Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational Physics 226 (2007) 1211–1218
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Short Note

Error dynamics: Beyond von Neumann analysis

T.K. Sengupta [a,*], A. Dipankar [b], P. Sagaut [b]

[a] Department of Aerospace Engineering, I.I.T. Kanpur, UP 208016, India
[b] Department of Mechanical Engineering, LMM – UPMC/CNRS, Paris, France

Received 1 July 2006; received in revised form 18 April 2007; accepted 4 June 2007
Available online 14 June 2007

1. Introduction

The propagation of a signal in a continuous medium and the associated evolution of error is of prime importance in many applications of applied physics. There have been many efforts in analyzing error dynamics, using method attributed to von Neumann [1,2], that is readily applied for linear equations and in quasi-linearized form for non-linear equations. The main assumption for linear problems is that the error and the signal follow the same dynamics. While this appears intuitively correct, the main aim behind this work is to show that this is not correct for discrete computing due to dispersion or phase error or when the numerical method is not strictly neutrally stable.

We demonstrate the above with the help of the linear advection equation. For the analysis of space-time discretization schemes, the linear advection equation as a model that represents many flows and wave phenomena is used,

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0, \quad c > 0 \tag{1}$$

As a mathematical equation, Eq. (1) is non-dispersive that convects the initial solution to the right, at a group velocity c equal to the phase speed at all times. Hence, this equation gives a basis for testing numerical methods for solution accuracy, error propagation and most importantly the dispersion error – as in [3–7].

We represent the unknown by its Fourier transform at the jth node of a uniformly spaced discrete grid of spacing h as, $u(x_j,t) = \int U(k,t)\mathrm{e}^{\mathrm{i}kx_j}\,\mathrm{d}k$ and the exact spatial derivative at the same node is given by, $\left[\frac{\partial u}{\partial x}\right]_{\text{exact}} = \int \mathrm{i}kU\mathrm{e}^{\mathrm{i}kx_j}\,\mathrm{d}k$. While solving Eq. (1) by discrete methods, the spatial derivative u'_j (denoted by a prime) can be shown [3,5] as equivalent to

$$[u'_j]_{\text{numerical}} = \int \mathrm{i}k_{\text{eq}}U\mathrm{e}^{\mathrm{i}kx_j}\,\mathrm{d}k \tag{2}$$

Numerically in a finite-domain the same derivative is estimated from [5], $\{u'\} = \frac{1}{h}[C]\{u\}$. One can obtain an appropriate matrix $[C]$ for finite-domain non-periodic problems, with the dimension of the matrix corresponding to the number of nodes, and which implies that the derivative at the jth node is evaluated as $u'_j = \frac{1}{h}\sum_{l=1}^{N} C_{jl}u_l$, where $u_l = u(x_l,t) = \int U(k,t)\mathrm{e}^{\mathrm{i}kx_l}\,\mathrm{d}k$ is the value of the function at the lth node and N is the total number of nodes used for discretization. Using the spectral representation, we can alternatively write the numerical derivative as,

* Corresponding author. Tel.: +91 512 2597945; fax: +91 512 2597561.
E-mail addresses: tksen@iitk.ac.in (T.K. Sengupta), sagaut@lmm.jussieu.fr (P. Sagaut).

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2007.06.001

$$u'_j = \int \frac{1}{h} \sum C_{jl} U(k,t) e^{ik(x_l - x_j)} e^{ikx_j} \, dk \tag{3}$$

Comparing Eq. (3) with (2), we note that

$$[ik_{eq}]_j = \frac{1}{h} \sum_{l=1}^{N} C_{jl} e^{ik(x_l - x_j)} \tag{4}$$

Although in physical plane computations, C_{jl}s are real, $[k_{eq}]_j$ is in general complex, with a real part that represents the numerical phase and an imaginary part that represents the numerical dissipation added by the choice of method determining the entries of the matrix $[C]$.

Other important numerical properties are obtained via the spectral representation in Eq. (1), that gives

$$\int \left[\frac{\partial U}{\partial t} + \frac{c}{h} \sum U C_{jl} e^{ik(x_l - x_j)} \right] e^{ikx_j} \, dk = 0 \tag{5}$$

Since the above equation is true for all wave numbers, the integrand must be zero for any k. We note that the present analysis in the physical plane for non-periodic problems, is different from the von Neumann analysis that is strictly valid for periodic problems only applied in the spectral plane for normal modes. The implicit condition of Eq. (5), can be reinterpreted as,

$$\frac{dU}{U} = -\left[\frac{c dt}{h} \right] \sum_{l=1}^{N} C_{jl} e^{ik(x_l - x_j)} \tag{6}$$

We note that the first factor on the right-hand side is nothing but the CFL number (N_c). Since the right-hand side of Eq. (6) is node-dependent, we can express the left hand side in terms of the nodal numerical amplification factor (G_j),

$$G_j = G(x = x_j) = 1 - N_c \sum_{l=1}^{N} C_{jl} e^{ik(x_l - x_j)} \tag{7}$$

for the Euler time discretization scheme. Similarly, one can obtain G_j for other time discretization schemes and for the four-stage Runge–Kutta time integration scheme, this has been obtained in [3,8] and [9] as,

$$G_j = 1 - A_j + \frac{A_j^2}{2} - \frac{A_j^3}{6} + \frac{A_j^4}{24} \tag{8}$$

where $A_j = N_c \sum_{l=1}^{N} C_{jl} e^{ik(x_l - x_j)}$. A similar relation for general class of the Runge–Kutta methods is given in [3, page 52]. The present relation is for the fourth order Runge–Kutta scheme obtained for any nodes of a non-periodic problem. Such full-domain analyses for some explicit and implicit numerical discretization schemes are available in [7]. While the amplification factor can be a source of error, additional error can arise due to dispersion and that is described next. If we represent the initial condition for Eq. (1) as given by

$$u(x_j, t = 0) = u_j^0 = \int A_0(k) e^{ikx_j} \, dk \tag{9}$$

then the general solution at any arbitrary time can be obtained as,

$$u_j^n = \int A_0(k) [|G_j|]^n e^{i(kx_j - n\beta_j)} \, dk \tag{10}$$

where $|G_j| = (G_{rj}^2 + G_{ij}^2)^{1/2}$ and $\tan(\beta_j) = -\frac{G_{ij}}{G_{rj}}$, with G_{rj} and G_{ij} as the real and imaginary parts of G_j, respectively. Thus, the phase of the solution is determined by $n\beta_j = kc_N t$, where c_N is the numerical phase speed. Although the physical phase speed is a constant for all wave numbers, this analysis shows that the numerical phase speed is wave number dependent i.e. the numerical solution is dispersive, in contrast with the non-dispersive nature of Eq. (1). The implications of this simple difference can be very important, as demonstrated below.

The general numerical solution of Eq. (1) is denoted as,

$$\bar{u}_N = \int A_0[|G|]^{t/\Delta t} e^{ik(x-c_N t)}\, dk \tag{11}$$

The fact that $c_N \neq c$ and is a function of k has been used to explain shortcomings of some multi-time step integration methods in [9]. Since the numerical dispersion relation is now given as $\omega_N = c_N k$, instead of $\omega = c\ k$, the non-dimensional phase speed and group velocity at the jth discrete node can be expressed as

$$\left[\frac{c_N}{c}\right]_j = \frac{\beta_j}{\omega \Delta t} \tag{12}$$

$$\left[\frac{V_{gN}}{c}\right]_j = \frac{1}{hN_c}\frac{d\beta_j}{dk} \tag{13}$$

The characteristics of some well known numerical methods have already been reported in [5–8] in terms of the amplification rates, numerical phase speed and group velocity here in Eqs. (7), (8), (12) and (13). The main purpose of the present work is to show the consequence of such dispersion to explain error growth for Eq. (1) and thereby draw general observations for error dynamics of linear systems as shown next.

If we define the computation error as $e(x,t) = u(x,t) - \bar{u}_N$, then we can obtain the governing equation for its dynamics in the following manner. Using Eq. (11) one obtains

$$\frac{\partial \bar{u}_N}{\partial x} = \int ikA_0[|G|]^{t/\Delta t} e^{ik(x-c_N t)}\, dk \tag{14}$$

and

$$\frac{\partial \bar{u}_N}{\partial t} = -\int ikc_N A_0[|G|]^{t/\Delta t} e^{ik(x-c_N t)}\, dk + \int \frac{Ln|G|}{\Delta t} A_0[|G|]^{t/\Delta t} e^{ik(x-c_N t)}\, dk \tag{15}$$

Eq. (15) can be simplified to yield,

$$\frac{\partial \bar{u}_N}{\partial t} + c_N \frac{\partial \bar{u}_N}{\partial x} = \int \frac{dc_N}{dk}\left[\int ik' A_0[|G|]^{t/\Delta t} e^{dk'(x-c_N t)}\, dk'\right] dk + \int \frac{Ln|G|}{\Delta t} A_0[|G|]^{t/\Delta t} e^{ik(x-c_N t)}\, dk \tag{16}$$

Thus, the error propagation equation is given by

$$\frac{\partial e}{\partial t} + c\frac{\partial e}{\partial x} = -c\left[1 - \frac{c_N}{c}\right]\frac{\partial \bar{u}_N}{\partial x} - \int \frac{dc_N}{dk}\left[\int ik' A_0[|G|]^{t/\Delta t} e^{ik'(x-c_N t)}\, dk'\right] dk - \int \frac{Ln|G|}{\Delta t} A_0[|G|]^{t/\Delta t} e^{ik(x-c_N t)}\, dk \tag{17}$$

This is the correct error propagation equation as opposed to that obtained using the assumption made in von Neumann analysis, where the right-hand side is identically taken to be zero on the premise that $c_N \equiv c$ i.e. there are no dispersion errors and the numerical method is perfectly neutral so that the last term on the right-hand side of Eq. (17) is also identically zero. Some combinations of numerical parameters for solving Eq. (1) lead to numerical instability, with an error growing faster than predicted by von Neumann error analysis. In [9], this has been studied by solving Eq. (1) using finite element and finite difference methods. Mismatch between error estimate by von Neumann analysis and numerical solution was attributed to dispersion of numerical methods and noting the error to dramatically increase when the numerical solution displays sharp spatial variation. This can now be explained as the effect of the first term on the right-hand side of Eq. (17). In [10–12], phase error ($c - c_N$) was reported for monochromatic wave for periodic problems. Its effect on signal error was not reported – but it was conjectured that the phase error can be reduced by grid refinement, while keeping N_c same. Similarly, in [4] a high order compact scheme was used for the numerical solution of Eq. (1) and attention was focused on spatial grid requirements to obtain a level of desired accuracy.

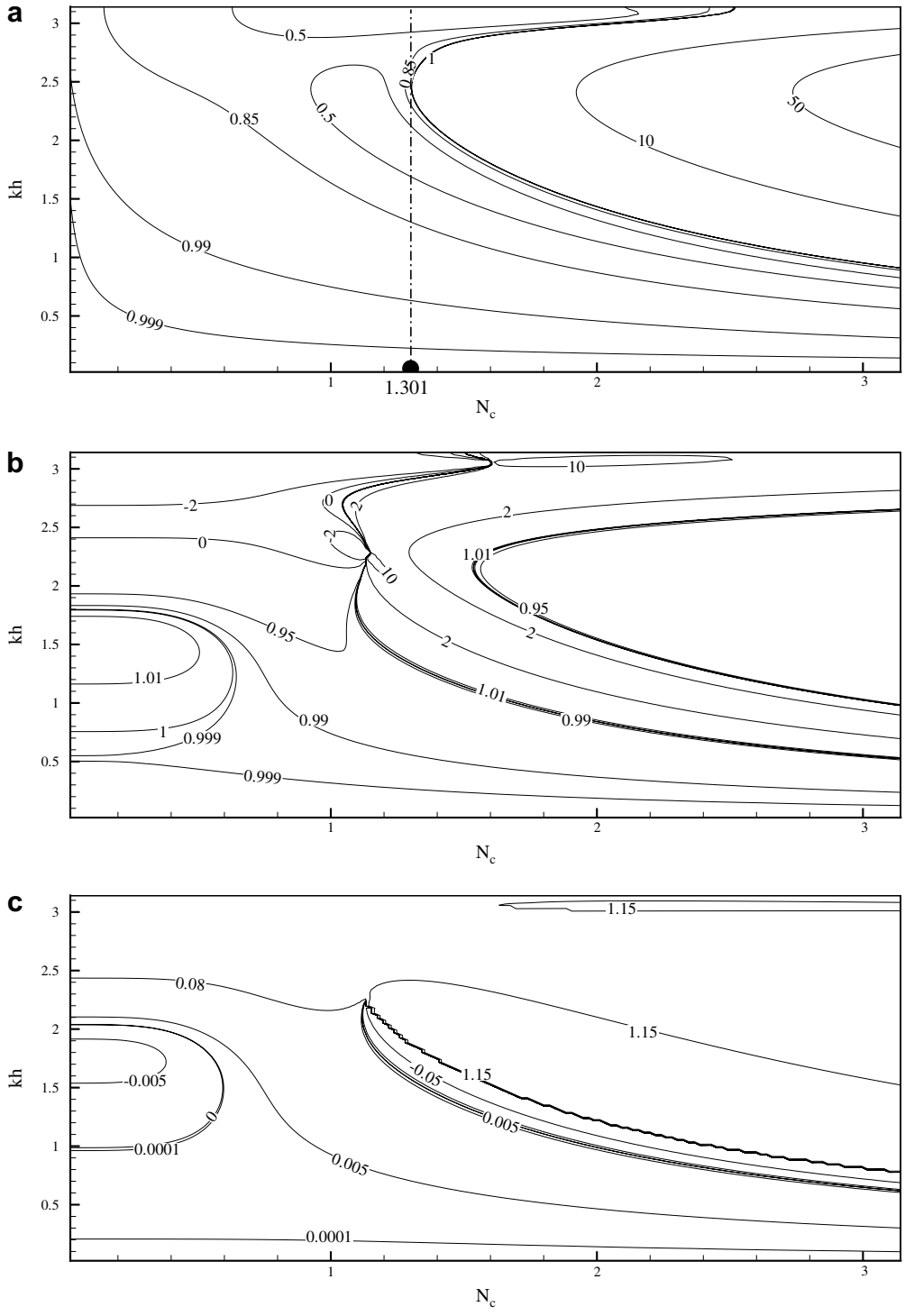

Fig. 1. Contour plots showing (a) Numerical amplification factor ($|G|$); (b) Scaled numerical group velocity ($V_{g\ N}/c$) and (c) ($1\text{-}c_N/c$) at the interior node, using RK_4 time marching and OUCS3 ($\eta = 2$) spatial discretization schemes in Eqs. (8), (12), (13).

2. Results and discussion

First of all, we obtain the properties of the compact scheme OUCS3 [6] for a central node when the four-stage Runge–Kutta time integration strategy is used for solving Eq. (1). In Fig. 1, $|G|$, V_{gN}/c and $(1 - c_N/c)$ are plotted as contours in the indicated ranges of kh and N_c for an interior node. It is noted from Fig. 1a for numerical amplification contours, that the scheme is stable for $N_c \leqslant 1.301$ i.e. $|G| \leqslant 1$. The scheme is neutrally stable for very small values of N_c and a limited range of kh – a property absolutely essential for direct simulation. One also notes that the last term on the right-hand side of Eq. (17) vanishes for neutrally stable case. In Fig. 1b, the scaled numerical group velocity contours display significant dispersion effects that would invalidate long time integration results – even when neutral stability is ensured by computing with vanishingly small time steps. In fact, above $kh \geqslant 2.4$ the numerical solution will travel in the wrong direction, as $V_{gN} \leqslant 0$ for

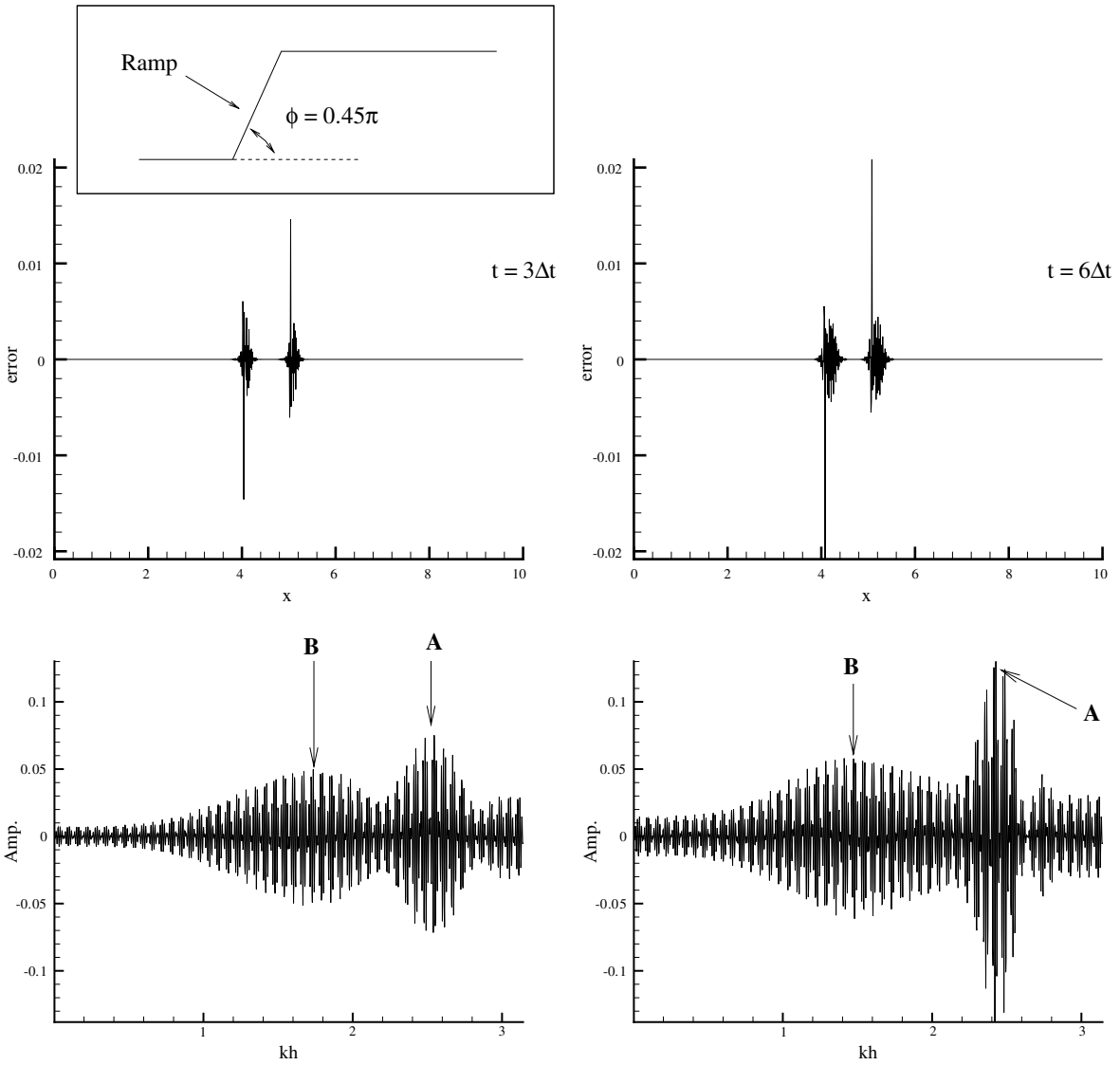

Fig. 2. Error evolution shown in the top panels (for the ramp function of the inset) and their FFT (in the bottom panels) are shown at the indicated times. Note the error components marked in the figures due to numerical instability (marked A) and phase error (marked B) – as a consequence of the properties of the numerical method shown in Fig. 3.

$N_c \cong 0$. In Eq. (17), we note that the first term on the right-hand side affects the error evolution via the numerical property $(1 - \frac{c_N}{c})$. In Fig. 1c, the contours of this quantity is shown plotted. This contribution will be significant when the solution has large non-negligible slope. Figs. 1b and c indicate the effects of dispersion error, that cannot be simply eliminated or reduced by grid refinement – as suggested in [4,10] and [11].

To establish the new error propagation equation, careful design of a test case is made for solving Eq. (1) with $c = 1$. The computed results are for $N_c = 4/3$, where most of the nodes experience numerical instability over a narrow range of kh. We consider the propagation of a ramp function (shown as an inset in Fig. 2) in the domain given by $0 \leqslant x \leqslant 30$, where the foot of the ramp is initially located at $x = 4$ and the ramp angle is given by $\phi = 0.45\pi$. The spatial grid is defined by $h = 0.01$ and the time step is chosen from $N_c = 4/3$. Discrete jumps, at the beginning and end of the ramp, give rise to the Gibbs' phenomenon via the error evolution as defined here. The error calculated at $t = 3\Delta t$ and $6\Delta t$ (from the numerical and exact solution) is shown in the top frames of Fig. 2. Immediately below these frames the Fourier transforms of the error are shown. This behaviour of the error evolution is explained next with the help of the numerical properties shown in Fig. 1.

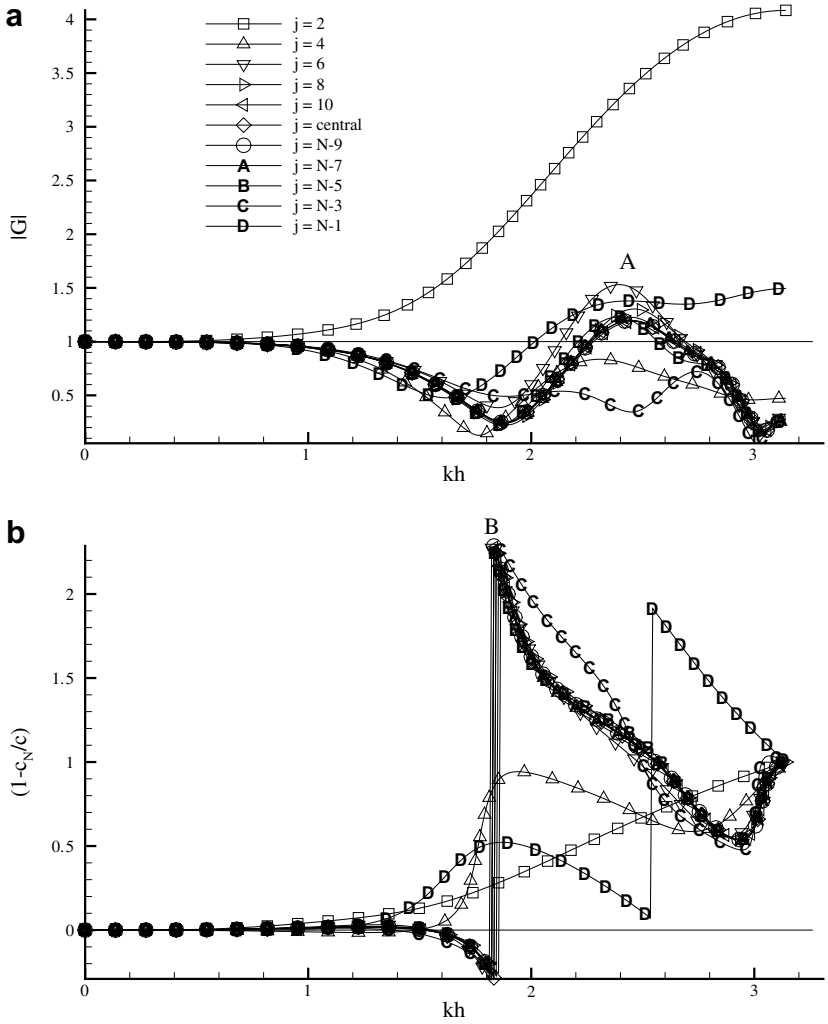

Fig. 3. Variation of (a) $|G|$ and (b) $(1\text{-}c_N/c)$ with kh is shown for the indicated nodes, for $N_c = 4/3$. The point A in (a) indicates the where the method is most unstable and the point B in (b) indicates where the method has maximum phase error in Eq. (17).

For the particular CFL number ($N_c = 4/3$) in Fig. 3a, $|G_j|$s are plotted for $0 \leqslant kh \leqslant \pi$ for different nodes – with the subscripts (j) indicating node number. The analysis results used here is based on Eqs. (8), (12) and (13) and the method was originally introduced in [6,7] for discrete computing methods with any type of boundary conditions. For the compact scheme OUCS3, it is noted [7,8] that the method displays directionality of the derivatives. This is true of all implicit methods for derivative evaluation, with nodes near the inflow behaving differently from the nodes near the outflow. In Fig. 3a, this is seen for the nodes at $j = 2, 4, N - 3$ and $N - 1$. It is also seen that the maximum numerical instability suffered in the interior is for $kh \cong 2.4$ – marked as A in the figure. In Fig. 3b, $(1 - c_N/c)$ is plotted as a function of kh for different nodes for $N_c = 4/3$. For most of the interior nodes, this quantity attains a maximum value at $kh \cong 1.8$ – marked as B in the figure, with a discrete jump in the value of $(1 - c_N/c)$.

In the bottom frames of Fig. 2 – for the Fourier transform – one can clearly identify two distinct peaks. These are also marked A and B, as they correspond to the two maxima of Fig. 3a and b. The point A, at the higher wave number produces error due to numerical instability (at $kh \cong 2.4$) for interior nodes. The point B contributes to the error due to phase error, arising from the first term on the right-hand side of Eq. (17) at $kh = 1.8$, for which this forcing is maximum. This numerical experiment demonstrates the correctness of the error evolution Eq. (17), as compared to what is obtained by putting all the right-hand side terms equal to zero – an assumption made in classical von Neumann analysis. For this linear problem, according to von Neumann

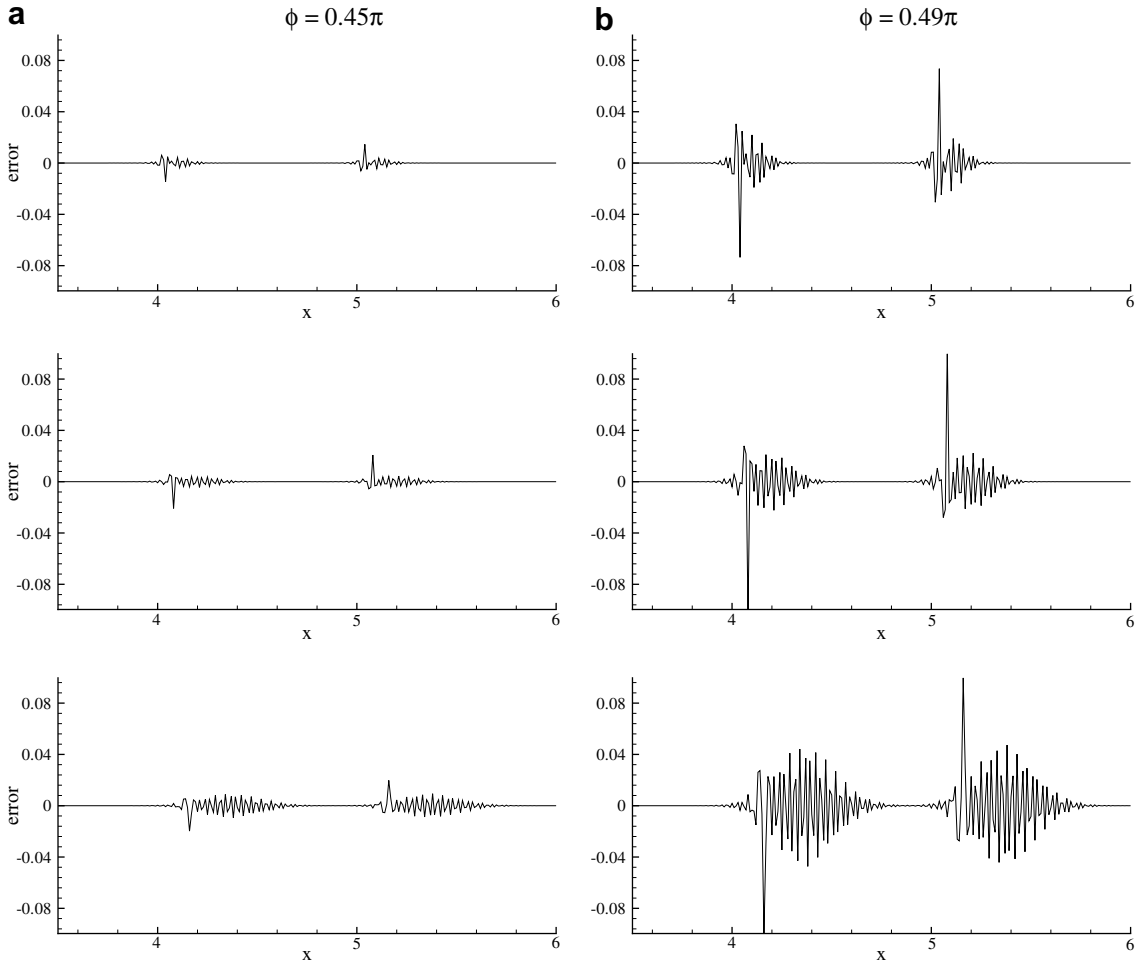

Fig. 4. Error evolution compared for the computed solutions obtained for $N_c = 4/3$ at the indicated times, for Eq. (1) at different ramp angles (ϕ) as indicated.

analysis the ramp solution should not have shown the peak at B, while it would have shown numerical instability at A.

To further emphasize the role of the first forcing term on the right-hand side of Eq. (17), another case is computed with $\phi = 0.49\pi$, so that the term $\frac{\partial \bar{u}_N}{\partial x}$ is increased from a value of 0.024679 to 0.02687, within the ramp. At $t = 0$, the ramp is located, as before, between $x = 4$ and 5. In Fig. 4, these two cases are compared for $t = 3\Delta t$, $6\Delta t$ and $12\Delta t$. Both the cases blow up eventually due to numerical instability – possessing identical numerical properties for the same grid and identical time steps. However, for the case of $\phi = 0.49\pi$, increased forcing due to the first term on the right-hand side of Eq. (17) causes larger error at all time instants.

3. Summary

The most significant result of the present exercise is to identify a correct error propagation equation that also accounts for dispersion error. It is shown that when this is included, the correct error evolution equation (17) for the scalar advection equation is not the one given by the traditional method due to von Neumann that assumes the signal and error to satisfy the same equation. Eq. (17) is obtained here for the first time, showing the dispersion error to change the actual constant phase speed to a variable numerical phase speed c_N. This error $c(1 - c_N/c)$, gives rise to additional forcing on the right-hand side of Eq. (17) dependent upon this error times the spatial gradient of numerical solution. The propagation of a ramp function is studied to demonstrate how the spatial gradient of signal affects error dynamics. In the process, we explain the genesis and evolution of Gibbs' phenomenon.

Acknowledgements

We thank the reviewers for encouraging comments and help with the presentation of the contents of this paper. The authors gratefully acknowledge the support of Indo-French Centre for the Promotion of Advanced Research (IFCPAR or CEFIPRA), New Delhi for the conduct of this research vide their sanctioned project (No. IFC/3401-1). The first author like to acknowledge all his students who have actively participated in the related research in the recent times.

References

[1] J.G. Charney, R. Fjørtoft, J. von Neumann, Tellus 2 (1950) 237.
[2] J. Crank, P. Nicholson, Proc. Cambridge Philos. Soc. 43 (1947) 50.
[3] R. Vichnevetsky, J.B. Bowles, Fourier Analysis of Numerical Approximations of Hyperbolic Equations, SIAM Stud. Appl. Math., vol. 5, SIAM, Philadelphia, USA, 1982.
[4] D.W. Zingg, SIAM J. Sci. Comput 22 (2) (2000) 476.
[5] T.K. Sengupta, Fundamentals of Computational Fluid Dynamics, Univ. Press, Hyderabad, India, 2004.
[6] T.K. Sengupta, G. Ganeriwal, S. De, J. Comput. Phys. 192 (2) (2003) 677.
[7] T.K. Sengupta, S.K. Sircar, A. Dipankar, J. Sci. Comput. 26 (2) (2006) 151.
[8] A. Dipankar, T.K. Sengupta, J. Comput. Phys. 215 (1) (2006) 245.
[9] R. Vichnevetsky, B. Peiffer, in: R. Vichnevetsky (Ed.), Advances in Computer Methods for Partial Differential Equations, vol.53, AICA, Ghent, Belgium, 1975.
[10] H. Kreiss, J. Oliger, Tellus 24 (1972) 199.
[11] B. Swartz, B. Wendroff, SIAM J. Numer. Anal. 11 (5) (1974) 979.
[12] L.N. Trefethen, SIAM Rev. 24 (2) (1982) 113.