) - \frac{\Delta x}{2}\,\partial_x u(t^n, 0)\,,$$

where $\approx$ means "equal up to $O(\Delta x^2)$", and we then use (1.1) to get:

$$u^n_0 \approx u(t^n, 0) + \frac{\Delta x}{2\,a}\,\partial_t u(t^n, 0) \,=\, g(t^n) + \frac{\Delta x}{2\,a}\,g'(t^n)\,.$$

The last term $\Delta x/(2\,a)\,g'(t^n)$ in the previous equality is precisely the correction that is required to recover the second order accuracy when dealing with the Lax-Wendroff scheme (compare with (1.6)). More generally speaking, we could have pushed further the above Taylor expansion and obtained as a final (formal !) result that u^n_0 should be "close" (whatever that means !) to some quantity of the form:

$$\sum_{\kappa=0}^{K} \frac{\Delta x^\kappa}{\kappa\,!\,a^\kappa}\,\alpha_\kappa\,g^{(\kappa)}(t^n)\,,$$

where K is a truncation order and $\alpha_0, \dots, \alpha_K$ are numerical constants.

The general form of the Dirichlet boundary conditions that we consider below is:

$$u^n_\ell \,=\, \sum_{\kappa=0}^{K} \frac{\Delta x^\kappa}{\kappa\,!\,a^\kappa}\,\alpha_{\kappa,\ell}\,g^{(\kappa)}(t^n)\,, \quad n \in \mathbb{N}\,, \quad \ell = 1 - r, \dots, 0\,,$$

where the $\alpha_{\kappa,\ell}$'s are numerical constants which will play a role (together with the truncation order K) in the consistency analysis. There are two main choices which we discuss in this article. The first one is given in [TS10, VS15]:

$$\alpha_{\kappa,\ell} \,:=\, \left(\frac{1}{2} - \ell\right)^\kappa\,, \quad \kappa \in \mathbb{N}\,, \quad \ell = 1 - r, \dots, 0\,,$$

and is relevant if u^n_j is eventually compared in the convergence analysis with $u(t^n, (x_{j-1} + x_j)/2)$, u being the exact solution (1.2). The other possible choice we advocate is:

$$\alpha_{\kappa,\ell} \,:=\, \frac{(-1)^\kappa}{\kappa + 1}\,\left(\ell^{\kappa+1} - (\ell - 1)^{\kappa+1}\right)\,, \quad \kappa \in \mathbb{N}\,, \quad \ell = 1 - r, \dots, 0\,, \tag{1.9}$$

and is relevant if u^n_j is eventually compared (as in our main result, which is Theorem 1.2 below) in the convergence analysis with the average of $u(t^n, \cdot)$ on the cell (x_{j-1}, x_j). The truncation order K is discussed with our main result in the following paragraph.

1.3 Results

We assume that the approximation (1.7) is consistent with the transport operator and that it defines a stable procedure on $\ell^2(\mathbb{Z})$.

Assumption 1.1 (Consistency and stability without any boundary). *The coefficients $a_{-r}, ..., a_p$ in (1.7) satisfy $a_{-r}\, a_p \neq 0$ (normalization), and for some integer $k \geq 1$, there holds:*

$$\forall\, m = 0, \dots, k\,, \quad \sum_{\ell=-r}^{p} \ell^m\, a_\ell \;=\; (-\lambda\, a)^m\,, \quad \text{(consistency of order } k\text{)}, \tag{1.10}$$

$$\sup_{\theta\in[0,2\pi]} \left| \sum_{\ell=-r}^{p} a_\ell\, e^{i\,\ell\,\theta} \right| \;\leq\; 1\,, \quad (\ell^2\text{-stability on } \mathbb{Z}). \tag{1.11}$$

For the Lax-Wendroff scheme (1.3), we have $p = r = 1$ if $\lambda\, a \neq 1$, the integer k equals 2, and (1.11) holds if and only if $\lambda\, a \leq 1$. In Theorem 1.2 below and all what follows, the velocity $a > 0$, the length $L > 0$, the parameter $\lambda = \Delta t/\Delta x$ and the extrapolation order $k_b \in \mathbb{N}$ at the outflow boundary are given. Subsequent constants may depend on them. The integer $k \geq 1$ is also fixed such that assumption 1.1 holds. We consider the initial condition (1.4) and its evolution by the numerical scheme (1.7), (1.8), the inflow ghost cell values being given by:

$$u_\ell^n \;=\; \sum_{\kappa=0}^{k-1} \frac{\Delta x^\kappa}{(\kappa+1)\,!\,(-a)^\kappa} \left(\ell^{\kappa+1} - (\ell-1)^{\kappa+1}\right) g^{(\kappa)}(t^n)\,, \quad n \in \mathbb{N}\,, \quad \ell = 1-r, \dots, 0\,. \tag{1.12}$$

Of course, prescribing (1.12) is meaningful only if g is sufficiently smooth (say, $g \in \mathscr{C}^{k-1}$). One could push further the Taylor expansion in (1.12) and consider higher order correctors but it would require further smoothness on g and it would eventually not improve our convergence result below, so fixing the truncation order $K = k - 1$ seems to be the most convenient choice. Our main convergence result is the extension of the main result in [CL20] to the case of nonzero boundary data.

Theorem 1.2 (Main convergence result). *Under assumption 1.1, there exists a constant $C > 0$ such that for any final time $T \geq 1$, any integer $J \in \mathbb{N}^*$, any data $f \in H^{k+1}((0,L))$ and $g \in H^{k+1}((0,T))$ satisfying the compatibility requirements at $t = x = 0$:*

$$\forall\, m = 0, \dots, k\,, \quad f^{(m)}(0) \;=\; (-a)^{-m}\, g^{(m)}(0)\,,$$

the solution (u_j^n) to (1.4), (1.7), (1.8), (1.12) satisfies:

$$\sup_{0\leq n\leq T/\Delta t} \sup_{1\leq j\leq J} \left| u_j^n - \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} u(t^n, x)\, \mathrm{d}x \right| \leq C\, T\, e^{C\,T/L}\, \Delta x^{\min(k,k_b)-1/2} \left(\|f\|_{H^{k+1}((0,L))} + \|g\|_{H^{k+1}((0,T))} \right), \tag{1.13}$$

with u the exact solution to (1.1), whose expression is given by (1.2).

Actually, the constant C in (1.13) is independent of $L \geq 1$, which is consistent with the convergence result we shall prove below for the half-space problem on $\mathbb{R}^+$ with inflow at $x = 0$. As in [CL20], the loss of $1/2$ in the rate of convergence of Theorem 1.2 looks somehow artificial and is mostly a matter of passing from the $\ell_n^\infty \ell_j^2$ topology to $\ell_{n,j}^\infty$. Our next result examines a situation where the optimal convergence rate $\min(k, k_b)$ can be obtained. In order to simplify (and shorten) the proof of Theorem 1.3, we only examine here the case of a half-space with extrapolation outflow conditions. The extension of the techniques to the case of an interval is left to the interested reader.

Theorem 1.3 (Optimal rate of convergence for the outflow problem)**.** *Under assumption 1.1 and under the additional assumption 3.2 stated hereafter, there exists a constant $C > 0$ such that for any final time $T \geq 1$, any integer $J \in \mathbb{N}^*$, any data $f \in H^{k+1}((-\infty, L))$, the solution to the scheme:*

$$\begin{cases} u_j^0 = \dfrac{1}{\Delta x} \displaystyle\int_{x_{j-1}}^{x_j} f(x)\,\mathrm{d}x\,, & j \leq J\,, \\ (D_-^{k_b} u^n)_{J+\ell} = 0\,, & 0 \leq n \leq T/\Delta t\,, \quad \ell = 1, \dots, p\,, \\ u_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, u_{j+\ell}^n\,, & 0 \leq n \leq T/\Delta t - 1\,, \quad j \leq J\,, \end{cases} \tag{1.14}$$

satisfies the error estimate

$$\sup_{0 \leq n \leq T/\Delta t} \sup_{j \leq J} \left| u_j^n - \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} f(x - a\,t^n)\,\mathrm{d}x \right| \leq C\,T\,\Delta x^{k_b}\,\|f\|_{H^{k+1}((0,L))}\,,$$

as long as $k_b < k$.

In other words, the technical assumption 3.2 hereafter, which is verified on many examples such as the Lax-Wendroff and $O3$ schemes, allows to recover the optimal rate $k_b = \min(k_b, k)$ in the case $k_b < k$. Of course, one would also like to improve the rate $\min(k_b, k) - 1/2$ in the case $k_b = k$, which is clearly the most natural choice. However, in that case, both the interior and boundary consistency errors scale like Δx^k and, in the framework of assumption 1.1, stability in the interior domain is available only in the ℓ_j^2 topology, so it is quite difficult to derive the convergence rate k in the ℓ_j^∞ topology. Theorem 1.3 already indicates that combining the approach of [CL20] with other techniques (here, boundary layer expansions) may improve some results. We hope to deal with the case $k_b = k$ in the future.

2 Convergence analysis for the inverse Lax-Wendroff method

This Section is devoted to the proof of Theorem 1.2. In order to shorten the exposition, we shall use some results of [CL20] as a black box and we refer the interested reader to [CL20] for more details. Following [CL20], we shall prove Theorem 1.2 by using a stability estimate for (1.7), (1.8), (1.12) and a superposition argument, which amounts to considering separately two half-space problems: one in which there is only inflow at $x = 0$, and one for which there is only outflow at $x = L$. The novelty here is the nonzero inflow source term so we first deal with that case.

2.1 Convergence analysis on a half-line for the inflow problem

We focus here on the inflow source term, and therefore start by proving the main convergence estimate that is the new ingredient for the proof of Theorem 1.2.

Theorem 2.1 (Convergence estimate for the inflow problem)**.** *Under assumption 1.1, there exists a constant $C > 0$ such that for any final time $T \geq 1$, for any $J \in \mathbb{N}^*$, for any initial condition $f \in H^{k+1}((0, +\infty))$ and boundary source term $g \in H^{k+1}((0, T))$ satisfying the compatibility conditions:*

$$\forall\, m = 0, \dots, k\,, \quad f^{(m)}(0) = (-a)^{-m}\, g^{(m)}(0)\,, \tag{2.1}$$

the solution $(u_j^n)_{j\geq 1-r, n\in\mathbb{N}}$ *to the numerical scheme:*

$$\begin{cases} u_j^0 \,=\, \dfrac{1}{\Delta x}\displaystyle\int_{x_{j-1}}^{x_j} f(x)\,\mathrm{d}x\,, & j\geq 1\,, \\ u_\ell^n \,=\, \displaystyle\sum_{\kappa=0}^{k-1} \frac{\Delta x^\kappa}{(\kappa+1)\,!\,(-a)^\kappa}\left(\ell^{\kappa+1}-(\ell-1)^{\kappa+1}\right) g^{(\kappa)}(t^n)\,, & 0\leq n\leq T/\Delta t\,, \quad \ell = 1-r,\dots,0\,, \\ u_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, u_{j+\ell}^n\,, & 0\leq n\leq T/\Delta t-1\,, \quad j\geq 1\,, \end{cases} \tag{2.2}$$

satisfies:

$$\sup_{0\leq n\leq T/\Delta t}\left(\sum_{j\geq 1}\Delta x\left(u_j^n-\frac{1}{\Delta x}\int_{x_{j-1}}^{x_j} u(t^n,x)\,\mathrm{d}x\right)^2\right)^{1/2} \leq\, C\,T\,\Delta x^k\left(\|f\|_{H^{k+1}((0,+\infty))}+\|g\|_{H^{k+1}((0,T))}\right),$$

where u *is the exact solution to the half-line transport problem:*

$$\begin{cases} \partial_t u + a\,\partial_x u \,=\, 0\,, & t\in(0,T)\,,\ x\geq 0\,, \\ u(0,x) \,=\, f(x)\,, & x\geq 0\,, \\ u(t,0) \,=\, g(t)\,, & t\in(0,T)\,. \end{cases} \tag{2.3}$$

Proof. For convenience, we first extend g into a function $g_\flat\in H^{k+1}((0,+\infty))$ and then define:

$$\forall\, x\in\mathbb{R}\,, \quad f_\sharp(x) \,:=\, \begin{cases} f(x)\,, & \text{if } x>0, \\ g_\flat(-x/a)\,, & \text{if } x<0. \end{cases}$$

Since f and g satisfy the compatibility conditions (2.1), we have $f_\sharp\in H^{k+1}(\mathbb{R})$, and the exact solution u to (2.3) is given by:

$$\forall\,(t,x)\in[0,T]\times(0,+\infty)\,, \quad u(t,x) \,=\, f_\sharp(x-a\,t)\,.$$

Let us now define:

$$\forall\, j\in\mathbb{Z}\,, \quad \forall\, n\in\mathbb{N}\,, \quad w_j^n \,:=\, \frac{1}{\Delta x}\int_{x_{j-1}}^{x_j} f_\sharp(x-a\,t^n)\,\mathrm{d}x\,,$$

which corresponds to the cell average of the exact solution to (2.3). With $(u_j^n)_{j\geq 1-r, 0\leq n\leq T/\Delta t}$ the solution to the numerical scheme (2.2), we define the error $\varepsilon_j^n := u_j^n - w_j^n$, that is a solution to:

$$\begin{cases} \varepsilon_j^0 \,=\, 0\,, & j\geq 1\,, \\ \varepsilon_\ell^n \,=\, u_\ell^n - w_\ell^n\,, & 0\leq n\leq T/\Delta t\,, \quad \ell = 1-r,\dots,0\,, \\ \varepsilon_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\,\varepsilon_{j+\ell}^n + \Delta t\, e_j^{n+1}\,, & 0\leq n\leq T/\Delta t-1\,, \quad j\geq 1\,. \end{cases} \tag{2.4}$$

The interior consistency error $(e_j^{n+1})_{j\geq 1, 0\leq n\leq T/\Delta t-1}$ is easily estimated by means of the Cauchy-

Schwarz inequality and Fourier analysis:

$$\begin{aligned}\sum_{j\geq 1}\Delta x\,(e_j^{n+1})^2 &= \frac{\Delta x}{\Delta t^2}\sum_{j\geq 1}\left(w_j^{n+1}-\sum_{\ell=-r}^{p}a_\ell\,w_{j+\ell}^n\right)^2\\ &= \frac{1}{\Delta x\,\Delta t^2}\sum_{j\geq 1}\left(\int_{x_{j-1}}^{x_j}\left(f_\sharp(x-a\,t^n-a\,\Delta t)-\sum_{\ell=-r}^{p}a_\ell\,f_\sharp(x-a\,t^n+\ell\,\Delta x)\right)\,\mathrm{d}x\right)^2\\ &\leq \frac{1}{\Delta t^2}\int_{\mathbb{R}}\left(f_\sharp(x-a\,t^n-a\,\Delta t)-\sum_{\ell=-r}^{p}a_\ell\,f_\sharp(x-a\,t^n+\ell\,\Delta x)\right)^2\,\mathrm{d}x\\ &\leq \frac{1}{2\,\pi\,\Delta t^2}\int_{\mathbb{R}}\left|\mathrm{e}^{-i\,a\,\lambda\,\Delta x\,\xi}-\sum_{\ell=-r}^{p}a_\ell\,\mathrm{e}^{i\,\ell\,\Delta x\,\xi}\right|^2\left|\widehat{f_\sharp}(\xi)\right|^2\mathrm{d}\xi\;\leq\;C\,\Delta x^{2\,k}\int_{\mathbb{R}}\xi^{2\,(k+1)}\left|\widehat{f_\sharp}(\xi)\right|^2\mathrm{d}\xi\,,\end{aligned}$$

where the final inequality comes from assumption 1.1 and the fact that the ratio $\Delta t/\Delta x$ is constant. Going back to the definition of $f_\sharp$, we have obtained the bound:

$$\sup_{0\leq n\leq T/\Delta t-1}\left(\sum_{j\geq 1}\Delta x\,(e_j^{n+1})^2\right)^{1/2}\;\leq\;C\,\Delta x^k\left(\|f\|_{H^{k+1}((0,+\infty))}+\|g\|_{H^{k+1}((0,T))}\right),\tag{2.5}$$

for some constant C that is independent of the final time $T\geq 1$ and the data f and g.

We now turn to the boundary errors in (2.4), and wish to estimate the following quantities:

$$\sum_{0\leq n\leq T/\Delta t-1}\Delta t\left(u_\ell^n-w_\ell^n\right)^2,\quad \ell=1-r,\ldots,0\,.$$

Let us consider an integer n such that $0\leq n\leq T/\Delta t-1$. From the definition of w_ℓ^n, $\ell\leq 0$, we have:

$$\begin{aligned}u_\ell^n-w_\ell^n &= \sum_{\kappa=0}^{k-1}\frac{\Delta x^\kappa}{(\kappa+1)\,!\,(-a)^\kappa}\left(\ell^{\kappa+1}-(\ell-1)^{\kappa+1}\right)g^{(\kappa)}(t^n)-\frac{1}{\Delta x}\int_{x_{\ell-1}}^{x_\ell}g_\flat(t^n-x/a)\,\mathrm{d}x\\ &= -\frac{(-a)^{-k}}{\Delta x}\int_{x_{\ell-1}}^{x_\ell}x^k\int_0^1\frac{y^{k-1}}{(k-1)\,!}\,g_\flat^{(k)}\left(t^n-\frac{x\,y}{a}\right)\mathrm{d}y\,\mathrm{d}x\,,\end{aligned}$$

where we have used the Taylor formula[3]. By the Cauchy-Schwarz inequality, we get:

$$\left(u_\ell^n-w_\ell^n\right)^2\;\leq\;\frac{C}{\Delta x}\int_{x_{\ell-1}}^{x_\ell}\int_0^1 x^{2\,k}y^{2\,(k-1)}\,g_\flat^{(k)}\left(t^n-\frac{x\,y}{a}\right)^2\mathrm{d}y\,\mathrm{d}x\,,$$

and we now apply the change of variables $(x,y)\to(x\,y,x)$ to get:

$$\left(u_\ell^n-w_\ell^n\right)^2\;\leq\;\int_{x_{\ell-1}}^{x_\ell}\left(\int_v^0|v|\,|u|^{2\,(k-1)}\,g_\flat^{(k)}\left(t^n-\frac{u}{a}\right)^2\mathrm{d}u\right)\mathrm{d}v\,.$$

[3]This is precisely at this point of the analysis that the definition of the coefficients $\alpha_{\kappa,\ell}$ in the inverse Lax-Wendroff method arises. Our choice in (1.12) is motivated by the fact that we compare the numerical solution with the cell average of the exact solution.

Restricting to $\ell = 1-r, \dots, 0$, we have:

$$\sum_{\ell=1-r}^{0} \left(u_\ell^n - w_\ell^n\right)^2 \le C\,\Delta x^{2\,k-1} \int_0^{r\,\Delta x/a} g_\flat^{(k)}(t^n+\tau)^2\,\mathrm{d}\tau\,.$$

Summing now with respect to n, we end up with the estimate:

$$\sum_{0\le n\le T/\Delta t-1} \Delta t \sum_{\ell=1-r}^{0} (\varepsilon_\ell^n)^2 \le C\,\Delta x^{2\,k} \int_0^{+\infty} g_\flat^{(k)}(t)^2\,\mathrm{d}t \le C\,\Delta x^{2\,k}\,\|g\|^2_{H^k((0,T))}\,, \tag{2.6}$$

for some constant C that is independent of the final time $T \ge 1$ and the data f and g.

We now apply the main stability estimate for the error problem (2.4), for which we refer to the seminal papers [GT78, GT81] and to the more recent works [CG11, CL20]:

$$\sup_{0\le n\le T/\Delta t} \left(\sum_{j\ge 1} \Delta x\,(\varepsilon_j^n)^2\right)^{1/2} \le C \left\{ T \sup_{1\le n\le T/\Delta t} \left(\sum_{j\ge 1} \Delta x\,(e_j^n)^2\right)^{1/2} + \left(\sum_{0\le n\le T/\Delta t-1} \Delta t \sum_{\ell=1-r}^{0} (\varepsilon_\ell^n)^2\right)^{1/2} \right\}.$$

The conclusion of Theorem 2.1 then comes from the combination of the estimates (2.5) and (2.6). □

2.2 Convergence estimate for the outflow problem

We recall the convergence estimate obtained in [CL20] for the complementary half-space problem where extrapolation conditions are prescribed at the outflow boundary. We refer the reader to [CL20] for the details.

Theorem 2.2 (Convergence estimate for the outflow problem [CL20]). *Under assumption 1.1, there exists a constant $C > 0$ such that for any final time $T \ge 1$, for any $J \in \mathbb{N}^*$ and for any initial condition $f \in H^{k+1}((-\infty, L))$, the solution $(u_j^n)_{j\le J+p, 0\le n\le T/\Delta t}$ to the numerical scheme:*

$$\begin{cases} u_j^0 = \dfrac{1}{\Delta x}\displaystyle\int_{x_{j-1}}^{x_j} f(x)\,\mathrm{d}x\,, & j \le J\,, \\ (D_-^{k_b} u^n)_{J+\ell} = 0\,, & 0 \le n \le T/\Delta t\,, \quad \ell = 1, \dots, p\,, \\ u_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, u_{j+\ell}^n\,, & 0 \le n \le T/\Delta t - 1\,, \quad j \le J\,, \end{cases}$$

satisfies:

$$\sup_{0\le n\le T/\Delta t} \left(\sum_{j\le J} \Delta x \left(u_j^n - \frac{1}{\Delta x}\int_{x_{j-1}}^{x_j} f(x - a\,t^n)\,\mathrm{d}x\right)^2\right)^{1/2} \le C\,T\,\Delta x^{\min(k,k_b)}\,\|f\|_{H^{k+1}((-\infty,L))}\,.$$

2.3 Proof of Theorem 1.2

It remains to combine the convergence estimates of Theorems 2.1 and 2.2 to prove Theorem 1.2. We use a slight modification of the superposition argument in [CL20] in order to cope with the nonzero incoming condition, but we basically follow the same lines. Let us consider a final time $T \geq 1$ and some data $f \in H^{k+1}((0,L))$, $g \in H^{k+1}((0,T))$ that satisfy the compatibility conditions stated in Theorem 1.2. We consider some function $\chi \in \mathscr{C}^\infty(\mathbb{R})$ such that $\chi(x) = 0$ if $x \leq 1/3$ and $\chi(x) = 1$ if $x \geq 2/3$. We then decompose the initial condition f as:

$$\forall\, x \in (0,L)\,, \quad f(x) \,=\, (1-\chi(x/L))\, f(x) + \chi(x/L)\, f(x)\,.$$

Since $(1-\chi(\cdot/L))\, f$ vanishes on $(2\,L/3, L)$, we can extend it by zero to the interval $(L, +\infty)$ and thus consider $(1-\chi(\cdot/L))\, f$ as an element of $H^{k+1}((0,+\infty))$. Furthermore, the functions $(1-\chi(\cdot/L))\, f$ and g satisfy the same compatibility conditions as f and g at $t = x = 0$. We can thus apply Theorem 2.1 to the sequence $(v_j^n)_{j\geq 1-r, 0\leq n\leq T/\Delta t}$ that is defined as the solution to the numerical scheme:

$$\begin{cases} v_j^0 \,=\, \dfrac{1}{\Delta x}\displaystyle\int_{x_{j-1}}^{x_j} (1-\chi(x/L))\, f(x)\,\mathrm{d}x\,, & j \geq 1\,, \\ v_\ell^n \,=\, \displaystyle\sum_{\kappa=0}^{k-1} \frac{\Delta x^\kappa}{(\kappa+1)\,!\,(-a)^\kappa} \left(\ell^{\kappa+1} - (\ell-1)^{\kappa+1}\right) g^{(\kappa)}(t^n)\,, & 0 \leq n \leq T/\Delta t\,, \quad \ell = 1-r, \dots, 0\,, \\ v_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, v_{j+\ell}^n\,, & 0 \leq n \leq T/\Delta t - 1\,, \quad j \geq 1\,. \end{cases}$$

We obtain the estimate:

$$\sup_{0\leq n\leq T/\Delta t} \left(\sum_{j\geq 1} \Delta x \left(v_j^n - \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} v(t^n, x)\,\mathrm{d}x \right)^2 \right)^{1/2} \leq C\, T\, \Delta x^k \left(\|f\|_{H^{k+1}((0,L))} + \|g\|_{H^{k+1}((0,T))} \right), \tag{2.7}$$

where v is the exact solution to the transport problem:

$$\begin{cases} \partial_t v + a\, \partial_x v \,=\, 0\,, & t \in (0,T)\,, \quad x \geq 0\,, \\ v(0,x) \,=\, (1-\chi(x/L))\, f(x)\,, & x \geq 0\,, \\ v(t,0) \,=\, g(t)\,, & t \in (0,T)\,. \end{cases}$$

Similarly, we can view $\chi(\cdot/L)\, f$ as an element of $H^{k+1}((-\infty, L))$ that vanishes on $(-\infty, L/3)$. Theorem 2.2 then shows that the solution $(w_j^n)_{j\leq J+p, 0\leq n\leq T/\Delta t}$ to the numerical scheme:

$$\begin{cases} w_j^0 \,=\, \dfrac{1}{\Delta x}\displaystyle\int_{x_{j-1}}^{x_j} \chi(x/L)\, f(x)\,\mathrm{d}x\,, & j \leq J\,, \\ (D_-^{k_b} w^n)_{J+\ell} \,=\, 0\,, & 0 \leq n \leq T/\Delta t\,, \quad \ell = 1, \dots, p\,, \\ w_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, w_{j+\ell}^n\,, & 0 \leq n \leq T/\Delta t - 1\,, \quad j \leq J\,, \end{cases}$$

satisfies:

$$\sup_{0\le n\le T/\Delta t}\left(\sum_{j\le J}\Delta x\left(w_j^n-\frac{1}{\Delta x}\int_{x_{j-1}}^{x_j}\chi((x-a\,t^n)/L)\,f(x-a\,t^n)\,\mathrm{d}x\right)^2\right)^{1/2}$$
$$\le\ C\,T\,\Delta x^{\min(k,k_b)}\,\|f\|_{H^{k+1}((0,L))}\,. \quad (2.8)$$

Using the support property of the function χ and the fact that the scheme (1.7) is explicit with a finite stencil, we find that for all time iteration n up to the threshold:

$$N\ :=\ \min\left(\mathrm{E}\left(\frac{J/3-k_b}{r}\right)\,,\,\mathrm{E}\left(\frac{\mathrm{E}(J/3)}{p}\right)\right)\,,$$

there holds:

$$w_{1-r}^n\ =\ \cdots\ =\ w_0^n\ =\ 0\,,\quad v_{J+1-k_b}^n\ =\ \cdots\ =\ v_{J+p}^n\ =\ 0\,.$$

In particular, the solution $(u_j^n)_{1-r\le j\le J+p,0\le n\le T/\Delta t}$ to (1.7), (1.4), (1.8), (1.12) satisfies:

$$\forall\, n\ =\ 0,\dots,N\,,\quad \forall\, j\ =1-r,\dots,J+p\,,\quad u_j^n\ =\ v_j^n+w_j^n\,.$$

Combining then the error estimates (2.7) and (2.8), we obtain:

$$\sup_{0\le n\le N}\left(\sum_{1\le j\le J}\Delta x\left(u_j^n-\frac{1}{\Delta x}\int_{x_{j-1}}^{x_j}u(t^n,x)\,\mathrm{d}x\right)^2\right)^{1/2}$$
$$\le\ C_1\,T\,\Delta x^{\min(k,k_b)}\left(\|f\|_{H^{k+1}((0,L))}+\|g\|_{H^{k+1}((0,T))}\right), \quad (2.9)$$

where u is the exact solution to (1.1).

It remains, as in [CL20], to iterate in time the error estimate (2.9). We follow again the argument in [CL20]. For any time iteration n between N and $2\,N$, we split the solution $(u_j^n)_{1-r\le j\le J+p,0\le n\le T/\Delta t}$ to (1.7), (1.4), (1.8), (1.12) as the sum of the solution to the problem:

$$\begin{cases}\tilde u_j^N\ =\ \dfrac{1}{\Delta x}\displaystyle\int_{x_{j-1}}^{x_j}u(t^N,x)\,\mathrm{d}x\,, & 1\le j\le J\,,\\ (D_-^{k_b}\tilde u^{N+n})_{J+\ell}\ =\ 0\,, & 0\le n\le N\,,\quad \ell=1,\dots,p\,,\\ \tilde u_\ell^{N+n}\ =\displaystyle\sum_{\kappa=0}^{k-1}\frac{\Delta x^\kappa}{(\kappa+1)\,!\,(-a)^\kappa}\left(\ell^{\kappa+1}-(\ell-1)^{\kappa+1}\right)g^{(\kappa)}(t^{N+n})\,, & 0\le n\le N\,,\quad \ell=1-r,\dots,0\,,\\ \tilde u_j^{N+n+1}\ =\displaystyle\sum_{\ell=-r}^{p}a_\ell\,\tilde u_{j+\ell}^{N+n}\,, & 0\le n\le N-1\,,\quad 1\le j\le J\,,\end{cases}$$

and of the (presumably small) solution to the 'error' problem:

$$\begin{cases}\varepsilon_j^N\ =\ u_j^N-\dfrac{1}{\Delta x}\displaystyle\int_{x_{j-1}}^{x_j}u(t^N,x)\,\mathrm{d}x\,, & 1\le j\le J\,,\\ (D_-^{k_b}\varepsilon^{N+n})_{J+\ell}\ =\ 0\,, & 0\le n\le N\,,\quad \ell=1,\dots,p\,,\\ \varepsilon_\ell^{N+n}\ =\ 0\,, & 0\le n\le N\,,\quad \ell=1-r,\dots,0\,,\\ \varepsilon_j^{N+n+1}\ =\displaystyle\sum_{\ell=-r}^{p}a_\ell\,\varepsilon_{j+\ell}^{N+n}\,, & 0\le n\le N-1\,,\quad 1\le j\le J\,.\end{cases}$$

Since the initial condition $u(\cdot - a\,t^N)$ and the boundary source term $g(t^N + \cdot)$ satisfy the compatibility conditions at the corner $t = x = 0$, we can apply the first step of the proof (leading to the error estimate (2.9)) for the $(\tilde{u}_j^{N+n})$ part, and we apply the stability estimate of [CL20, Proposition 4.1] for the (ε_j^{N+n}) part. This leads to the second error estimate:

$$\sup_{N\le n\le 2\,N}\left(\sum_{1\le j\le J}\Delta x\left(u_j^n-\frac{1}{\Delta x}\int_{x_{j-1}}^{x_j}u(t^n,x)\,\mathrm{d}x\right)^2\right)^{1/2}$$
$$\le\ C_1\,(1+C_2)\,T\,\Delta x^{\min(k,k_b)}\left(\|f\|_{H^{k+1}((0,L))}+\|g\|_{H^{k+1}((0,T))}\right),$$

and, more generally, to:

$$\sup_{\mu\,N\le n\le(\mu+1)\,N}\left(\sum_{1\le j\le J}\Delta x\left(u_j^n-\frac{1}{\Delta x}\int_{x_{j-1}}^{x_j}u(t^n,x)\,\mathrm{d}x\right)^2\right)^{1/2}$$
$$\le\ C_1\left(\sum_{\nu=0}^{\mu}C_2^\nu\right)T\,\Delta x^{\min(k,k_b)}\left(\|f\|_{H^{k+1}((0,L))}+\|g\|_{H^{k+1}((0,T))}\right).$$

The end of the proof is the same as in [CL20] and we refer the interested reader to that reference for the details.

3 High order outflow boundary layer analysis

In the present section, we explain how the analysis of [BC17], which dealt with the case of the Dirichlet boundary condition at the outflow boundary, can be extended to the case of high order extrapolation (1.8). The goal is to obtain an accurate description of the numerical solution close to the outflow boundary by means of a boundary layer expansion. The leading order term in the expansion corresponds to the exact solution to the transport equation. However, this leading order term does not satisfy the extrapolation condition (1.8), leading to a consistency error of magnitude $O(\Delta x^{k_b})$ on the boundary. Under some mild structural assumption on the numerical scheme (1.7), we show below that this $O(\Delta x^{k_b})$ error on the boundary gives rise to a boundary layer term which scales as $O(\Delta x^{k_b+1/2})$ in the ℓ_j^2 norm. This gain of a factor $\Delta x^{1/2}$ enables us to recover the optimal convergence rate k_b in the maximum norm on the whole spatial domain for $k_b < k$.

3.1 An introductive example

Let us go back for a while to the case of the Lax-Wendroff scheme (1.3), which we consider here on the left half space:

$$u_j^{n+1}\ =\ u_j^n-\frac{\lambda\,a}{2}\,(u_{j+1}^n-u_{j-1}^n)+\frac{(\lambda\,a)^2}{2}\,(u_{j+1}^n-2\,u_j^n+u_{j-1}^n),\quad n\in\mathbb{N},\quad j\le J.$$

At the outflow boundary, we impose the first order extrapolation condition (which corresponds to $k_b = 1$ while $k = 2$ for the Lax-Wendroff scheme):

$$u_{J+1}^n\ =\ u_J^n,\quad n\in\mathbb{N}.$$

We start with some smooth initial condition f defined on $(-\infty, L)$ which we project as a piecewise constant function:

$$u_j^0 := \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} f(x)\,\mathrm{d}x\,, \quad j \le J\,.$$

The exact solution to the transport equation on $(-\infty, L)$ with initial condition f is $u(t,x) = f(x - a\,t)$ (recall $a > 0$). Hence the consistency analysis of the Lax-Wendroff scheme indicates that u_j^n reads:

$$u_j^n \;=\; \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} f(x - a\,t^n)\,\mathrm{d}x + \text{error}\,, \tag{3.1}$$

where the first term in the expansion on the right hand side yields an $O(\Delta x^2)$ consistency error in the interior domain, but also an $O(\Delta x)$ consistency error on the boundary. If we wish to push forward the above expansion, we need to take into account the boundary consistency error and introduce a corrector which will hopefully not alter the interior consistency error. This can be achieved by observing that the sequence:

$$v_j := \kappa^j\,, \quad j \in \mathbb{Z}\,, \qquad \kappa := -\frac{1 + \lambda\,a}{1 - \lambda\,a}\,,$$

is kept unchanged by the Lax-Wendroff scheme on $\mathbb{Z}$, and belongs to $\ell^2(-\infty, J)$ (we assume $0 < \lambda\,a < 1$ so $|\kappa| > 1$). Hence, to remove the boundary consistency error, we can add a corrector on the right hand side of (3.1) in the following way:

$$u_j^n \;=\; \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} f(x - a\,t^n)\,\mathrm{d}x + \Delta x\,w^n\,v_{j-J} + \text{error}\,, \tag{3.2}$$

where w^n is defined in such a way that the two first terms on the right hand side satisfy the first order extrapolation condition, that is:

$$w^n := -\frac{1}{\Delta x\,(\kappa - 1)} \int_{x_{J-1}}^{x_J} \frac{f(x + \Delta x - a\,t^n) - f(x - a\,t^n)}{\Delta x}\,\mathrm{d}x\,, \quad n \in \mathbb{N}\,.$$

If f is sufficiently smooth, then w^n is $O(1)$ and the first corrector on the right hand side of (3.2) is $O(\Delta x)$ in ℓ^∞_j. Note however that the ℓ^2 norm (in space) of this boundary layer corrector scales as $\Delta x^{3/2}$ since the sequence (κ^j) is square integrable on $(-\infty, 0)$. Another important observation at this point is that defining w^n requires the real number κ not to equal 1. This fact follows here from a mere verification but it is a general consequence of the analysis in [Gol77] of the Lopatinskii determinant associated with the boundary condition (1.8) (see also the proof of Lemma 3.5 below).

At this stage, the error analysis amounts to studying the system satisfied by the sequence:

$$\left(u_j^n - \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} f(x - a\,t^n)\,\mathrm{d}x - \Delta x\,w^n\,\kappa^{j-J} \right)_{j \le J+1, n \in \mathbb{N}}\,,$$

the main point being that there is no boundary forcing term, and since the added boundary layer corrector is $O(\Delta x^{3/2})$ in ℓ^2, the initial condition and interior consistency errors will be $O(\Delta x^{3/2})$. Overall, the stability estimate for the Lax-Wendroff scheme with first order extrapolation at the boundary yields the convergence estimate:

$$\sup_{n \le T/\Delta t} \left\| u_j^n - \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} f(x - a\,t^n)\,\mathrm{d}x - \Delta x\,w^n\,v_{j-J} \right\|_{\ell^2(-\infty, J)} \;=\; O(\Delta x^{3/2})\,.$$

By the triangle inequality, we thus obtain:

$$\sup_{n\leq T/\Delta t}\left\| u_j^n - \frac{1}{\Delta x}\int_{x_{j-1}}^{x_j} f(x-a\,t^n)\,\mathrm{d}x\right\|_{\ell^2(-\infty,J)} \;=\; O(\Delta x^{3/2})\,,$$

and this immediately gives the uniform convergence estimate:

$$\sup_{j\leq J, n\leq T/\Delta t}\left| u_j^n - \frac{1}{\Delta x}\int_{x_{j-1}}^{x_j} f(x-a\,t^n)\,\mathrm{d}x\right| \;=\; O(\Delta x)\,.$$

The above brief sketch is made complete and rigorous below in the general framework of Theorem 1.3.

3.2 Discrete steady states

Formalizing somehow the previous example in a more general framework, let us now introduce the following definition.

Definition 3.1 (Steady state for the numerical scheme)**.** *A sequence $(v_j)_{j\in\mathbb{Z}}$ is called a (discrete) steady state for the scheme (1.7) if it is kept unchanged by the time iteration process on $\mathbb{Z}$, that is, if it satisfies:*

$$\forall\, j\in\mathbb{Z}\,,\quad \sum_{\ell=-r}^{p} a_\ell\, v_{j+\ell} \;=\; v_j\,. \tag{3.3}$$

In order to characterize the discrete steady states, it is natural to introduce the characteristic polynomial:

$$A(X) \;:=\; \sum_{\ell=-r}^{p} a_\ell\, X^{\ell+r} - X^r\,. \tag{3.4}$$

From the consistency property (1.10), any constant sequence is a discrete steady state for (1.7), the same property being available for the continuous model (namely, the transport operator). However, the discrete nature of the differentiation operator involved in the numerical scheme (1.7) allows the existence of many other discrete steady states. The latter play an important role when considering the half-space problem with some discrete boundary conditions.

From the non-characteristic assumption $a>0$, it follows that among the roots of A, $X=1$ is always a simple root. Let us now introduce the whole set of (pairwise distinct) roots of A together with their multiplicities through the factorization of A in $\mathbb{C}[X]$:

$$A(X) \;=\; a_p \prod_{\sigma=1}^{\tau} (X-\kappa_\sigma)^{\mu_\sigma}\,. \tag{3.5}$$

Clearly, looking at the degree of the polynomial A, one has the equality

$$\sum_{\sigma=1}^{\tau} \mu_\sigma \;=\; r+p\,.$$

For convenience, we order the roots of A with decreasing modulus:

$$|\kappa_1| \;\geq\; |\kappa_2| \;\geq\; \cdots \;\geq\; |\kappa_\tau|\,.$$

To make the analysis more intelligible, we will work under the following assumption, which was already present in [BC17].

Assumption 3.2. *The characteristic polynomial A defined in (3.4) has a unique root (equal to 1) on the unit circle $\mathbb{S}^1 = \{z \in \mathbb{C}\,,\ |z| = 1\}$. In other words, we assume:*

$$\bigcup_{\sigma=1}^{\tau} \{\kappa_\sigma\} \cap \mathbb{S} \;=\; \{1\}\,. \tag{3.6}$$

As observed on the above example of the Lax-Wendroff scheme, the steady states we are looking at should decrease rapidly as j tends to $-\infty$, so that they provide with a localized correction (near the boundary) to the usual convergence analysis and belong to $\ell^2(-\infty, J)$. We are therefore only concerned with those roots of A that have modulus larger than 1. Lemma 3.3 below gives the precise number of such roots (counted with their multiplicities). We refer to [BC17, Lemma 2.1] for the proof.

Lemma 3.3 (Unstable roots of A [BC17])**.** *Under assumptions 1.1 and 3.2, letting $\kappa_1, \dots, \kappa_{\tau_+}$ be the roots of A that belong to $\mathbb{U} := \{z \in \mathbb{C}\,,\ |z| > 1\}$ with their corresponding multiplicities $\mu_1, \dots, \mu_{\tau_+}$, then one has*

$$\sum_{\sigma=1}^{\tau_+} \mu_\sigma \;=\; p\,. \tag{3.7}$$

A direct consequence of Lemma 3.3 is the following description of steady states for (1.7) that belong to $\ell^2(-\infty, J)$. The proof follows from the standard description of the set of solutions to the recurrence relation (3.3).

Lemma 3.4. *The set of discrete steady states for the scheme (1.7) that belong to $\ell^2(-\infty, J)$ is the finite dimensional linear subspace spanned by the p linearly independent sequences $\rho^{(\sigma,\nu)}$:*

$$\rho_j^{(\sigma,\nu)} \;:=\; (j-J)^\nu\, \kappa_\sigma^{j-J}\,, \quad j \in \mathbb{Z}\,, \quad 1 \le \sigma \le \tau_+\,, \quad 0 \le \nu < \mu_\sigma\,. \tag{3.8}$$

Equivalently, such discrete steady states in $\ell^2(-\infty, J)$ read:

$$v_j \;=\; \sum_{\sigma=1}^{\tau_+} p_\sigma(j)\, \kappa_\sigma^{j-J}\,, \quad j \in \mathbb{Z}\,, \tag{3.9}$$

where $p_\sigma \in \mathbb{C}_{\mu_\sigma - 1}[X]$ for all index $1 \le \sigma \le \tau_+$.

Let us detail the parametrization of the set of (stable) discrete steady states on the two main examples we are concerned with. For the Lax-Wendroff scheme (1.3), one has:

$$A(X) \;=\; -\frac{\lambda\, a\, (1 - \lambda\, a)}{2}\, X^2 + (1 - (\lambda\, a)^2)\, X + \frac{\lambda\, a\, (\lambda\, a + 1)}{2}\,.$$

The (two simple) roots of A are 1 and:

$$\kappa \;:=\; -\frac{1 + \lambda\, a}{1 - \lambda\, a}\,,$$

with $\kappa \in \mathbb{U}$ assuming, as usual, $0 < \lambda\, a < 1$. For the half space problem on $(-\infty, J)$, κ is therefore the unique stable root, and 1 counts as an unstable root (see [BC17]). In particular, assumption 3.2 is satisfied. The set of solutions to (3.3) that belong to $\ell^2(-\infty, J)$ is the one-dimensional subspace spanned by the sequence $(\kappa^{j-J})_{j \in \mathbb{Z}}$.

Let us now consider the so-called $O3$ scheme, which is a convex combination of the Lax-Wendroff and Beam-Warming schemes, see [Str62, Des08]. We now have $p = 1$ and $r = 2$, and the scheme reads:

$$u_j^{n+1} = -\frac{\lambda a\,(1-(\lambda a)^2)}{6}\,u_{j-2}^n + \frac{\lambda a\,(1+\lambda a)\,(2-\lambda a)}{2}\,u_{j-1}^n + \frac{(1-(\lambda a)^2)\,(2-\lambda a)}{2}\,u_j^n - \frac{\lambda a\,(1-\lambda a)\,(2-\lambda a)}{6}\,u_{j+1}^n\,, \quad (3.10)$$

with, again, $0 < \lambda a < 1$. Assumption 1.1 is then satisfied (with $k = 3$). The roots of the corresponding characteristic polynomial A are:

$$\kappa_\pm := \frac{-(1+\lambda a)\,(5-2\lambda a) \pm \sqrt{(1+\lambda a)\,(33-15\lambda a)}}{2\,(1-\lambda a)\,(2-\lambda a)}\,, \quad \kappa_0 := 1\,,$$

each of them being simple. The root κ_- is the only one in $\mathbb{U}$ and κ_+ belongs to the open unit disk $\mathbb{D}$, which is consistent with Lemma 3.4 ($p = 1$). In particular, assumption 3.2 is satisfied.

3.3 The boundary layer expansion. Proof of Theorem 1.3

We now start proving Theorem 1.3, and for that, we consider some initial condition $f \in H^{k+1}((-\infty, L))$ which, for convenience, we extend to the whole real line $\mathbb{R}$ as an element of $H^{k+1}(\mathbb{R})$. Our aim is to compare the solution to the scheme (1.14) (which is set on a half line) with the piecewise constant projection of the exact solution to the transport equation. We thus introduce the notation:

$$\omega_j^n := \frac{1}{\Delta x}\int_{x_{j-1}}^{x_j} f(x - a\,t^n)\,\mathrm{d}x\,, \quad j \le J + p\,, \quad n \in \mathbb{N}\,.$$

The consistency analysis in [CL20] of the scheme (1.14) amounts to considering the numerical scheme satisfied by the error $(u_j^n - \omega_j^n)$. It is proved in [CL20] that the resulting boundary consistency errors have size $O(\Delta x^{k_b})$, while the interior consistency errors have size $O(\Delta x^k)$. Here we have $k_b < k$ so the worst term is on the boundary. Following the arguments in [BC17], we are therefore going to introduce a boundary layer corrector in order to remove the boundary consistency error, up to introducing new initial and interior consistency errors, whose size will be proven to be $O(\Delta x^{k_b+1/2})$ hence the final result of Theorem 1.3. Let us make this argument precise.

The consistent expansion of the numerical solution (u_j^n) takes the form of a corrected version of (ω_j^n), involving now a boundary layer term $(v_j^n) \in \ell^2(-\infty, J)$ as for the above introductive example. The aim is to reduce the magnitude, at the boundary, of the following error:

$$\varepsilon_j^n := \omega_j^n - u_j^n + \Delta x^{k_b}\,v_j^n\,, \quad j \le J + p\,, \quad n \in \mathbb{N}\,. \quad (3.11)$$

The definition of $(v_j^n)_{j\le J+p, n\in\mathbb{N}}$ is chosen so as to correct the consistency error at the boundary. The simplest way to do so consists in chosing $(v_j^n)_{j\le J+p, n\in\mathbb{N}}$ so as to get precisely in the ghost cells the relations $(D_-^{k_b}\varepsilon^n)_{J+\ell} = 0$, $\ell = 1, \dots, p$. From now on, we formulate the problem in such a way to normalize the generating sequences according to the value of J. In view of Lemma 3.4, the problem to be solved writes:

$$v_j^n = \sum_{\sigma=1}^{\tau_+}\sum_{\nu=0}^{\mu_\sigma-1} z_{\sigma,\nu}^n\,\rho_j^{(\sigma,\nu)}\,, \quad j \le J+p\,, \quad n \in \mathbb{N}\,, \quad (3.12)$$

$$(D_-^{k_b} v^n)_{J+\ell} = -\frac{1}{\Delta x^{k_b}}\,(D_-^{k_b}\omega^n)_{J+\ell}\,, \quad \ell = 1, \dots, p\,, \quad n \in \mathbb{N}\,, \quad (3.13)$$

where the sequences $\rho^{(\sigma,\nu)}$ are defined in (3.8). Equivalently to (3.12), we can look for the boundary layer corrector $(v_j^n)_{j\leq J+p,n\in\mathbb{N}}$ under the form

$$v_j^n \;=\; \sum_{\sigma=1}^{\tau_+} p_{n,\sigma}(j-J)\,\kappa_\sigma^{j-J}\,, \tag{3.14}$$

where $p_{n,\sigma}\in\mathbb{C}_{\mu_\sigma-1}[X]$ for all index $1\leq\sigma\leq\tau_+$. The existence of the corrector (v_j^n) is given by the following result. We recall that in the framework of Theorem 1.3, there holds $k_b<k$.

Lemma 3.5. *Consider the initial condition $f\in H^{k+1}((-\infty,L))$ extended to the whole real line $\mathbb{R}$. Then the boundary layer problem (3.12)-(3.13) admits a unique solution $(v_j^n)_{j\leq J+p,n\in\mathbb{N}}$, and this solution satisfies the estimate:*

$$\sup_{n\in\mathbb{N}}\left(\sum_{j\leq J}\Delta x\,(v_j^n)^2\right)^{1/2} \leq C\,\Delta x^{1/2}\,\|f\|_{H^{k_b+1}((-\infty,L))}\,, \tag{3.15}$$

where the constant $C>0$ is independent of $\Delta x>0$, J, L and f.

Proof. Let us fix some integer $n\in\mathbb{N}$. The solution $(u_j^n)_{j\leq J+p}$ to (1.14) solves the homogeneous boundary condition (1.8), thus equivalently to (3.13) one has to find the vector of coordinates $z\in\mathbb{C}^p$ solution to the linear system $A_{k_b}\,z+b=0$ where $b=\Delta x^{-k_b}\,((D_-^{k_b}\omega^n)_{J+\ell})_{1\leq\ell\leq p}$, and the $p\times p$ matrix A_{k_b} is defined as follows:

$$A_{k_b} \;:=\; \begin{pmatrix} (D_-^{k_b}\rho^{(1)})_1 & \dots & (D_-^{k_b}\rho^{(p)})_1 \\ \vdots & & \vdots \\ (D_-^{k_b}\rho^{(1)})_p & \dots & (D_-^{k_b}\rho^{(p)})_p \end{pmatrix},$$

where we have relabeled the sequences $\rho^{(\sigma,\nu)}$, $\sigma=1,\dots,\tau_+$, $\nu=0,\dots,\mu_\sigma-1$ as $\rho^{(1)},\dots,\rho^{(p)}$ in order to make the definition of A_{k_b} easier to read. The latter matrix is somehow the k_bth-order discrete derivative of the so-called confluent Vandermonde matrix. It seems possible to compute the determinant of A_0, see [HJ94], and then to extend this result to higher values of k_b but we prefer to avoid such complicated computations. From the identity of dimensions, we shall just prove that the matrix A_{k_b} is one-to-one, in other words we shall prove that the problem (3.12)-(3.13), or equivalently (3.14)-(3.13), admits a trivial kernel.

Dealing with discrete derivatives of products of polynomial and/or geometric sequences, the divided difference algebra appears as a suitable tool in our analysis. For more details we refer the interested reader to [Ste39, Pop40, dB05]. For consistency in the notation, we recall hereafter the recursive definition of divided differences, but specified for the case of consecutive integer abscissae. Being given a sequence of complex numbers $(w_j)_{j\in\mathbb{Z}}$, one defines:

$$\begin{aligned} &w[j] \;:=\; w_j\,, \quad j\in\mathbb{Z}\,, \\ &w[j-k,\dots,j] \;:=\; \frac{1}{m}\Big(w[j-m+1,\dots,j]-w[j-m,\dots,j-1]\Big)\,, \quad j\in\mathbb{Z}\,,\ m\in\mathbb{N}^\star\,. \end{aligned} \tag{3.16}$$

The quantity $(D_-^{k_b}w)_j$ is directly related to the divided difference $w[j-k_b,\dots,j]$ by the equality:

$$(D_-^{k_b}w)_j \;=\; k_b!\,w[j-k_b,\dots,j]\,, \quad j\in\mathbb{Z}\,. \tag{3.17}$$

Importantly, we may also use the Leibniz formula for divided differences of a product of two sequences:

$$(w\,\tilde{w})[j-k_b,\dots,j] \;=\; \sum_{m=0}^{k_b} w[j-k_b,\dots,j-m]\,\tilde{w}[j-m,\dots,j]\,, \quad j\in\mathbb{Z}\,. \tag{3.18}$$

In terms of the D_- operator, using the relation (3.17), the Leibniz formula (3.18) rewrites under the more recognizable form:

$$(D_-^{k_b}(w\,\tilde{w}))_j \;=\; \sum_{m=0}^{k_b} \binom{k_b}{m} (D_-^{k_b-m}w)_{j-m}\,(D_-^m\tilde{w})_j\,, \quad j\in\mathbb{Z}\,.$$

Let us continue with the representation formula (3.9) of the solution to the boundary layer problem. Looking at the kernel of the linear problem (3.13), we have to find polynomials $(p_\sigma)_{1\le\sigma\le\tau_+}$ with respective degrees less than or equal to $(\mu_\sigma-1)_{1\le\sigma\le\tau_+}$, satisfying the set of equations:

$$\sum_{\sigma=1}^{\tau_+}\sum_{m=0}^{k_b} p_\sigma[\ell-k_b,\dots,\ell-m]\,\kappa_\sigma[\ell-m,\dots,\ell] \;=\; 0\,, \quad 1\le\ell\le p\,,$$

where we denote, with a slight abuse in the notation, κ_σ for the corresponding geometric sequence $(\kappa_m)_{m\in\mathbb{Z}}$, for any $\sigma=1,\dots,\tau_+$. Actually, from the identity (3.17) and by induction on the integer m (or using (3.16)), it is easy to prove that the m-th order divided difference of κ_σ is given by:

$$\kappa_\sigma[\ell-m,\dots,\ell] \;=\; \frac{1}{m!}\,(D_-^m\kappa_\sigma)_\ell \;=\; \frac{1}{m!}\,(1-\kappa_\sigma^{-1})^m\,\kappa_\sigma^\ell\,, \quad 1\le\ell\le p\,.$$

Let us introduce, for any integer σ and any polynomial p_σ with degree less than or equal to $\mu_\sigma-1$, the following polynomial Q_σ also with degree less than or equal to $\mu_\sigma-1$:

$$Q_\sigma(X) \;:=\; \sum_{k=0}^{k_b} (1-\kappa_\sigma^{-1})^k\,p_\sigma[X-k_b,\dots,X-k]\,. \tag{3.19}$$

With these notations, the equations to solve now equivalently read:

$$\sum_{\sigma=1}^{\tau_+} Q_\sigma(\ell)\,\kappa_\sigma^\ell \;=\; 0, \quad 1\le\ell\le p\,. \tag{3.20}$$

Actually, the above set of equations (3.20) exactly corresponds to the generalized Lagrange-Hermite interpolation problem, which is known to be invertible. Thus one necessarily has $Q_\sigma=0$ for any $\sigma=1,\dots,\tau_+$. It then only remains to deduce that any of the polynomials p_σ is also zero.

Observe that for any integer k with $0\le k\le k_b$, from the divided difference algebra, the polynomial $p_\sigma[X-k_b,\dots,X-k]$ has degree less than $\mu_\sigma-(k_b-k)$, see (3.17). Thus the highest degree polynomial involved in the sum (3.19) is $p_\sigma[X-k_b]$ (for $k=k_b$). Since we know that Q_σ is zero, then p_σ is also necessarily zero (consider the highest degree coefficient). The injectivity of the boundary layer problem (3.14)-(3.13) is proved, and the matrix A_{k_b} is therefore invertible.

As a consequence, there exist some uniquely determined coefficients $(\beta_{\sigma,\nu,\ell})$ that depend only on the considered scheme and on the extrapolation order k_b (but neither on the initial condition f nor on the time index n), such that the solution to (3.12)-(3.13) has the form:

$$v_j^n \;=\; \Delta x^{-k_b} \sum_{\sigma=1}^{\tau_+}\sum_{\nu=0}^{\mu_\sigma-1}\sum_{\ell=1}^{p} \beta_{\sigma,\nu,\ell}\,(D_-^{k_b}\omega^n)_{J+\ell}\,\rho_j^{(\sigma,\nu)}\,. \tag{3.21}$$

Using now triangular inequalities, we obtain, for some constant $C > 0$, the upper bound:

$$(v_j^n)^2 \,\leq\, C\,\Delta x^{-2\,k_b} \sum_{\ell=1}^{p} \big((D_-^{k_b}\omega^n)_{J+\ell}\big)^2 \sum_{\sigma=1}^{\tau_+} \sum_{\nu=0}^{\mu_\sigma-1} (\rho_j^{(\sigma,\nu)})^2\,, \quad j \leq J\,.$$

On the one side, we recall the definition (3.8) of the sequences $\rho^{(\sigma,\nu)}$ in Lemma 3.4, hence the estimate:

$$\left(\sum_{j\leq J} \Delta x \sum_{\sigma=1}^{\tau_+} \sum_{\nu=0}^{\mu_\sigma-1} (\rho_j^{(\sigma,\nu)})^2\right)^{1/2} \,\leq\, C\,\sqrt{\Delta x}\,, \tag{3.22}$$

where the constant $C > 0$ is independent of J and Δx. On the other side, from [CL20, Lemma 3.6] and the continuity of the reflection operator from $H^{k_b+1}((-\infty, L))$ to $H^{k_b+1}(\mathbb{R})$, we have the upper bound:

$$\big|(D_-^{k_b}\omega^n)_{J+\ell}\big| \,\leq\, C\,\Delta x^{k_b}\,\|f\|_{H^{k_b+1}((-\infty,L))}\,, \quad \ell = 1,\dots,p\,, \quad n \in \mathbb{N}\,,$$

and thus the required estimate (3.15) follows. □

The interested reader will find in [Gol77] a similar argument to the one developed in the proof of Lemma 3.5. In [Gol77], the analysis of the determinant of the matrix A_{k_b} arises from the verification of the so-called Uniform Kreiss-Lopatinskii Condition (a condition whose significance is based on the work [GKS72]). Let us now prove Theorem 1.3. The error $(\varepsilon_j^n)_{j\leq J+p, n\in\mathbb{N}}$ introduced in (3.11), and fully defined through Lemma 3.5, satisfies the following set of equations[4]:

$$\begin{cases} \varepsilon_j^0 \,=\, \Delta x^{k_b}\, v_j^0\,, & j \leq J\,, \\ (D_-^{k_b}\varepsilon^n)_{J+\ell} \,=\, 0\,, & 0 \leq n \leq T/\Delta t\,, \quad \ell = 1,\dots,p\,, \\ \varepsilon_j^{n+1} \,=\, \displaystyle\sum_{\ell=-r}^{p} a_\ell\, \varepsilon_{j+\ell}^n + \Delta t\, F_j^{n+1}\,, & 0 \leq n \leq T/\Delta t - 1\,, \quad j \leq J\,. \end{cases} \tag{3.23}$$

Here above, the consistency error F_j^{n+1} consists of two terms: a first one coming from the usual interior consistency error denoted e_j^{n+1}, and a second one coming from the time evolution of the boundary layer corrector denoted δ_j^{n+1}. In other words, we split $F_j^{n+1} = e_j^{n+1} + \delta_j^{n+1}$ with:

$$e_j^{n+1} \,:=\, \frac{1}{\Delta t}\left(\omega_j^{n+1} - \sum_{\ell=-r}^{p} a_\ell\, \omega_{j+\ell}^n\right), \quad \text{and} \quad \delta_j^{n+1} \,:=\, \frac{\Delta x^{k_b}}{\Delta t}\left(v_j^{n+1} - \sum_{\ell=-r}^{p} a_\ell\, v_{j+\ell}^n\right).$$

Considering the scheme (3.23), the error $(\varepsilon_j^n)_{j\leq J+p, 0\leq n\leq T/\Delta t}$ obeys the stability estimate applicable in the case of the homogeneous extrapolation boundary condition, see [CL20, Proposition 3.4]:

$$\sup_{0\leq n\leq T/\Delta t} \sum_{j\leq J} \Delta x\,(\varepsilon_j^n)^2 \,\leq\, C\left\{\sum_{j\leq J} \Delta x\,(\varepsilon_j^0)^2 + T^2 \sup_{1\leq n\leq T/\Delta t} \sum_{j\leq J} \Delta x\,(F_j^n)^2\right\}. \tag{3.24}$$

It therefore remains to estimate the initial and interior consistency errors in (3.23):

[4] Here we use $u_j^0 = \omega_j^0$ for $j \leq J$.

- The initial consistency error. Estimating the initial condition $(\varepsilon_j^0)_{j\le J}$ directly follows from the estimate (3.15) in Lemma 3.5:

$$\sum_{j\le J}\Delta x\,(\varepsilon_j^0)^2 \;\le\; C\,\Delta x^{2\,k_b+1}\,\|f\|^2_{H^{k_b+1}((-\infty,L))}\,.$$

- The interior consistency error. I. Estimating the interior consistency error (e_j^n) related to the projected exact solution (ω_j^n) has already been achieved in [CL20] so we just report the result:

$$\sup_{1\le n\le T/\Delta t}\sum_{j\le J}\Delta x\,(e_j^n)^2 \;\le\; C\,\Delta x^{2\,k}\,\|f\|^2_{H^{k+1}((-\infty,L))}\,.$$

- The interior consistency error. II. Estimation of the new error term related to (δ_j^n). Observe that, first due to the steady states decomposition from Lemma 3.4, and then using successively (3.12) and (3.21), the interior consistency error arising from the boundary layer corrector rewrites as:

$$\delta_j^{n+1} \;=\; \frac{\Delta x^{k_b}}{\Delta t}\,(v_j^{n+1}-v_j^n) \;=\; \frac{1}{\Delta t}\sum_{\sigma=1}^{\tau_+}\sum_{\nu=0}^{\mu_\sigma-1}\sum_{\ell=1}^{p}\beta_{\sigma,\nu,\ell}\,(D_-^{k_b}(\omega^{n+1}-\omega^n))_{J+\ell}\,\rho_j^{(\sigma,\nu)}\,.$$

Thus, from Cauchy-Schwarz inequalities, there exists a constant C such that

$$\sum_{j\le J}\Delta x\,(\delta_j^{n+1})^2 \;\le\; C\,\Delta x\sup_{\ell=1,\dots,p}\left(D_-^{k_b}\left(\frac{\omega^{n+1}-\omega^n}{\Delta t}\right)_{J+\ell}\right)^2\,.$$

In the above formula, the discrete in time derivative of ω_j^n rewrites, for any $j\le J+p$ as

$$\begin{aligned}\frac{\omega_j^{n+1}-\omega_j^n}{\Delta t} &= \frac{1}{\Delta x}\int_{x_{j-1}}^{x_j}\frac{f(x-at^n-a\Delta t)-f(x-at^n)}{\Delta t}\,\mathrm{d}x\\ &= \frac{1}{\Delta x}\int_{x_{j-1}}^{x_j}\underbrace{\frac{1}{\Delta t}\int_{x-at^n}^{x-at^n-a\Delta t}f'(y)\,\mathrm{d}y}_{=:F(x)}\,\mathrm{d}x\,.\end{aligned}$$

Since $f\in H^{k+1}(\mathbb{R})$ with $k>k_b$, we have at least $f\in H^{k_b+2}(\mathbb{R})$ and therefore $F\in H^{k_b+1}(\mathbb{R})$ with

$$\|F^{(k_b+1)}\|_{L^2(\mathbb{R})} \;\le\; a^2\,\|f^{(k_b+2)}\|_{L^2(\mathbb{R})}\,,$$

from which we deduce, using again [CL20, Lemma 3.6]:

$$\left|D_-^{k_b}\left(\frac{\omega^{n+1}-\omega^n}{\Delta t}\right)_{J+\ell}\right| \;\le\; C\,\Delta x^{k_b}\,\|f\|_{H^{k_b+2}(\mathbb{R})}\,,\quad \ell=1,\dots,p\,.$$

Thus, using again the upper bound (3.22), the above estimate and the H^{k_b+2}-continuity of the extension operator, we end up with:

$$\sup_{1\le n\le T/\Delta t}\sum_{j\le J}\Delta x\,(\delta_j^n)^2 \;\le\; C\,\Delta x^{2\,k_b+1}\,\|f\|^2_{H^{k+1}((-\infty,L))}\,.$$

Let us now come back to the stability estimate (3.24) and use the three above consistency estimates to get (recall $T \geq 1$ and $k_b < k$):

$$\sup_{0 \leq n \leq T/\Delta t} \left(\sum_{j \leq J} \Delta x \, (\varepsilon_j^n)^2 \right)^{1/2} \leq C \, T \, \Delta x^{k_b + 1/2} \, \| f \|_{H^{k+1}((-\infty, L))} \, .$$

From the constructive formula for the boundary layer corrector (v_j^n), we have derived the bound (3.15) which, by the triangle inequality, yields the convergence estimate (recall $\varepsilon_j^n = \omega_j^n - u_j^n + \Delta x^{k_b} \, v_j^n$):

$$\sup_{0 \leq n \leq T/\Delta t} \left(\sum_{j \leq J} \Delta x \, (u_j^n - \omega_j^n)^2 \right)^{1/2} \leq C \, T \, \Delta x^{k_b + 1/2} \, \| f \|_{H^{k+1}} \, .$$

Using now the (crude) estimate:

$$\sup_{j \leq J} |b_j| \, \leq \, \Delta x^{-1/2} \left(\sum_{j \leq J} \Delta x \, b_j^2 \right)^{1/2} ,$$

we complete the proof of Theorem 1.3.

4 Numerical experiments

4.1 The Lax-Wendroff scheme

We report in this paragraph on various numerical experiments with the Lax-Wendroff scheme (1.3) (which corresponds to $p = r = 1$). Assumption 1.1 is satisfied provided that $\lambda \, a \leq 1$, and the order k equals 2. In all what follows, we choose $a = 1$ and $\lambda = 5/6$. The interval length is $L = 6$ and the final time T equals 8. The initial condition is $f(x) = \sin x$ and the boundary source term is $g(t) = -\sin t$ so that the exact solution to (1.1) is $u(t, x) = \sin(x - t)$. With the values of J reported in Table 1 below, we implement the Lax-Wendroff scheme (1.3) with the following numerical boundary conditions:

$$\begin{aligned} u_{J+1}^n \, &= \, u_J^n \, , \quad \text{(first order outflow extrapolation condition)}, \\ u_0^n \, &= \, \begin{cases} -\sin t^n \, , & \text{(Dirichlet inflow condition (1.6))}, \\ -\sin t^n - (\Delta x/2) \, \cos t^n \, , & \text{(inverse Lax-Wendroff inflow condition (1.12))}. \end{cases} \end{aligned}$$

The errors, as measured in the statement of Theorem 1.2, are reported in Table 1 below for each of the two cases (either the Dirichlet inflow condition (1.6) or the inverse Lax-Wendroff inflow condition (1.12)). In either case, the observed convergence rate is 1 since increasing J by a factor 2 decreases the error of the same factor 2. This behavior is fully justified by Theorem 1.3 since we have $k_b < k$ here.

We now turn to the second order outflow extrapolation condition:

$$\begin{aligned} u_{J+1}^n \, &= \, 2 \, u_J^n - u_{J-1}^n \, , \quad \text{(second order outflow extrapolation condition (1.5))}, \\ u_0^n \, &= \, \begin{cases} -\sin t^n \, , & \text{(Dirichlet inflow condition (1.6))}, \\ -\sin t^n - (\Delta x/2) \, \cos t^n \, , & \text{(inverse Lax-Wendroff inflow condition (1.12))}. \end{cases} \end{aligned}$$

Number of cells J	Dirichlet inflow condition	Inverse Lax-Wendroff inflow condition
1000	$4.1 \cdot 10^{-3}$	$5.1 \cdot 10^{-4}$
2000	$2.1 \cdot 10^{-3}$	$2.5 \cdot 10^{-4}$
4000	$1.1 \cdot 10^{-3}$	$1.3 \cdot 10^{-4}$
8000	$5.3 \cdot 10^{-4}$	$6.3 \cdot 10^{-5}$

Table 1: The $\ell^\infty_{n,j}$ error for the Lax-Wendroff scheme (1.3) with first order outflow extrapolation and either the Dirichlet, or inverse Lax-Wendroff, inflow condition.

The errors, as measured in the statement of Theorem 1.2, are reported in Table 2 below for each of the two cases (either the Dirichlet inflow condition (1.6) or the inverse Lax-Wendroff inflow condition (1.12)). For the Dirichlet inflow condition, the observed convergence rate is 1 again (despite the more accurate outflow treatment), but one recovers the convergence rate 2 with the inverse Lax-Wendroff inflow condition (1.12). However, proving rigorously that this numerical scheme converges with the rate 2 in the maximum norm might be very difficult (it might actually even be wrong !), even for smooth data, since the Lax-Wendroff scheme is known to be unstable in $\ell^\infty(\mathbb{Z})$. Improving the convergence rate 3/2 of Theorem 1.2 in the case of the Lax-Wendroff scheme with second order extrapolation outflow condition is left to a future work.

Number of cells J	Dirichlet inflow condition	Inverse Lax-Wendroff inflow condition
1000	$3.7 \cdot 10^{-3}$	$1.2 \cdot 10^{-5}$
2000	$1.8 \cdot 10^{-3}$	$2.9 \cdot 10^{-6}$
4000	$9.3 \cdot 10^{-4}$	$7.3 \cdot 10^{-7}$
8000	$4.7 \cdot 10^{-4}$	$1.8 \cdot 10^{-7}$

Table 2: The $\ell^\infty_{n,j}$ error for the Lax-Wendroff scheme (1.3) with second order outflow extrapolation (1.5) and either the Dirichlet or inverse Lax-Wendroff inflow condition.

4.2 The $O3$ scheme

Let us now consider the $O3$ scheme (3.10), which is implemented by considering the recurrence:

$$u_j^{n+1} \;=\; a_{-2}\, u_{j-2}^n + a_{-1}\, u_{j-1}^n + a_0\, u_j^n + a_1\, u_{j+1}^n\,, \quad n \in \mathbb{N}\,, \quad j = 1, \dots, J\,,$$

with:

$$a_{-2} := -\frac{\lambda\, a}{6}\left(1 - (\lambda\, a)^2\right), \quad a_{-1} := \frac{\lambda\, a}{2}\,(1 + \lambda\, a)\,(2 - \lambda\, a)\,,$$
$$a_0 := \frac{1}{2}\left(1 - (\lambda\, a)^2\right)(2 - \lambda\, a)\,, \quad a_1 := -\frac{\lambda\, a}{6}\,(1 - \lambda\, a)\,(2 - \lambda\, a)\,.$$

The reader can verify that assumption 1.1 is satisfied provided that $\lambda\, a \leq 1$, and the order k equals 3 ($r = 2$ and $p = 1$ here). To maintain third order accuracy, we implement the latter scheme with the following boundary conditions:

$u_{J+1}^n \;=\; 3\, u_J^n - 3\, u_{J-1}^n + u_{J-2}^n\,,$ (third order outflow extrapolation condition, $k_b = 3$),
$u_0^n \;=\; -\sin t^n - (\Delta x/2)\, \cos t^n + (\Delta x^2/6)\, \sin t^n\,,$ (inverse Lax-Wendroff inflow condition (1.12)),
$u_{-1}^n \;=\; -\sin t^n - (3\,\Delta x/2)\, \cos t^n + (7\,\Delta x^2/6)\, \sin t^n\,,$ (inverse Lax-Wendroff inflow condition (1.12)).

The measured errors are reported in Table 3 below. They correspond to a rate of convergence 3. Let us observe that the $O3$ scheme is known to be stable in $\ell^\infty(\mathbb{Z})$, see [Tho65, Des08], hence there is a genuine hope of proving rigorously that this rate of convergence does indeed hold (for smooth compatible data). Such a justification is also left to a future work.

Number of cells J	Inverse Lax-Wendroff inflow condition
1000	$2.1 \cdot 10^{-8}$
2000	$2.6 \cdot 10^{-9}$
4000	$3.3 \cdot 10^{-10}$

Table 3: The $\ell^\infty_{n,j}$ error for the $O3$ scheme (1.3) with third order outflow extrapolation and the inverse Lax-Wendroff inflow condition.

References

[BC17] B. Boutin and J.-F. Coulombel. Stability of finite difference schemes for hyperbolic initial boundary value problems: numerical boundary layers. *Numer. Math. Theory Methods Appl.*, 10(3):489–519, 2017.

[BGS07] S. Benzoni-Gavage and D. Serre. *Multidimensional hyperbolic partial differential equations*. Oxford Mathematical Monographs. Oxford University Press, 2007.

[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*, to appear, 2020.

[dB05] C. de Boor. Divided differences. *Surv. Approx. Theory*, 1:46–69, 2005.

[DDJ18] G. Dakin, B. Després, and S. Jaouen. Inverse Lax-Wendroff boundary treatment for compressible Lagrange-remap hydrodynamics on Cartesian grids. *J. Comput. Phys.*, 353:228–257, 2018.

[Des08] B. Després. Finite volume transport schemes. *Numer. Math.*, 108(4):529–556, 2008.

[FY13] F. Filbet and C. Yang. An inverse Lax-Wendroff method for boundary conditions applied to Boltzmann type models. *J. Comput. Phys.*, 245:43–61, 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.

[HJ94] R. A. Horn and C. R. Johnson. *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[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.

[Pop40] T. Popoviciu. Introduction à la théorie des différences divisées. *Bull. Math. Soc. Roumaine Sci.*, 42(1):65–78, 1940.

[ST17] C.-W. Shu and S. Tan. Inverse Lax-Wendroff procedure for numerical boundary treatment of hyperbolic equations. In *Handbook of numerical methods for hyperbolic problems*, volume 18 of *Handb. Numer. Anal.*, pages 23–52. Elsevier/North-Holland, 2017.

[Ste39] J. F. Steffensen. Note on divided differences. *Danske Vid. Selsk. Mat.-Fys. Medd.*, 17(3):12, 1939.

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

[Tho65] V. Thomée. Stability of difference schemes in the maximum-norm. *J. Differential Equations*, 1:273–292, 1965.

[TS10] S. Tan and C.-W. Shu. Inverse Lax-Wendroff procedure for numerical boundary conditions of conservation laws. *J. Comput. Phys.*, 229(21):8144–8166, 2010.

[VS15] F. Vilar and C.-W. Shu. Development and stability analysis of the inverse Lax-Wendroff boundary treatment for central compact schemes. *ESAIM Math. Model. Numer. Anal.*, 49(1):39–67, 2015.