Finite Volume Transport Schemes

Bruno Després

Finite Volume Transport Schemes

Bruno Després*

December 22, 2006

Résumé On analyse des schémas Volume Fini de transport d'ordre arbitraire en dimension un d'espace. Le résultat principal est une preuve de stabilité dans L^1 et L^∞ pour les schémas d'ordre impair. On obtient des estimations de convergence numérique optimales pour des données BV. De manière générale cela permet de comprendre que les schémas de transport d'ordre impair ont un meilleur comportement que les schémas d'ordre pair vis à vis du phénomène de Gibbs. Quelques expériences numériques valident les développements théoriques. Les oscillations importantes du schéma de Lax-Wendroff avec un petit nombre de Courant sont interprétées comme la résultante de la non stabilité dans L^1 de ce schéma. Un schéma Volume Fini d'ordre trois est présenté, qui se révèle stable dans L^1 et présente en pratique peu d'oscillations.

Abstract

We analyze arbitrary order linear finite volume transport schemes and show asymptotic stability in L^1 and L^∞ for odd order schemes in dimension one. It gives sharp fractional order estimates of convergence for BV solutions. It shows odd order finite volume advection schemes are better than even order finite volume schemes. Therefore the Gibbs phenomena is controled for odd order finite volume schemes. Numerical experiments sustain the theoretical analysis. In particular the oscillations of the Lax-Wendroff scheme for small Courant numbers are correlated with its non stability in L^1. A scheme of order three is proved to be stable in L^1 and L^∞.

1 Introduction

We address the numerical analysis in L^1, L^2 and L^∞ of the high order finite volume schemes recently derived in [3, 10, 9] for discretization of the advection equation $\partial_t u + a\partial_x u = 0$. We shall assume $a > 0$ for simplicity. Some low and high order finite volume transport schemes are shown in table 1. In our context a finite volume scheme is based on some generalization of the upwind scheme in

*Research Director at CEA, CEA/DIF, 91680 Bruyères le Chatel, BP 12, France and Associate Member of the JLL Lab, University Paris VI, 175 rue du Chevaleret, 75013 Paris, France

the spirit of the Lax-Wendroff scheme. Obtaining optimal results of convergence in L^d is a consequence of their stability in L^d. By definition

$$||v||_{L^d} = \left(\int_{\mathbb{R}} |v(x)|^d dx\right)^{\frac{1}{d}} = \left(\Delta x \sum |v_j|^d\right)^{\frac{1}{d}}, \quad 1 \leq d < \infty,$$

and $||v||_{L^\infty} = \sup |v(x)| = \sup |v_j|$. In practice $d = 1$, 2 or ∞ are the most interesting cases. A classical theorem of Godunov states first order linear schemes are the only linear ones that satisfy the maximum principle. Asymptotic stability in L^1 may be useful to evaluate the oscillating behaviour of high order schemes.

Definition 1. A-stability We say a scheme is A-stable (asymptotic stability) in L^d if there exists a bound $K > 0$ which does not depend on Δx, ν such that

$$||u^n||_{L^d} \leq K||u^0||_{L^d}, \quad \forall n.$$

A-stability is more stringent than the usual uniform stability of Lax [8, 7] or Godunov [4], for which the constant K may depend on some final time $T > 0$ and the estimate is valid for $n\Delta t \leq T$. For advection the requirement of A-stability is nevertheless very natural since it is a property of the exact solution.

As usual the analysis in L^2 is based on Fourier analysis, see [5]. All standard advection schemes are A-stable in L^2 (with $K = 1$) and therefore have optimal convergence properties in L^2. For example we will prove the following result. Consider an initial data in $L^\infty \cap BV$ function. The order of convergence in L^2 of the numerical solution to the exact solution is

$$||u^n - v^n||_{L^2} \leq \left(C||u||_{L^\infty}^{\frac{1}{2}} |u|_{BV}^{\frac{1}{2}}\right) \times \left(\Delta x^a T^b + \Delta x^{\frac{1}{2}}\right) \tag{1}$$

with $a = \frac{p}{2(p+1)} < \frac{1}{2}$ and $b = \frac{1}{2(p+1)}$ where $p \in \mathbb{N}^*$ is the order of the scheme. The $\Delta x^{\frac{1}{2}}$ comes from the interpolation error between point-wise solutions and averaged solutions and does not depend on the time T. For $T > 0$, the error in space is Δx^a with $a < \frac{1}{2}$. In L^2 we do not see any fundamental difference between odd and even order schemes.

Scheme	$\alpha_\nu(\theta\Delta x)$	p
Upwind	$(1-\nu) + \nu e^{\mathbf{i}\nu\theta\Delta x}$	1
Lax-Wendroff (LW)	$(1-\nu^2) + \frac{\nu+\nu^2}{2} e^{\mathbf{i}\theta\Delta x} + \frac{\nu^2-\nu}{2} e^{-\mathbf{i}\theta\Delta x}$	2
Beam-Warming (BW)	$\left(1 - \frac{3}{2}\nu + \frac{1}{2}\nu^2\right) + (\nu - \nu^2) e^{\mathbf{i}\theta\Delta x} + \frac{\nu^2-\nu}{2} e^{\mathbf{i}2\theta\Delta x}$	2
Order 3 ($O3$)	$O3 = (1-\alpha)LW + \alpha BW$ with $\alpha = \frac{1+\nu}{3}$	3

Table 1: The amplification factor of Fourier modes is $\alpha_\nu(\theta\Delta x)$. The order of the scheme is p. The Courant number is $\nu = a\frac{\Delta t}{\Delta x}$.

The question we address in this work can by summed up as follows. Is it possible for some schemes of table 1 to be A-stable in L^1 or L^∞ with $p \geq 2$? If so, what are the consequences ?

1.1 A-stability in L^1

Theorem 6 is the main contribution of this work. It is a proof of A-stability in L^1 (and L^∞) of high order finite volume schemes provided the formal order is odd

$$p = 2k + 1$$

and some technical conditions on the amplification factor are true. See the theorem for a precise statement. **In particular the O3 scheme of table 1 is A-stable in L^1 and L^∞ under CFL.** The Lax-Wendroff and Beam-Warming schemes are not A-stable in L^1 and L^∞. It is probably possible to generalize some parts of this work to non linear equations (such one is the Burgers equation [2]).

1.2 Optimal convergence for BV solutions

A-stability in L^1 has two major consequences. The first consequence is an optimal result of convergence for BV solutions.

Theorem 2. *Assume $u_0 \in L^\infty \cap BV$. Assume $p = 2k+1$ is odd and the scheme is A-stable in L^1. Then*

$$||u^n - u(n\Delta t)||_{L^1} \leq C_p |u|_{BV} \left(\Delta x^a T^b + \Delta x\right) \tag{2}$$

with $a = \frac{p}{p+1}$ and $b = \frac{1}{p+1}$.

This estimate is sharper than (1). Using very high order schemes as in advocated in [3, 9, 10] means p is large. In this case $\frac{p}{p+1}$ is very close to 1. We consider this is optimal because an error of order one is what we get by a one cell translation of the Heavyside function. In a nutshell: very high odd order advection schemes have nearly optimal order of convergence in L^1 even for discontinuous function. Perhaps even more important for applications is the very small dependence with respect to the time T since $\frac{1}{p+1}$ is close to zero for large p. This means that the difference between the true solution and the numerical solution does not evolve significantly in time.

1.3 Dispersion and the Gibbs phenomena

A second important consequence of A-stability in L^1 is an explanation of the low dispersion encountered with such schemes. Dispersion is in practice related to oscillations that appear when a group of Fourier modes move at different velocities. As time increases the Fourier modes decouple and the oscillations appear.

Consider a A-stable in L^1 linear scheme. The iteration operator of the scheme may be considered as an infinite band matrix. For infinite band matrices the L^1 norm is also the L^∞ norm (see section 4). Since the O3 scheme of table 1 is A-stable in L^1, it is also A-stable in L^∞. Therefore oscillations cannot grow indefinitely for this scheme.

The structure of the oscillations is also controlled. Assume the error between the exact solution and the numerical solution tends to zero in L^1 with respect to the mesh size Δx, as in theorem 2. Then the Fourier transform of the error is bounded in L^∞. It means the error evaluated on Fourier modes tends to zero in a much stronger space, L^∞, compared with the usual L^2 space used for the error analysis.

Numerical experiments sustain this analysis. Numerical tests show the dispersion coefficient can outperform the diffusion coefficient for the Lax-Wendroff scheme (for small Courant numbers) and for the Beam-Warming scheme (for large Courant numbers) which are not A-stable in L^1. In this case huge oscillations may pollute the solution for the Lax-Wendroff and Beam-Warming schemes. A typical numerical experiment consists in the numerical advection of a Heavyside function. The Gibbs phenomena (oscillations near the discontinuity) is under control with the O3 scheme which is odd order.

1.4 Notations and plan

Any universal constant independent of n, ν, Δx, T, $\ldots$, may be referred to as C. The constant may depend on p the order of the scheme. There is only one constant, denoted as C_2, which is written differently from the others. This is because of its importance in the proof of A-stability in L^1.

This work is organized as follows. Section 2 is devoted to a review of some standard second order explicit schemes and to the construction of a third order scheme. These schemes are characterized with their amplification factor for Fourier modes. Then in section 3 we prove the convergence in L^2 for smooth solutions and for BV solutions. Section 4 corresponds to the main contribution of this work: we analyze the stability in L^1 and L^∞ using new estimates for the amplification factors. Section 5 is the proof of theorem 2. Our final section 6 is devoted to numerical experiments.

2 Some high order explicit schemes

We review some well known high order schemes. All these schemes may written in the Finite Volume setting

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + a\frac{u_{j+\frac{1}{2}}^n - u_{j-\frac{1}{2}}^n}{\Delta x} = 0. \tag{3}$$

It is sufficient to define the flux $u_{j+\frac{1}{2}}^n$ in function of the neighboring values to completely define the scheme. The Courant number is

$$\nu = a\frac{\Delta t}{\Delta x}.$$

2.1 The Lax-Wendroff scheme

The scheme [7] is widely used in practical computations and is the base of more specific methods. The Lax-Wendroff flux is

$$u^n_{j+\frac{1}{2}} = u^n_j + \frac{1}{2}(1-\nu)(u^n_{j+1} - u^n_j). \tag{4}$$

The final scheme writes $u^{n+1}_j = (1-\nu^2)u^n_j + \frac{\nu+\nu^2}{2}u^n_{j-1} + \frac{\nu^2-\nu}{2}u^n_{j+1}$. It is convenient to look at Fourier modes. So let assume that

$$u^n_j = \lambda^n e^{\mathbf{i}\theta j \Delta x}, \quad \mathbf{i}^2 = -1, \quad \theta \in \mathbb{R}.$$

Plugging this anzatz in (3-4) one gets the law

$$\lambda^{n+1} = \left((1-\nu^2) + \frac{\nu+\nu^2}{2}e^{\mathbf{i}\theta\Delta x} + \frac{\nu^2-\nu}{2}e^{\mathbf{i}\theta\Delta x}\right)\lambda^n.$$

Therefore the amplification factor of the Fourier mode is

$$\alpha^{\mathrm{lw}}_\nu(\theta\Delta x) = (1-\nu^2) + \frac{\nu+\nu^2}{2}e^{\mathbf{i}\theta\Delta x} + \frac{\nu^2-\nu}{2}e^{-\mathbf{i}\theta\Delta x}. \tag{5}$$

The L^2 stability of the Lax-Wendroff scheme under CFL $\nu \le 1$ is shown by the relation

$$\left|\alpha^{\mathrm{lw}}_\nu(\theta\Delta x)\right|^2 = 1 - 4\nu^2(1-\nu^2)\sin^4\left(\frac{\theta\Delta x}{2}\right). \tag{6}$$

This amplification factor is by construction an approximation of the amplification factor of the exact equation which is $\alpha^{\mathrm{ex}}_\nu(\theta\Delta x) = e^{\mathbf{i}\nu\theta\Delta x}$. The second order formal accuracy is recovered looking at

$$\beta^{\mathrm{lw}}_\nu(\theta\Delta x) = \alpha^{\mathrm{lw}}_\nu(\theta\Delta x)\alpha^{\mathrm{ex}}_\nu(\theta\Delta x)^{-1} = \alpha^{\mathrm{lw}}_\nu(\theta\Delta x)e^{-\mathbf{i}\nu\theta\Delta x}.$$

Standard Taylor expansions show that

$$\beta^{\mathrm{lw}}_\nu(\theta\Delta x) - 1 = -i\frac{\nu(1-\nu^2)}{6}(\theta\Delta x)^3 - \frac{\nu^2(1-\nu^2)}{8}(\theta^4\Delta x) + O((\theta\Delta x)^5). \tag{7}$$

The second order formal accuracy of the Lax-Wendroff scheme is equivalent to say that $\beta^{\mathrm{lw}}_\nu(\theta\Delta x)$ is an approximation of 1 at the third order. Since the coefficient of the leading term is an imaginary number, one says that dispersion is dominant. We can compare with the same quantity but for the upwind scheme

$$\beta^{\mathrm{up}}_\nu(\theta\Delta x) - 1 = \left((1-\nu) + \nu e^{\mathbf{i}\theta\Delta x}\right)\frac{1+\nu}{3}e^{-\mathbf{i}\nu\theta\Delta x} - 1$$

$$= -\frac{\nu(1-\nu)}{2}(\theta\Delta x)^2 + O((\theta\Delta x)^3)$$

which is a first order scheme and is diffusion dominant because the leading order of the Taylor expansion is real and negative.

2.2 The Beam-Warming scheme

The Beam-Warming flux [11] is quite close to the Lax-Wendroff flux. It is

$$u^n_{j+\frac{1}{2}} = u^n_j + \frac{1}{2}(1-\nu)(u^n_j - u^n_{j-1}). \tag{8}$$

The scheme writes $u^{n+1}_j = \left(1 - \frac{3}{2}\nu + \frac{1}{2}\nu^2\right) u^n_j + (\nu-\nu^2)u^n_{j-1} + \frac{\nu^2-\nu}{2}u^n_{j-2}$. The amplification factor is

$$\alpha^{\mathrm{bw}}_\nu(\theta\Delta x) = \left(1 - \frac{3}{2}\nu + \frac{1}{2}\nu^2\right) + (\nu-\nu^2)e^{\mathbf{i}\theta\Delta x} + \frac{\nu^2-\nu}{2}e^{\mathbf{i}2\theta\Delta x}. \tag{9}$$

Tedious calculations show that

$$\left|\alpha^{\mathrm{bw}}_\nu(\theta\Delta x)\right|^2 = 1 - 4\nu(1-\nu)^2(2-\nu)\sin^4(\frac{\theta\Delta x}{2}) \tag{10}$$

from which we deduce the Beam-Warming scheme is stable in L^2 under CFL $\nu \leq 2$. Note that $\alpha^{\mathrm{lw}}_\nu(\theta\Delta x) = \alpha^{\mathrm{up}}_\nu(\theta\Delta x) - \frac{\nu^2-\nu}{2}(1-e^{\mathbf{i}\theta\Delta x})^2 e^{-\mathbf{i}\theta\Delta x}$ and $\alpha^{\mathrm{bw}}_\nu(\theta\Delta x) = \frac{1+\nu}{3}\alpha^{\mathrm{up}}_\nu(\theta\Delta x) - \frac{\nu^2-\nu}{2}(1-e^{\mathbf{i}\theta}\Delta x)^2$. Hence we deduce that

$$\alpha^{\mathrm{bw}}_\nu(\theta\Delta x) = \alpha^{\mathrm{lw}}_\nu(\theta\Delta x) - \frac{\nu^2-\nu}{2}(1-e^{\mathbf{i}\theta\Delta x})^2(1-e^{-\mathbf{i}\theta\Delta x}).$$

So one has the relation $\alpha^{\mathrm{bw}}_\nu(\theta\Delta x) = \alpha^{\mathrm{lw}}_\nu(\theta\Delta x) + 4\mathbf{i}\nu(1-\nu)e^{\mathbf{i}\frac{\theta\Delta x}{2}}\sin^3(\frac{\theta\Delta x}{2})$. Define $\beta^{\mathrm{bw}}_\nu(\theta\Delta x) = \alpha^{\mathrm{bw}}_\nu(\theta\Delta x)e^{-\mathbf{i}\nu\theta\Delta x}$. Then

$$\beta^{\mathrm{bw}}_\nu(\theta\Delta x) = \beta^{\mathrm{lw}}_\nu(\theta\Delta x) + 4\mathbf{i}\nu(1-\nu)e^{\mathbf{i}(1-2\nu)\frac{\theta\Delta x}{2}}\sin^3(\frac{\theta\Delta x}{2}). \tag{11}$$

The Taylor expansion of $\beta^{\mathrm{bw}}_\nu(\theta\Delta x)$ is

$$\beta^{\mathrm{bw}}_\nu(\theta\Delta x) = \beta^{\mathrm{lw}}_\nu(\theta\Delta x) + 4\mathbf{i}\nu(1-\nu)\left((1+\mathbf{i}(1-2\nu)\frac{\theta}{2}\right)\frac{(\theta\Delta x)^3}{8} + O((\theta\Delta x)^5),$$

that is

$$\beta^{\mathrm{bw}}_\nu(\theta\Delta x) - 1 = \mathbf{i}\frac{\nu(\nu-1)(\nu-2)}{6}(\theta\Delta x)^3 - \frac{\nu(1-\nu)^2(2-\nu)}{8}(\theta\Delta x)^4 + O(\theta^5). \tag{12}$$

For the Beam-Warming scheme dispersion is also dominant. The sign of the third order contribution is opposite to the sign of the third order contribution of the Lax-Wendroff scheme.

2.3 A third order scheme

The development of a third order scheme is quite easy from the Beam-Warming and the Lax-Wendroff schemes. The definition of the scheme is

$$O3 = (1-\alpha)LW + \alpha BW$$

where $\alpha \in \mathbb{R}$ is a number chosen to zeroing the third order contribution in $\beta_\nu^{O3} = (1-\alpha)\beta_\nu^{\text{lw}} + \alpha\beta_\nu^{\text{bw}}$. By inspection of formulas (7) and (11) the definition of α is $-\frac{\nu(1-\nu^2)}{6} + \alpha\frac{\nu(1-\nu)}{2} = 0$. It gives

$$\alpha = \frac{1+\nu}{3}.$$

One has by construction $\alpha_\nu^{O3}(\theta\Delta x) = (1-\alpha)\alpha_\nu^{\text{lw}}(\theta\Delta x) + \alpha\alpha_\nu^{\text{bw}}(\theta\Delta x)$. Assume $\nu \in [0,1]$. Then $\alpha \in [0,1]$ which means the third order scheme is an average of the Lax-Wendroff scheme and the Beam-Warming scheme. Since we also have $\left|\alpha_\nu^{\text{lw}}(\theta\Delta x)\right| \le 1$ and $\left|\alpha_\nu^{\text{bw}}(\theta\Delta x)\right| \le 1$, then

$$\left|\alpha_\nu^{O3}(\theta\Delta x)\right| \le 1.$$

Therefore the amplification coefficient of the $O3$ scheme is always less than one in modulus, which gives the stability under CFL. One finally has

$$\beta_\nu^{O3}(\theta\Delta x) - 1 = -\frac{\nu(1-\nu)(1+\nu)(2-\nu)}{2} 4(\theta\Delta x)^4 + O((\theta\Delta x)^5). \tag{13}$$

The fourth order contribution is dominant in (13). Most presumably this third order scheme is the same as in [3, 10, 9].

2.4 General schemes

Since we are interested with the numerical analysis of high order and very high order schemes for transport schemes that have been introduced in the recent works [3, 10, 9], we shall need general notations.

We begin with some standard Fourier considerations. Let define the generating function as follows $f^n(\theta) = \Delta x \sum_j u_j^n e^{\mathbf{i}\theta j\Delta x}$. By construction one has

$$u_j^n = \frac{1}{2\pi}\int_{\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} f^n(\theta)e^{-\mathbf{i}\theta j\Delta x}d\theta.$$

The Parseval identity is $||u^n||_{L^2}^2 = \Delta x\sum_j (u_j^n)^2 = \frac{1}{2\pi}\int_{\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} |f^n(\theta)|^2 d\theta$. We consider a scheme characterized by its amplification factor $\theta\Delta x \mapsto \alpha_\nu(\theta\Delta x)$. So one has $f^{n+1}(\theta) = \alpha_\nu(\theta\Delta x)f^n(\theta)$. Therefore

$$||u^n||_{L^2}^2 = \frac{1}{2\pi}\int_{\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} |\alpha_\nu(\theta\Delta x)|^{2n}|f^0(\theta)|^2 d\theta.$$

So the scheme is A-stable in L^2 if and only if

$$|\alpha_\nu(\theta\Delta x)| \le 1, \quad \forall\theta\Delta x \in \mathbb{R}. \tag{14}$$

These is the case once $\nu \in]0,1]$ for all schemes defined previously. Formula (14) shows that it is possible to consider only

$$\theta\Delta x \in [-\pi,\pi].$$

This is actually compatible with the periodicity of the function $\mu \mapsto \alpha_\nu(\mu)$ which is by construction a linear combination of periodic functions $\mu \mapsto e^{\mathbf{i}k\mu}$ where $k \in \mathbb{Z}$.

Expanding the function $\beta_\nu(\mu) = \alpha_\nu(\mu)e^{-\mathbf{i}\nu\mu}$ one gets

$$\beta_\nu(\mu) = 1 + C(\nu)\nu\mu^{p+1} + O(\mu^{p+2}).$$

In this formula $p \in \mathbb{N}$ is the order of the scheme. By construction $C(\nu)$ is a polynomial function of the CFL number ν. Moreover $C(\nu)$ is real for odd p's and is imaginary for even p's. This is because the functions $\alpha_\nu(\mu)$ and $e^{-\mathbf{i}\nu\mu}$ can be expanded in Taylor series where all coefficients are real for even orders and imaginary for odd orders. For convenience we can rewrite it as

$$|\beta_\nu(\mu) - 1| \leq C\nu(\mu)^{p+1} \tag{15}$$

where $\beta_\nu(\theta\Delta x) = \alpha_\nu(\theta\Delta x)e^{-\mathbf{i}\theta\nu\Delta x}$ and $C > 0$ is some constant.

The difference between odd and even orders motivates the definition of $q \in \mathbb{N}$ such that (under CFL)

$$|\beta_\nu(\mu)| \leq 1 - C_2\nu(|\mu)^{q+1}, \quad C_2 > 0, \tag{16}$$

where C_2 is some universal positive constant independent on ν and μ. The constant C_2 is the only constant that we shall write differently in this work because of its prominent role in the proof of A-stability in L^1 and L^∞. By inspection of the schemes above p and q may be a priori different. What we can infer from the inspection of the schemes above is that: if $p = 2k$ then $q = p + 1$; if $p = 2k + 1$ then $q = p$. The O3 scheme is such that $p = q = 3$. The LW and BW schemes are such that $p = 2$ and $q = 3$. For the LW scheme, the assumption (16) needs to be modified: C_2 is a linear function of ν because the fourth order coefficient depends on ν^2 in the expansion (7); $C_2(\nu) = 0$ for $\nu = 0$. It will be related with the oscillating behavior of the Lax-Wendroff for small CFL numbers.

We also assume

$$|\beta'_\nu(\mu)| \leq C\nu|\mu|^p \text{ and } |\beta''_\nu(\mu)| \leq C\nu|\mu|^{p-1} \tag{17}$$

for some universal constants which are independent on ν and θ. This last hypothesis (17) is of course compatible with (16).

The order of a scheme can also be characterized by the norm of the truncation error. Let define $v^n = (v_j^n)$

$$v_j^n = u(n\Delta t, j\Delta x)$$

where u is a smooth function. The truncation error is defined by

$$r_j^n = \frac{v_j^{n+1} - v_j^n}{\Delta t} + a\frac{v_{j+\frac{1}{2}}^n - v_{j-\frac{1}{2}}^n}{\Delta x}$$

where $v_j^n = u(n\Delta t, j\Delta x)$ is the exact solution and $v_{j+\frac{1}{2}}^n$ is the flux calculated with v^n. The L^∞ norm of the truncation error is bounded by

$$||r^n||_{L^\infty} \le C\Delta x^p ||u^{(p+1)}||_{L^\infty}. \tag{18}$$

This inequality is very classical and can be recovered as a consequence of (15). Indeed for Fourier modes one has

$$r_j^n = \frac{e^{\mathbf{i}\nu\theta\Delta x} - \alpha_\nu(\theta\Delta x)}{\Delta t} e^{\mathbf{i}j\theta\Delta x}.$$

Plugging (15) in this expression we get (18). This is why (18) and (15) are compatible.

Lemma 3. *One also has*

$$||r^n||_{L^2} \le C\Delta x^p ||u^{(p+1)}||_{L^2} \tag{19}$$

and

$$||r^n||_{L^1} \le C\Delta x^p ||u^{(p+1)}||_{L^1} \tag{20}$$

We give only a sketch of the proof. From (18) the truncation error is zero if u is a polynomial of order less than $p+1$: $r^n = 0$ for $u(0,x) = P_p(x) = \sum_{r=0}^p \alpha_r x^r$. Therefore the truncation error can be rewritten as

$$r_j^n = \frac{w_j^{n+1} - w_j^n}{\Delta t} + a\frac{w_{j+\frac{1}{2}}^n - w_{j-\frac{1}{2}}^n}{\Delta x}, \quad w^n = v^n - P_p, \ \forall P_p.$$

Here $w^{n+1} = u(n\Delta t, x - a\Delta t) - P_p(x - a\Delta t)$ is the difference between he exact solution at time $n+1$ and the exact solution issued form P_p. Therefore

$$|r_j^n| \le |a| \max_{j^-\Delta x \le x \le j^+\Delta x} \left(|(w^n)'(x)|\right) + C\frac{|a|}{\Delta x} \max_{j^-\Delta x \le x \le j^+\Delta x} \left(|w^n(x)|\right). \tag{21}$$

Here $j^+ - j^- \le R$ where $R \in \mathbb{N}$ is the size of the stencil of the scheme. The polynomial P_p is arbitrary. Let choose P_p as the Taylor series of the function u at point $j\Delta x$ truncated at order p

$$w^n(y) = u(y) - P_p(y) = \int_{j\Delta x}^y \frac{(y-z)^p}{p!} u^{(p+1)}(z)dz \tag{22}$$

Then for $y \in [j^-\Delta x, j^+\Delta x]$

$$|w^n(y)| \le C\Delta x^{p+\frac{1}{2}} ||u^{(p+1)}||_{L^2[j^-\Delta x, j^+\Delta x]}$$

and

$$|(w^n)'(y)| \le C\Delta x^{p-\frac{1}{2}} ||u^{(p+1)}||_{L^2[j^-\Delta x, j^+\Delta x]}.$$

Plugging in inequality (21) we get (19). Similarly (22) implies

$$|w^n(y)| \le C\Delta x^p ||u^{(p+1)}||_{L^1[j^-\Delta x, j^+\Delta x]}$$

and

$$|(w^n)'(y)| \le C\Delta x^{p-1} ||u^{(p+1)}||_{L^1[j^-\Delta x, j^+\Delta x]}.$$

Using this in (21) we get (20). It ends the proof.

3 L^2 analysis

We discretize the advection equation on $x \in \mathbb{R}$ with the scheme (3) and we assume a formula for the flux has been chosen. For example one can take the Lax-Wendroff, Beam-Warming or the Order 3 flux. These schemes are A-stable in L^2 under CFL.

3.1 Convergence in L^2 for smooth solutions

Define $v_j^n = u(n\Delta t, j\Delta x)$ the solution at time step $n\Delta t$ and node $j\Delta x$.

Theorem 4. *Consider the scheme defined by (3) and the point-wise initialization*

$$u_j^0 = u_0(j\Delta x).$$

Assume the amplification factor of the scheme satisfies (14-15). Then one has the estimate of convergence in L^2

$$||u^n - v^n||_{L^2} \leq \left(C||u^{p+1}||_{L^2}\right)\Delta x^p T, \quad n\Delta t \leq T. \tag{23}$$

C depends on p. Since the scheme is linear, stable under CFL $\nu \leq 1$ with $K = 1$ and consistent, we just add the errors

$$||u^n - v^n||_{L^2} \leq ||u^{n-1} - v^{n-1}||_{L^2} + \Delta t||r^{n-1}||_{L^2}$$

$$\leq ||u^{n-2} - v^{n-2}||_{L^2} + \Delta t||r^{n-2}||_{L^2} + \Delta t||r^{n-1}||_{L^2} \leq \cdots$$

$$\leq ||u^0 - v^0||_{L^2} + (n\Delta t)\max_p ||r^p||_{L^2}$$

$$\leq (n\Delta t)\max_p ||r^p||_{L^2} \leq C_p||u^{(p+1)}||_{L^2}\Delta x^p T.$$

It proves (23).

3.2 Convergence in L^2 for BV solutions

Next we adapt the estimate (23) to BV functions (which are not in H^1). For BV initial data, the point-wise initialization is not always possible. We shall use the mean initialization

$$u_j^0 = \frac{1}{\Delta x}\int_{(j-\frac{1}{2})\Delta x}^{(j+\frac{1}{2})\Delta x} u_0(x)dx. \tag{24}$$

We shall apply theorem 4 for a regularized initial data u_0^ε

$$u_0^\varepsilon(x) = \frac{1}{\varepsilon}\int_{\mathbb{R}} \varphi\left(\frac{x-y}{\varepsilon}\right)u_0(y)dy$$

where $x \mapsto \varphi(x)$ is a given C^∞ non negative function with compact support and such that $\int_{\mathbb{R}} \varphi(x)dx = 1$. Among standard inequalities let us stress

$$\|u_0^\varepsilon\|_{L^\infty} \leq \|u_0\|_{L^\infty} \text{ and } \|\left(u_0^\varepsilon\right)'\|_{L^1(\mathbb{R})} \leq |u_0|_{BV}. \tag{25}$$

One also has

$$||u_0^\varepsilon - u_0||_{L^2(\mathbb{R})} \le C\varepsilon^{\frac{1}{2}} ||u||_{L^\infty}^{\frac{1}{2}} |u|_{BV}^{\frac{1}{2}}. \tag{26}$$

It comes from

$$\int_{\mathbb{R}} (u_0^\varepsilon(x) - u(x))^2 dx = \int_{\mathbb{R}} \left(\frac{1}{\varepsilon} \int_{\mathbb{R}} \varphi\left(\frac{x-y}{\varepsilon}\right) (u_0(y) - u_0(x)) dy \right)^2 dx.$$

Since

$$\left\| \frac{1}{\varepsilon} \int_{\mathbb{R}} \varphi\left(\frac{\cdot - y}{\varepsilon}\right) (u_0(y) - u_0(\cdot)) dy \right\|_{L^1(\mathbb{R})} \le C\varepsilon |u|_{BV}$$

and

$$\left\| \frac{1}{\varepsilon} \int_{\mathbb{R}} \varphi\left(\frac{\cdot - y}{\varepsilon}\right) (u_0(y) - u_0(\cdot)) dy \right\|_{L^\infty(\mathbb{R})} \le C\|u\|_{L^\infty}$$

it proves (26). Third one has

$$|| \left(\partial_x^{r+1}\right) u_0^\varepsilon||_{L^2(\mathbb{R})} \le \frac{C}{\varepsilon^{r+\frac{1}{2}}} ||u||_{L^\infty}^{\frac{1}{2}} |u|_{BV}^{\frac{1}{2}}, \quad r = 0, 1, 2, \cdots. \tag{27}$$

It comes from

$$\begin{aligned} &\int_{\mathbb{R}} (\partial_x^{r+1} u_0^\varepsilon(x))^2 dx \\ &= \int_{\mathbb{R}} \left(\frac{1}{\varepsilon^{r+1}} \int_{\mathbb{R}} \varphi^{(r)} \left(\frac{x-y}{\varepsilon}\right) u_0'(y) dy \right) \\ &\times \left(\frac{1}{\varepsilon^{r+2}} \int_{\mathbb{R}} \varphi^{(r+1)} \left(\frac{x-y}{\varepsilon}\right) u_0(y) dy \right) dx. \end{aligned}$$

The first parenthesis is bounded in $L^1(\cdot)$ by $C\varepsilon^{-r}|u|_{BV}$. The second parenthesis is also a function of the x variable, and is bounded in L^∞ by $C\varepsilon^{-r-1}||u||_{L^\infty}$. Therefore the scalar product is bounded by $\varepsilon^{-2r-1}|u|_{BV}||u||_{L^\infty}$. It proves (27).

Theorem 5. *Let $v^n = (v_j^n)$ be the mean value in the cell of the exact solution: $v_j^n = \frac{1}{\Delta x} \int_{(j-\frac{1}{2})\Delta x}^{(j+\frac{1}{2})\Delta x} u(n\Delta t, x) dx$. Consider the scheme defined by (3) and the mean initialization. Assume the amplification factor satisfies (14-15). Then one has the estimate of convergence in L^1*

$$||u^n - v^n||_{L^1} \le C||u||_{L^\infty}^{\frac{1}{2}} |u|_{BV}^{\frac{1}{2}} \left(\Delta x^a T^b + \Delta x\right) \tag{28}$$

with $a = \frac{p}{2(p+1)}$ and $b = \frac{1}{2(p+1)}$.

To prove the theorem we introduce the regularized solution in the inequality

$$||u^n - v^n||_{L^2} \le ||u^n - (u^\varepsilon)^n||_{L^2} + ||(u^\varepsilon)^n - (v^\varepsilon)^n||_{L^2} + ||(v^\varepsilon)^n - v^n||_{L^2}, \tag{29}$$

where $(v^\varepsilon)_j^n = u^\varepsilon(n\Delta t, j\Delta x)$ is the point-wise exact regularized solution at time-space point $(n\Delta t, j\Delta x)$ and $(u^\varepsilon)^n$ is the numerical solution issued from regularized point-wise initial data $(u^\varepsilon)_j^0 = u_0^\varepsilon(j\Delta x)$. The first contribution in (29) is

$$||u^n - (u^\varepsilon)^n||_{L^2} \le ||u^0 - (u^\varepsilon)^0||_{L^2}.$$

One has $\left[u^0 - (u^\varepsilon)^0\right]_j = w_j + z_j$ where

$$w_j = \frac{1}{\Delta x}\int_{(j-\frac12)\Delta x}^{(j+\frac12)\Delta x} (u_0 - u_0^\varepsilon)(x)dx \text{ and } z_j = \frac{1}{\Delta x}\int_{(j-\frac12)\Delta x}^{(j+\frac12)\Delta x} (u_0^\varepsilon(x) - u_0^\varepsilon(j\Delta x))dx.$$

The L^2 norm of $w = (w_j)$ is bounded by the L^2 norm of the difference between the exact initial data and its regularization $u_0 - u_0^\varepsilon$ for which we use (26). So $||w||_{L^2} \le C\varepsilon^{\frac12}||u||_{L^\infty}^{\frac12}|u|_{BV}^{\frac12}$. Concerning $z = (z_j)$ one has

$$||z||_{L^\infty} \le 2||u_0^\varepsilon||_{L^\infty} \le 2||u_0||_{L^\infty}$$

and

$$|z_j| \le ||\left(u_0^\varepsilon\right)'||_{L^1[(j-\frac12)\Delta x,(j+\frac12)\Delta x]}$$

$$\Longrightarrow ||z||_{L^1} = \Delta x \sum_j |z_j| \le \Delta x||\left(u_0^\varepsilon\right)'||_{L^1(\mathbb{R})} \le \Delta x|u_0|_{BV}.$$

Therefore $||z||_{L^2} \le ||z||_{L^1}^{\frac12}||z||_{L^\infty}^{\frac12} \le C\Delta x^{\frac12}||u||_{L^\infty}^{\frac12}|u|_{BV}^{\frac12}$. So the first term in the right hand side of (29) is bounded by

$$||u^n - (u^\varepsilon)^n||_{L^2} \le ||u^0 - (u^\varepsilon)^0||_{L^2} \le C\left(\Delta x^{\frac12} + \varepsilon^{\frac12}\right)||u||_{L^\infty}^{\frac12}|u|_{BV}^{\frac12}. \tag{30}$$

The second term in the right hand side of (29) is the difference of the point value smooth solution $(v^\varepsilon)^n$ and the numerical solution $(u^\varepsilon)^n$ with a point-wise initialization. So we can apply theorem 4 with the bound (27). It gives

$$||(u^\varepsilon)^n - (v^\varepsilon)^n||_{L^2} \le \left(\frac{C}{\varepsilon^{p+\frac12}}||u||_{L^\infty}^{\frac12}|u|_{BV}^{\frac12}\right)\Delta x^p T. \tag{31}$$

The third term in the right hand side of (29) is the difference of a point-wise smooth solution and the mean value of the same smooth solution. So we can bound it like z, that is

$$||(v^\varepsilon)^n - v^n||_{L^2} \le C\Delta x^{\frac12}||u||_{L^\infty}^{\frac12}|u|_{BV}^{\frac12}. \tag{32}$$

Gathering (30-32) one gets

$$||u^n - v^n||_{L^2} \le \left(C||u||_{L^\infty}^{\frac12}|u|_{BV}^{\frac12}\right)\left(\varepsilon^{\frac12} + \frac{\Delta x^p T}{\varepsilon^{p+\frac12}} + \Delta x^{\frac12}\right).$$

The regularization parameter ε is arbitrary. An optimal value is

$$\varepsilon = \Delta x^{\frac{p}{p+1}}T^{\frac{1}{p+1}}.$$

The proof is ended.

4 L^1 analysis

We now turn to establish the stability in L^1. The norm in L^1 is $||u^n||_{L^1} = \Delta x \sum_j |u_j^n|$. Let $B = (b_{kl})_{lm}$ be a infinite band matrix

$$b_{kl} = a_j, \quad j = k - l.$$

So

$$||B||_{L^1} = \sup_l \left(\sum_k |b_{kl}| \right) = \sum_j |a_j|.$$

The norm in L^∞ is $||u^n||_{L^\infty} = \sup_j |u_j^n|$. By symmetry one also has

$$||B||_{L^\infty} = \sup_k \left(\sum_k |b_{kl}| \right) = \sum_j |a_j| = ||B||_{L^1}.$$

The L^∞ is equal to the L^1 norm for infinite band matrices.

4.1 A- stability in L^1

A- stability in L^2 does not imply necessarily A- stability in L^1. Our goal is to determine which schemes are A- stable in L^1. We consider a scheme defined by (3). The amplification factor is of the form (16-17) where q and p are the coefficients of (16) and (17).

Theorem 6. *If $p = q$ the scheme is A- stable in L^1.*

Applying one of the linear scheme presented above to this initial data is equivalent to the multiplication by B^n where B is the iteration transfer matrix. By construction B^n is a band matrix: $B^n = (b_{kl}^n)$ with $b_{ij}^n = a_j^n, \quad j = k - l$. The a_j^n are easy to compute from the amplification factor. One has

$$a_j^n = \frac{\Delta x}{2\pi} \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} \alpha_\nu(\theta \Delta x)^n e^{-\mathbf{i} j\theta\Delta x} d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} \alpha_\nu(\theta)^n e^{-\mathbf{i} j\theta} d\theta. \tag{33}$$

One has immediately

$$\left| a_j^n \right| \leq 1. \tag{34}$$

The proof consists in getting more estimates on $\left| a_j^n \right|$ and then to bound $\sum_j \left| a_j^n \right|$. The idea behind these estimates is that the transfer matrix is probably close to the shift operator for accurate schemes. So a_j^n must be close to one near the diagonal $j \approx \nu n$ (estimate (35)) while a_j^n must be close to zero away from the same diagonal (estimate (39)).

First step consists in a simple bound for $\left| a_j^n \right|$. From (16) one gets

$$\left| a_j^n \right| \leq \frac{1}{\pi} \int_0^{\pi} (1 - C_2 \nu \mu^{q+1})^n d\mu, \quad C_2 > 0.$$

Let perform the change of variable $\mu = \frac{\varphi}{(C_2\nu n)^{\frac{1}{q+1}}}$. Then

$$\left|a_j^n\right| \leq \frac{1}{\pi\left(C_2\nu n\right)^{\frac{1}{q+1}}} \int_0^{\pi(C_2\nu n)^{\frac{1}{q+1}}} \left(1 - \frac{\varphi^{q+1}}{n}\right)^n d\varphi$$

$$\leq \frac{1}{\pi\left(C_2\nu n\right)^{\frac{1}{q+1}}} \int_0^{\pi(C_2\nu n)^{\frac{1}{q+1}}} e^{-C_2\nu\mu^{q+1}} d\varphi \leq \frac{1}{\pi\left(C_2\nu n\right)^{\frac{1}{q+1}}} \int_0^{\infty} e^{-\mu^{q+1}} d\varphi$$

where we have used $\left(1+\frac{z}{n}\right)^n \leq e^z$. Therefore one has the first estimate

$$\left|a_j^n\right| \leq \frac{C}{(\nu n)^{\frac{1}{q+1}}}. \tag{35}$$

Second step One also has

$$a_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \beta_\nu(\theta)^n e^{\mathbf{i}n(\nu-\alpha)\theta} d\theta \tag{36}$$

where $\alpha = \frac{j}{n}$. By construction the function $\theta \mapsto \beta_\nu(\theta)$ is periodic $\beta_\nu(\theta + 2\pi) = \beta_\nu(\theta)$. Standard integration by parts in Fourier integrals shows that

$$a_j^n = \frac{1}{2\pi\left[\mathbf{i}n(\nu-\alpha)\right]^2} \int_{-\pi}^{\pi} \frac{d^2}{d\theta^2}\left(\beta_\nu(\theta)^n\right) e^{\mathbf{i}n(\nu-\alpha)\theta} d\theta \tag{37}$$

One has

$$a_j^n = -\frac{1}{2\pi n^2(\nu-\alpha)^2}$$

$$\times \int_{-\pi}^{\pi} \left(n\beta_\nu''(\theta)\beta_\nu(\theta)^{n-1} + n(n-1)\beta_\nu'(\theta)^2\beta_\nu(\theta)^{n-2}\right) e^{\mathbf{i}n(\nu-\alpha)\theta} d\theta$$

and

$$|a_j^n| \leq \frac{1}{2\pi n^2(\nu-\alpha)^2} \int_{-\pi}^{\pi} \left|n\beta_\nu''(\theta)\beta_\nu(\theta)^{n-1} + n(n-1)\beta_\nu'(\theta)^2\beta_\nu(\theta)^{n-2}\right| d\theta.$$

The derivatives of β_ν are bounded by hypothesis (17). So

$$|a_j^n| \leq \frac{1}{\pi n^2(\nu-\alpha)^2} \int_0^{\pi} \left(n\nu\theta^{p-1}(1 - C_2\nu\theta^{q+1})^{n-1}\right.$$

$$\left. + n(n-1)C\nu^2\theta^{2p}(1 - C_2\nu\theta^{q+1})^{n-2}\right) d\theta.$$

Next we perform the same change of variable as in the first step $\theta = \frac{\varphi}{(C_2\nu n)^{\frac{1}{q+1}}}$. We get for the first term that appear in the right hand side of $|a_j^n|$

$$\int_0^{\pi} n\nu\theta^{p-1}(1 - C_2\nu\theta^{q+1})^{n-1} d\theta$$

$$= (n\nu)^{1-\frac{p}{q+1}} \int_0^{(C_2\nu n)^{\frac{1}{q+1}}\pi} \varphi^{p-1}\left(1-\frac{\varphi^{q+1}}{n}\right)^{n-1} d\varphi$$

$$\leq (n\nu)^{1-\frac{p}{q+1}} \int_0^{(C_2\nu n)^{\frac{1}{q+1}}\pi} \varphi^{p-1} e^{-\frac{n-1}{n}\varphi^{q+1}} d\varphi$$

Provided $n \geq 2$ then $\frac{n-1}{n} \geq \frac{1}{2}$. Therefore

$$\int_0^{(C_2\nu n)^{\frac{1}{q+1}}\pi} \varphi^{p-1} e^{-\frac{n-1}{n}\varphi^{q+1}} d\varphi \leq \int_0^{\infty} \varphi^{p-1} e^{-\frac{1}{2}\varphi^{q+1}} d\varphi$$

which is a convergent integral. So we can write

$$\int_0^{\pi} nC\nu\theta^{p-1}(1-C_2\nu\theta^{q+1})^{n-1} d\theta \leq C(n\nu)^{1-\frac{p}{q+1}}$$

for some constant C which does not depend on n and ν. Using the same method we get that

$$\int_0^{\pi} n(n-1)C\nu^2\theta^{2p}(1-C_2\nu\theta^{q+1})^{n-2} d\theta \leq C(\nu n)^{2-\frac{2p+1}{q+1}}.$$

So

$$|a_j^n| \leq \frac{C}{n^(\nu-\alpha)^2}\left((n\nu)^{1-\frac{p}{q+1}} + (n\nu)^{2-\frac{2p+1}{q+1}}\right).$$

So

$$\left|a_j^n\right| \leq \frac{C}{n^2(\nu-\alpha)^2}\left((n\nu)^{1-\frac{p}{q+1}} + (n\nu)^{2-\frac{2p+1}{q+1}}\right) \leq \frac{C}{n^2(\nu-\alpha)^2}(n\nu)^{\frac{2(q-p)+1}{q+1}}. \quad (38)$$

Let us now assume that $p = q$

$$\left|a_j^n\right| \leq \frac{C}{n^2(\nu-\alpha)^2}(n\nu)^{\frac{1}{p+1}}. \quad (39)$$

This is our second estimate for $|a_j^n|$

End of the proof It remains to estimate $\sum_j |a_j^n|$ to get a bound for the L^1 norm of the operator. One has

$$\sum_j |a_j^n| = \sum_{j:\ |j-n\nu|\leq N} |a_j^n| + \sum_{j:\ |j-n\nu|>N} |a_j^n|.$$

For the first sum we use estimate (35) except perhaps for one of the a_j^n for which we us $|a_j^n| \leq 1$ (see (33)). We get

$$\sum_{j:\ |j-n\nu|\leq N} |a_j^n| \leq 1 + 2N \times \frac{C}{(n\nu)^{\frac{1}{p+1}}}.$$

For the second sum we use (39) recalling that $\alpha = \frac{j}{n}$. We get

$$\sum_{j:\ |j-n\nu|>N} |a_j^n| \le C(n\nu)^{\frac{1}{p+1}} \left(2 \sum_{j>N} \frac{1}{j^2} \right).$$

The series is bounded by

$$\sum_{j>N} \frac{1}{j^2} \le \int_N \frac{dx}{x^2} = \frac{1}{N}.$$

Therefore

$$\sum_{j:\ |j-n\nu|>(n\nu)^{\frac{1}{p+1}}} |a_j^n| \le 1 + 2N \times \frac{C}{(n\nu)^{\frac{1}{p+1}}} + C(n\nu)^{\frac{1}{p+1}} \frac{1}{N}$$

for a constant which does not depend on n and ν. The optimal value of N is $N = (n\nu)^{\frac{1}{p+1}}$. Finally it proves that $\sum_j |a_j^n|$ is bounded uniformly with respect to n and ν. It ends the proof.

Remark 7. If $q > p$ then estimate (38) is not sharp enough and it is not possible to bound the series uniformly. The case $q = p + 1$ deserve some particular interest because it corresponds to the Lax Wendroff scheme for ν away from zero. Assume for convenience that ν is bounded away from 0 and 1 that all constants are uniformly bounded (see the discussion in the numerical section), one looses a $n^{\frac{2}{q+1}} = n^{\frac{2}{p+3}}$ in (38). So finally one gets

$$\text{If } q = p+1 \text{ then } \sum_j |a_j^n| \le 1 + C\frac{N}{n^{\frac{1}{q+1}}} + C\frac{n^{\frac{3}{q+1}}}{N} \le Cn^{\frac{1}{q+1}} \tag{40}$$

by taking the optimal value $N = n^{\frac{2}{q+1}}$. For the Lax-Wendroff and Beam-Warming schemes $q = 3$. So one gets (for ν away from 0 and 1) a bound

$$\sum_j |a_j^n| \le Cn^{\frac{1}{4}} \tag{41}$$

which reveals a increasing L^1 norm of the power of the iteration operator. For the Lax Wendroff scheme C_2 is not uniformly bounded from below for $\nu \to 0$. It means the approximation (41) is not valid for small Courant numbers.

5 Convergence in L^1 (proof of theorem 2)

For the proof of theorem 2 we use the same method as in the proof of theorem 5. First the order of convergence is Δx^p for a solution which is differentiable

$p+1$ times in L^1. Then we consider a solution only in BV and we regularize it. This shows the error is upper bounded by

$$||u^n - v^n||_{L^1} \leq C\left(\varepsilon + \frac{\Delta x^p}{\varepsilon^p} T + \Delta x\right)|u|_{BV}$$

where Δx is the interpolation error and is an order in L^1 for a solution in BV. Now the optimal value of ε is $\varepsilon = \Delta^{\frac{p}{p+1}} T^{\frac{1}{p+1}}$. It ends the proof of the theorem.

Remark 8. Using the same method of the proof but for the Lax-Wendroff and Beam-Warming schemes (and assuming (41) is true) one gets an error

$$||u^n - v^n||_{L^1} \leq C\left(\varepsilon + \frac{\Delta x^2}{\varepsilon^2} n^{\frac{1}{4}} T + \Delta x\right)|u|_{BV}.$$

The optimal value of ε is now

$$\varepsilon = \Delta x^{\frac{2}{3} - \frac{1}{12}} T^{\frac{1}{2}}, \quad (n = \frac{T}{\Delta t} = \frac{\Delta x T}{a}).$$

It means we loose a (small) factor $\frac{1}{12}$ with respect to the optimal value which should be $\frac{2}{3}$. One more times this is not true for small Courant number for the Lax Wendroff scheme.

Remark 9. The schemes which are classically used for wave computations are non dissipative [1, 6]. With our notations it means $q = \infty$. Most probably these schemes are not stable in L^1 neither in L^∞ even in dimension one.

6 Numerical experiments

We study the stability and convergence in L^1 by means of numerical tests and comparisons with the theoretical results. All numerical tests have done with periodic boundary conditions.

First the initial condition is a (numerical) Dirac profile. That the discrete initialization is one in one cell and zero in all other cells. The L^1 norm of the upwind, Lax-Wendroff, Beam-Warming and order three schemes are displayed in figure 1 (Lax-Wendroff) and 2 (order three). For this CFL ν the results for the Beam-Warming scheme are quite similar to those of the Lax-Wendroff scheme. For the simulation with 800 cells, a very good approximation is

$$t \mapsto Ct^\alpha, \quad \alpha \approx 0.1058$$

which is smaller than the theoretical bound $\alpha_{\text{theor}} = \frac{1}{4} = 0.25$ (see (40) with $q = 3$). On the other hand the A-stability of the third order scheme is evident on figure 2. Similar results have been observed at any order $1 \leq p \leq 20$ for the schemes defined in [3].

Newt we study the convergence of these schemes to the numerical solution issued from the discontinuous initial data

$$u_0(x) = 1 \text{ for } 0 < x < 0.5, \quad u_0(x) = 1 \text{ for } 0.5 < x < 1$$

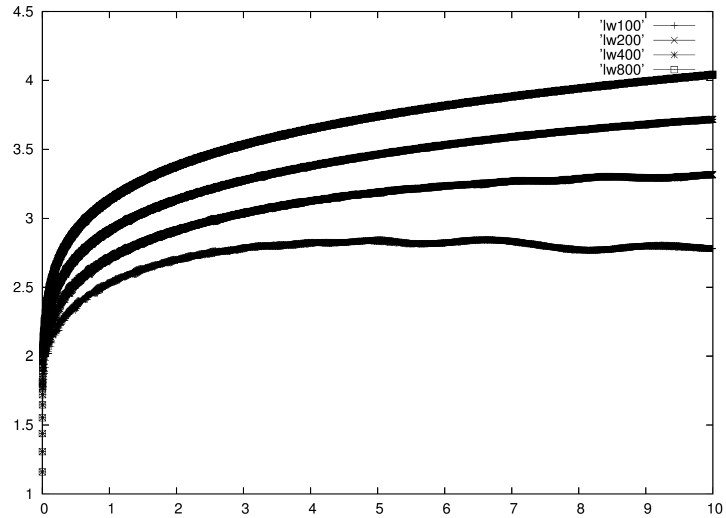

Figure 1: Non A-stability in L^1 of the Lax-Wendroff scheme. The norm increases with respect to T and $\frac{1}{\Delta x}$. Computations done with 100, 200, 400 and 800 cells on aperiodic domain. Final time $T = 10$. The CFL number is $\nu = 0.2$.

with periodic boundary conditions. In table 2 are given the errors and experimental order of convergence in the L^1 norm. All schemes seem to converge with the order $\frac{p}{p+1}$ where $p = 1$ for Upwind, $p = 2$ for Lax-Wendroff and Beam-Warming and $p = 3$ for $O3$. For Upwind and $O3$ this is exactly the theoretical prediction. For Lax-Wendroff and Beam-Warming it can appear as quite surprising. However it is somehow compatible with remark 8 which shows the loss of convergence order is bounded by $\frac{1}{12}$. This is probably too small to be able to distinguish it in our numerical tests.

Table 3 shows the numerical orders of convergence in L^2 are compatible with the theory.

Table 4 shows the numerical orders of convergence in L^1 and L^2 for a small Courant number $\nu = 0.001$. In such a case the dispersion of the Lax-Wendroff scheme $-\frac{\nu(1-\nu^2)}{6}$ outperforms the diffusion $\frac{\nu^2(1-\nu^2)}{8}$ because there is a ν^2 in the diffusion coefficient. The hypotheses of remark 7 are no more true ($C_2 = 0$ with our notations (16)). In this case the numerical order of convergence of the Lax-Wendroff scheme is clearly lower than $\frac{2}{3}$. It is remarka le to notice that the order of convergence in L^2 ($\frac{1}{3}$) is compatible with the theory. The O3 scheme convergence rate is in agreement with the theory both in L^1 and L^2.

One has the same behavior for the Beam-Warming scheme and the O3 scheme for large Courant numbers. In the numerical test of table 5 the Courant number is $\nu = 0.99$. The final time must be large enough for letting the oscillations. The explanation is the same: for large $\nu \approx 1$ the dispersion $\frac{\nu(\nu-1)(\nu-2)}{6}$

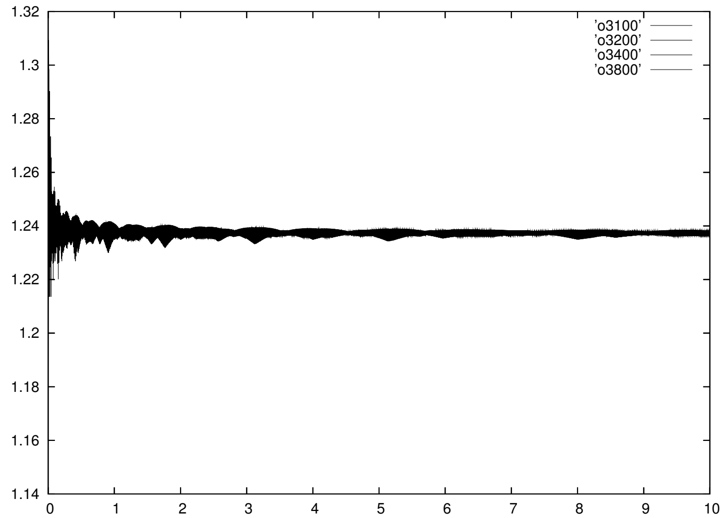

Figure 2: A-stability in L^1 of the order-3 scheme. The norm is uniformly bounded with respect to T and $\frac{1}{\Delta x}$. Computations done with 100, 200, 400 and 800 cells on a periodic domain. Final time $T = 10$. The CFL number is $\nu = 0.2$.

is much greater than diffusion $-\frac{\nu(1-\nu)^2(2-\nu)}{8}$, see equation (12).

Finally we show the numerical profiles figure in 3. LW and $O3$ for a artificially small Courant number $\nu = 0.01$. Non optimal rate of convergence for LW of (tables 4 is correlated with huge oscillations.

Cells	Upwind	LW	BW	O3
100	0.142533	0.105236	0.090885	0.039297
200	0.100855	0.070092	0.060767	0.023352
400	0.071349	0.046024	0.040658	0.013926
800	0.050457	0.030256	0.026857	0.008292
1600	0.035679	0.019958	0.017683	0.004943
order	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{3}{4}$

Table 2: Errors and order of convergence in L^1. CFL=0.2

The author warmly thanks Phillipe Villedieu for many fruitful discussions about the in-and-outs of high order finite volume transport schemes.

100	Upwind	LW	BW	O3
100	0.204370	0.165566	0.154501	0.100967
200	0.171888	0.132580	0.125728	0.078042
400	0.144545	0.104283	0.103494	0.059947
800	0.121560	0.083612	0.083504	0.046356
1600	0.102227	0.067617	0.066693	0.035915
order	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{3}{8}$

Table 3: Errors and order of convergence in L^2. CFL=0.2

cells	LW (L^1)	LW (L^2)	O3 (L^1)	O3 (L^2)
100	0.136120	0.183949	0.040989	0.102097
800	0.056499	0.094048	0.008625	0.047275
order	$\frac{1}{3} < 0.42 < \frac{2}{3}$	$\frac{1}{3}$	$\frac{3}{4}$	$\frac{3}{8}$

Table 4: Errors and order of convergence in L^1 and L^2. CFL=0.001. The convergence of LW is non optimal in L^1

References

[1] C. Bailly and C. Bogey, Numerical solution of acoustic propagation problems using linearized Euler equations, AIAA Journal, 38(1), 22-29, 2000.

[2] A. V. Porubov, D. Bouche and G. Bonnaud, Asymptotic solutions of differential approximation may reflect the features of the schemes of non-linear advection equation, preprint (communicated by D. Bouche).

[3] S. Delpino and H. Jourdren, Arbitrary high-order schemes for linear advection and waves equations: application to hydrodynamics and aeroacoustic, Comptes Rendus Acad. Sciences, I, 2006

[4] S. Godunov and Ryaben'kii, Introduction to the theory of difference schemes, Fizmatgiz, 1962.

[5] A. Iserles and G. Strang, The optimal accuracy of difference schemes, Trans. of the AMS, Vol. 277, 2, 198, 779–803.

[6] A. Laurent, Joly, Patrick and Q.H. Tran, Construction and analysis of higher order finite difference schemes for the 1D wave equation. (English) Comput. Geosci. 4, No.3, 207-249 (2000).

[7] P.D. Lax and B. Wendroff, On the stability of difference schemes, Comm. Pure and Appl. Math., 15 1962, 363–371.

[8] R.D. Richtmeyer and K.W. Morton, Difference methods for initial value problems, 2nd edition, John Wiley and Sons, 1967.

cells	BW (L^1)	BW (L^2)	O3 (L^1)	O3 (L^2)
100	0.131108	0.182966	0.041018	0.102122
800	0.050060	0.091611	0.008641	0.047280
order	$\frac{1}{3} < 0.42 < \frac{2}{3}$	$\frac{1}{3}$	$\frac{3}{4}$	$\frac{3}{8}$

Table 5: Errors and order of convergence in L^1 and L^2. CFL=0.99. The convergence of BW is non optimal in L^1

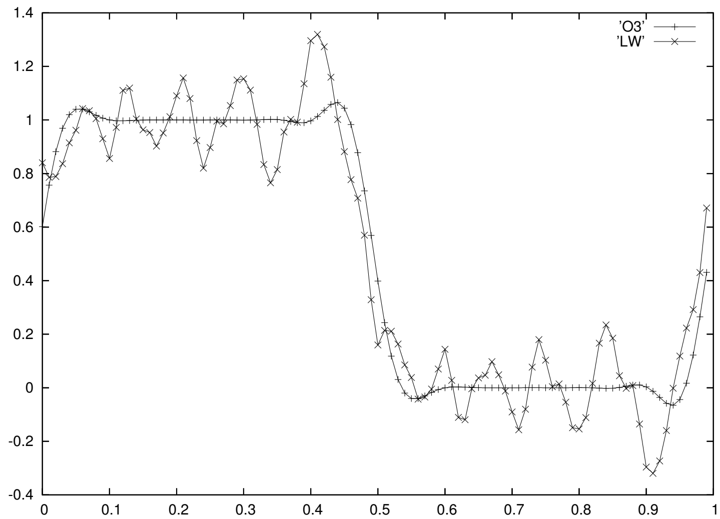

Figure 3: Huge and non smooth oscillations for the Lax-Wendroff scheme for very small CFL number ($\nu = 0.01$). The oscillation of O3 are under control. $T = 1$, 100 cells

[9] T. Schwartzkopff, M. Dumbser and C.D. Munz, Fast high order ADER schemes for linear hyperbolic equations and their numerical dispersion and dissipation, JCP, 197, 532–538, 2004.

[10] V.A. Titarev and E.T. Toro, High order ADER schemes for scalar advection-diffusion-reaction equations, J. Comp. Fluid Dynamics, 1–6, Vol 12, 1, 2003.

[11] R.F. Warming and R.M. Beam, Recent advances in the development of implicit schemes for the equations of gas dynamics, Seventh International Conf. on Numerical Methods in Fluid Dynamics, 429–433, Lecture Notes in Physics, Springer, 1981.