

The Neumann numerical boundary condition for transport equations

Jean-François Coulombel, Frédéric Lagoutière

▶ **To cite this version:**

Jean-François Coulombel, Frédéric Lagoutière. The Neumann numerical boundary condition for transport equations. Kinetic and Related Models , 2020, 13 (1), pp.1-32. 10.3934/krm.2020001 . hal-01902551v2

HAL Id: hal-01902551
https://hal.science/hal-01902551v2

Submitted on 5 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Neumann numerical boundary condition for transport equations

Jean-François Coulombel* & Frédéric Lagoutière†

November 5, 2018

Abstract

In this article, we show that prescribing homogeneous Neumann type numerical boundary conditions at an outflow boundary yields a convergent discretization in ℓ^∞ for transport equations. We show in particular that the Neumann numerical boundary condition is a stable, local, and absorbing numerical boundary condition for discretized transport equations. Our main result is proved for explicit two time level numerical approximations of transport operators with arbitrarily wide stencils. The proof is based on the energy method and bypasses any normal mode analysis.

AMS classification: 65M12, 65M06, 65M20.
Keywords: transport equations, numerical schemes, Neumann boundary condition, stability, convergence.

1 Introduction

It is a well-known fact that transport equations do not require prescription of any boundary condition at an *outflow* boundary, that is, when the transport velocity is outgoing with respect to the boundary of the spatial domain. This can be easily understood by integrating the equation along the characteristics. However, many discretizations of the transport equation involve a stencil that includes cells of the numerical grid that are located in the *downstream* region. Such discretizations necessitate the prescription of numerical boundary conditions at an outflow boundary [GKO95, Str04], even though the underlying partial differential operator does not require any boundary condition for determining the solution.

The construction and analysis of transparent *versus* absorbing numerical boundary conditions for wave propagation problems now has a very long history, going back at least to the fundamental contribution by Engquist and Majda [EM77], see also among numerous other works [Hag99, Hal82, Hig86, Hig94, GKO95, AAB+08] and references therein. Our goal in this article is to rigorously justify the commonly acknowledged fact that enforcing Neumann type boundary conditions at an outflow boundary "*does the job*", in the simple one-dimensional case with a constant velocity (both on the half-line and on an interval). In several respects, the result is definitely not new. Stability estimates for Neumann numerical boundary conditions were stated -though without proof- for instance by Kreiss [Kre66], and a detailed proof of the

*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 BoND, ANR-13-BS01-0009, and by the ANR project NABUCO, ANR-17-CE40-0025.

†Université de Lyon, Université Claude Bernard Lyon 1, Institut Camille Jordan (CNRS UMR5208), 43 boulevard du 11 novembre 1918, F-69622 Villeurbanne Cedex, France. Email: lagoutiere@math.univ-lyon1.fr. Research of the author was supported by ANR project BoND, ANR-13-BS01-0009, and by the ANR project NABUCO, ANR-17-CE40-0025.

latter result was later provided by Goldberg [Gol77]. This result now enters the more general framework of [GT81]. However, the approach in [Kre66, Gol77, GT81] is rather elaborate since it relies first on the verification of the so-called Uniform Kreiss-Lopatinskii Condition (that is, in the present context of numerical schemes, a refined version of the Godunov-Ryabenkii condition [GKS72, GKO95]), and then on the application of deep general results which show that the latter condition is sufficient for the derivation of optimal semigroup estimates. Such general results first arose in [Kre66] and were later proved in further generality by Kreiss, Osher and followers [Kre68, Osh69, Wu95, CG11]. When combined with the trace estimates provided by the fulfillment of the Uniform Kreiss-Lopatinskii Condition and the general convergence result of Gustafsson [Gus75], one gets a complete -though lengthy (!) and somehow unclear as far as the topology is concerned- justification that enforcing Neumann type numerical boundary conditions at an outflow boundary yields a convergent scheme for discretized transport equations.

Our goal here is to bypass all the arguments of those previous works that were based on the normal mode analysis and to obtain a more direct convergence result *in the* ℓ^∞ *topology* with arguments that are as elementary as possible. Our approach is based on the energy method and discrete integration by parts rather than on the Laplace transform. We hope that some of our arguments might be useful to deal with more involved problems such as multidimensional hyperbolic systems. The main gain when enforcing the Neumann numerical boundary condition with respect, for instance, to the Dirichlet numerical boundary condition is to obtain convergence results in ℓ^∞, while in the case of the Dirichlet boundary condition at an outflow boundary, despite unconditional stability properties [GT81], boundary layer phenomena allow at best for convergence results in ℓ^p, $p < +\infty$, only [KL68, CHG01, BC17]. Our main result below gives a rate of convergence in ℓ^∞ for such discretizations of the transport equations. We do not claim that our rate is optimal, but we do not assume the discretization of the transport equation to be stable in $\ell^\infty(\mathbb{Z})$ either, so if we do not reach optimality we are certainly not too far from it. We plan to study the more favorable case of ℓ^∞ stable schemes in the future, with the aim of improving the rate of convergence.

The rest of this article is organized as follows. In Section 2, we introduce some notation and state our main convergence result for Neumann numerical boundary conditions at an outflow boundary. Based on the standard approach of numerical analysis, our convergence result relies on accurate stability estimates and a consistency analysis. Our main stability estimate, based on a new elementary approach, is stated and proved in Section 3. (A crucial discrete integration by parts lemma is given in Appendix A.) The concluding arguments for proving our main result are given in Section 4.

2 Main result

In this article, we consider the following transport problem on an interval. We are given a fixed constant velocity $a > 0$, an interval length $L > 0$ and consider the problem:

$$\begin{cases} \partial_t u + a\,\partial_x u = 0\,, & t \geq 0\,,\ x \in (0,L)\,, \\ u(0,x) = u_0(x)\,, & x \in (0,L)\,, \\ u(t,0) = 0\,, & t \geq 0\,, \end{cases} \tag{2.1}$$

with, at least, $u_0 \in L^2((0,L))$. Actually, further regularity and compatibility requirements will be enforced later on, but let us stick to that simple framework for the moment. We could consider an inhomogeneous Dirichlet boundary condition at $x = 0$ in (2.1), but we rather stick to this simpler case in order to highlight the main novelty of our approach which rather focuses on the downstream boundary $x = L$ of the interval.

The solution to (2.1) is given by the method of characteristics, which yields the explicit representation

formula:

$$\forall\,(t,x)\in\mathbb{R}^+\times(0,L)\,,\quad u(t,x)=u_0(x-a\,t)\,, \tag{2.2}$$

where it is understood in (2.2) that the initial condition u_0 has been extended by zero to $\mathbb{R}^-$ (no extension is needed on $(L,+\infty)$ since a is positive and therefore $x-a\,t<L$ for all relevant values of t and x).

It may seem a too much trivial problem to approximate the problem (2.1) for which an explicit solution is given, but one should keep in mind that such a representation formula ceases to be available for hyperbolic *systems* in *several* space dimensions, and our goal is to develop analytical tools which do not rely on the fact that (2.1) is a *one-dimensional scalar* problem. We therefore consider from now on an approximation of (2.1) by means of a finite difference scheme. We are given a positive integer J, that is meant to be large, and define accordingly the space step Δx and the grid points (x_j) by:

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

The time step Δt is then defined as $\Delta t:=\lambda\,\Delta x$ where $\lambda>0$ is a *fixed* constant that is tuned so that our main Assumption 2.1 below is satisfied. The interval $(0,L)$ is divided in J cells (x_{j-1},x_j), $j=1,\dots,J$, as depicted in Figure 2.1. In what follows, we use the notation $t^n:=n\,\Delta t$, $n\in\mathbb{N}$, and u_j^n will play the role of an approximation for the solution u to (2.1) at time t^n on the cell (x_{j-1},x_j) (or at the mid-point $(x_{j-1}+x_j)/2$). We do not wish to discriminate between finite difference or finite volume schemes for (2.1), so rather than deriving this or that type of numerical scheme, we consider a linear iteration for the u_j^n's that reads in the *interior* domain:

$$u_j^{n+1}=\sum_{\ell=-r}^{p}a_\ell\,u_{j+\ell}^n\,,\quad n\in\mathbb{N}\,,\quad j=1,\dots,J\,. \tag{2.3}$$

In (2.3), r,p are fixed nonnegative integers, and the coefficients a_ℓ, $\ell=-r,\dots,p$ may only depend on the ratio λ and the velocity a. Most of the usual linear explicit schemes, such as the upwind, Rusanov, Lax-Friedrichs and Lax-Wendroff schemes, can be put in that form. The interior domain corresponds to the indices $j=1,\dots,J$. However, in (2.3), the determination of $(u_1^{n+1},\dots,u_J^{n+1})$ requires the prior knowledge of $(u_{1-r}^n,\dots,u_{J+p}^n)$, which corresponds to a larger set of cells. In what follows the cells (x_{j-1},x_j) with $j=1-r,\dots,0$ and $j=J+1,\dots,J+p$ will be referred to as "ghost cells". They are depicted in red in Figure 2.1 (in that example, $p=r=2$).

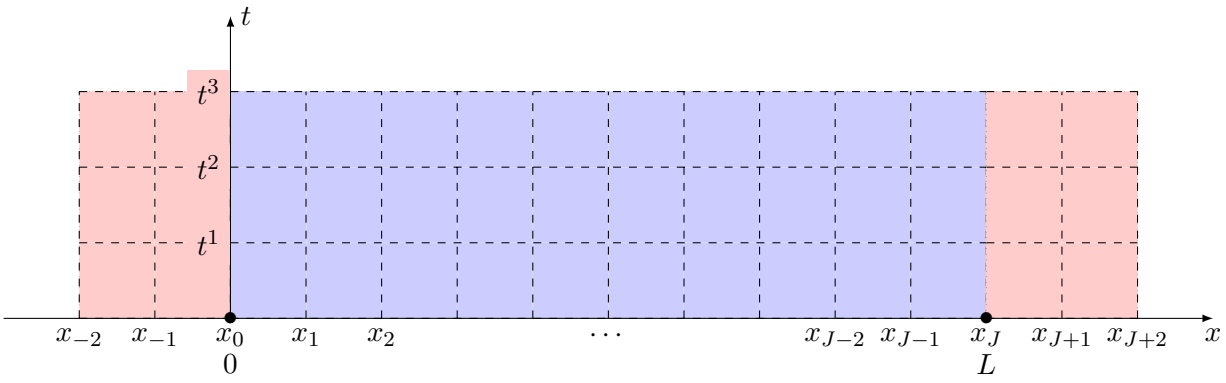


Figure 2.1: The mesh on $\mathbb{R}^+\times(0,L)$ in blue, and the "ghost cells" in red ($r=p=2$ here).

Before describing the numerical boundary conditions we enforce for (2.3), let us state our main and in fact only assumption on the coefficients in (2.3).

Assumption 2.1 (Consistency and stability)**.** *The coefficients $a_{-r},\dots,a_p$ in (2.3) satisfy $a_{-r}\,a_p\neq 0$ (normalization), and for some integer $k\geq 1$, there holds:*

$$\forall\, m=0,\dots,k\,,\quad \sum_{\ell=-r}^{p}\ell^m\,a_\ell=(-\lambda\,a)^m\,,\qquad\qquad \text{(consistency of order } k)\,, \tag{2.4}$$

$$\sup_{\theta\in[0,2\,\pi]}\left|\sum_{\ell=-r}^{p}a_\ell\,\mathrm{e}^{i\,\ell\,\theta}\right|\leq 1\,,\qquad\qquad (\ell^2\text{-stability})\,. \tag{2.5}$$

Provided that the relations (2.4) are satisfied for $m=0$ (conservativity) and $m=1$ (consistency of order 1) with $a>0$, the stability assumption (2.5) implies $r\geq 1$, which we assume from now on. Though we view this observation here, as in [Str62], as a *necessary*condition for *stability*, the condition $r\geq 1$ is also known to be *necessary* for *convergence* by comparing the numerical and continuous dependency domains, see [CFL28]. Let us observe that (2.5) is a necessary and sufficient condition for stability of the iteration process (2.3) on $\ell^2(\mathbb{Z})$ in a strong sense, meaning here that the map

$$(v_j)_{j\in\mathbb{Z}}\longmapsto\left(\sum_{\ell=-r}^{p}a_\ell\,v_{j+\ell}\right)_{j\in\mathbb{Z}}$$

is a contraction (it has norm 1) as an operator on $\ell^2(\mathbb{Z})$. However, (2.5) is not sufficient to yield stability in $\ell^\infty(\mathbb{Z})$ for (2.3), see [Tho65, Hed66]. Note that through the dependence of the a_ℓ with respect to $\lambda=\Delta t/\Delta x$, (2.5) is usually intended to be true only under a so-called Courant-Friedrichs-Lewy condition asking for λ to be less than some constant depending on the scheme and the velocity a. (Indeed, the Bernstein inequality for trigonometric polynomials implies $\lambda\,|a|\leq\max(p,r)$, see [Str62].) We postpone the extension of our work to $\ell^\infty(\mathbb{Z})$ stable discretizations to a future work since the methods to be used in that framework will definitely need to be different from the more "Hilbertian" techniques we use here.

Let us observe that Assumption 2.1 does not include any *dissipative* behavior for (2.3), meaning that we do not assume a bound of the form:

$$\forall\,\theta\in[-\pi,\pi]\,,\quad\left|\sum_{\ell=-r}^{p}a_\ell\,\mathrm{e}^{i\,\ell\,\theta}\right|\leq 1-c\,\theta^{2\,q}\,,$$

for some suitable integer q and positive constant c. In that respect, the framework of Assumption 2.1 is more general than the works [Kre68, Gol77, GKS72, Osh69] and following works that are based on these pioneering results. We thus expect that our approach may be useful to deal with multidimensional problems in which dissipativity is most of the time excluded (or restrictive).

If the interior cells of the grid are labeled, as in (2.3), by $j\in\{1,\dots,J\}$, the numerical approximation of (2.1) requires, for passing from one time index n to the next, prescribing r numerical boundary conditions on the left of the interval (that is, close to $x=0$), and p numerical boundary conditions on the right (that is, close to $x=L$). In other words, we need to prescribe the value of the approximate solution (u_j^n) in the ghost cells located at the boundary of the interior domain. For simplicity, and in order to be consistent with the continuous problem (2.1), we prescribe Dirichlet homogeneous boundary conditions in conjunction with (2.3) on the left of the interval $(0,L)$:

$$u_\ell^n=0\,,\quad \ell=1-r,\dots,0\,. \tag{2.6}$$

On the right of the interval $(0, L)$, there is nothing to be done if $p = 0$, that is, in the case of an *upwind* discretization, for in that case, given the vector $(u_1^n, \dots, u_J^n)$, the vector $(u_1^{n+1}, \dots, u_J^{n+1})$ is entirely determined by (2.6) and (2.3), so we can iterate the scheme (2.3)-(2.6) starting from some initial data $(u_1^0, \dots, u_J^0)$ to any positive time level n. We therefore assume from now on $p \geq 1$, which is the interesting case where the numerical discretization of (2.1) necessitates an outflow numerical boundary condition while the continuous problem does not "obviously" provide with one. In this article, we shall prescribe Neumann type numerical boundary conditions (these are called *extrapolation* numerical boundary conditions in [Gol77]). For ease of writing, we introduce the difference operator in space which acts on vectors $(v_j)_{j=1-r,\dots,J+p}$ as follows:

$$\forall\, j = 2 - r, \dots, J + p\,, \quad (D_- v)_j := v_j - v_{j-1}\,.$$

Higher order difference operators D_-^m, $m \geq 2$, are defined accordingly by iterating D_-. Then given a fixed integer $k_b \in \mathbb{N}$ (b stands for "boundary"), we prescribe the following numerical boundary condition in conjunction with (2.3):

$$(D_-^{k_b} u^n)_{J+\ell} = 0\,, \quad \ell = 1, \dots, p\,. \tag{2.7}$$

If $k_b = 0$, this corresponds to prescribing homogeneous Dirichlet boundary conditions, while if $k_b = 1$, this corresponds to the standard Neumann numerical boundary condition:

$$u_{J+1}^n = \cdots = u_{J+p}^n := u_J^n\,.$$

The iteration (2.3), (2.6), (2.7) thus proceeds as follows, see Figure 2.2 for an illustration. Given the vector $(u_1^n, \dots, u_J^n)$ for some time level n, one first determines the ghost values $(u_{1-r}^n, \dots, u_0^n, u_{J+1}^n, \dots, u_{J+p}^n)$ by (2.6) and (2.7). The *new* vector $(u_1^{n+1}, \dots, u_J^{n+1})$ is then determined by applying (2.3). It is assumed that $J \geq 1$ in order to make the space step $\Delta x = L/J$ meaningful and to have at least one cell in the interval $(0, L)$.

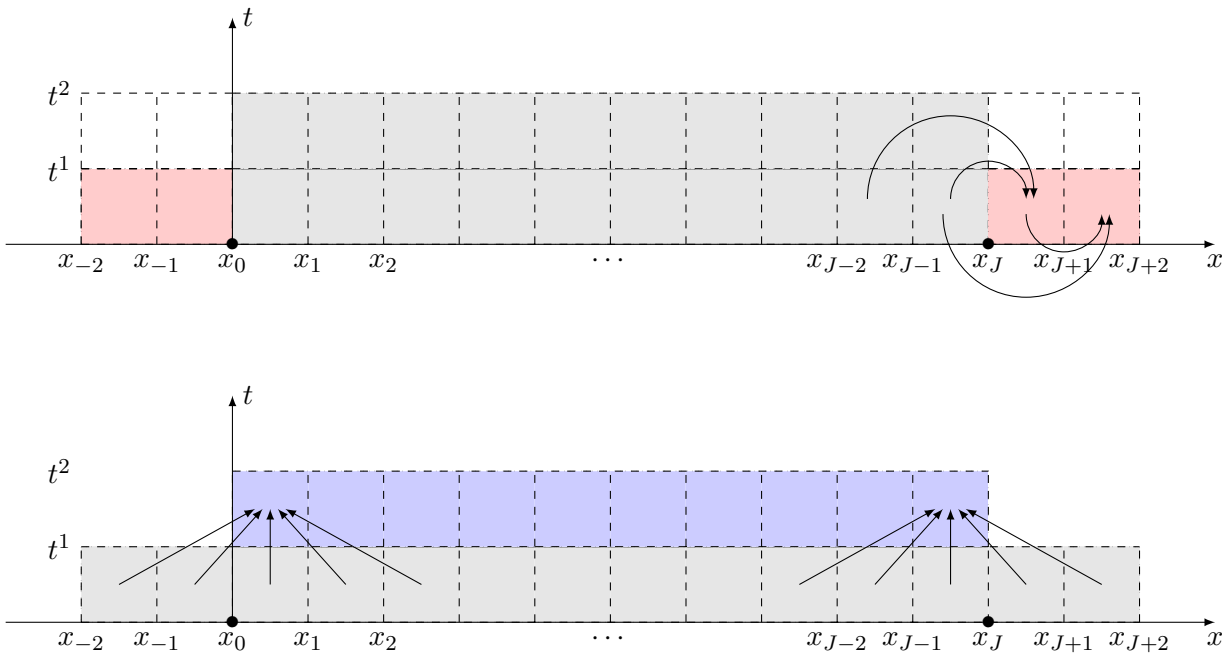


Figure 2.2: Top: updating iteratively the ghost values at the outflow boundary ($r = p = k_b = 2$). Bottom: updating the numerical approximation in the interior.

Let us emphasize that the same procedure (2.7) is applied at each ghost cell close to the outflow boundary $x = L$. In the terminology of [GT78, GT81], the numerical boundary conditions are of *translatory* type.

The scheme (2.3), (2.6), (2.7) is initialized with the piecewise constant projection of the initial condition for (2.1), that is, for the interior cells:

$$\forall\, j = 1, \dots, J\,, \quad u_j^0 := \frac{1}{\Delta x} \int_{(j-1)\,\Delta x}^{j\,\Delta x} u_0(y)\,\mathrm{d}y\,. \tag{2.8}$$

Our main result is the following convergence estimate for the scheme (2.3), (2.6), (2.7) supplemented with the initial condition (2.8).

Theorem 2.2. *Let $a > 0$, let $k \in \mathbb{N}^*$ and $k_b \in \mathbb{N}$, let $\lambda > 0$ and assume that the coefficients in (2.3) satisfy Assumption 2.1 with integer k. Then there exists $C > 0$ such that for any $T > 0$ and $J \in \mathbb{N}^*$, for any $L \geq 1$, and for any $u_0 \in H^{k+1}((0, L))$ satisfying the compatibility requirements at $x = 0$:*

$$\forall\, m = 0, \dots, k\,, \quad u_0^{(m)}(0) = 0\,,$$

the solution (u_j^n) to (2.3), (2.6), (2.7) with initial datum (2.8) satisfies:

$$\sup_{0 \leq n \leq T/\Delta t} \sup_{j=1,\dots,J} \left| u_j^n - \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} u_0(y - a\,t^n)\,\mathrm{d}y \right| \leq C\,\left(\sqrt{T} + T\right) \mathrm{e}^{C\,T/L}\,\Delta x^{\min(k,k_b)-1/2}\,\|u_0\|_{H^{k+1}((0,L))}\,, \tag{2.9}$$

where $\Delta x = L/J$ and $\Delta t = \lambda\,\Delta x$. It is understood in (2.9) that u_0 is extended by 0 to $\mathbb{R}^-$.

The restriction to numerical schemes with two time levels only in (2.3) is necessary since, for instance, the Neumann numerical boundary condition (2.7) is known to yield violent instabilities when used in conjunction with the leap-frog scheme [Tre84]. We postpone the study of outflow numerical boundary conditions for general multistep schemes to a future work.

The integer k_b in (2.7) prescribes the approximation order of the numerical outflow boundary condition. In particular, Theorem 2.2 shows that the Neumann numerical boundary condition ($k_b = 1$) gives a local, stable hence convergent way to approximate the exact discrete transparent boundary condition for (2.3) which is nonlocal in time (see [Cou17] for a general derivation of discrete transparent boundary conditions).

There is a "loss" of $1/2$ in the somehow expected rate of convergence $\min(k, k_b)$ in (2.9). We emphasize once again that Assumption 2.1 does not imply stability of (2.3) in $\ell^\infty(\mathbb{Z})$ and therefore even on the whole real line, the rate of convergence k in ℓ^∞ cannot be attained in general. With the additional technical complexity of dealing with numerical boundary conditions, we view this loss of $1/2$ as a minor disagreement of our method.

Applying estimate (2.9) to the "theoretical" final time $T := L/a$ after which the exact solution to the transport equation (2.1) becomes zero (everything has flowed out of the interval $(0, L)$ through the outflow boundary $x = L$), we obtain:

$$\sup_{j=1,\dots,J} \left| u_j^{N+1} \right| \lesssim \Delta x^{\min(k,k_b)-1/2}, \quad N := \mathrm{E}\left(\frac{L}{a\,\Delta t} \right),$$

meaning that the numerical approximation of the solution to (2.1) has become uniformly small on the interval $(0, L)$ at time level $N+1$. At later times, stability estimates for the numerical scheme (2.3), (2.6), (2.7), which we shall prove below, assert that the solution (u_j^n) to (2.3), (2.6), (2.7) remains "small" (at least on any given finite time interval, J being large).

3 Preliminary analysis on a half-line

In this section, we show a preliminary result, which is Theorem 3.5 below, that is entirely analogous to our main result, Theorem 2.2 above, except that the space domain is a half-line with an outgoing velocity at the boundary. The result is interesting on its own since a byproduct of our analysis is the verification of the so-called Uniform Kreiss Lopatinskii Condition though we completely bypass the normal mode analysis that is commonly used to verify it [Gol77]. The convergence analysis on the half-line is more classical. In addition to its own interest, which somehow focuses on the Neumann boundary condition without coupling it to Dirichlet boundary condition, Theorem 3.5 below will be a major building block in the proof of Theorem 2.2.

3.1 Stability estimates for Neumann numerical outflow boundary conditions

In this paragraph, we prove Theorem 3.1 below that provides us with stability estimates for the Neumann numerical boundary condition on a half-line. Theorem 3.1 is a key tool for proving stability estimates for the scheme (2.3), (2.6), (2.7) on a finite interval, which in turn yields the convergence result of Theorem 2.2. Let us recall that Theorem 3.1 below is already known to hold true thanks to the joint results of [Kre66, Gol77, Kre68] and in a more general setting [Wu95, CG11, Cou13] (see also references therein). It roughly means that the Neumann numerical boundary conditions satisfy the Uniform Kreiss Lopatinskii Condition at an outflow boundary. However we emphasize that Assumption 2.1 does not imply "dissipativity" for (2.3), which already prohibits using the main results in [Kre68, GKS72] and

therefore strongly advocates using the energy method, as we shall do, whenever possible. Independently of the subtleties of the assumptions in those works, we give here an elementary proof of the stability result in [Kre66, Gol77] without any Laplace transform nor "GKS" type arguments, and with more easily checkable assumptions. Our integration by parts approach was probably already well-known for three point schemes and the classical Neumann boundary condition $D_- u = 0$, but our main contribution is to show that the energy method can be used in order to deal with numerical schemes with an arbitrarily wide stencil and for any extrapolation order at the boundary.

Before going on, let us fix the space domain that we consider. Since we deal with a constant coefficient linear problem, by translation invariance, there is no loss of generality in considering the half-line $(-\infty, L)$. With the same positive constant $\lambda > 0$ as in the previous section, we consider $J \in \mathbb{N}^*$, and some space and time steps $\Delta x = L/J$ and $\Delta t = \lambda \Delta x$, assuming that J is large enough to ensure that $\Delta t \in (0, 1]$. We also keep the notation $t^n := n \Delta t$, $n \in \mathbb{N}$ and $x_j := j \Delta x$, $j \in \mathbb{Z}$. The grid and the associated ghost cells are depicted in Figure 3.1. Our result can be stated as follows.

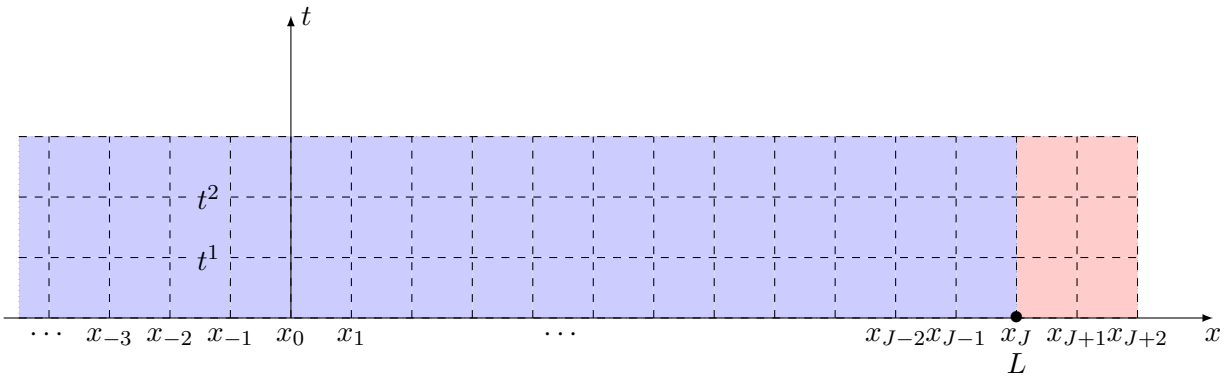


Figure 3.1: The mesh on $\mathbb{R}^+ \times (-\infty, L)$ in blue, and the "ghost cells" in red ($p = 2$ here).

Theorem 3.1. *Let $a > 0$, let $k \geq 1$ and $k_b \in \mathbb{N}$, let $\lambda > 0$. Then there exists a constant $C > 0$ such that for all initial data $(f_j)_{j \leq J} \in \ell^2$ and for all boundary source terms $(g^n_{J+1})_{n \geq 0}, \dots, (g^n_{J+p})_{n \geq 0}$ verifying the growth condition:*

$$\forall \Gamma > 0\,, \quad \sum_{n \geq 0} \mathrm{e}^{-2\Gamma n} \left((g^n_{J+1})^2 + \cdots + (g^n_{J+p})^2 \right) < +\infty\,,$$

the solution $(u^n_j)_{j \leq J+p, n \in \mathbb{N}}$ to the scheme:

$$\begin{cases} u^0_j = f_j\,, & j \leq J\,, \\ (D_-^{k_b} u^n)_{J+\ell} = g^n_{J+\ell}\,, & n \in \mathbb{N}\,, \quad \ell = 1, \dots, p\,, \\ u^{n+1}_j = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, u^n_{j+\ell}\,, & n \in \mathbb{N}\,, \quad j \leq J\,, \end{cases} \tag{3.1}$$

where the coefficients a_ℓ satisfy Assumption 2.1 with integer k, satisfies the estimate:

$$\sup_{n\in\mathbb{N}}\left(\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{j\le J}\Delta x\,(u_j^n)^2\right)+\sum_{n\ge 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{\ell=1-r-k_b}^{p}(u_{J+\ell}^n)^2 \le C\left\{\sum_{j\le J}\Delta x\,(f_j)^2+\sum_{n\ge 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{\ell=1}^{p}(g_{J+\ell}^n)^2\right\} \tag{3.2}$$

for any $\gamma>0$, and for any Δt and Δx such that $\Delta t/\Delta x=\lambda$ and $\Delta t\in(0,1]$. In particular, the numerical boundary conditions in (3.1) satisfy the Uniform Kreiss Lopatinskii Condition.

Before proving Theorem 3.1, let us recall that we always assume the ratio $\Delta t/\Delta x$ to be constant. This will be used several times below and is reminiscent of the scale invariance properties of the underlying continuous problem. Observe also that in (3.2), the larger the integer k_b the better the trace estimate on the left hand side behaves. In particular, the numerical boundary conditions in (3.1) involve the values $u_{J+1-k_b}^n,\dots,u_{J+p}^n$, and we get "for free" in (3.2) not only the control of those terms but also the extra control of $u_{J+1-r-k_b}^n,\dots,u_{J-k_b}^n$ (recall $r\ge 1$). The fact that (3.2) implies the Uniform Kreiss Lopatinskii Condition is not our main focus here, so instead of recalling many definitions, we rather refer the interested reader to the review [Cou13].

Proof. We shall use Assumption 2.1 in the proof below only for $k=1$, that is, we make the "minimal" consistency requirements for the scheme (2.3). Unlike [Kre66, Gol77], the proof of Theorem 3.1 is done by *induction* with respect to the index $k_b\in\mathbb{N}$ and relies on the *energy method*. Let us start with the case $k_b=0$, which corresponds to Dirichlet numerical boundary conditions.

- The case $k_b=0$. We consider the numerical scheme:

$$\begin{cases} u_j^0=f_j\,, & j\le J\,,\\ u_{J+\ell}^n=g_{J+\ell}^n\,, & n\in\mathbb{N}\,,\quad \ell=1,\dots,p\,,\\ u_j^{n+1}=\displaystyle\sum_{\ell=-r}^{p}a_\ell\,u_{j+\ell}^n\,, & n\in\mathbb{N}\,,\quad j\le J\,. \end{cases} \tag{3.3}$$

A straightforward proof of the stability estimate (3.2) for $k_b=0$ was achieved in [CG11] (even in some cases of multidimensional systems), see also [GT81] for an earlier general result based on the theory of [GKS72]. We reproduce here the short proof of (3.2) for the scheme (3.3) for the sake of completeness.

Assumption 2.1 implies by the Plancherel Theorem that the mapping:

$$(v_j)_{j\in\mathbb{Z}}\in\ell^2\longmapsto\left(\sum_{\ell=-r}^{p}a_\ell\,v_{j+\ell}\right)_{j\in\mathbb{Z}}\in\ell^2\,, \tag{3.4}$$

is a contraction, in the sense that the operator norm is not larger than 1. Let us now consider the solution $(u_j^n)_{j\le J+p,n\in\mathbb{N}}$ to (3.3) at some time index $n\in\mathbb{N}$. We extend the sequence $(u_j^n)_{j\le J+p}$ by 0 for $j\ge J+p+1$ and still denote $u^n\in\ell^2(\mathbb{Z})$ the resulting sequence. Let us then define

$$\forall\, j\in\mathbb{Z}\,,\quad v_j^{n+1}:=\sum_{\ell=-r}^{p}a_\ell\,u_{j+\ell}^n\,,$$

so that $v_j^{n+1} = u_j^{n+1}$ for $j \le J$, and $v_j^{n+1} = 0$ if $j \ge J+p+r+1$. Observe that, due to the boundary conditions in (3.3), we do not necessarily have $v_j^{n+1} = u_j^{n+1}$ for $j = J+1, \dots, J+p$, nor for $j = J+p+1, \dots, J+p+r$ (extending also $(u_j^{n+1})_{j \le J+p}$ by 0 for $j \ge J+p+1$). Using Assumption 2.1, we get:

$$\sum_{j \le J} \Delta x \, (u_j^{n+1})^2 + \sum_{j=J+1}^{J+p+r} \Delta x \, (v_j^{n+1})^2 = \sum_{j \in \mathbb{Z}} \Delta x \, (v_j^{n+1})^2 \le \sum_{j \in \mathbb{Z}} \Delta x \, (u_j^n)^2 = \sum_{j \le J+p} \Delta x \, (u_j^n)^2 \,,$$

since (3.4) is a contraction. Equivalently, taking the boundary conditions of (3.3) into account, we get:

$$\underbrace{\sum_{j \le J} \Delta x \, (u_j^{n+1})^2 - \sum_{j \le J} \Delta x \, (u_j^n)^2}_{\text{Discrete time derivative}} + \underbrace{\Delta x \sum_{\ell=1}^{p+r} (v_{J+\ell}^{n+1})^2}_{\text{Trace term}} \le \underbrace{\Delta x \sum_{\ell=1}^{p} (g_{J+\ell}^n)^2}_{\text{Source term}} \,. \tag{3.5}$$

We now derive a bound from below for the trace term arising on the left hand side of (3.5). The real numbers $v_{J+\ell}^{n+1}$, $\ell = 1, \dots, p+r$, depend linearly on $u_{J+1-r}^n, \dots, u_{J+p}^n$. The coefficients in each linear combination are taken among the a_ℓ's. Hence the quantity

$$\sum_{\ell=1}^{p+r} (v_{J+\ell}^{n+1})^2$$

can be seen as a nonnegative quadratic form in the variables $u_{J+1-r}^n, \dots, u_{J+p}^n$. It is also rather easy to see that this quadratic form is *positive definite* for we have $v_{J+p+r}^{n+1} = a_{-r} \, u_{J+p}^n$, and, therefore, if $v_{J+1}^{n+1} = \cdots = v_{J+p+r}^{n+1} = 0$, then we first have $u_{J+p}^n = 0$ and recursively we can also show $u_{J+p-1}^n = \cdots = u_{J+1-r}^n = 0$. Hence there exists a fixed constant $c_0 > 0$, that only depends on the (fixed) coefficients a_ℓ in (2.3), such that:

$$\sum_{\ell=1}^{p+r} (v_{J+\ell}^{n+1})^2 \ge c_0 \sum_{\ell=1-r}^{p} (u_{J+\ell}^n)^2 \,.$$

Reporting in (3.5) and using that $\Delta t / \Delta x = \lambda$ is a fixed positive constant, we get for some constant $c > 0$ ($c = c_0/\lambda$ is suitable):

$$\sum_{j \le J} \Delta x \, (u_j^{n+1})^2 - \sum_{j \le J} \Delta x \, (u_j^n)^2 + c \, \Delta t \sum_{\ell=1-r}^{p} (u_{J+\ell}^n)^2 \le \frac{1}{\lambda} \, \Delta t \sum_{\ell=1}^{p} (g_{J+\ell}^n)^2 \,. \tag{3.6}$$

We now apply the following discrete Gronwall type lemma (with the positive parameter $\Gamma := \gamma \, \Delta t$), see [CG11] for repeated use of such summation arguments.

Lemma 3.2. *Let $(\mathscr{G}_n)_{n \ge 0}$ be a sequence of nonnegative real numbers such that:*

$$\forall \, \Gamma > 0 \,, \quad \sum_{n \ge 0} \mathrm{e}^{-2 \, \Gamma \, n} \, \mathscr{G}_n < +\infty \,.$$

Let $(\mathscr{U}_n)_{n \ge 0}$, $(\mathscr{B}_n)_{n \ge 0}$ be two sequences of nonnegative real numbers such that:

$$\forall \, n \in \mathbb{N} \,, \quad \mathscr{U}_{n+1} - \mathscr{U}_n + \mathscr{B}_n \le \mathscr{G}_n \,.$$

Then there holds for all $\Gamma > 0$:

$$\sup_{n \in \mathbb{N}} \mathrm{e}^{-2 \, \Gamma \, n} \, \mathscr{U}_n + \sum_{n \ge 0} \mathrm{e}^{-2 \, \Gamma \, n} \, \mathscr{B}_n \le \mathscr{U}_0 + \sum_{n \ge 0} \mathrm{e}^{-2 \, \Gamma \, n} \, \mathscr{G}_n \,.$$

The proof of Lemma 3.2 is straightforward and therefore omitted. We apply Lemma 3.2 to (3.6) and derive the estimate:

$$\sup_{n\in\mathbb{N}}\left(\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{j\le J}\Delta x\,(u_j^n)^2\right)+\sum_{n\ge 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{\ell=1-r}^{p}(u_{J+\ell}^n)^2$$
$$\le C\left\{\sum_{j\le J}\Delta x\,(f_j)^2+\sum_{n\ge 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{\ell=1}^{p}(g_{J+\ell}^n)^2\right\}, \quad (3.7)$$

which is (3.2) for $k_b=0$.

We emphasize that when dealing with the case $k_b=0$, we have only used the stability condition (2.5) of Assumption 2.1, and we have never used (2.4) (not even for $m=0$). This is consistent with the result of [GT81] which proves that the Dirichlet boundary condition satisfies the Uniform Kreiss Lopatinskii Condition independently of the nature of the boundary (inflow or outflow).

• The induction argument. We now assume that the stability estimate (3.2) is valid up to some index $k_b\in\mathbb{N}$, and consider the following numerical scheme:

$$\begin{cases} u_j^0=f_j\,, & j\le J\,,\\ (D_-^{k_b+1}u^n)_{J+\ell}=g_{J+\ell}^n\,, & n\in\mathbb{N}\,,\quad \ell=1,\dots,p\,,\\ u_j^{n+1}=\displaystyle\sum_{\ell=-r}^{p}a_\ell\,u_{j+\ell}^n\,, & n\in\mathbb{N}\,,\quad j\le J\,. \end{cases} \quad (3.8)$$

We use the induction assumption by defining the following sequence:

$$\forall\, n\in\mathbb{N}\,,\quad \forall\, j\le J+p\,,\quad w_j^n:=(D_-\,u^n)_j=u_j^n-u_{j-1}^n\,,$$

which satisfies:

$$\begin{cases} w_j^0=f_j-f_{j-1}\,, & j\le J\,,\\ (D_-^{k_b}w^n)_{J+\ell}=g_{J+\ell}^n\,, & n\in\mathbb{N}\,,\quad \ell=1,\dots,p\,,\\ w_j^{n+1}=\displaystyle\sum_{\ell=-r}^{p}a_\ell\,w_{j+\ell}^n\,, & n\in\mathbb{N}\,,\quad j\le J\,. \end{cases}$$

Applying the stability estimate (3.2) for the index k_b and omitting one of the two nonnegative terms on the left hand side of (3.2), we have already derived the preliminary estimate:

$$\sum_{n\ge 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{\ell=1-r-k_b}^{p}(u_{J+\ell}^n-u_{J+\ell-1}^n)^2\le C\left\{\sum_{j\le J}\Delta x\,(f_j-f_{j-1})^2+\sum_{n\ge 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{\ell=1}^{p}(g_{J+\ell}^n)^2\right\}$$
$$\le C\left\{\sum_{j\le J}\Delta x\,(f_j)^2+\sum_{n\ge 0}\Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\sum_{\ell=1}^{p}(g_{J+\ell}^n)^2\right\}. \quad (3.9)$$

Let us now turn back to the numerical scheme (3.8) to which we are going to apply the so-called energy method. For a given time index $n\in\mathbb{N}$, we compute:

$$\sum_{j\le J}\Delta x\,(u_j^{n+1})^2-\sum_{j\le J}\Delta x\,(u_j^n)^2=\sum_{j\le J}\Delta x\left\{2\,u_j^n\Big(\sum_{\ell=-r}^{p}a_\ell\,u_{j+\ell}^n-u_j^n\Big)+\Big(\sum_{\ell=-r}^{p}a_\ell\,u_{j+\ell}^n-u_j^n\Big)^2\right\}. \quad (3.10)$$

We use the following discrete integration by parts result, whose proof is postponed to Appendix A.

Lemma 3.3. *Let $a > 0$ and let Assumption 2.1 be satisfied for some $k \in \mathbb{N}^*$. Then there exist a unique quadratic form $\mathscr{Q}$ on $\mathbb{R}^{p+r}$ and some real coefficients $d_1, \dots, d_{p+r}$ such that, for any real numbers $v_{j-r}, \dots, v_{j+p}$, there holds*

$$\begin{aligned} 2\, v_j \Big(\sum_{\ell=-r}^{p} a_\ell\, v_{j+\ell} - v_j \Big) + \Big(\sum_{\ell=-r}^{p} a_\ell\, v_{j+\ell} - v_j \Big)^2 &= \sum_{\ell=1}^{p+r} d_\ell\, (v_{j+\ell-r} - v_{j-r})^2 \\ &+ \mathscr{Q}\big(v_{j+2-r} - v_{j+1-r}, \dots, v_j - v_{j-1}, v_j, v_{j+1} - v_j, \dots, v_{j+p} - v_{j+p-1}\big) \\ &- \mathscr{Q}\big(v_{j+1-r} - v_{j-r}, \dots, v_{j-1} - v_{j-2}, v_{j-1}, v_j - v_{j-1}, \dots, v_{j+p-1} - v_{j+p-2}\big)\,. \end{aligned}$$

Furthermore, the quadratic form $\mathscr{Q}$ satisfies

$$\mathscr{Q}(0, \dots, 0, \underbrace{1}_{r\text{-th entry}}, 0, \dots, 0) = \sum_{\ell=-r}^{p} \ell\, a_\ell = -\lambda\, a < 0\,.$$

We apply Lemma 3.3 to the right hand side in (3.10). The sum with respect to $j \le J$ makes the $\mathscr{Q}$ terms a telescopic sum, and we are left with the energy balance:

$$\begin{aligned} \sum_{j \le J} \Delta x\, (u_j^{n+1})^2 - \sum_{j \le J} \Delta x\, (u_j^n)^2 &= \sum_{j \le J} \sum_{\ell=1}^{p+r} d_\ell\, \Delta x\, (u_{j+\ell-r}^n - u_{j-r}^n)^2 \\ &+ \Delta x\, \mathscr{Q}(u_{J+2-r}^n - u_{J+1-r}^n, \dots, u_J^n - u_{J-1}^n, \underbrace{u_J^n}_{r\text{-th entry}}, u_{J+1}^n - u_J^n, \dots, u_{J+p}^n - u_{J+p-1}^n)\,. \end{aligned} \tag{3.11}$$

Let us start with the "boundary term" on the right hand side of (3.11). Thanks to the property of $\mathscr{Q}$ given in Lemma 3.3, we know that the coefficient of $(u_J^n)^2$ in the expression

$$\mathscr{Q}(u_{J+2-r}^n - u_{J+1-r}^n, \dots, u_J^n - u_{J-1}^n, u_J^n, u_{J+1}^n - u_J^n, \dots, u_{J+p}^n - u_{J+p-1}^n)$$

is $-\lambda\, a < 0$. Hence, by repeatedly applying the Young inequality for the cross terms in the quadratic form $\mathscr{Q}$, we get, for a suitable constant $C > 0$ (recall the relation $\lambda\, \Delta x = \Delta t$):

$$\begin{aligned} \sum_{j \le J} \Delta x\, (u_j^{n+1})^2 - \sum_{j \le J} \Delta x\, (u_j^n)^2 + \frac{a}{2}\, \Delta t\, (u_J^n)^2 \\ \le \sum_{\ell=1}^{p+r} d_\ell \sum_{j \le J} \Delta x\, (u_{j+\ell-r}^n - u_{j-r}^n)^2 + C\, \Delta t \sum_{\ell=2-r}^{p} (u_{J+\ell}^n - u_{J+\ell-1}^n)^2\,. \end{aligned} \tag{3.12}$$

It remains to estimate the "bulk" term on the right hand side of (3.12), which encodes the "ℓ^2-dissipation" of the discretization (2.3) of the transport equation. More precisely, if we start from a sequence $v \in \ell^2(\mathbb{Z})$ and consider its image by the contraction (3.4), we can rewrite Equation (3.10) for v and use the decomposition given in Lemma 3.3 to derive the inequality:

$$\forall\, v \in \ell^2(\mathbb{Z})\,, \quad \sum_{j \in \mathbb{Z}} \Delta x \left(\sum_{\ell=-r}^{p} a_\ell\, v_{j+\ell} \right)^2 - \sum_{j \in \mathbb{Z}} \Delta x\, (v_j)^2 = \sum_{\ell=1}^{p+r} d_\ell \sum_{j \in \mathbb{Z}} \Delta x\, (v_{j+\ell-r} - v_{j-r})^2 \le 0\,. \tag{3.13}$$

Let us now consider the sequence $(u_j^n)_{j\le J+p} \in \ell^2$ for the given integer $n \in \mathbb{N}$. We extend the latter sequence as an ℓ^2 sequence on the whole set of integers $\mathbb{Z}$ by symmetry with respect to the index $J+p$:

$$\forall\, j \in \mathbb{Z}\,, \quad v_j := \begin{cases} u_j^n\,, & \text{if } j \le J+p\,, \\ u^n_{2\,(J+p)-j}\,, & \text{if } j \ge J+p\,. \end{cases}$$

We then compute

$$\underbrace{\sum_{\ell=1}^{p+r} d_\ell \sum_{j\in\mathbb{Z}} \Delta x\,(v_{j+\ell-r} - v_{j-r})^2}_{\le 0 \text{ by } (3.13)}$$

$$= 2\sum_{\ell=1}^{p+r} d_\ell \sum_{j\le J} \Delta x\,(u^n_{j+\ell-r} - u^n_{j-r})^2 + \sum_{\ell=1}^{p+r} d_\ell\,\Delta x \sum_{j=J+1}^{J+2\,(p+r)-\ell-1} (v_{j+\ell-r} - v_{j-r})^2. \quad (3.14)$$

For $\ell = 1, \dots, p+r$ and $j = J+1, \dots, J+2\,(p+r)-\ell-1$, $v_{j+\ell-r}$ and v_{j-r} on the right hand side of (3.14) are taken among the real numbers $u^n_{J+1-r}, \dots, u^n_{J+p}$. We therefore obtain the upper bound:

$$\left| \sum_{\ell=1}^{p+r} d_\ell\,\Delta x \sum_{j=J+1}^{J+2\,(p+r)-\ell-1} (v_{j+\ell-r} - v_{j-r})^2 \right| \le C\,\Delta t \sum_{\ell=2-r}^{p} (u^n_{J+\ell} - u^n_{J+\ell-1})^2\,.$$

Using the latter bound in (3.14), we get the estimate

$$\sum_{\ell=1}^{p+r} d_\ell \sum_{j\le J} \Delta x\,(u^n_{j+\ell-r} - u^n_{j-r})^2 \le C\,\Delta t \sum_{\ell=2-r}^{p} (u^n_{J+\ell} - u^n_{J+\ell-1})^2\,,$$

which simplifies (3.12) into:

$$\sum_{j\le J} \Delta x\,(u_j^{n+1})^2 - \sum_{j\le J} \Delta x\,(u_j^n)^2 + \frac{a}{2}\,\Delta t\,(u_J^n)^2 \le C\,\Delta t \sum_{\ell=2-r}^{p} (u^n_{J+\ell} - u^n_{J+\ell-1})^2\,.$$

We now apply the summation argument of Lemma 3.2 to the latter inequality and derive the estimate:

$$\sup_{n\in\mathbb{N}} \left(\mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j\le J} \Delta x\,(u_j^n)^2 \right) + \sum_{n\ge 0} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,(u_J^n)^2$$

$$\le C \left\{ \sum_{j\le J} \Delta x\,(f_j)^2 + \sum_{n\ge 0} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{\ell=2-r}^{p} (u^n_{J+\ell} - u^n_{J+\ell-1})^2 \right\}.$$

We then combine the latter inequality with the preliminary estimate (3.9), which yields:

$$\sup_{n\in\mathbb{N}} \left(\mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{j\le J} \Delta x\,(u_j^n)^2 \right) + \sum_{n\ge 0} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t}\,(u_J^n)^2$$

$$\le C \left\{ \sum_{j\le J} \Delta x\,(f_j)^2 + \sum_{n\ge 0} \Delta t\,\mathrm{e}^{-2\,\gamma\,n\,\Delta t} \sum_{\ell=1}^{p} (g^n_{J+\ell})^2 \right\}. \quad (3.15)$$

At this stage, we have almost proved that (3.2) holds up to the index k_b+1. Indeed, if we combine the trace estimates provided by (3.9) and (3.15), we get a full control of the trace of (u^n), that is of $(u^n_{J+\ell})_{n\geq 0}$ at $\ell = -r-k_b, \dots, p$:

$$\sum_{n\geq 0} \Delta t\, \mathrm{e}^{-2\gamma n \Delta t} \sum_{\ell=-r-k_b}^{p} (u^n_{J+\ell})^2 \leq C \left\{ \sum_{j\leq J+p} \Delta x\, (f_j)^2 + \sum_{n\geq 1} \Delta t\, \mathrm{e}^{-2\gamma n \Delta t} \sum_{\ell=1}^{p} (g^n_{J+\ell})^2 \right\}. \tag{3.16}$$

Combining with (3.15), we have completed the proof of (3.2) for the index k_b+1, which also completes the proof of Theorem 3.1. □

3.2 Convergence estimates for Neumann outflow numerical boundary conditions

In the previous paragraph, we have proved the stability estimate (3.2) in order to highlight the fact that our method automatically yields the verification of the Uniform Kreiss Lopatinskii Condition. However, the exponential weights arising in (3.2) and the fact that no "interior" source term is considered in (3.1) make this estimate hardly applicable as such in view of the convergence analysis below. We therefore state a slightly weakened but more practical version of Theorem 3.1 which will help us proving Theorem 3.5 below.

Proposition 3.4. *Let $a>0$, let $k\geq 1$ and $k_b\in\mathbb{N}$, let $\lambda>0$. Then there exists a constant $C>0$ such that for all initial data $(f_j)_{j\leq J}\in\ell^2$, for any $N\in\mathbb{N}^*$, for all boundary source terms $(g^n_{J+1})_{0\leq n\leq N-1},\dots,(g^n_{J+p})_{0\leq n\leq N-1}$, and all interior source terms $(F^n_j)_{j\leq J, 1\leq n\leq N}$ with $(F^n_j)_{j\leq J}\in\ell^2$ for all $n=1,\dots,N$, the solution $(u^n_j)_{j\leq J, 0\leq n\leq N}$ to the scheme:*

$$\begin{cases} u^0_j = f_j\,, & j\leq J\,, \\ (D_-^{k_b} u^n)_{J+\ell} = g^n_{J+\ell}\,, & n=0,\dots,N-1\,, \quad \ell=1,\dots,p\,, \\ u^{n+1}_j = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, u^n_{j+\ell} + \Delta t\, F^{n+1}_j\,, & n=0,\dots,N-1\,, \quad j\leq J\,, \end{cases} \tag{3.17}$$

where the coefficients a_ℓ satisfy Assumption 2.1 with integer k, satisfies the estimate:

$$\sup_{n=0,\dots,N} \sum_{j\leq J} \Delta x\, (u^n_j)^2 \leq C \left\{ \sum_{j\leq J} \Delta x\, (f_j)^2 + (N\,\Delta t)^2 \sup_{n=1,\dots,N} \sum_{j\leq J} \Delta x\, (F^n_j)^2 + \sum_{n=0}^{N-1} \Delta t \sum_{\ell=1}^{p} (g^n_{J+\ell})^2 \right\} \tag{3.18}$$

for any Δt and Δx such that $\Delta t/\Delta x = \lambda$ and $\Delta t\in(0,1]$.

Proof. By linearity of (3.17), it is sufficient to examine separately the cases $F\equiv 0$ (no interior source term), and $f\equiv 0$, $g\equiv 0$ (interior forcing only). In the case $F\equiv 0$, the estimate (3.18) is directly obtained by:

- first extending the boundary source terms $(g^n_{J+\ell})_{\ell=1,\dots,p}$ by 0 for $n>N-1$, which does not affect the solution to (3.17) for $j\leq J$ at time steps earlier than N,
- then passing to the limit $\gamma\to 0$ in (3.2) (and forgetting about the nonnegative trace estimate on the left hand side of (3.2)).

It therefore remains to examine the case $f \equiv 0$, $g \equiv 0$, which is done by using the Duhamel formula. Namely, Theorem 3.1 shows that the solution $(v_j^n)_{j\le J+p, n\in\mathbb{N}}$ to the numerical scheme with homogeneous boundary conditions:

$$\begin{cases} v_j^0 = f_j\,, & j \le J\,, \\ (D_-^{k_b} v^n)_{J+\ell} = 0\,, & n \in \mathbb{N}\,, \quad \ell = 1, \dots, p\,, \\ v_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, v_{j+\ell}^n\,, & n \in \mathbb{N}\,, \quad j \le J\,, \end{cases}$$

satisfies the uniform in time bound

$$\sup_{n\in\mathbb{N}} \sum_{j\le J} \Delta x\, (v_j^n)^2 \le C \sum_{j\le J} \Delta x\, (f_j)^2\,.$$

Writing v^n under the form $\mathscr{S}^n f$, this means that the operator $\mathscr{S}$, which is the generator of the discrete semigroup $(\mathscr{S}^n)_{n\in\mathbb{N}}$, is power bounded on $\ell^2(-\infty, J)$. At this stage, it remains to observe that the solution to (3.17) in the case $f \equiv 0$, $g \equiv 0$, can be written with the Duhamel formula:

$$\forall\, n = 0, \dots, N\,, \quad u^n = \sum_{m=1}^{n} \Delta t\, \mathscr{S}^{n-m} F^m\,.$$

Since $\mathscr{S}$ is power bounded on ℓ^2, we end up with:

$$\sup_{n=0,\dots,N} \left(\sum_{j\le J} \Delta x\, (u_j^n)^2 \right)^{1/2} \le C \sum_{n=1}^{N} \Delta t \left(\sum_{j\le J} \Delta x\, (F_j^n)^2 \right)^{1/2} \le C\, N\, \Delta t \sup_{n=1,\dots,N} \left(\sum_{j\le J} \Delta x\, (F_j^n)^2 \right)^{1/2}.$$

This completes the proof of (3.18). □

We are now ready to prove our convergence result for the Neumann boundary condition at an outflow boundary.

Theorem 3.5. *Let $a > 0$, let $k \in \mathbb{N}^*$ and $k_b \in \mathbb{N}$, let $\lambda > 0$. Then there exists $C > 0$ such that for any $T > 0$ and $J \in \mathbb{N}^*$, for any $L \ge 1$, and for any $u_0 \in H^{k+1}((-\infty, L))$, the solution $(u_j^n)_{j\le J, n\in\mathbb{N}}$ to*

$$\begin{cases} u_j^0 = u_j^{\mathrm{in}} := \dfrac{1}{\Delta x} \int_{x_{j-1}}^{x_j} u_0(y)\, \mathrm{d}y\,, & j \le J\,, \\ (D_-^{k_b} u^n)_{J+\ell} = 0\,, & 0 \le n \le T/\Delta t - 1\,, \quad \ell = 1, \dots, p\,, \\ u_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, u_{j+\ell}^n\,, & 0 \le n \le T/\Delta t - 1\,, \quad j \le J\,, \end{cases} \tag{3.19}$$

where the coefficients a_ℓ satisfy Assumption 2.1 with integer k, satisfies, with $k_0 := \min(k, k_b)$:

$$\sup_{0\le n\le T/\Delta t} \left(\sum_{j\le J} \Delta x \left(u_j^n - \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} u_0(y - a\, t^n)\, \mathrm{d}y \right)^2 \right)^{1/2} \le C \left(\sqrt{T} + T\right) \Delta x^{k_0} \left\| u_0 \right\|_{H^{k_0+1}((-\infty, L))}\,, \tag{3.20}$$

where $\Delta x = L/J$ and $\Delta t = \lambda\, \Delta x$ (assumed to be less than 1). In particular, the following global ℓ^∞ convergence estimate is satisfied:

$$\sup_{0\le n\le T/\Delta t} \sup_{j\le J} \left| u_j^n - \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} u_0(y - a\, t^n)\, \mathrm{d}y \right| \le C \left(\sqrt{T} + T\right) \Delta x^{k_0 - 1/2} \left\| u_0 \right\|_{H^{k_0+1}((-\infty, L))}\,. \tag{3.21}$$

Proof. To pass from (3.20) to (3.21) is very crude, and simply relies on the inequality:

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

for any real sequence $(b_j)_{j \le J}$. We therefore focus on the derivation of the $\ell_n^\infty(\ell_j^2)$ estimate (3.20) which unsurprisingly relies on the stability estimate (3.18) provided by Proposition 3.4. In the proof below, we assume $k_b \le k$, so that the limiting order of convergence arises from the numerical boundary conditions in (3.19) and not from the discretization of the transport equation. The proof in the case $k_b > k$ is quite similar and we leave the corresponding modifications to the interested reader. Since the validity of Assumption 2.1 for some integer $k \ge 1$ implies the validity of the relations (2.4) for the restricted subset of indices $m = 0, \dots, k_b$, we can even assume without loss of generality $k_b = k$.

We now denote by U_0 the extension of u_0 as a function in $H^{k+1}(\mathbb{R})$ by the linear continuous reflexion operator of [DL85], and define, for any $j \in \mathbb{Z}$,

$$w_j^n := \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} U_0(y - a\, t^n)\, \mathrm{d}y\,, \tag{3.22}$$

to be the cell average of the exact solution $((t,x) \longmapsto U_0(x - a\,t))$ to the transport equation over $\mathbb{R}$.

For the rest of this proof, we will denote the integer part of $T/\Delta t$ by N_T. Then for all $n = 0, \dots, N_T$ and $j \le J + p$, we define the error $\varepsilon_j^n := u_j^n - w_j^n$. By definition of the sequences (u_j^n) and (w_j^n), the error $(\varepsilon_j^n)_{j \le J, n = 0, \dots, N_T}$ is a solution to:

$$\begin{cases} \varepsilon_j^0 = \varepsilon_j^{\mathrm{in}}\,, & j \le J\,, \\ (D_-^{k_b} \varepsilon^n)_{J+\ell} = -(D_-^{k_b} w^n)_{J+\ell}\,, & 0 \le n \le N_T - 1\,, \quad \ell = 1, \dots, p\,, \\ \varepsilon_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, \varepsilon_{j+\ell}^n + \Delta t\, e_j^{n+1}\,, & 0 \le n \le N_T - 1\,, \quad j \le J\,, \end{cases} \tag{3.23}$$

where the initial condition $(\varepsilon_j^{\mathrm{in}})_{j \le J}$ in (3.23) satisfies $\varepsilon_j^{\mathrm{in}} = 0$ for any $j \le J$, and the interior consistency error $(e_j^n)_{j \le J, n = 1, \dots, N_T}$ in (3.23) is given by

$$e_j^n := -\frac{1}{\Delta t} \left(w_j^n - \sum_{\ell=-r}^{p} a_\ell\, w_{j+\ell}^{n-1} \right) \tag{3.24}$$

for any $1 \le n \le N_T$ and any $j \le J$. To prove the convergence estimate (3.21), the first task is to evaluate the size of the consistency errors in (3.23). The boundary errors in (3.23) are dealt with by the following elementary result.

Lemma 3.6. *Let $m, p \in \mathbb{N}$. There exists a constant $C > 0$ such that for any $\Delta x > 0$ and for any $v \in H^{m+1}((-\infty, L + p\,\Delta x))$, defining*

$$v_j := \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} v(y)\, \mathrm{d}y\,, \quad j \le J + p\,,$$

one has

$$\left| (D_-^m v)_{J+\ell} \right| \le C\, \Delta x^m\, \| v^{(m)} \|_{H^1((-\infty, L + p\,\Delta x))}\,, \quad \ell = 1, \dots, p\,.$$

Proof of Lemma 3.6. For $\ell = 1, \dots, p$, there holds

$$
\begin{aligned}
(D_-^m v)_{J+\ell} &= \sum_{m'=0}^{m} \binom{m}{m'} (-1)^{m-m'} \frac{1}{\Delta x} \int_{x_{J-m-1+\ell}}^{x_{J-m+\ell}} v(y + m'\,\Delta x)\,\mathrm{d}y \\
&= \sum_{m'=0}^{m} \binom{m}{m'} \frac{(-1)^{m-m'}}{(m-1)!} \frac{1}{\Delta x} \int_{x_{J-m-1+\ell}}^{x_{J-m+\ell}} \int_{y}^{y+m'\,\Delta x} v^{(m)}(z)\,(y + m'\,\Delta x - z)^{m-1}\,\mathrm{d}z\,\mathrm{d}y\,,
\end{aligned}
$$

where we have used Taylor's formula and cancellation properties of the binomial coefficients. Applying the Cauchy-Schwarz inequality to each double integral in the latter expression, we get

$$
\begin{aligned}
|(D_-^m v)_{J+\ell}| &\le C \sum_{m'=0}^{m} \left(\int_{x_{J-m-1+\ell}}^{x_{J-m+\ell}} \int_{y}^{y+m'\,\Delta x} v^{(m)}(z)^2\,(y + m'\,\Delta x - z)^{2\,m-2}\,\mathrm{d}z\,\mathrm{d}y \right)^{1/2} \\
&\le C\,\Delta x^{m-1/2} \left(\int_{x_{J-m-1+\ell}}^{x_{J+\ell}} v^{(m)}(z)^2\,\mathrm{d}z \right)^{1/2} \le C\,\Delta x^m\,\|v^{(m)}\|_{L^\infty(-\infty, L+p\Delta x)}\,.
\end{aligned}
$$

The proof follows by using the imbedding of H^1 in L^∞ in one space dimension. □

We now apply Lemma 3.6 with $m = k_b$ to evaluate the boundary errors in (3.23). We get:

$$
\begin{aligned}
\sum_{n=0}^{N_T-1} \Delta t \sum_{\ell=1}^{p} \left((D_-^{k_b} \varepsilon^n)_{J+\ell}\right)^2 &= \sum_{n=0}^{N_T-1} \Delta t \sum_{\ell=1}^{p} \left((D_-^{k_b} w^n)_{J+\ell}\right)^2 \\
&\le C\,N_T\,\Delta t\,\Delta x^{2\,k_b}\,\|U_0(\cdot - a t^n)\|^2_{H^{k_b+1}((-\infty, L+p\Delta x))} \\
&\le C\,T\,\Delta x^{2\,k_b}\,\|u_0\|^2_{H^{k_b+1}((-\infty, L))} \qquad (3.25)
\end{aligned}
$$

thanks to the continuity of the reflexion operator. Using (3.25) in the stability estimate (3.18) provided by Proposition 3.4, we already get

$$
\sup_{n=0,\dots,N_T} \sum_{j \le J} \Delta x\,(\varepsilon_j^n)^2 \le C \left\{ T^2 \sup_{n=1,\dots,N_T} \sum_{j \le J} \Delta x\,(e_j^n)^2 + \Delta x^{2\,k_b}\,T\,\|u_0\|^2_{H^{k_b+1}((-\infty, L))} \right\}, \qquad (3.26)
$$

so the remaining task is to evaluate the size of the interior consistency error $(e_j^n)_{j \le J}$.

By its definition (3.22), the value of w_j^n in each interior cell is the average of the extension of the exact solution, that is the cell average of $U_0(\cdot - a\,t^n)$. We thus have, for any $j \le J$ and $n \ge 1$:

$$
\begin{aligned}
e_j^n &= \frac{1}{\Delta t} \left(\frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} U_0(y - a\,\Delta t - a\,t^{n-1})\,\mathrm{d}y - \sum_{\ell=-r}^{p} \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} a_\ell\,U_0(y + \ell\,\Delta x - a\,t^{n-1})\,\mathrm{d}y \right) \\
&= \frac{1}{\Delta t\,\Delta x} \int_{x_{j-1}}^{x_j} \left(U_0(y - a\,\Delta t - a\,t^{n-1}) - \sum_{\ell=-r}^{p} a_\ell\,U_0(y + \ell\,\Delta x - a\,t^{n-1}) \right) \mathrm{d}y\,, \\
&= \frac{1}{\Delta t\,\Delta x} \int_{x_{j-1}}^{x_j} \left(U^{n-1}(y - a\,\Delta t) - \sum_{\ell=-r}^{p} a_\ell\,U^{n-1}(y + \ell\,\Delta x) \right) \mathrm{d}y
\end{aligned}
$$

where we defined $U^{n-1} \in H^{k+1}(\mathbb{R})$ as $U_0(\cdot - a\,t^{n-1})$. We also define the consistency error on the whole domain:

$$\forall\, j \in \mathbb{Z}\,, \quad E_j^n := \frac{1}{\Delta t\,\Delta x} \int_{x_{j-1}}^{x_j} \left(U^{n-1}(y - a\,\Delta t) - \sum_{\ell=-r}^{p} a_\ell\, U^{n-1}(y + \ell\,\Delta x) \right) \mathrm{d}y\,,$$

so that we have $e_j^n = E_j^n$ for all $j \le J$. We now use Fourier analysis and derive the estimate:

$$\begin{aligned}
\sum_{j\le J} \Delta x\,(e_j^n)^2 \le \sum_{j\in\mathbb{Z}} \Delta x\,(E_j^n)^2 &\le \frac{1}{\Delta t^2} \sum_{j\in\mathbb{Z}} \frac{1}{\Delta x} \left(\int_{x_{j-1}}^{x_j} \left(U^{n-1}(y - a\,\Delta t) - \sum_{\ell=-r}^{p} a_\ell\, U^{n-1}(y + \ell\,\Delta x) \right) \mathrm{d}y \right)^2 \\
&\le \frac{1}{\Delta t^2} \int_{\mathbb{R}} \left(U^{n-1}(y - a\,\Delta t) - \sum_{\ell=-r}^{p} a_\ell\, U^{n-1}(y + \ell\,\Delta x) \right)^2 \mathrm{d}y \\
&\le \frac{1}{2\,\pi\,\Delta t^2} \int_{\mathbb{R}} \left| \widehat{U^{n-1}}(\xi) \right|^2 \left| \mathrm{e}^{-i\,\lambda\,a\,\Delta x\,\xi} - \sum_{\ell=-r}^{p} a_\ell\, \mathrm{e}^{i\,\ell\,\Delta x\,\xi} \right|^2 \mathrm{d}\xi\,.
\end{aligned}$$

We now use Assumption 2.1 (recall $k_b = k$), to derive the bound:

$$\left| \mathrm{e}^{-i\,\lambda\,a\,\Delta x\,\xi} - \sum_{\ell=-r}^{p} a_\ell\, \mathrm{e}^{i\,\ell\,\Delta x\,\xi} \right| \le C\,(\Delta x\,|\xi|)^{k_b+1}\,,$$

uniformly with respect to Δx and $\xi \in \mathbb{R}$. Recalling that the ratio $\Delta t/\Delta x$ is kept fixed, we end up with the consistency estimate:

$$\sum_{j\le J} \Delta x\,(e_j^n)^2 \le C\,\Delta x^{2\,k_b}\,\|U^{n-1}\|^2_{H^{k_b+1}(\mathbb{R})} = C\,\Delta x^{2\,k_b}\,\|U_0\|^2_{H^{k_b+1}(\mathbb{R})} \le C\,\Delta x^{2\,k_b}\,\|u_0\|^2_{H^{k_b+1}((-\infty,L))}\,, \quad (3.27)$$

where we have used the fact that U^{n-1} has been constructed as an extension of $u_0(\cdot - a\,t^{n-1})$ boundedly with respect to the H^{k_b+1} norm. Going back to (3.26) and using the estimate (3.27), we finally derive

$$\sup_{n=0,\dots,N_T} \sum_{j\le J} \Delta x\,(\varepsilon_j^n)^2 \le C\,\Delta x^{2\,k_b}\,(T + T^2)\,\|u_0\|^2_{H^{k_b+1}((-\infty,L))}\,,$$

which completes the proof of (3.20). □

3.3 Convergence estimates for Dirichlet numerical boundary conditions

This short paragraph is devoted to the complementary boundary value problem on a half-line that can be extracted from (2.3), (2.6), (2.7), and that focuses on the (homogeneous) Dirichlet boundary condition (2.6) at the inflow boundary $x = 0$. Hence the half-line is now $\mathbb{R}^+$. To be consistent with the notations introduced in Section 2, we let $x_j := j\,\Delta x$, $j \in \mathbb{Z}$. The ghost cells correspond to the intervals $(x_{-r}, x_{1-r}), \dots, (x_{-1}, x_0)$, see Figure 3.2. The time step Δt is linked to the space step by keeping $\Delta t/\Delta x = \lambda$, λ being the same fixed positive constant as in Section 2. We prove the following result, that is the analogue of Theorem 3.5.

Proposition 3.7. *Let $a > 0$, let $k \in \mathbb{N}^*$, let $\lambda > 0$. There exists $C > 0$ such that for any $T > 0$ and $J \in \mathbb{N}^*$, for any $L \ge 1$ and for any $u_0 \in H^{k+1}((0,+\infty))$ satisfying the flatness conditions:*

$$u_0(0) = \cdots = u_0^{(k)}(0) = 0\,,$$

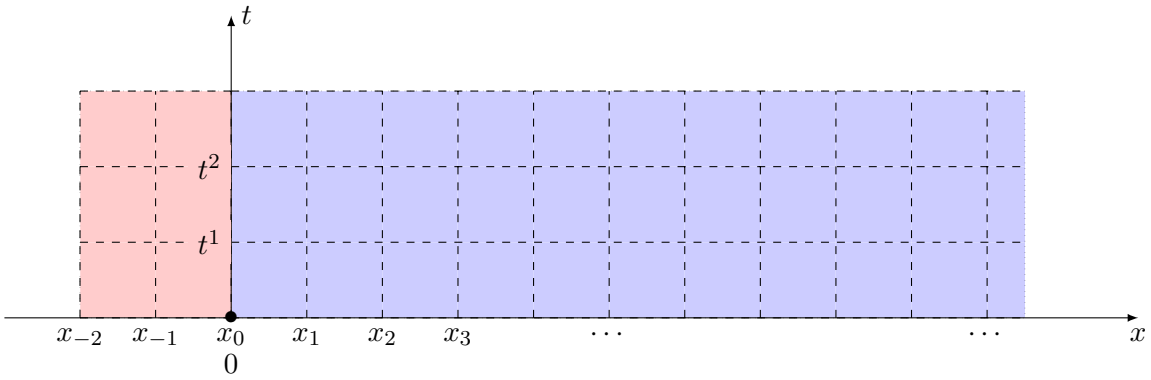


Figure 3.2: The mesh on $\mathbb{R}^+$ in blue, and the "ghost cells" in red ($r = 2$ here).

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

$$\begin{cases} u_j^0 = u_j^{\mathrm{in}} := \dfrac{1}{\Delta x}\int_{x_{j-1}}^{x_j} u_0(y)\,\mathrm{d}y\,, & j \geq 1, \\ u_\ell^n = 0\,, & 0 \leq n \leq T/\Delta t - 1\,, \quad \ell = 1-r,\dots,0\,, \\ u_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, u_{j+\ell}^n\,, & 0 \leq n \leq T/\Delta t - 1\,, \quad j \geq 1\,, \end{cases} \tag{3.28}$$

where the coefficients a_ℓ *satisfy Assumption 2.1 with integer* k*, satisfies*

$$\sup_{0\leq n\leq T/\Delta t} \left(\sum_{j\geq 1} \Delta x \left(u_j^n - \frac{1}{\Delta x}\int_{x_{j-1}}^{x_j} u_0(y - a\,t^n)\,\mathrm{d}y \right)^2 \right)^{1/2} \leq C\,T\,\Delta x^k\,\|u_0\|_{H^{k+1}((0,+\infty))} \tag{3.29}$$

where $\Delta x = L/J$ *and* $\Delta t = \lambda\,\Delta x$*. It is understood in (3.29) that the initial condition* u_0 *has been extended by* 0 *to* $\mathbb{R}^-$*. In particular, the following global* ℓ^∞ *convergence estimate is satisfied:*

$$\sup_{0\leq n\leq T/\Delta t}\ \sup_{j\geq 1} \left| u_j^n - \frac{1}{\Delta x}\int_{x_{j-1}}^{x_j} u_0(y - a\,t^n)\,\mathrm{d}y \right| \leq C\,T\,\Delta x^{k-1/2}\,\|u_0\|_{H^{k+1}((0,+\infty))}\,. \tag{3.30}$$

The proof follows more or less the exact same arguments as the proof of Theorem 3.5. One first defines the cell average of the exact solution $((t,x)\mapsto u_0(x-a\,t))$ to the transport equation in the half space $\mathbb{R}^+$ with homogeneous Dirichlet boundary condition at $x=0$. The flatness assumption on u_0 ensures that extending u_0 by 0 on $\mathbb{R}^-$ yields a function in $H^{k+1}(\mathbb{R})$. The error between the solution to (3.28) and the cell average of the exact solution will satisfy a system of the form (3.28) with *zero* initial condition and *homogeneous* Dirichlet boundary condition but with a nonzero interior forcing term (the so-called consistency error). Estimating this error is done by using a similar (and even easier) extension procedure as in the proof of Theorem 3.5 and Fourier analysis. We leave the details to the interested reader. The stability estimate for (3.28) is analogous to that of Proposition 3.4 and follows from Theorem 3.1 (in the case $k_b = 0$) since we have already observed that dealing with this case does not necessitate the consistency conditions (2.4) but only the stability assumption (2.5).

4 Proof of Theorem 2.2

4.1 Stability estimates on a finite interval

We now turn to the study of the numerical scheme (2.3), (2.6), (2.7), which is an iteration in a finite dimensional space and therefore really corresponds to a numerical scheme that can be implemented in practice. Let us recall that the space step Δx is given by $\Delta x = L/J$ with J a positive integer, the spatial grid is defined by means of the points $x_j := j\,\Delta x$, $j \in \mathbb{Z}$, and the time step is given by $\Delta t = \lambda\,\Delta x$ with λ a fixed positive constant. We first prove a stability estimate for (2.3), (2.6), (2.7), which will have various consequences.

Proposition 4.1. *Let $a > 0$, let $k \geq 1$ and $k_b \in \mathbb{N}$, let $\lambda > 0$. Then there exists a constant $C_0 > 0$ such that for all initial data $(f_1, \dots, f_J)$, the solution $(u_j^n)_{1 \leq j \leq J, n \in \mathbb{N}}$ to the scheme:*

$$\begin{cases} u_j^0 = f_j\,, & j = 1, \dots, J, \\ u_\ell^n = 0\,, & n \in \mathbb{N}\,, \quad \ell = 1 - r, \dots, 0\,, \\ (D_-^{k_b} u^n)_{J+\ell} = 0\,, & n \in \mathbb{N}\,, \quad \ell = 1, \dots, p\,, \\ u_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, u_{j+\ell}^n\,, & n \in \mathbb{N}\,, \quad j = 1, \dots, J\,, \end{cases} \tag{4.1}$$

where the coefficients a_ℓ satisfy Assumption 2.1 with integer k, satisfies

$$\forall\, n \in \mathbb{N}\,, \quad \left(\sum_{j=1}^{J} \Delta x\, (u_j^n)^2 \right)^{1/2} \leq C_0\, \mathrm{e}^{C_0\, n\, \Delta t / L} \left(\sum_{j=1}^{J} \Delta x\, (f_j)^2 \right)^{1/2}. \tag{4.2}$$

The estimate (4.2) is compatible with the limit $L \to +\infty$ (and Δx, Δt fixed) where we formally recover either the stability estimate (3.18) or its analogue for Dirichlet boundary conditions on a half-line, at least in the case of homogeneous boundary conditions and no interior source term. More important, the estimate (4.2) also has a consequence in terms of the spectral radius of the perturbed Toeplitz matrix associated with (4.1). In the absence of the outflow boundary at $x = L$, the numerical scheme (3.28) with homogeneous Dirichlet boundary conditions at the inflow boundary $x = 0$ corresponds to iterating the Toeplitz operator represented by the semi-infinite matrix:

$$\begin{pmatrix} a_0 & \cdots & a_p & 0 & \cdots & \cdots & \cdots \\ \vdots & \ddots & & \ddots & \ddots & & \\ a_{-r} & \cdots & a_0 & \cdots & a_p & \ddots & \\ 0 & \ddots & & \ddots & & \ddots & \ddots \\ \vdots & \ddots & \ddots & & \ddots & & \ddots \end{pmatrix}.$$

Acting on $\ell^2(\mathbb{N})$, the spectrum of this operator is known to contain at least the closed curve $\{\sum a_\ell \exp(i\,\ell\,\theta)$, $\theta \in [0, 2\,\pi]\}$, see [TE05, Chapter 7] and references therein for a precise statement. In particular, Assumption 2.1 and [TE05, Theorem 7.1] show that the spectrum of this operator lies inside the closed unit disk $\{z \in \mathbb{C}\,,\ |z| \leq 1\}$ and accumulates at the point 1. Truncating the above semi-infinite matrix to make it of size J corresponds to prescribing homogeneous Dirichlet boundary conditions at the outflow boundary. In

that case, the resulting operator is of the form $\Pi\, T\, \Pi$ with Π an orthogonal projection and T a contraction. In particular, prescribing homogeneous Dirichlet boundary conditions at the outflow boundary leaves the spectrum (consisting now of finitely many eigenvalues) within the closed unit disk.

Prescribing Neumann numerical boundary conditions as in (4.1) corresponds to first truncating the semi-infinite matrix to make it of size J, and then making "large" perturbations of the coefficients in the k_b last columns of the last p rows of the truncated matrix. For instance, if $r = p = 1$ and $k_b = 1$, resp. $k_b = 2$, the corresponding matrix reads:

$$\begin{pmatrix} a_0 & a_1 & 0 & \cdots & 0 \\ a_{-1} & a_0 & a_1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & a_0 & a_1 \\ 0 & \cdots & 0 & a_{-1} & a_0 + a_1 \end{pmatrix}, \quad \text{resp.} \quad \begin{pmatrix} a_0 & a_1 & 0 & \cdots & 0 \\ a_{-1} & a_0 & a_1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & a_{-1} & a_0 & a_1 \\ 0 & \cdots & 0 & a_{-1} - a_1 & a_0 + 2\, a_1 \end{pmatrix}.$$

It is not obvious at first sight that the spectrum of the matrix $\mathscr{A}_J \in \mathscr{M}_J(\mathbb{R})$ associated with the iteration (4.1) will not deviate from the closed unit disk at a distance $O(1)$, or even $O(J^{-\alpha})$ for some $\alpha \in (0,1)$, giving rise to instabilities for (4.1). However, the estimate (4.2) has an important consequence for the matrix $\mathscr{A}_J$ since it shows in particular that its ℓ^2 induced norm verifies

$$\forall\, n \in \mathbb{N}\,, \quad \|\mathscr{A}_J^n\|_2 \leq C\, \mathrm{e}^{C\, n\, \Delta t/L}\,,$$

so taking the $1/n$-th power and passing to the limit (recalling $\Delta t/L = \lambda/J$, $\lambda > 0$ constant), the spectral radius of $\mathscr{A}_J$ satisfies

$$\rho(\mathscr{A}_J) \leq \exp(C/J)\,.$$

Hence the spectrum of $\mathscr{A}_J$ does not deviate from the closed unit disk at a distance larger than $O(1/J)$. (The bound in (4.2) is even more precise since it even predicts the behavior of the ε-pseudospectrum of $\mathscr{A}_J$, see [TE05], prohibiting in particular Jordan blocks associated with eigenvalues of largest modulus.)

Proof of Proposition 4.1. The derivation of (4.2) follows from the finite speed of propagation of the numerical scheme (2.3). (Observe that our argument below does not extend to implicit discretizations of the transport equation.) More precisely, let us assume for simplicity that J is even. Let

$$N_0 := \min\left(\mathrm{E}\left(\frac{J}{2\,p}\right), \mathrm{E}\left(\frac{J/2 - k_b}{r}\right)\right).$$

Then for $n \leq N_0$, we can write the solution to (4.1) as the superposition of solutions to two initial boundary value problems of the form (3.19) and (3.28). Indeed, if the initial condition $(f_j = u_j^{\mathrm{in}})_{j \leq J}$ in (3.19) vanishes for $j \leq J/2$, then the solution to (3.19) satisfies the homogeneous Dirichlet boundary condition:

$$\forall\, n \leq N_0\,, \quad \forall\, \ell = 1 - r, \ldots, 0\,, \quad u_\ell^n = 0\,,$$

because the support of u^{in} is shifted of p cells to the left at each time iteration. In particular, the restriction of the solution to the half-line problem (3.19) with some initial condition u^{in} vanishing for $j \leq J/2$ satisfies the numerical scheme (4.1) up to $n = N_0$. Similarly, if the initial condition $(u_j^{\mathrm{in}})_{j \geq 1}$ in (3.28) vanishes for $j \geq J/2 + 1$, then the solution to the half line problem (3.28) satisfies:

$$\forall\, n \leq N_0\,, \quad \forall\, \ell = 1, \ldots, p\,, \quad (D_-^{k_b}\, u^n)_{J+\ell} = 0\,,$$

because the support of $u^{\rm in}$ is shifted of r cells to the right at each time iteration, and the definition of N_0 ensures

$$\frac{J}{2} + 1 + N_0\, r \leq J + 1 - k_b\,,$$

so the solution to (3.28) vanishes at all points that are used in the calculation of the finite differences $(D_-^{k_b}\, u^n)_{J+\ell}$. Truncating the initial condition for (4.1) to the left and to the right of the medium cell $J/2$, we can, thanks to the linearity of the scheme, combine the stability estimate (3.18) and the analogous one for Dirichlet boundary conditions to show that the solution to (4.1) satisfies:

$$\sup_{n=0,\dots,N_0} \sum_{j=1}^{J} \Delta x\, (u_j^n)^2 \leq C \sum_{j=1}^{J} \Delta x\, (f_j)^2\,,$$

with a constant C that is independent of L and J. It then remains to iterate the latter estimate on each interval of integers $\{0,\dots,N_0\}$, $\{N_0,\dots,2\,N_0\}$ and so on to derive (4.2). □

4.2 Convergence

We now prove Theorem 2.2, whose proof combines the convergence estimates given by Theorem 3.5 and Proposition 3.7, and the stability estimate of Proposition 4.1. Let us recall that we consider for the continuous problem transport an initial condition $u_0 \in H^{k+1}((0,L))$ that satisfies the flatness conditions:

$$u_0(0) = \cdots = u_0^{(k)}(0) = 0\,.$$

We therefore extend u_0 by zero on $\mathbb{R}^-$, considering now u_0 as an element of $H^{k+1}((-\infty,L))$ that vanishes on $\mathbb{R}^-$. Let us consider a $\mathscr{C}^\infty$ function χ on $\mathbb{R}$ that satisfies:

$$\chi(x) = \begin{cases} 0\,, & \text{if } x \leq 1/3, \\ 1\,, & \text{if } x \geq 2/3, \end{cases}$$

and then decompose u_0 as follows:

$$\forall\, x < L\,, \quad u_0(x) = \underbrace{\chi(x/L)\, u_0(x)}_{=:\, v_0(x)} + \underbrace{(1-\chi(x/L))\, u_0(x)}_{=:\, w_0(x)}\,.$$

The function v_0 belongs to $H^{k+1}((-\infty,L))$ and vanishes for $x \leq L/3$. Conversely, the function w_0 belongs to $H^{k+1}((-\infty,L))$ and vanishes for $x \geq 2\,L/3$. Moreover, using $L \geq 1$, there holds uniformly in L:

$$\|v_0\|_{H^{k+1}((-\infty,L))} \leq C\, \|u_0\|_{H^{k+1}((0,L))}\,, \quad \|w_0\|_{H^{k+1}((-\infty,L))} \leq C\, \|u_0\|_{H^{k+1}((0,L))}\,,$$

where C depends only on the fixed function χ (and on k).

Let $N \in \mathbb{N}$ denote the largest integer such that

$$N\, p \leq \frac{J}{3}\,, \quad \text{and} \quad N\, r \leq \frac{J}{3} - k_b\,.$$

We assume that J is large enough so that $N \geq 1$. Applying the same finite speed of propagation argument as in the proof of Proposition 4.1, we can write

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

where $(v_j^n)_{j\le J, n=0,\dots,N}$ satisfies an iteration of the form (3.19), namely:

$$\begin{cases} v_j^0 = v_j^{\rm in}\,, & j \le J\,, \\ (D_-^{k_b} v^n)_{J+\ell} = 0\,, & n = 0,\dots,N-1\,, \quad \ell = 1,\dots,p\,, \\ v_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, v_{j+\ell}^n\,, & n = 0,\dots,N-1\,, \quad j \le J\,, \end{cases} \tag{4.3}$$

and $(w_j^n)_{j\ge 1, n=0,\dots,N}$ satisfies an iteration of the form (3.28), namely:

$$\begin{cases} w_j^0 = w_j^{\rm in}\,, & j \ge 1 \\ w_\ell^n = 0\,, & n = 0,\dots,N-1\,, \quad \ell = 1-r,\dots,0\,, \\ w_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, w_{j+\ell}^n\,, & n = 0,\dots,N-1\,, \quad j \ge 1\,. \end{cases} \tag{4.4}$$

The initial conditions in (4.3) and (4.4) are defined as follows. Following (2.8), we define the discretized initial condition associated with v_0 for (4.3) by letting[1]:

$$\forall\, j \le J\,, \quad v_j^{\rm in} := \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} v_0(y)\,{\rm d}y\,.$$

To define the initial condition for (4.4), let us recall that $1-\chi(\cdot/L)$ vanishes on the interval $(2\,L/3, L)$ so we can extend w_0 by zero to $(L,+\infty)$ and view w_0 as an element of $H^{k+1}((0,\infty))$ that satisfies the flatness conditions of Proposition 3.7. We then define the initial condition for (4.4) by letting:

$$\forall\, j \ge 1\,, \quad w_j^{\rm in} := \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} w_0(y)\,{\rm d}y\,.$$

We can apply Theorem 3.5 to (4.3) and get the convergence estimate:

$$\begin{aligned} \sup_{0\le n\le N} \left(\sum_{j=1}^{J} \Delta x \left(v_j^n - \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} v_0(y - a\,t^n)\,{\rm d}y \right)^2 \right)^{1/2} \\ \le C\,(\sqrt{N\,\Delta t} + N\,\Delta t)\,\Delta x^{k_0}\,\|v_0\|_{H^{k+1}((-\infty,L))} \\ \le C\,(\sqrt{N\,\Delta t} + N\,\Delta t)\,\Delta x^{k_0}\,\|u_0\|_{H^{k+1}((0,L))}\,, \end{aligned} \tag{4.5}$$

with $k_0 := \min(k, k_b)$ and where the estimate is uniform with respect to L. Applying Proposition 3.7 to (4.4), we get the other convergence estimate:

$$\sup_{0\le n\le N} \left(\sum_{j=1}^{J} \Delta x \left(w_j^n - \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} w_0(y - a\,t^n)\,{\rm d}y \right)^2 \right)^{1/2} \le C\,N\,\Delta t\,\Delta x^k\,\|u_0\|_{H^{k+1}((0,L))}\,,$$

which we combine with (4.5) to derive:

$$\begin{aligned} \sup_{0\le n\le N} \left(\sum_{j=1}^{J} \Delta x \left(u_j^n - \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} u_0(y - a\,t^n)\,{\rm d}y \right)^2 \right)^{1/2} \\ \le C_1\,(\sqrt{N\,\Delta t} + N\,\Delta t)\,\Delta x^{k_0}\,\|u_0\|_{H^{k+1}((0,L))}\,. \end{aligned} \tag{4.6}$$

[1] Recall here that u_0 has been extended (by zero) to $\mathbb{R}^-$ so v_0 is well defined on $(-\infty, L)$.

Here we use a specific notation C_1 for the constant in (4.6) in order to emphasize its role in the concluding argument below (in the same way, we use a specific notation C_0 for the constant in the stability estimate (4.2)).

It remains more or less to iterate in time (4.6). Namely, at the time level N, (4.6) shows that we can write:

$$\forall\, j = 1, \dots, J\,, \quad u_j^N = \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} u_0(y - a\, t^N)\, \mathrm{d}y + \varepsilon_j\,,$$

with

$$\left(\sum_{j=1}^{J} \Delta x\, \varepsilon_j^2 \right)^{1/2} \leq C_1\, (\sqrt{N\, \Delta t} + N\, \Delta t)\, \Delta x^{k_0}\, \|u_0\|_{H^{k+1}((0,L))}\,.$$

To iterate the process, we define $(u_j^{N,n})_{1 \leq j \leq J, 0 \leq nN}$ by

$$\begin{cases} u_j^{N,0} = \dfrac{1}{\Delta x} \displaystyle\int_{x_{j-1}}^{x_j} u_0(y - a t^N)\, \mathrm{d}y\,, & j = 1, \dots, J, \\ u_\ell^{N,n} = 0\,, & 0 \leq n \leq N-1\,, \quad \ell = 1-r, \dots, 0\,, \\ (D_-^{k_b} u^{N,n})_{J+\ell} = 0\,, & 0 \leq n \leq N-1\,, \quad \ell = 1, \dots, p\,, \\ u_j^{N,n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, u_{j+\ell}^{N,n}\,, & 0 \leq n \leq N-1\,, \quad j = 1, \dots, J\,. \end{cases}$$

Note that $u_0(\cdot - a\, t^N)$ satisfies the flatness conditions of Theorem 2.2, and that:

$$\|u_0(\cdot - a t^N)\|_{H^{k+1}((0,L))} \leq \|u_0\|_{H^{k+1}((0,L))}\,.$$

Thus, on the one side, (4.6) applies and we have

$$\sup_{0 \leq n \leq N} \left(\sum_{j=1}^{J} \Delta x \left(u_j^{N,n} - \frac{1}{\Delta x} \int_{x_{j-1}}^{x_j} u_0(y - a(t^N + t^n)\, \mathrm{d}y \right)^2 \right)^{1/2} \leq C_1\, (\sqrt{N\, \Delta t} + N\, \Delta t)\, \Delta x^{k_0}\, \|u_0\|_{H^{k+1}((0,L))}\,.$$

On the other side, the sequence $(\delta_j^n := u_j^{N+n} - u_j^{N,n})_{1 \leq j \leq J, 0 \leq nN}$ satisfies:

$$\begin{cases} \delta_j^0 = \varepsilon_j\,, & j = 1, \dots, J, \\ \delta_\ell^n = 0\,, & 0 \leq n \leq N-1\,, \quad \ell = 1-r, \dots, 0\,, \\ (D_-^{k_b} \delta^n)_{J+\ell} = 0\,, & 0 \leq n \leq N-1\,, \quad \ell = 1, \dots, p\,, \\ \delta_j^{n+1} = \displaystyle\sum_{\ell=-r}^{p} a_\ell\, \delta_{j+\ell}^n\,, & 0 \leq n \leq N-1\,, \quad j = 1, \dots, J \end{cases}$$

so that, thanks to the stability estimate (4.2) of Proposition 4.1, for any $n \leq N$,

$$\left(\sum_{j=1}^{J} \Delta x\, (\delta_j^n)^2 \right)^{1/2} \leq C_0\, \mathrm{e}^{C_0\, n\, \Delta t/L} \left(\sum_{j=1}^{J} \Delta x\, (\varepsilon_j)^2 \right)^{1/2} \leq C_0\, \mathrm{e}^{C_0\, n\, \Delta t/L}\, C_1\, (\sqrt{N\, \Delta t} + N\, \Delta t)\, \Delta x^{k_0}\, \|u_0\|_{H^{k+1}((0,L))}\,.$$

Finally, thanks to the triangle inequality,

$$\sup_{0\le n\le N}\left(\sum_{j=1}^{J}\Delta x\left(u_j^{N+n}-\frac{1}{\Delta x}\int_{x_{j-1}}^{x_j}u_0(y-a(t^N+t^n)\,\mathrm{d}y\right)^2\right)^{1/2}$$
$$\le C_1(1+C_0\,\mathrm{e}^{C_0\,N\,\Delta t/L})\,(\sqrt{N\,\Delta t}+N\,\Delta t)\,\Delta x^{k_0}\,\|u_0\|_{H^{k+1}((0,L))}\,.$$

Applying iteratively the same argument, we end up with

$$\forall\, m\in\mathbb{N}\,,\quad \sup_{0\le n\le m\,N}\left(\sum_{j=1}^{J}\Delta x\left(u_j^{n}-\frac{1}{\Delta x}\int_{x_{j-1}}^{x_j}u_0(y-a\,t^n)\,\mathrm{d}y\right)^2\right)^{1/2}$$
$$\le C_1\sum_{\mu=0}^{m-1}\left(C_0\,\mathrm{e}^{C_0\,N\,\Delta t/L}\right)^{\mu}(\sqrt{N\,\Delta t}+N\,\Delta t)\,\Delta x^{k_0}\,\|u_0\|_{H^{k+1}((0,L))}$$
$$\le C_1\,C_0^m\,\mathrm{e}^{C_0\,m\,N\,\Delta t/L}\,(\sqrt{N\,\Delta t}+N\,\Delta t)\,\Delta x^{k_0}\,\|u_0\|_{H^{k+1}((0,L))}\,,$$

where we have assumed $C_0\ge 2$ without loss of generality. It remains to choose $m:=\mathrm{E}(N\Delta t/T)+1$, which by definition of N is uniformly bounded with respect to J (N scales like $c\,J$ with $c>0$ constant, and Δt scales like c'/J), and we end up with:

$$\sup_{0\le n\le T/\Delta t}\left(\sum_{j=1}^{J}\Delta x\left(u_j^{n}-\frac{1}{\Delta x}\int_{x_{j-1}}^{x_j}u_0(y-a\,t^n)\,\mathrm{d}y\right)^2\right)^{1/2}\le C\,(\sqrt{T}+T)\,\mathrm{e}^{C\,T/L}\,\Delta x^{k_0}\,\|u_0\|_{H^{k+1}((0,L))}\,.$$

The convergence estimate (2.9) of Theorem 2.2 follows by a direct lower bound for the norm on the left hand side.

5 Numerical experiments

In this section, we illustrate with numerical examples the results proved in the paper. We consider the second order in time and space Lax-Wendroff scheme, that writes

$$\frac{u_j^{n+1}-u_j^n}{\Delta t}+a\,\frac{u_{j+1}^n-u_{j-1}^n}{2\,\Delta x}=\frac{a^2\,\Delta t}{2}\,\frac{u_{j+1}^n-2\,u_j^n+u_{j-1}^n}{\Delta x^2},\quad n\in\mathbb{N}\,,\quad j=1,\dots,J\,,\tag{5.1}$$

which means, in the present formalism, that $r=p=1$ and $a_{-1}=(a^2\,\lambda^2+a\,\lambda)/2$, $a_0=1-a^2\,\lambda^2$, $a_1=(a^2\,\lambda^2-a\,\lambda)/2$ and $k=2$. This scheme is stable (over $\mathbb{Z}$) and second order accurate provided that $\lambda\,|a|\le 1$, and we choose $a=1$ and $\lambda=0.7$ in the numerical simulations reported below.

We compare the results with the first order Neumann outflow condition ($k_b=1$) and the second order one ($k_b=2$), for various initial conditions. Namely, we consider $L=1$ for the length of the space interval and we choose as initial data,

$$u^{0,1}(x)=((x-1/2)^+)^3,\quad x\in(0,1)\,,$$
$$u^{0,2}(x)=((x-1/2)^+)^{2.6},\quad x\in(0,1)\,,$$
$$u^{0,3}(x)=((x-1/2)^+)^{2.5},\quad x\in(0,1)\,,$$

which satisfy $u^{0,1} \in H^3$, $u^{0,2} \in H^3$ and $u^{0,3} \in H^{3-\epsilon}$ for any $\epsilon > 0$. The final time is $T = 0.5$ (after T the exact solution vanishes). The convergence result provided in this paper applies for the initial data $u^{0,1}$ and $u^{0,2}$, and forecasts that the l^∞ error is bounded by $C\,\Delta x^{1.5}$ if $k_b = 2$ and $C\,\Delta x^{0.5}$ if $k_b = 1$.

Figure 5.1 represents the initial condition $u^{0,2}$ on a grid with 40 cells on $(0, 1)$. On figure 5.2 and 5.3, we plot the numerical solutions (at different times) obtained with $k_b = 0$ (homogeneous Dirichlet outflow condition), $k_b = 1$ (homogeneous Neumann outflow condition) and $k_b = 2$ (homogeneous "second order" Neumann outflow condition). As expected, the Dirichlet condition shows a larger boundary layer, and, especially at time 0.2625, the solution with $k_b = 2$ is much nearer to the exact solution than the others.

Let us now analyze more precisely the error of the schemes. In the following, the computed "error" components actually are not the $u_j^n - \int_{x_{j-1}}^{x_j} u_0(y - at^n)\,\mathrm{d}y/\Delta x$: for simplicity, they are replaced with the local values $u_j^n - u_0((x_{j-1} + x_j)/2 - at^n)$. If $u_0 \in H^2$, $|u_0((x_{j-1} + x_j)/2 - at^n) - \int_{x_{j-1}}^{x_j} u_0(y - at^n)\,\mathrm{d}y/\Delta x|$ can be bounded by a constant times Δx^2, thus the present formula will not disrupt our evaluation of the order of the scheme, which cannot be greater than 2. By observing Table 1, we see that:

- with $k_b = 2$ the order of the scheme indeed seems to be 2,
- while with $k_b = 1$, it seems to be 1.

In Table 2 with a less smooth (but still in H^3) initial datum,

- with $k_b = 2$ the order of the scheme seems to belong to $(1.7, 1.75)$,
- and with $k_b = 1$, it seems to be 1.

At last, Table 3 suggests that:

- with $k_b = 2$ the order of the scheme be in $(1.65, 1.7)$,
- and with $k_b = 1$, the order still be 1.

Clearly the convergence result we have obtained here seem to be non-optimal and there is hope, for numerical schemes such as (5.1) that fit into the framework of [BC17], to fill this discrepancy by using numerical boundary layer expansions. However, one clearly observes that the convergence order does depend not only on the smoothness of the solution but also on the order of the homogeneous Neumann condition for the outflow boundary.

In order to illustrate the discussion about the stability of the scheme (4.1) and the properties of the matrix $\mathscr{A}_J \in \mathscr{M}_J(\mathbb{R})$ (Section 4.1), we report in Table 4 the spectral radius and the l^2 induced norm of this matrix (both are computed approximately thanks to the functions `eig` and `norm` of the library numpy.linalg in the Python language), with respect to the boundary condition and the number of cells. We observe that in all cases the spectral radius is smaller than 1, but that with the second order Neumann condition, the norm of the matrix is larger than 1. This means that stability estimates cannot be obtained by showing that the ℓ^2-norm does not increase, hence the need for some well-designed analytical tool (here we have used an induction argument with respect to k_b combined with the finite speed of propagation).

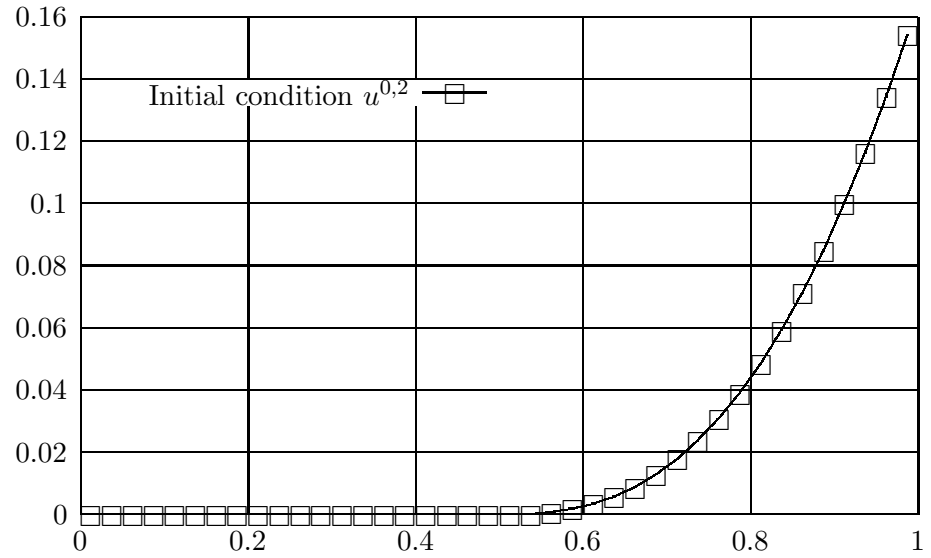


Figure 5.1: Initial condition $u^{0,2}$ with 40 cells in $(0, 1)$.

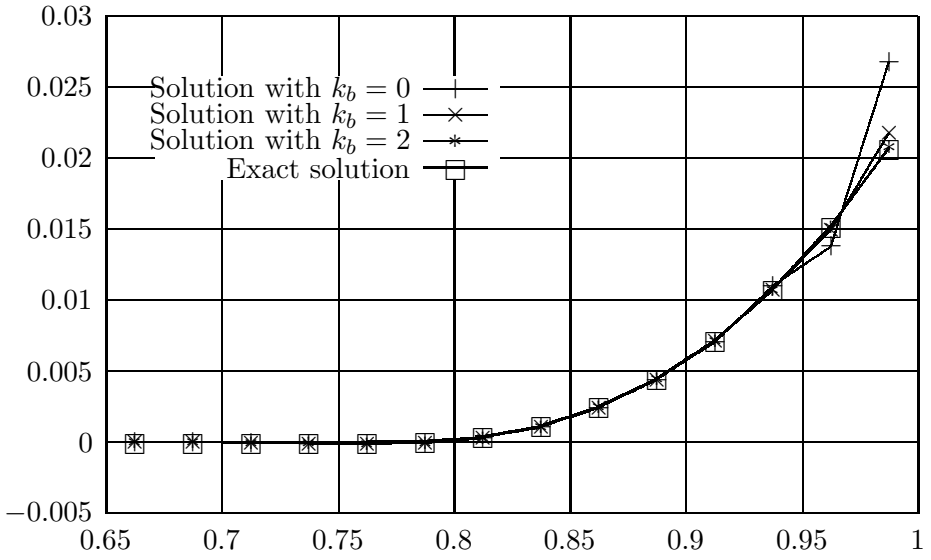


Figure 5.2: Numerical and exact solutions at time $t = 0.2625$ with initial condition $u^{0,2}$ (40 cells).

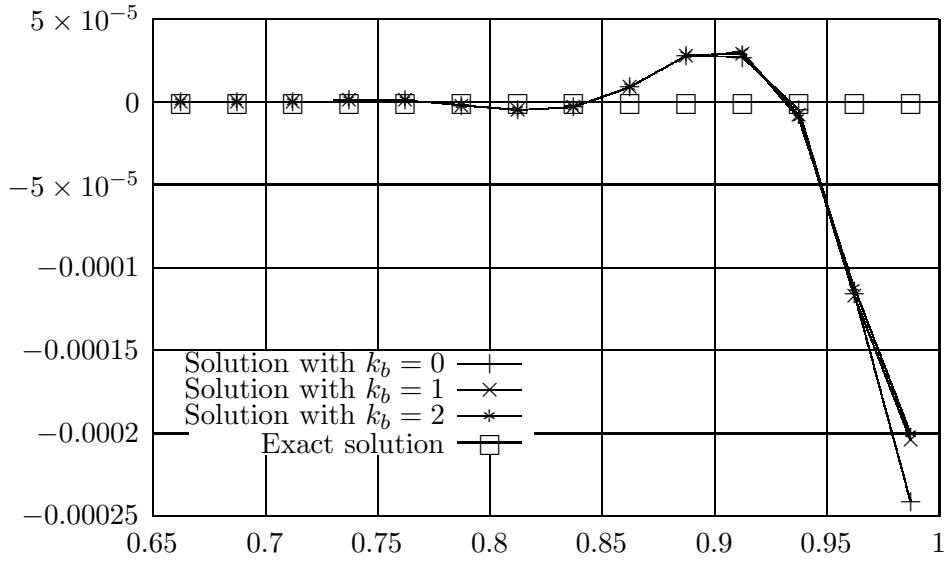


Figure 5.3: Numerical and exact solutions at time $t = 0.5075$ with initial condition $u^{0,2}$ (40 cells).

Number of cells J	Measured error with $k_b = 2$	Measured error with $k_b = 1$
10	0.0025305	0.00833660625
20	0.0008281875	0.00491559140625
40	0.0002314921875	0.00262908841699
80	0.0000609287109375	0.0013994637865
160	0.0000156141357422	0.000720704311203
320	0.00000397348640443	0.000365563075521
640	0.00000100290833469	0.00018408024467
1280	0.000000251919175326	0.0000923642961781

Table 1: Error for datum $u^{0,1}$.

Number of cells J	Measured error with $k_b = 2$	Measured error with $k_b = 1$
10	0.00280385837572	0.0102978586289
20	0.000825428449649	0.00578352637669
40	0.000252680165957	0.00308529222599
80	0.0000781474537246	0.00161972927959
160	0.0000236164563317	0.000828965994226
320	0.00000711489098145	0.000419239010994
640	0.00000213591643874	0.000210806199835
1280	0.000000643052172999	0.000105699491246

Table 2: Error for datum $u^{0,2}$.

Number of cells J	Measured error with $k_b = 2$	Measured error with $k_b = 1$
10	0.00284239926561	0.0108072887024
20	0.00091837995271	0.00600083229228
40	0.000301806292425	0.00319976806911
80	0.0000975906472619	0.00167418222795
160	0.0000308167600202	0.000855523358729
320	0.00000972494981438	0.00043235384157
640	0.00000308448727156	0.000217323081725
1280	0.000000971185766911	0.000108947852591

Table 3: Error for datum $u^{0,3}$.

J	Spectral radius, $k_b = 1$	l^2 norm, $k_b = 1$	Spectral radius, $k_b = 2$	l^2 norm, $k_b = 2$
20	0.7100	0.9999	0.7098	1.0035
80	0.74300	0.9999	0.7513	1.0035
320	0.9208	0.9999	0.9212	1.0035
1280	0.9817	0.9999	0.9805	1.0035

Table 4: Spectral radii and l^2 induced norms of the linear operator associated with the scheme.

A A discrete integration by parts lemma

In this appendix, we prove the following result, of which Lemma 3.3 is an immediate corollary as explained below.

Lemma A.1. *Let $S \in \mathscr{M}_m(\mathbb{R})$, $m \geq 2$, be a real symmetric matrix satisfying*

$$\sum_{i,j=1}^{m} S_{ij} = 0\,.$$

Then there exists a unique real symmetric matrix $\widetilde{S}$ of size $m-1$, and some unique real numbers $d_1, \dots, d_{m-1}$, such that:

$$S = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \widetilde{S} & \\ 0 & & & \end{pmatrix} - \begin{pmatrix} & & & 0 \\ & & \widetilde{S} & \vdots \\ & & & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix} + d_1 \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ -1 & 1 & \vdots & & \vdots \\ 0 & \cdots & 0 & & \vdots \\ \vdots & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} + \cdots + d_{m-1} \begin{pmatrix} 1 & 0 & \cdots & 0 & -1 \\ 0 & 0 & & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & & 0 & 0 \\ -1 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Proof. The proof is completely elementary. The vector space of real symmetric matrices of size m satisfying the condition

$$\sum_{i,j=1}^{m} S_{ij} = 0\,,$$

has dimension $m\,(m+1)/2 - 1 = (m-1)\,(m+2)/2$. By standard linear algebra, it is therefore sufficient to prove that if a real symmetric matrix $\widetilde{S}$ of size $m-1$ and some real numbers $d_1, \dots, d_{m-1}$ satisfy:

$$\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \widetilde{S} & \\ 0 & & & \end{pmatrix} - \begin{pmatrix} & & & 0 \\ & & \widetilde{S} & \vdots \\ & & & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix} + d_1 \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ -1 & 1 & \vdots & & \vdots \\ 0 & \cdots & 0 & & \vdots \\ \vdots & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} + \cdots + d_{m-1} \begin{pmatrix} 1 & 0 & \cdots & 0 & -1 \\ 0 & 0 & & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & & 0 & 0 \\ -1 & 0 & \cdots & 0 & 1 \end{pmatrix} = 0\,, \quad \text{(A.1)}$$

then $\widetilde{S} = 0$ and $d_1 = \cdots = d_{m-1} = 0$. Let us therefore assume that $\widetilde{S}$ and $d_1, \dots, d_{m-1}$ satisfy (A.1). Then considering the upper right coefficient, we first get $d_{m-1} = 0$. Considering then the last column, we get $\widetilde{S}_{i,m-1} = 0$ for all $i = 1, \dots, m-1$. The proof follows by induction on m. ∎

Let us now explain how Lemma A.1 gives the result claimed in Lemma 3.3. On $\mathbb{R}^{p+r+1}$, with vectors written under the form $(v_{-r},\dots,v_p)$, we consider the quadratic form

$$(v_{-r},\dots,v_p)\longmapsto 2\,v_0\Big(\sum_{\ell=-r}^{p}a_\ell\,v_\ell-v_0\Big)+\Big(\sum_{\ell=-r}^{p}a_\ell\,v_\ell-v_0\Big)^2,$$

Because of (2.4) for $m=0$, the vector $(1,\dots,1)$ belongs to the isotropic cone of this quadratic form. Hence we can decompose the real symmetric matrix S associated with this quadratic form by using Lemma A.1. We get a decomposition of the form

$$2\,v_0\Big(\sum_{\ell=-r}^{p}a_\ell\,v_\ell-v_0\Big)+\Big(\sum_{\ell=-r}^{p}a_\ell\,v_\ell-v_0\Big)^2=\sum_{\ell=1}^{p+r}d_\ell\,(v_{\ell-r}-v_{-r})^2+\widetilde{\mathscr{Q}}\big(v_{1-r},\dots,v_p\big)-\widetilde{\mathscr{Q}}\big(v_{-r},\dots,v_{p-1}\big),$$

for some suitable real quadratic form $\widetilde{\mathscr{Q}}$ on $\mathbb{R}^{p+r}$. It only remains to make the invertible change of variables

$$(v_{1-r},\dots,v_p)\longmapsto(v_{2-r}-v_{1-r},\dots,v_0-v_{-1},v_0,v_1-v_0,\dots,v_p-v_{p-1}),$$

in the argument of $\widetilde{\mathscr{Q}}$, which modifies this quadratic form into some $\mathscr{Q}$, and we obtain the result of Lemma 3.3 as announced. The value of $\mathscr{Q}$ on the r-th vector of the canonical basis of $\mathbb{R}^{p+r}$ is obtained by considering the specific sequence:

$$\forall\, j\in\mathbb{Z},\quad v_j:=j,$$

and by identifying the dominant term in j.

References

[AAB+08] X. Antoine, A. Arnold, C. Besse, M. Ehrhardt, and A. Schädle. A review of transparent and artificial boundary conditions techniques for linear and nonlinear Schrödinger equations. *Commun. Comput. Phys.*, 4(4):729–796, 2008.

[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.

[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.

[CHG01] C. Chainais-Hillairet and E. Grenier. Numerical boundary layers for hyperbolic systems in 1-D. *M2AN Math. Model. Numer. Anal.*, 35(1):91–106, 2001.

[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.

[Cou17] J.-F. Coulombel. Transparent numerical boundary conditions for evolution equations: derivation and stability analysis. *Ann. Fac. Sci. Toulouse Math. (6)*, 26, 2017.

[DL85] R. Dautray and J.-L. Lions. *Analyse mathématique et calcul numérique pour les sciences et les techniques. Tome 3.* Collection du Commissariat à l'Énergie Atomique: Série Scientifique. Masson, Paris, 1985.

[EM77] B. Engquist and A. Majda. Absorbing boundary conditions for the numerical simulation of waves. *Math. Comp.*, 31(139):629–651, 1977.

[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.

[Hag99] T. Hagstrom. Radiation boundary conditions for the numerical simulation of waves. In *Acta numerica, 1999*, volume 8 of *Acta Numer.*, pages 47–106. Cambridge Univ. Press, 1999.

[Hal82] L. Halpern. Absorbing boundary conditions for the discretization schemes of the one-dimensional wave equation. *Math. Comp.*, 38(158):415–429, 1982.

[Hed66] G. W. Hedstrom. Norms of powers of absolutely convergent Fourier series. *Michigan Math. J.*, 13:393–416, 1966.

[Hig86] R. L. Higdon. Absorbing boundary conditions for difference approximations to the multidimensional wave equation. *Math. Comp.*, 47(176):437–459, 1986.

[Hig94] R. L. Higdon. Radiation boundary conditions for dispersive waves. *SIAM J. Numer. Anal.*, 31(1):64–100, 1994.

[KL68] H.-O. Kreiss and E. Lundqvist. On difference approximations with wrong boundary values. *Math. Comp.*, 22:1–12, 1968.

[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.

[Osh69] S. Osher. Systems of difference equations with general homogeneous boundary conditions. *Trans. Amer. Math. Soc.*, 137:177–201, 1969.

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

[Str04] J. C. Strikwerda. *Finite difference schemes and partial differential equations*. Society for Industrial and Applied Mathematics (SIAM), 2004.

[TE05] L. N. Trefethen and M. Embree. *Spectra and pseudospectra*. Princeton University Press, 2005. The behavior of nonnormal matrices and operators.

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

[Tre84] L. N. Trefethen. Instability of difference models for hyperbolic initial-boundary value problems. *Comm. Pure Appl. Math.*, 37(3):329–367, 1984.

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