Non-linear CFL Conditions Issued from the von Neumann Stability Analysis for the Transport Equation

Erwan Deriaz, Pierre Haldenwang

Non-linear CFL Conditions Issued from the von Neumann Stability Analysis for the Transport Equation

Erwan Deriaz [1] , **Pierre Haldenwang** [2]

Abstract
This paper presents a theory of the possible non-linear stability conditions encountered in the simulation of convection dominated problems. Its main objective is to study and justify original CFL-like stability conditions thanks to the von Neumann stability analysis. In particular, we exhibit a wide variety of stability conditions of the type $\Delta t \leq C \Delta x^{\alpha}$ with Δt the time step, Δx the space step, and α a rational number within the interval [1, 2]. Numerical experiments corroborate these theoretical results.

Keywords CFL condition · von Neumann stability · Transport equation · Runge–Kutta schemes · Finite differences · Turbulence

Introduction

This paper prospects the stability conditions coming from the von Neumann stability analysis of the transport equation with various finite difference discretizations. It provides a large variety of CFL-like stability conditions. These conditions observed in finite difference methods [9] also apply to other numerical methods such as Discontinuous Galerkin [25], Continuous Finite Elements [7,13] as well as Cartesian Level Set methods [10].

In the following we denote the time step by Δt, the space step by Δx and the velocity by a. When a is omitted then Δx stands for $\Delta x/a$. The CFL condition, a physical criterion from the founding paper [4] based on the influence areas of the points, asserts that the numerical

✉ Erwan Deriaz
erwan.deriaz@univ-lorraine.fr

Pierre Haldenwang
haldenwang@l3m.univ-mrs.fr

[1] Institut Jean Lamour, Campus Artem, 2 allée André Guinier, BP 50840, 54011 Nancy Cedex, France

[2] Laboratoire de Mécanique, Modélisation et Procédés Propres, 38 rue Frédéric Joliot-Curie, 13451 Marseille Cedex 20, France

simulation of a transport phenomenon at speed a has to satisfy $\Delta t \leq C\,\Delta x/a$ with C a constant close to 1. It is a necessary condition for the numerical stability of explicit schemes.

Nevertheless, the linear CFL condition may not be sufficient to ensure the numerical stability. Thanks to the von Neumann stability analysis [1], and allowing a controlled exponential growth of the error, it is possible to establish stability conditions of the type $\Delta t \leq C\,\Delta x^{\alpha}$ with $\alpha \in [1, 2]$ a rational number. This exponent α is given by $\alpha = \frac{p(2q-1)}{q(2p-1)}$ with p and q integers such that $q \geq p > 0$. This result was announced but not justified in [6]. Here we detail the mathematical premise. Then we also prove and test these non-linear CFL conditions.

The organization of the paper is the following: in the first part we briefly review the stability domain of the time schemes, mainly studied for the ordinary differential equation theory and comprehensively presented in references [14,15], in the second part we discuss the spectra of the finite difference approximations, in the third part, we combine these two elements to make the stability condition $\Delta t \leq C\,\Delta x^{\alpha}$ appear, and in the fourth part we test this condition numerically.

1 Stability Domain of the Time Schemes

Many of the time schemes used in Partial Differential Equations [11] were developed in the frame of Ordinary Differential Equations [14,15]. We will briefly review the usual von Neumann stability analysis of the schemes for ODE's.

Let us consider the differential equation

$$\partial_t u = F(u), \quad t \geq 0. \tag{1.1}$$

To solve this equation numerically, one can use an explicit one step scheme—i.e. a scheme of Runge–Kutta type:

$$u_{(i)} = u_n + \Delta t \sum_{j=0}^{i-1} a_{ij} F(u_{(j)}), \quad \text{for } 0 \leq i \leq s, \quad \text{and } u_{n+1} = u_{(s)}, \tag{1.2}$$

where u_n is an approximation of the solution u at time $n\,\Delta t$.
Or, one can use an explicit multistep scheme [20] which is written:

$$u_{n+1} = \sum_{i=0}^{r} a_i u_{n-i} + \Delta t \sum_{i=0}^{s} b_i F(u_{n-i}). \tag{1.3}$$

Depending on the choice of the coefficients a_{ij} or a_i and b_i, the numerical scheme will have a certain accuracy and certain stability features.

In order to derive the stability properties of these schemes it is usual to consider the differential equation $\partial_t u = \zeta\, u$ i.e. the case $F(u) = \zeta\, u$ with $\zeta \in \mathbb{C}$. Then the scheme (1.2) satisfies

$$u_{n+1} = G(z) u_n, \tag{1.4}$$

where $z = \zeta\,\Delta t$. The polynomial function G is called the amplification factor.

For multistep schemes, the definition of the amplification factor has to be more general. Let

$$X_n = \begin{pmatrix} u_n \\ u_{n-1} \\ \vdots \\ u_{n-s} \end{pmatrix}, \tag{1.5}$$

then the scheme (1.3) corresponds to $X_{n+1} = M(z)X_n$, where $M(z)$ is a $(s+1) \times (s+1)$ matrix. As $X_{i+n} = M(z)^n X_i$, we will be interested in the eigenvalues of $M(z)$ and particularly in the one with maximal norm. Hence, in the multistep case, $|G(z)|$ stands for the maximum of the norms of the eigenvalues.

In the theory of ordinary differential equations [14], the von Neumann stability domain gathers all the points $z \in \mathbb{C}$ where the amplification factor is less than one:

$$\mathcal{D} = \{z \in \mathbb{C}, |G(z)| \leq 1\}. \tag{1.6}$$

In Fig. 1 we represented the stability domains of the first four Runge–Kutta schemes and of a Runge–Kutta scheme of order 5 taken from the numerical analysis manual [5].

When we have to deal with partial differential equations, the semi-discretization of the equations (only in space, and not in time), and their diagonalization (when it is possible) stem to multiple ordinary differential equations. For instance, the transport equation on the domain $\mathbb{T} = [0, 1]$ with periodic boundary conditions

$$\partial_t u + a\partial_x u = 0, \quad (t, x) \in [0, T] \times \mathbb{T}$$

for $a \in \mathbb{R}$ given and $u : \mathbb{T} \to \mathbb{R}$ the unknown, when using a Fourier spectral discretization in space stems to

$$\partial_t \hat{u} + ia\xi\hat{u} = 0, \quad t \in [0, T] \tag{1.7}$$

for each wave number $\xi \in 2\pi[-N/2, N/2]$ of the discretization

$$u(t, x) = \sum_{k=-N/2}^{N/2} \hat{u}(t, k)\, e^{2i\pi kx}$$

in the usual Fourier basis. So with this discretization, the time scheme is stable for the transport equation if $-ia\xi \in \mathcal{D}$ for $\xi \in 2\pi[-N/2, N/2]$.

Hence the position of the domain $\mathcal{D}$ with respect to the y-axis near zero dictates the behavior of the schemes for simulating transport-dominated equations. If $\partial\mathcal{D}$ goes to the right (RK3, RK4) then the scheme is dissipative and stable when the stability domain $\mathcal{D}$ encloses the segment $ia\pi\,\Delta t[-N, N]$. Therefore it is stable under the linear CFL condition $\Delta t \leq C\Delta x/a$. If $\partial\mathcal{D}$ goes to the left (Euler, RK2, RK5) then the scheme is excitatory. Therefore it calls a stronger CFL condition $\Delta t \leq C(\Delta x/a)^\alpha$ with $\alpha > 1$ [3,6,8].

In Fig. 2 we show the stability domains of the first five Adams–Bashforth schemes. According to the behavior of the tangencies along the y-axis, we can say that the schemes AB3 and AB4 are stable under the usual linear CFL condition while the schemes Euler, AB2 and AB5 are not.

Figures 1 and 2 plot the boundary of the stability domain $\mathcal{D}$, Eq. (1.6), for various classical schemes. This curve corresponds to the set $\partial\mathcal{D} = \{z \in \mathbb{C} \text{ s.t. } G(z) = e^{i\theta}, \theta \in [-\pi, \pi]\}$. We invert the relation $G(z) = e^{i\theta}$ to parameterize the curve $\partial\mathcal{D}$ for small values of $|z|$:

$$z = G^{-1}(e^{i\theta}), \quad \theta \in \mathbb{R} \text{ close to } 0. \tag{1.8}$$

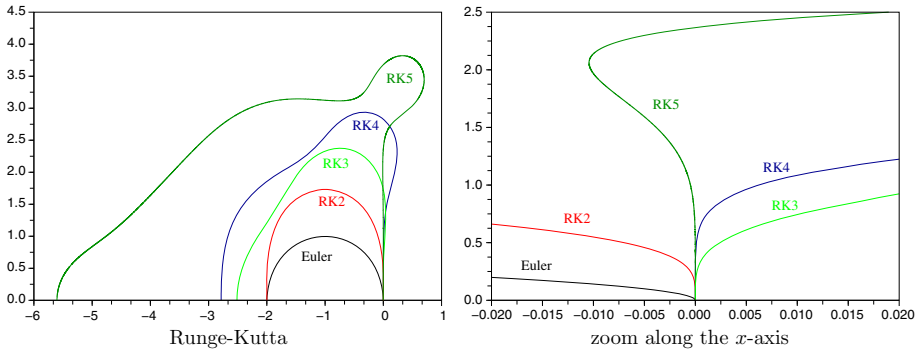

Fig. 1 Von Neumann stability domains for first five Runge–Kutta schemes. Left: on the half-plan. Right: zooming in the x direction. The fifth order Runge–Kutta scheme is taken from [5]

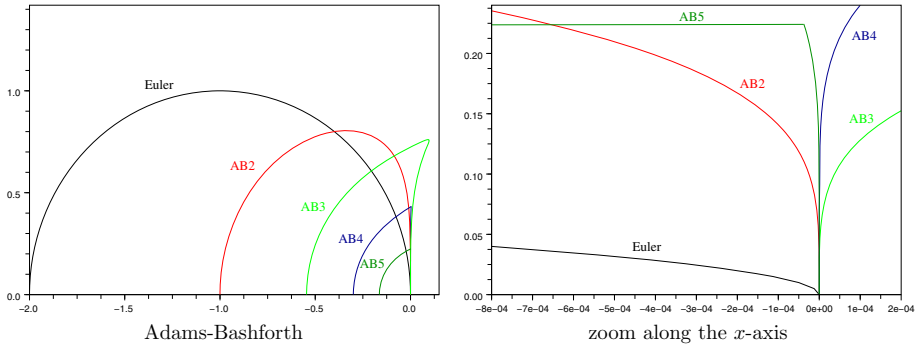

Fig. 2 Von Neumann stability domains for the first five Adams–Bashforth schemes. Left: on the half-plan. Right: zooming in the x direction

In a neighborhood of zero, the boundary of $\mathcal{D}$ satisfies:

$$z = i(\theta + o(\theta)) - T_{\mathcal{D}}\theta^{2p} + o(\theta^{2p}), \quad \theta \in \mathbb{R} \text{ close to } 0. \tag{1.9}$$

The parameters $p \in \mathbb{N}$ and $T_{\mathcal{D}} \in \mathbb{R}$ are computed from the coefficients of the polynomial $G(z)$ for the one step schemes Eq. (1.2), or by writing directly $e^{i\theta}$ as an eigenvalue of $M(z)$ for the multistep schemes Eq. (1.3). We refer to [6] for more details.

2 Space Differentiation

2.1 Spectrum of the Discrete Differentiation

Finite differences originate from the approximation of functions by polynomials, with a first appearance in Newton's notes as early as the seventeenth century [17]. It was the first approximation used to solve Partial Differential Equations numerically, particularly in computational fluid dynamics [19]. Following [22] which introduces to finite difference approximations, we extract the important features of the numerical approximations regarding the CFL stability. These are necessary to unwrap the non-linear CFL stability criteria.

We consider the problem of approximating the pointwise differentiation of a function. This operation modifies the spectrum of the differential operator—contrarily to a spectral discretization. Given a smooth function $u : \mathbb{R} \to \mathbb{R}$ discretized on a regular space grid $\{k\,\Delta x, k \in \mathbb{Z}\}$, we approximate $\partial_x u(x)$ for $x \in \Delta x\mathbb{Z}$ using the points $\{x - m\,\Delta x, \ldots, x + n\,\Delta x\}$ for $m, n \in \mathbb{N}$—i.e. the stencil [-m,n]—by:

$$\mathcal{A}\,u(x) = \frac{1}{\Delta x} \sum_{k=-m}^{n} a_k u(x + k\,\Delta x) \tag{2.1}$$

where the operator $\mathcal{A}$ denotes a discrete differentiation approximating ∂_x.

The spectrum of $\mathcal{A}$ can be obtained by taking $u(x) = e^{\frac{i\xi x}{\Delta x}}$, then $\mathcal{A}\,u(x) = \frac{1}{\Delta x} A(\xi)\,u(x)$ with

$$A(\xi) = \sum_{k=-m}^{n} a_k e^{ik\xi}, \quad \xi \in [-\pi, \pi]. \tag{2.2}$$

We plot this spectrum on the plan of complex numbers by considering the curve

$$\mathcal{S} = \{A(\xi), \xi \in [-\pi, \pi]\} \subset \mathbb{C}. \tag{2.3}$$

Since $a_k \in \mathbb{R}$ for all k, we have $A(-\xi) = \overline{A(\xi)}$. Hence the curve presents an axial symmetry along the x-axis.

To be consistent, the discrete differentiation $\mathcal{A}$ has to satisfy $\mathcal{A}\,u(x) = \partial_x u(x) + O(\Delta x)$ i.e.

$$\begin{cases} \sum_k a_k = 0 \\ \sum_k k\,a_k = 1 \end{cases} \quad \text{which corresponds to} \quad \begin{cases} A(0) = 0 \\ A'(0) = i \end{cases}. \tag{2.4}$$

The other interesting properties are listed below:

- the first and most important issue is the order of the approximation. An approximation $\mathcal{A}$ is said to be consistent of order $\beta > 0$ if it satisfies

$$\exists C > 0, \quad \forall u \in C_0^\infty, \quad \|\partial_x u - \mathcal{A}\,u\|_{L^2} \leq C\,\Delta x^\beta\,\|\partial^{\beta+1} u\|_{L^2}, \tag{2.5}$$

- another important issue affects the stability with respect to the time integration: the upwind or downwind feature of a scheme. A necessary condition for the time integration to be stable is that the spectrum of the differentiation have no strictly positive real part i.e. $\mathcal{S} \subset \mathbb{C}^- = \{\xi \in \mathbb{C}, \mathrm{Re}\,\xi \leq 0\}$, see Fig. 3. For centered finite differences, we have $\mathcal{S} \subset i\mathbb{R}$.
- and finally, another important issue, essential for the present article, regards the conservativity of the space discretization. The conservativity of a differentiation is linked to the tangency of its spectrum $\mathcal{S}$ to the y-axis:

$$A(\xi) = i\,(\xi + o(\xi)) - T_{\mathcal{S}}\,\xi^{2q} + o(\xi^{2q}) \tag{2.6}$$

with $q \in \mathbb{N}^*$ and $T_{\mathcal{S}} > 0$.

The link between the tangency expression (2.6) and the conservativity goes as follows: let us assume that we are using a differentiation $\mathcal{A}$ satisfying (2.6), then the method of lines (i.e. considering exclusively the space discretization) transforms the transport equation $\partial_t u = \partial_x u$, into $\partial_t \hat{u}(t, \xi) = \Delta x^{-1} A(\xi\,\Delta x)\hat{u}(t, \xi)$. The solution of this equation is given by $\hat{u}(t, \xi) = \exp(\Delta x^{-1} A(\xi\,\Delta x)t)\hat{u}(0, \xi)$. So, for $\Delta x \to 0$, according to (2.6), the damping of the frequency ξ may be approximated by

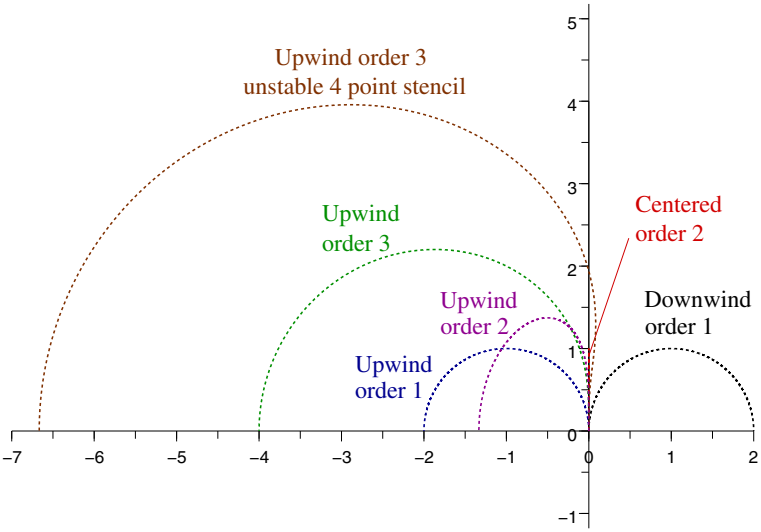

Fig. 3 Spectra of several finite difference schemes for space differentiation. These spectra have to be compared with the domains of stability of the temporal schemes Figs. 1 and 2

$$|\hat{u}(t,\xi)| = \exp(-T_{\mathcal{S}}\,\xi^{2q}\,\Delta x^{2q-1}\,t + o(\Delta x^{2q-1}))|\hat{u}(0,\xi)| \tag{2.7}$$

$$= (1 - T_{\mathcal{S}}\,\xi^{2q}\,\Delta x^{2q-1}\,t + o(\Delta x^{2q-1}))|\hat{u}(0,\xi)|. \tag{2.8}$$

The multiplicative factor goes to 1 as fast as $T_{\mathcal{S}}\,\Delta x^{2q-1}$ goes to zero. Hence a larger q allows a better conservation of the energy. Notwithstanding the time scheme, the frequency ξ will be damped after a typical time $t_\xi = 1/T_{\mathcal{S}}\,\xi^{2q}\,\Delta x^{2q-1}$.

This concerns the upwind differentiations as for instance the first order upwind scheme with $a_0 = -1, a_1 = 1$ in (2.1) which is known to dissipate a lot. As we can see on Fig. 6, the dark blue curve only retains a few coarsest Fourier modes. Symmetric schemes have their spectra included in the y-axis (see Fig. 3 for second order centered scheme spectrum). Hence they conserve the energy exactly [23]. On the other hand, as we can see on Fig. 8 they tend to create spurious oscillations.

2.2 Most Conservative Differentiations

For further experiments, we look for discretized differentiations $\mathcal{A}$ maximizing the tangency of the spectrum to the y-axis, i.e. maximizing the number q in the expression (2.6). From (2.2), we deduce:

$$A(\xi) = \sum_{k=-m}^{n} a_k e^{ik\xi} = \sum_{k=-m}^{n} a_k \sum_{\ell\geq 0} \frac{(ik\xi)^\ell}{\ell!} = \sum_{\ell\geq 0} \frac{i^\ell}{\ell!} \left(\sum_{k=-m}^{n} a_k k^\ell \right) \xi^\ell. \tag{2.9}$$

Maximizing the order of the approximation is equivalent to solving the system of equations

$$\sum_{k=-m}^{n} a_k = 0 \tag{2.10}$$

$$\sum_{k=-m}^{n} k\, a_k = 1 \tag{2.11}$$

$$2 \le \ell \le s, \quad \sum_{k=-m}^{n} k^{\ell} a_k = 0 \tag{2.12}$$

for $s = m + n$, so this provides an order s approximation. A finite differentiation is stable if its spectrum shows no strictly positive real part. If s is odd—i.e. with an even number of points—one can notice that there is only one stencil providing an order s stable finite differentiation, and if s is even, that there are two of them. For each s, one stable instance corresponds to an upwind scheme, and the other—for s even—to a centered scheme.

In order to maximize the tangency of $\mathcal{S}$ to the y-axis, we extract the real part of the polynomial expansion of $A(\xi)$ (2.9). As a result, only the equalities (2.4) and (2.12) with even parameter $\ell = 2\ell'$ enter into play:

$$\ell = 0, \quad \sum_{k=-m}^{n} a_k = 0 \tag{2.13}$$

$$\ell = 1, \quad \sum_{k=-m}^{n} k\, a_k = 1 \tag{2.14}$$

$$2 \le 2\ell' \le 2(q-1), \quad \sum_{k=-m}^{n} k^{2\ell'} a_k = 0 \tag{2.15}$$

If $a_{-k} = -a_k$, then the scheme is centered and $q = +\infty$. If it is not the case, then we have an expansion of the type:

$$A(\xi) = i\,\xi + o(i\,\xi) + \left(\frac{(-1)^q}{(2q)!} \sum_{k=-m}^{n} k^{2q} a_k \right) \xi^{2q} + o(\xi^{2q}). \tag{2.16}$$

From (2.9), the real part of $A(\xi)$ is given by:

$$f(\xi) = \mathrm{Re}(A(\xi)) = a_0 + \sum_{k=1}^{N} (a_k + a_{-k}) \cos(k\,\xi) \tag{2.17}$$

with $N = \max(n, m)$ and (a_k) completed by $a_k = 0$ when k does not remain to the stencil. In the following we consider that $N = n$. We maximize the number of vanishing derivatives of f at zero, by taking

$$f(\xi) = -K_{m,n}(1 - \cos\xi)^n, \tag{2.18}$$

with $K_{m,n}$ a positive constant for $n \ge m$. With this choice, $q = n$ in the expansion (2.16).

Remark 1 This form for the real part of $A(\xi)$ maximizes the tangency to the y-axis. And, since it does not contain any strictly positive value, it provides a stable upwind differentiation.

For $m = 0$, this leads to the unique solution $K_{0,n} = \frac{2^{n-1}}{C_{2n-2}^{n-1}}$, $a_0 = -\frac{C_{2n}^{n}}{2\,C_{2n-2}^{n-1}}$ and $a_k = (-1)^{k+1} \frac{C_{2n}^{n+k}}{C_{2n-2}^{n-1}}$ for $k \ge 1$ [12].

For the further numerical experiments, we will be using the case $m = n-1$ with maximum order of approximation (i.e. order $s = 2n - 1$), leading to the coefficients of Table 1 for a velocity going from the right to the left: $\partial_t u + a\,\partial_x u = 0$ with $a < 0$.

Table 1 Upwind finite difference coefficients $(a_k)_{k\in[-m,n]}$ for several stencils $[-n, n+1]$. The resulting orders of approximation are equal to $2q-1$

k / q	-4	-3	-2	-1	0	1	2	3	4	5
1					-1	1				
2				$-\frac{1}{3}$	$-\frac{1}{2}$	1	$-\frac{1}{6}$			
3			$\frac{1}{20}$	$-\frac{1}{2}$	$-\frac{1}{3}$	1	$-\frac{1}{4}$	$\frac{1}{30}$		
4		$-\frac{1}{105}$	$\frac{1}{10}$	$-\frac{3}{5}$	$-\frac{1}{4}$	1	$-\frac{3}{10}$	$\frac{1}{15}$	$-\frac{1}{140}$	
5	$\frac{1}{504}$	$-\frac{1}{42}$	$\frac{1}{7}$	$-\frac{2}{3}$	$-\frac{1}{5}$	1	$-\frac{1}{3}$	$\frac{2}{21}$	$-\frac{1}{56}$	$\frac{1}{630}$

Remark 2 For a finite number of points N with periodic boundary conditions, the Fourier modes of the simulation $e^{\frac{i\xi x}{\Delta x}}$ form a discrete set with $\{\xi = \pi\frac{k}{N},\ -N+1 \le k \le N\}$. We plotted the spectra of the discretized operators $\mathcal{A}$ applied to these modes in Fig. 3.

3 Stability Condition $\Delta t \le C\,\Delta x^{\alpha}$ with $\alpha > 1$

The partial differential equation (1.7) has solutions of the type $u(t,x) = f(x-at)$. So in particular for $\xi \in \mathbb{R}$ we have as a solution, the wave function

$$u(t,x) = u_0 e^{i\xi(x-at)}.$$

Discretized with Δx and Δt, $x_k = k\Delta x$ and $t_n = n\Delta t$, for $u_n(x_k) = u(t_n, x_k) = e^{i\xi(x_k - at_n)}$ the numerical scheme provides $u_{n+1} = g(\xi, \Delta t, \Delta x)u_n$. Following [21], the scheme is stable under the condition $|g(\xi, \Delta t, \Delta x)| \le 1 + C\Delta t$ with $C \ge 0$ a constant which limits the exponential growth of the numerical solution:

$$\|u_n\| \le |g(\xi, \Delta t, \Delta x)|^n \|u_0\| \le (1 + C\Delta t)^{(t_n/\Delta t)}\, \|u_0\| \le e^{C\,t_n}\|u_0\|. \tag{3.1}$$

We concentrate on the module of $g(\xi, \Delta t, \Delta x)$ but numerical accuracy also depends on its phase $-i\xi\omega(\xi, \Delta t, \Delta x)\Delta t$ defined by

$$g(\xi, \Delta t, \Delta x) = |g(\xi, \Delta t, \Delta x)| e^{-i\xi\omega(\xi, \Delta t, \Delta x)\Delta t}$$

and how close $\omega(\xi, \Delta t, \Delta x)$ remains to a.

Considering Eq. (3.1), we notice that the case $C > 0$ is more permissive than the case $C = 0$. The case $C > 0$ is actually encountered in numerical simulations and allows to explain more accurately some actual stability phenomena. Then we deduce the following result which links the non linear CFL condition to the stability domain of the time scheme and to the spectrum of the discrete differentiation applied to the space variable:

Theorem 3.1 *Let us consider a numerical solver for the transport equation $\partial_t u + a\,\partial_x u = 0$ using finite differences with time step Δt and space step Δx. If, close to zero, the boundary of the stability domain of the time integrator—noted $\mathcal{D}$ in (1.6)—satisfies:*

$$\partial\mathcal{D}: \quad z = i(\theta + o(\theta)) - T_{\mathcal{D}}\theta^{2p} + o(\theta^{2p}) \tag{3.2}$$

for $p \in \mathbb{N}$ and $T_{\mathcal{D}} > 0$, and if the spectrum $\mathcal{S} = \{A(\xi), \xi \in [-\pi, \pi]\}$ of the discrete differentiation—from (2.3)—satisfies:

$$\mathcal{S}: \quad z = i(\theta + o(\theta)) - T_{\mathcal{S}}\theta^{2q} + o(\theta^{2q}) \tag{3.3}$$

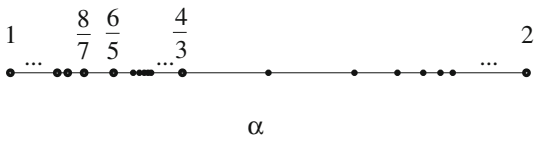

Fig. 4 Possible locations of the exponent α

for $q \in \mathbb{N}$ and $T_{\mathcal{S}} > 0$, then the von Neumann stability condition

$$\|u_{n+1}\|_{L^2} \leq (1 + C\Delta t)\|u_n\|_{L^2} \tag{3.4}$$

of the numerical solution (u_n) is given by

$$\Delta t \leq C' \Delta x^{\alpha} \tag{3.5}$$

with $C' > 0$ a constant independent of Δt and Δx and with

- $\alpha = 1$ *if* $q \leq p$ *(stable under linear CFL condition),*
- $\alpha = \frac{p(2q-1)}{(2p-1)q} \in (1, 2]$ *if* $q > p$*, and in this case, the constant C' from (3.5) is written:*

$$C' = C^{\frac{q-p}{q(2p-1)}} \frac{T_{\mathcal{S}}^{\frac{p}{q(2p-1)}}}{T_{\mathcal{D}}^{\frac{1}{2p-1}}} \left(\left(\frac{p}{q}\right)^{\frac{p}{q-p}} - \left(\frac{p}{q}\right)^{\frac{q}{q-p}} \right)^{-\frac{q-p}{q(2p-1)}}, \tag{3.6}$$

with p and $T_{\mathcal{D}}$ from (3.2), q and $T_{\mathcal{S}}$ from (3.3) and C from (3.4).

Remark 1 If $q = +\infty$ (centered schemes, spectral method) it makes appear the case presented in [6] for conservative space discretizations: $\alpha = \frac{2p}{2p-1}$.

Remark 2 The localization of all possible values of α is plotted in Fig. 4. It consists in accumulations from the left to the points of the sequence $(\frac{2p}{2p-1})_{p\geq 1}$. This sequence itself has an accumulation point at 1, from the right.

Remark 3 This result also applies to other schemes than finite differences. Following Cockburn et al. [2], a Discontinuous Galerkin method combining a second order upwind scheme in space ($q = 2$) with a forward Euler scheme in time ($p = 1$) turns out to be stable under the condition $\Delta t \leq C\Delta x^{\alpha}$ (3.5) with the exponent $\alpha = \frac{3}{2}$, which explains why this numerical scheme "is stable if $\frac{\Delta t}{\Delta x}$ is of order $\sqrt{\Delta x}$".

Proof We prove Theorem 3.1 directly by computation. It also corresponds to geometrical considerations: the spectrum $\frac{\Delta t}{\Delta x} \times \mathcal{S}$ must not get out of the domain of stability $\mathcal{D}$ by more than $C\Delta t$, see Fig. 5.

Geometrically, if $q \leq p$ and $\mathcal{S} \cap i\mathbb{R} = \{0\}$ then for $\tau = \frac{\Delta x}{\Delta t}$ sufficiently small the spectrum $\tau\mathcal{S}$ fits inside the stability domain $\mathcal{D}$.

Let us consider the case $q > p$. The demonstration relies on the same computations as the demonstration of Theorem 3.1 of [6]. The tangency of the stability domain to the y-axis, Eq. (3.2), corresponds to an amplification factor

$$G(z) = \beta_0 + \beta_1 z + \beta_2 z^2 + \cdots \quad \text{such that} \quad |G(i\,\zeta)|^2 = 1 + 2T_{\mathcal{D}}\zeta^{2p} + o(\zeta^{2p}) \tag{3.7}$$

for $\zeta \in \mathbb{R}$ close to zero. We assume the convergence of this power series, hence $G(z) = G_0(z) + o(|z|^{2p})$ with G_0 a polynomial.

We are interested in the case that deviates a little from the spectral case $z \in i\,\mathbb{R}$, and we need to compute $|g(\xi, \Delta t, \Delta x)| = |G(\frac{\Delta t}{\Delta x} A(\xi))|$ for $A(\xi) = i\xi - T_{\mathcal{S}}\xi^{2p} + o() \notin i\,\mathbb{R}$. We

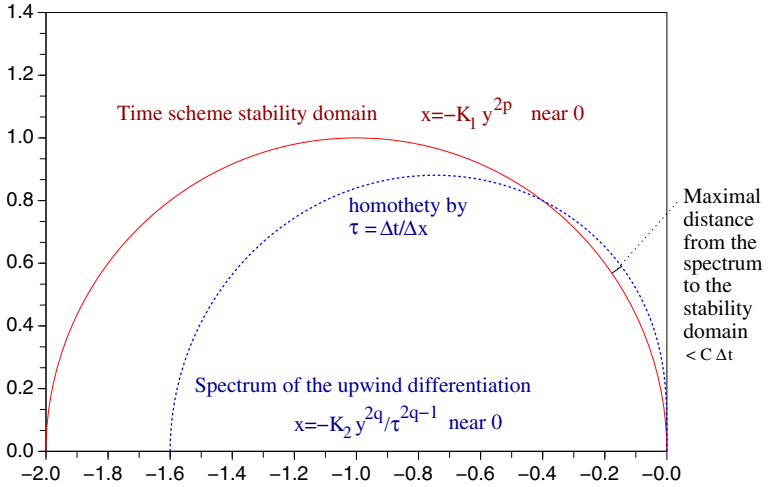

Fig. 5 Stability domain $\mathcal{D}$ of the time scheme and spectrum $\mathcal{S}$ of the discretized finite difference space operator

put $\tau = \frac{\Delta t}{\Delta x}$. Then $\tau A(\xi) = a + i\,b$ with $a = \tau Re(A(\xi)) = -T_{\mathcal{S}}\tau\,\xi^{2q} + o(\tau\,\xi^{2q})$ and $b = \tau Im(A(\xi)) = \tau\,\xi + o(\tau\,\xi)$ for $\xi \to 0$. As $\tau \to 0$ and $\xi \to 0$, a and b tend to zero with $a = o(b)$. Given that $\beta_0 = \beta_1 = 1$, we can write

$$G(a + i\,b) = 1 + (a + i\,b) + \beta_2(a + i\,b)^2 + \cdots + \beta_{2p}(a + i\,b)^{2p} + o(|z|^{2p}) \quad (3.8)$$

then

$$|G(a + i\,b)|^2 = (1 + a + \beta_2 a^2 - \beta_2 b^2)^2 + (b + 2\beta_2 a\,b)^2 + \cdots + o(b^{2p}) \quad (3.9)$$

$$= 1 + 2a - 2\beta_2 b^2 + \cdots + b^2 + \cdots + o(a) + o(b^{2p}) \quad (3.10)$$

As all the terms $a^{m+1}b^n$ are negligible with respect to a for $m + n > 0$, then

$$|G(a + i\,b)|^2 = |1 + (a + i\,b) + \beta_2(i\,b)^2 + \beta_3(i\,b)^3 + \cdots|^2 + o(a) + o(b^{2p}) \quad (3.11)$$

$$= (1 + a - \beta_2 b^2 + \beta_4 b^4 + \cdots)^2 + (b - \beta_3 b^3 + \cdots)^2 + o(a) + o(b^{2p}) \quad (3.12)$$

$$= 1 + 2a + (1 - 2\beta_2)b^2 + \cdots + \left(\sum_{j=0}^{2\ell}(-1)^{\ell+j}\beta_j\beta_{2\ell-j}\right)b^{2\ell} + \cdots$$

$$+ o(a) + o(b^{2p}) \quad (3.13)$$

$$= 1 + 2a + 2T_{\mathcal{D}}b^{2p} + o(a) + o(b^{2p}) \quad (3.14)$$

Now, using the fact that $a = -T_{\mathcal{S}}\tau\,\xi^{2q} + o(\tau\,\xi^{2q})$, $\xi \in [-\pi, \pi]$, and $b = \tau\,\xi + o(\tau\,\xi)$, we look for the maximal value of:

$$f(\tau, \xi) = 1 - 2T_{\mathcal{S}}\tau\,\xi^{2q} + 2T_{\mathcal{D}}\tau^{2p}\xi^{2p} = |G(\tau A(\xi))|^2 + o(\tau\,\xi^{2q}) + o(\tau^{2p}\xi^{2p}) \quad (3.15)$$

This maximum is reached for

$$\xi = \xi_0 = \left(\frac{p\,T_{\mathcal{D}}}{q\,T_{\mathcal{S}}}\right)^{\frac{1}{2(q-p)}}\tau^{\frac{2p-1}{2(q-p)}} \quad (3.16)$$

then

$$|G(\tau\, A(\xi_0))|^2 = 1 + 2\frac{T_{\mathcal{D}}^{\frac{q}{q-p}}}{T_{\mathcal{S}}^{\frac{p}{q-p}}}\left(\left(\frac{p}{q}\right)^{\frac{p}{q-p}} - \left(\frac{p}{q}\right)^{\frac{q}{q-p}}\right)\tau^{\frac{p(2q-1)}{q-p}} + o(\tau^{\frac{p(2q-1)}{q-p}}). \quad (3.17)$$

Hence, knowing that $\tau = \frac{\Delta t}{\Delta x}$, the von Neumann stability condition $|G(\tau\, A(\xi_0))|^2 \leq 1 + 2C\,\Delta t$ is given by

$$\left(\frac{\Delta t}{\Delta x}\right)^{\frac{p(2q-1)}{q-p}} \leq C\frac{T_{\mathcal{S}}^{\frac{p}{q-p}}}{T_{\mathcal{D}}^{\frac{q}{q-p}}}\left(\left(\frac{p}{q}\right)^{\frac{p}{q-p}} - \left(\frac{p}{q}\right)^{\frac{q}{q-p}}\right)^{-1}\Delta t \quad (3.18)$$

i.e. by the condition $\Delta t \leq C'\,\Delta x^{\alpha}$ (3.5) with $\alpha = \frac{p(2q-1)}{(2p-1)q}$ and C' given by (3.6). □

4 Numerical Experiment

We test the stability predictions (3.5) of Theorem 3.1 on the simplest case of hyperbolic equation. On the interval [0, 1] with periodic boundary conditions denoted $\mathbb{T}$, we simulate the transport of a function $f : \mathbb{T} \to \mathbb{R}$ by a constant speed, i.e. we solve $\partial_t u = -\partial_x u$ for $(t, x) \in [0, 5] \times \mathbb{T}$ where

$$\begin{aligned} u :&[0, 5] \times \mathbb{T} \to \mathbb{R} \\ &(t, x) \quad \mapsto u(t, x) \end{aligned}$$

with initial condition $u(0, x) = f(x)$ and periodic boundary condition $u(t, 0) = u(t, 1)$. During this process, the function f makes five laps, and the exact solution at time $t = 5$ is: $u(5, x) = f(x)$.

This (rather simple) experiment accounts of numerous simulations where the transport dominates the stability properties of the numerical solution. We apply an explicit scheme of Runge–Kutta type with time step Δt to the PDE $\partial u = F\,u$:

$$u_{n+1} = u_n + b_1\Delta t\; F\,(u_n + b_2\Delta t\; F\,(u_n + \ldots b_{s-1}\Delta t\; F\,(u_n + b_s\Delta t\; F\,u_n)\ldots)) \quad (4.1)$$

where the parameters $(b_r)_{1\leq r\leq s}$ are chosen according to the desired amplification factor. The coefficients β_ℓ of the amplification factor (3.7) are given by:

$$\beta_\ell = \prod_{k=1}^{\ell} b_k.$$

We couple this time integration with the finite differences from Sect. 2.1. And we use the spacing $\Delta x = \frac{1}{N}$ between the N points of the discretization. So it allows to observe the effects of the parameters p—regarding the time integration—and q—regarding the finite differences—on the stability condition $\Delta t_{\max}(\Delta x)$ and to evidence the exponents of Table 2.

For simplicity reasons, we apply second order Runge–Kutta schemes which maximize p and whose coefficients are given by Table 3.

We simulate the transport of the function $f(x) = x$ which presents a discontinuity at $x = 0$ by a constant velocity. Then, in Fig. 6, we represent the Fourier transform of the numerical solution at time $t = 5$ obtained using the second order Runge–Kutta integration scheme in conjunction with several finite difference approximations. Three typical cases appear, using a linear CFL condition:

Table 2 Table of the exponents $\alpha = \frac{p(2q-1)}{q(2p-1)}$ expected for the non-linear CFL stability condition $\Delta t \leq \Delta x^{\alpha}$

q / p	1	2	3	4	5	$+\infty$
1	1	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{7}{4}$	$\frac{9}{5}$	2
2	1	1	$\frac{10}{9}$	$\frac{7}{6}$	$\frac{6}{5}$	$\frac{4}{3}$
3	1	1	1	$\frac{21}{20}$	$\frac{27}{25}$	$\frac{6}{5}$
4	1	1	1	1	$\frac{36}{35}$	$\frac{8}{7}$

Table 3 Table of coefficients b_r forming the numerical schemes (4.1) with parameters $p = 1, 2, 3, 4$ regarding the tangency property (3.2)

r / p	1	2	3	4
1	1			
2	1	$\frac{1}{2}$		
3	1	$\frac{1}{2}$	$\frac{1}{4}$	
4	1	$\frac{1}{2}$	$\frac{2-\sqrt{2}}{2}$	$\frac{2-\sqrt{2}}{4}$

- the numerical scheme is stable but the high frequencies are dumped, e.g. second order Runge–Kutta time scheme with a second order upwind space differentiation ($p = 2$ and $q = 2$, see Fig. 6), or every third or fourth order Runge–Kutta time schemes,
- an instability occurs at high frequency $\xi = \pi \frac{N}{2}$, e.g. Euler or Runge–Kutta 2 in association with centered finite differences (see Figs. 6, 8),
- an instability occurs at low frequency $\xi = C\, N^{-1/2q}\, N$, $q \geq 2$ (see Fig. 6), e.g. Euler with upwind of order two, Runge–Kutta 2 with upwind of order greater or equal to four (see Figs. 7, 8). It can be checked that an order $2n$ scheme has a tangency parameter q at least equal to $n + 1$ as indicated in Theorem 3.2 of [6].

In Figs. 6, 7 and 8, the number of points was chosen equal to $N = 271$, and the time step equal to $\frac{5}{6}\Delta t_{\max}$ with $\Delta t_{\max}$ coming from the second experiment equation (4.2).

In Fig. 6 we can observe the spectra of the transported function with six different finite-difference schemes. The corresponding physical results are plotted in Figs. 7 and 8. We remark that all the numerical solutions present spurious oscillations except when using the upwind order one differentiation. While the centered finite differences allow the best conservativity as seen in Fig. 6, they also alter the aspect of the solution dramatically as seen in Fig. 8, whatever explicit time scheme we choose. This is due to the Gibbs phenomenon and is not related to any stability issue.

Now, let us check the relevance of the stability condition (3.5). For different space steps $\Delta x = \frac{1}{N}$, we compute the maximal time steps $\Delta t_{\max}$ such that

$$\forall \Delta t \leq \Delta t_{\max},\quad \forall n \leq \frac{5}{\Delta t},\quad \|u_n\|_{\mathrm{TV}} \leq 4\, \|u_{\frac{1}{\Delta t}}\|_{\mathrm{TV}}. \tag{4.2}$$

where the constant 4 is arbitrary, and the solution at time $t = 1$, $u_{\frac{1}{\Delta t}}$ chosen so that the Gibbs phenomenon does not interfere with stability issues.

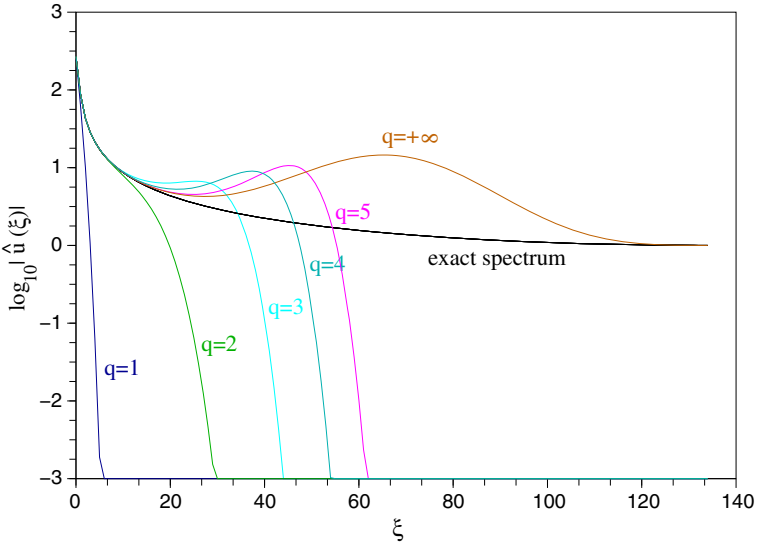

Fig. 6 Fourier transform $\hat{u}(5, \xi)$ of the numerical solution obtained with a second order Runge–Kutta time scheme with $p = 2$ and various finite difference differentiations: upwind schemes of orders 1, 3, 5, 7 and 9 ($q = 1, 2, 3, 4, 5$), and centered scheme ($q = +\infty$)

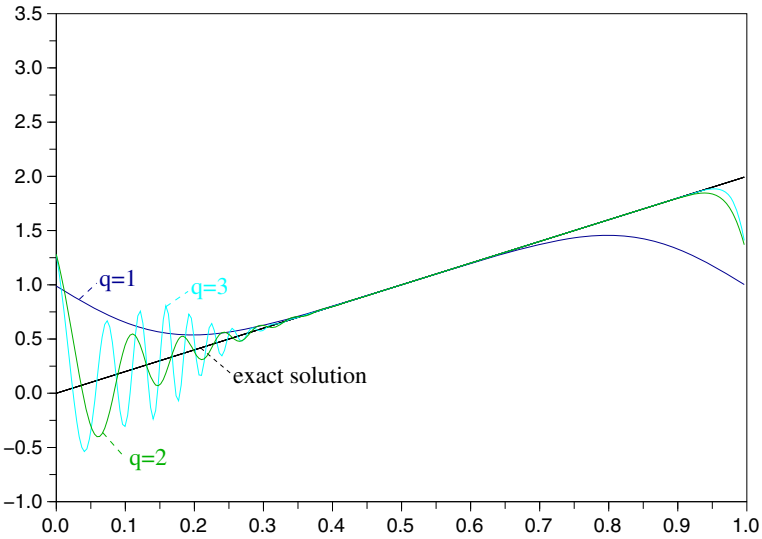

Fig. 7 Physical result of the transport of a discontinuous function using 2nd order Runge–Kutta and upwind of orders 1, 3 and 5 finite differences ($q = 1, 2, 3$) after five loops: $u(5, x)$

We plot the results for various values of p and q in Figs. 9, 10 and 11, taking as an initial condition the continuous function:

$$f(x) = \begin{cases} 5x & \text{for} \quad x \leq \frac{1}{5}, \\ 2 - 5x & \text{for} \quad x \in]\frac{1}{5}, \frac{2}{5}], \\ 0 & \text{for} \quad x > \frac{2}{5} \end{cases} \tag{4.3}$$

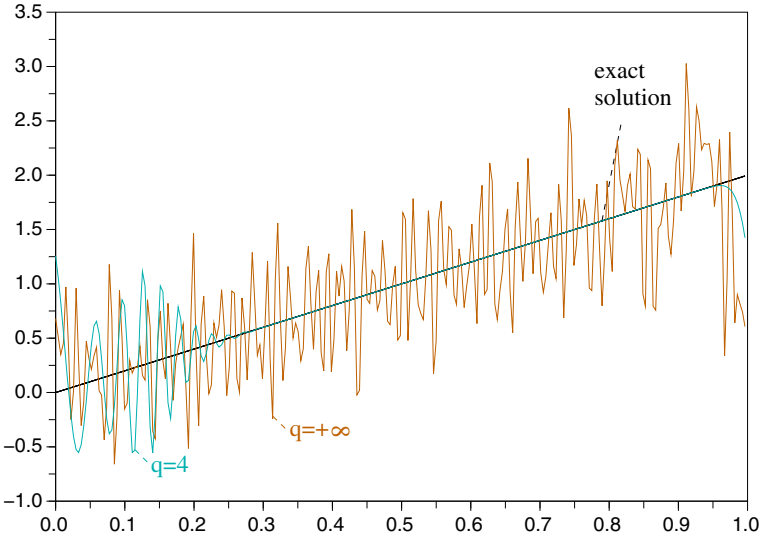

Fig. 8 Physical result of the transport of a discontinuous function using 2nd order Runge–Kutta and 7th order upwind ($q = 4$) and centered ($q = +\infty$) finite differences

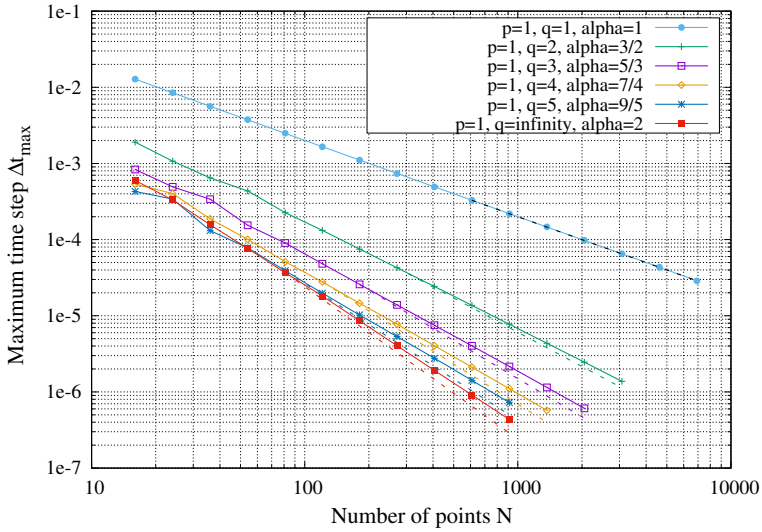

Fig. 9 Numerical experiment with $p = 1$ and various q. The x-axis figures the number N of points in the space discretization, the y-axis, the time step $\Delta t_{\max}$ in log-scale. The dotted lines represent the theoretical slopes $\Delta t = O(N^{-\alpha})$

This experiment corroborates our approach: although the initial condition and the divergence criterion influence the slopes i.e. the exponent α, these are located only slightly above the predicted ones, and their relative positions fit our predictions very well (Figs. 9, 10, 11). Actually, the smaller and smaller initial value of the most excited mode when N increases explains the shifting of the slope.

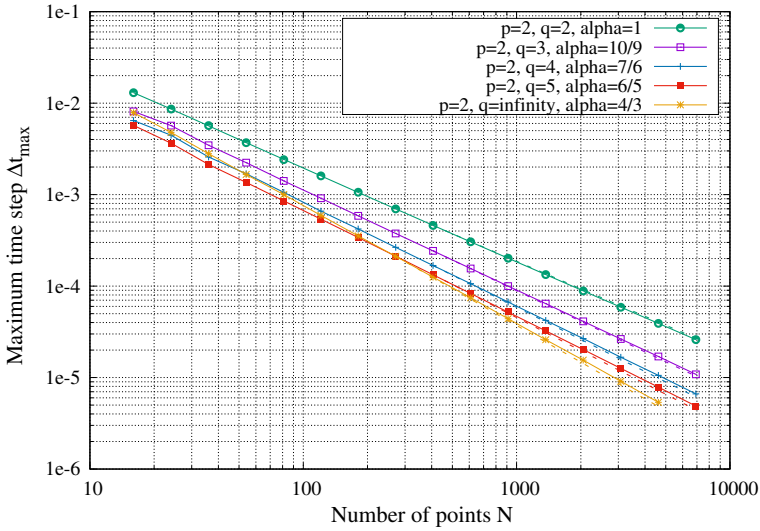

Fig. 10 Numerical experiment with $p = 2$ and various q

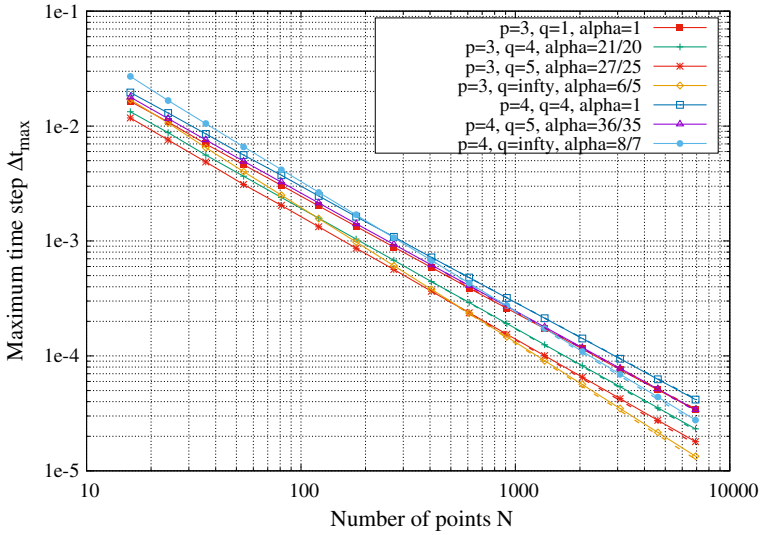

Fig. 11 Numerical experiment with $p = 3, 4$ and various q

Conclusion

The present paper explores in details the non-linear CFL conditions which may appear in transport dominated simulations. It completes the previous paper [6] by proving and testing the upwind Von Neumann stability condition: $\Delta t \leq C\ \Delta x^{\alpha}$ with $C > 0$ and $\alpha \in [1, 2]$. The numerical experiments meet our predictions accurately.

Although some CFL instabilities or more precisely exotic CFL stability conditions had been observed previously [19], our precise stability assertions—Theorem 3.1—had not been

clearly established earlier [18,24]. These results help to select a time integration scheme: for instance for someone using an upwind of second order, it is useless—from a stability point of view—to go higher than the second order for the time integration. They also explain some possible instabilities appearing in explicit hyperbolic numerical simulations: at high frequencies using centered discretizations [19], and at low frequencies for upwind schemes [2].

As presented in [8], it also helps to select a suitable time step for simulating the turbulence if one is interested in capturing precisely the spectrum which characterizes the auto-similarity of the turbulence e.g. the $-5/3$ Kolmogorov Law [16]. Then the non linear CFL condition applies even though it is not necessary for the stability. It also explains why unstable second order time schemes are so widely in use in Computational Fluid Dynamics [20,23]. The excitatory time schemes such as Runge–Kutta 2 or Adams–Bashforth 2 are more adapted to this kind of experiments because when they do not reconstruct the spectrum correctly (as Δt is not small enough) then they crash, while dissipative schemes such as Runge–Kutta 3 or Adams–Bashforth 3 do not crash but provide with biased spectra.

Data Availibility The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Charney, J.G., Fjörtoft, R., von Neumann, J.: Numerical integration of the barotropic vorticity equation. Tellus **2**, 237–254 (1950)
2. Cockburn, B., Karniadakis, G.E., Shu, C.-W. (eds.): Discontinuous Galerkin Methods: Theory, Computation and Applications. Springer, Berlin (2000)
3. Cockburn, B., Shu, C.-W.: Runge–Kutta discontinuous Galerkin methods for convection-dominated problems. J. Sci. Comput. **16**, 173–261 (2001)
4. Courant, R., Friedrichs, K., Lewy, H.: On the partial difference equations of mathematical physics. IBM J. (1967). Translation from a paper originally appeared in Mathematische Annalen **100**, 32–74, (1928)
5. Crouzeix, M., Mignot, A.L.: Analyse numérique des équations différentielles. Masson editor (1992)
6. Deriaz, E.: Stability conditions for the numerical solution of convection-dominated problems with skew-symmetric discretizations. SIAM J. Numer. Anal. **50**(3), 1058–1085 (2012)
7. Deriaz, E., Després, B., Faccanoni, G., Gostaf, K.P., Imbert-Gérard, L.-M., Sadaka, G., Sart, R.: Magnetic equations with FreeFem++: the Grad–Shafranov equation & the current hole. ESAIM Proc. **32**, 76–94 (2011)
8. Deriaz, E., Kolomenskiy, D.: Stabilité sous condition CFL non linéaire. ESAIM Proc. **35**, 114–121 (2012)
9. Deriaz, E., Perrier, V.: Direct numerical simulation of turbulence using divergence-free wavelets. SIAM Multiscale Model. Simul. **7**(3), 1101–1129 (2008)
10. Gallinato, O., Poignard, C.: Superconvergent second order Cartesian method for solving free boundary problem for invadopodia formation. J. Comput. Phys. **339**, 412–431 (2017)
11. Godlewski, E., Raviart, P.A.: Numerical Approximation of Hyperbolic Systems of Conservation Laws. Springer, Berlin (1996)
12. Gustafsson, B.: High Order Difference Methods for Time Dependent PDE. Springer Series in Computational Mathematics, vol. 38. Springer, Berlin (2008)
13. Grote, M.J., Mehlin, M., Mitkova, T.: Methods and algorithms for scientific computing Runge–Kutta-based explicit local time-stepping methods for wave propagation. SIAM J. Sci. Comput. **37**(2), A747–A775 (2015)
14. Hairer, E., Nørsett, S.P., Wanner, G.: Solving Ordinary Differential Equations. I. Nonstiff Problems. Springer Series in Computational Mathematics, vol. 8. Springer, Berlin (1987). Second revised edition 1993
15. Hairer, E., Wanner, G.: Solving Ordinary Differential Equations. II. Stiff and Differential-Algebraic Problems. Springer Series in Computational Mathematics, vol. 14. Springer, Berlin (1991). Second revised edition 1996
16. Kolmogorov, A.N.: Dissipation of energy in locally isotropic turbulence. Dokl. Akad. Nauk SSSR **32**, 16–18 (1941)

17. Kolmogorov, A.N., Yushkevich, A.P. (eds.): Mathematics of the 19th Century, vol. 3 (1998)
18. Levy, D., Tadmor, E.: From semidiscrete to fully discrete: stability of Runge–Kutta schemes by the energy method. SIAM Rev. **40**(1), 40–73 (1998)
19. Mohan Rai, M., Moin, P.: Direct simulations of turbulent flow using finite-difference schemes. J. Comput. Phys. **96**(1), 15–53 (1991)
20. Peyret, R.: Spectral Methods for Incompressible Viscous Flow, vol. 148. Springer, Berlin (2002)
21. Strikwerda, J.C.: Finite Difference Schemes and Partial Differential Equations, 2nd edn. SIAM, Philadelphia (2004)
22. Trefethen, L.N.: Finite difference and spectral methods for ordinary and partial differential equations. Unpublished text (1996). http://people.maths.ox.ac.uk/trefethen/pdetext.html
23. Verstappen, R.W.C.P., Veldman, A.E.P.: Symmetry-preserving discretization of turbulent flow. J. Comput. Phys. **187**(1), 343–368 (2003)
24. Wesseling, P.: Principles of Computational Fluid Dynamics. Springer, Berlin (2001)
25. Yuan, X., Zhang, Q., Shu, C., Wang, H.: The L^2-norm stability analysis of Runge–Kutta discontinuous Galerkin methods for linear hyperbolic equations. SIAM J. Numer. Anal. **57**(4), 1574–1601 (2019)