

ON THE GROUP VELOCITY OF SYMMETRIC AND UPWIND NUMERICAL SCHEMES

SMADAR KARNI*

Department of Mathematics, The University of Michigan, Ann Arbor, MI 48109, U.S.A.

SUMMARY

Dissipative numerical approximations to the linear advection equation are considered with respect to their behaviour in the limit of weak dissipation. The context is wave propagation under typical far-field conditions where grids are highly stretched and waves are underresolved. Three classes of schemes are analysed: explicit two-level (i) symmetric and (ii) upwind schemes of optimal accuracy are considered as well as (iii) (symmetric) Runge–Kutta schemes. In the far-field the dissipation of all schemes diminishes. Group speeds of high-frequency modes assume the incorrect sign and may admit ‘backward’ wave propagation if waves are not damped. A fundamental difference arises between the symmetric and upwind cases owing to the different rates at which the dissipation diminishes. In the upwind case, while the amount of damping per time step diminishes, the accumulative damping remains exponential in time. In the symmetric case the accumulative damping tends to unity, yielding in practice non-dissipative schemes. In this light, parasitic modes constitute much less of a problem in the upwind case than in the symmetric case. Numerical tests confirm these findings.

KEY WORDS Linear advection Numerical approximation Dissipation Group velocity

INTRODUCTION

Consider the linear advection equation

$$u_t = au_x, \quad (1)$$

where $a > 0$. Define a grid by $x_j = jh$, $t_n = nk$ with h and k the grid partitions in space and time respectively and consider the two-level explicit numerical scheme to approximate (1):

$$u_j^{n+1} = \sum_{k=-l}^r c_k u_{j+k}^n. \quad (2)$$

The scheme has a stencil that stretches l points to the left and r points to the right of u_j , with an overall width of $l + r + 1$ points. The coefficients c_k are constant and depend on the Courant–Friedrichs–Levy (CFL) number $v = ak/h$. Given the width of the stencil, the maximum order of accuracy attainable is $p = l + r$ and the coefficients $c_k(v)$ are uniquely determined. Following standard stability analysis, let $u_j^n = Z^n e^{i\zeta jh}$, where $Z = e^{i\omega k}$ is the amplification factor. Upon substitution into (2), one obtains the dispersion relation

$$Z = \sum_{k=-l}^r c_k e^{i\zeta kh}. \quad (3)$$

* Present address: Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, U.S.A.

For Cauchy stability the scheme must satisfy $|z| \leq 1$ for all $\xi \in \mathbb{R}$. This places a restriction on the allowable size of ν , which for a given geometry implies a restriction on the allowable time step for stable integration.

We focus our attention on schemes which are stable at small time steps, i.e. schemes which include $\nu \approx 0$ in their stability region. Other schemes are related to those through a simple sliding of the numerical stencil and a corresponding adjustment of the coefficients c_k and CFL stability range. If a scheme (2) is to be stable at small CFL numbers, its order of accuracy p is restricted not only by the width of the stencil, $p \leq l + r$, but also by $p \leq 2 \min(l, r + 1)$.¹ This indicates that there is no advantage in biasing the stencil of the scheme too strongly in the upwind direction. In fact, if the scheme is fully upwinded (i.e. with $r = 0$), the above restriction implies that it is at most second-order-accurate.^{1,2}

Wherever applicable, the numerical propagation speed of the wave is governed by the group velocity

$$G(\xi) = -\frac{d\omega}{d\xi} = -\frac{a}{\nu Z} \sum_{k=-1}^r kc_k e^{i\xi kh} \quad (4)$$

and depends on the wave number ξ . Thus, numerical wave propagation is dispersive even though (1) is not. It is well known that certain numerical schemes for the linear advection equation admit wave propagation not only at incorrect speeds but also in the incorrect direction. Typically these are high-frequency waves and are underresolved relative to the grid. They arise owing to the finiteness of the grid and are regarded as numerical parasitic modes. Group velocity is a useful tool in analysing discrete wave propagation and confirms the admittance of such spurious waves. Strictly speaking, the concept of group velocity is applicable only to non-dissipative wave propagation. Examples of non-dissipative schemes that admit ‘backward’ wave propagation include the (generalized) leap-frog¹ and Crank–Nicolson⁵ schemes.

In this paper we employ the concept of group velocity to analyse *dissipative* wave propagation in the limit of vanishing dissipation. Though not strictly valid, group velocity analysis still yields good predictions and proves a useful tool.

The work presented in this paper was prompted by a study of discrete wave propagation in the far field. Far fields are often characterized by highly stretched numerical grids, with respect to which waves are underresolved and tend to appear as high-frequency waves. CFL numbers in the far field tend to become small owing to grid stretching, hence our focus on schemes which are stable at small CFL numbers. The combination of small local CFL numbers and local high frequencies is unfortunate, yielding undesirable wave behaviour. In Reference 6 the dissipative Lax–Wendroff (LW) scheme was analysed under such typical far-field conditions. It was shown that as ν decreases in the far field, the dissipation of the scheme diminishes. At the same time more and more waves tend to appear as high frequencies relative to the stretched grids. Group velocity analysis, justifiable on the grounds of weak dissipation, shows that high-frequency saw-tooth waves propagate ‘backward’ and persist longer at smaller CFL numbers. As a consequence, waves originally travelling out into the far field may undergo internal reflection due to the stretching of the grid even when the grid expansion is very mild. This pathological behaviour was shown to have serious implications on computational aspects such as the time accuracy of transient solutions and their convergence rate to steady state and on the rate of energy absorption at artificial boundaries.⁶

In this paper we show that this pathological behaviour is not unique to the LW scheme but is in fact more general. We consider general explicit two-level schemes of even/odd order. In the even-order case we consider symmetric schemes with $l = r$. Other even-order schemes are

equivalent through a simple shift of stencil. Similarly, odd-order schemes are centred at half-grid points and it is sufficient to consider schemes with $l = r + 1$ 'centred' at the half-grid point $j - \frac{1}{2}$. Other odd-order schemes are equivalent.

For upwind schemes of optimal accuracy the dissipation is shown to diminish when $\nu \rightarrow 0$ and examples are shown where the group velocity may assume the wrong sign, admitting waves to propagate 'backwards'. In this case, while the amount of dissipation per time step diminishes, accumulative wave damping remains exponential in *time* and overall the scheme maintains enough dissipation to kill the spurious waves. For symmetric schemes, when the CFL number diminishes, the dissipation vanishes. Using group velocity analysis, we show that high-frequency waves always propagate 'backwards' and in the limit $\nu \rightarrow 0$ are not dissipated at all, yielding practically non-dissipative schemes. In this light, parasitic modes constitute a much bigger problem in symmetric schemes than they do in upwind schemes. We also analyze the class of symmetric Runge–Kutta (RK) schemes. Generally these schemes are dissipative, but they are non-dissipative for high-frequency saw-tooth-like waves at all CFL numbers. The group velocity of high-frequency modes is shown to assume the incorrect sign, admitting 'backward' propagation of such waves which persist without being dissipated at all. We conclude this paper by presenting simple numerical experiments to support these findings.

ANALYSIS

Symmetric schemes

In the symmetric case we have $l = r = s$ and the stencil of (2) stretches over $2s + 1$ grid points. The maximal order of accuracy is $p = 2s$ and is attained by choosing c_k that satisfy

$$\sum_{k=-s}^s k^p c_k = \nu^p, \quad p = 0, 1, \dots, 2s. \quad (5)$$

Equation (5) is a linear non-homogeneous system of the Vandermonde type which can be solved using Cramer's rule:

$$c_k(\nu) = \prod_{\substack{j=-s \\ j \neq k}}^s \frac{\nu - j}{k - j}. \quad (6)$$

Using the generalized factorial notation

$$(\nu - j)_{n+1} = (\nu - j)(\nu - j + 1) \cdots (\nu - j + n),$$

equation (6) simplifies to read

$$c_k(\nu) = (-1)^{s+k} \frac{(\nu - s)_{2s+1}}{(s - k)!(s + k)!(\nu - k)}. \quad (7)$$

Equation (7) leads to the following.

- (a) $c_k(\nu) = c_{-k}(-\nu)$.
- (b) The amplification factor (3)

$$Z(\xi h = \pi) = \sum_{k=-s}^s (-1)^k c_k(\nu)$$

is real and is a symmetric function of the CFL number ν .

(c) Assuming Cauchy stability, for ν sufficiently small and for some positive constant c ,

$$Z \approx 1 - c\nu^2. \quad (8)$$

(d) The group velocity of the mode $\xi h = \pi$ is

$$G = -\frac{a}{\nu Z} \sum_{k=-s}^s (-1)^k k c_k(\nu),$$

which together with (7) yields

$$G = \frac{a}{Z} \frac{(1-\nu)_s (\nu+1)_s}{(2s)!} \sum_{k=1}^s \binom{2s}{s+k} \frac{2k^2}{k^2 - \nu^2}. \quad (9)$$

It follows from (c) that for sufficiently small ν , Z is positive. For $\nu \leq 1$ all the ν -dependent terms in (9) are positive, yielding a *positive* group velocity and implying that high-frequency waves propagate 'backwards'.

(e) Since at small CFL numbers waves propagate through fewer grid cells, the practical implication of the fact that the dissipation diminishes has to be assessed after a fixed time, T , rather than a fixed number of time steps N . We take the double limit $\nu \rightarrow 0$, $N \rightarrow \infty$, holding the time $T = Nk = N\nu h/a$ fixed. Assuming $a = 1$ and $h = 1$, the mode $\xi h = \pi$ will have dissipated by

$$Z^{T/\nu} \approx (1 - c\nu^2)^{(1/c\nu^2)c\nu T} \approx e^{-c\nu T} \xrightarrow{|\nu| \ll 1} 1, \quad (10)$$

and in the limit of small CFL number the mode is not dissipated at all.

An example of a scheme in this class is the LW scheme

$$u_j^{n+1} = u_j^n + \frac{\nu}{2} (u_{j+1}^n - u_{j-1}^n) + \frac{\nu^2}{2} (u_{j+1}^n - 2u_j^n + u_{j-1}^n). \quad (11)$$

Upwind schemes

In the upwind case $r > l$ and in general the amplification factor of the mode $\xi h = \pi$ is real and of the form

$$Z = 1 - c\nu + O(\nu^2),$$

where c is some positive scheme-dependent constant. For example, the first-, second- and third-order upwind schemes are

$$\begin{aligned} u_j^{n+1} &= (1-\nu)u_j^n + \nu u_{j+1}^n, \\ u_j^{n+1} &= \frac{(1-\nu)(2-\nu)}{2} u_j^n + \nu(2-\nu)u_{j+1}^n - \frac{\nu(1-\nu)}{2} u_{j+2}^n, \\ u_j^{n+1} &= \frac{\nu(1-\nu)(2-\nu)}{6} u_{j-1}^n + \frac{(1+\nu)(1-\nu)(2-\nu)}{2} u_j^n \\ &\quad + \frac{(1+\nu)\nu(2-\nu)}{2} u_{j+1}^n - \frac{(1+\nu)\nu(1-\nu)}{6} u_{j+2}^n. \end{aligned} \quad (12)$$

The respective amplification factors $Z(\xi h = \pi)$ are

$$Z(\pi) = 1 - 2\nu, \quad Z(\pi) = 1 - 4\nu + 2\nu^2, \quad Z(\pi) = 1 - \frac{4}{3}\nu - 2\nu^2 + \frac{4}{3}\nu^3.$$

The group velocities are respectively

$$G = a/Z, \quad G = a(3 - 2\nu)/Z, \quad G = a(5 - 2\nu^2)/3Z$$

and may assume the wrong sign if ν is small. The fundamental difference between the symmetric and upwind cases is the way the damping accumulates with time. After a fixed time $T = Nk$ high frequencies will have dissipated by

$$Z^{T/\nu} \approx (1 - c\nu)^{T/\nu} \xrightarrow{|\nu| \ll 1} e^{-Tc}, \tag{13}$$

maintaining a sufficient amount of dissipation to damp the spurious waves exponentially in time. This is unlike the symmetric case where the accumulative damping factor tends to unity and spurious waves tend to persist. One implication is that in the upwind case high-frequency parasitic modes do not constitute a severe computational problem, since even at small CFL numbers they are damped exponentially in time. These different asymptotic behaviours are confirmed by numerical experiments.

RK schemes

Consider the (s)-stage RK algorithm to approximate (1):

$$\begin{aligned} u_j^{(1)} &= u_j^n + \lambda_s \frac{\nu}{2} (K - K^{-1})u_j^n, \\ u_j^{(2)} &= u_j^n + \lambda_{s-1} \frac{\nu}{2} (K - K^{-1})u_j^{(1)}, \\ &\vdots \\ u_j^{(s)} &= u_j^n + \lambda_1 \frac{\nu}{2} (K - K^{-1})u_j^{(s-1)}, \\ u_j^{n+1} &= u_j^{(s)}, \end{aligned} \tag{14}$$

where K is the shift operator $Ku_j^n = u_{j+1}^n$ and λ_k are arbitrary positive constants. We note here that for (14) to be a consistent approximation, we must have $\lambda_1 = 1$. Let $\Lambda_k = (\nu/2)^k \prod_{l=1}^k \lambda_l$. Then (14) may be written explicitly as

$$u_j^{n+1} = \left(I + \sum_{k=1}^s \Lambda_k \sum_{m=0}^k (-1)^m K^{k-2m} \right) u_j^n = \sum_{k=-s}^s c_k(\nu) u_{j+k}^n. \tag{15}$$

The following can be shown

(a)

$$c_k(\nu) = \begin{cases} c_{-k}(\nu), & k = \text{even} \\ -c_{-k}(\nu), & k = \text{odd} \end{cases}$$

and for $k \geq 0$,

$$c_k(\nu) = \sum_{n=0}^{\lfloor (s-k)/2 \rfloor} (-1)^n \Lambda_{k+2n} \binom{k+2n}{n}, \tag{16}$$

where $[x]$ denotes the integer part of x .

(b) $Z(\xi h = \pi) = 1$ and the group velocity strictly applies.

(c) Using (a) and (b),

$$G(\xi h = \pi) = -\frac{a}{v} \left(\sum_{\substack{k=-s \\ k=\text{odd}}}^s + \sum_{\substack{k=-s \\ k=\text{even}}}^s \right) (-1)^k k c_k(v).$$

The even-order terms cancel out and the odd-order terms yield

$$G = \frac{2a}{v} \sum_{\substack{k=1 \\ k=\text{odd}}}^s k c_k(v), \tag{17}$$

with $c_k(v)$ given by (16). For specific choices of (s) we obtain

(i) $s = 1$ ($s = 2$)

$$G(\pi) = \frac{2a}{v} \Lambda_1 \binom{1}{0},$$

(ii) $s = 3$ ($s = 4$)

$$G(\pi) = \frac{2a}{v} \left\{ \Lambda_1 \binom{1}{0} + \Lambda_3 \left[3 \binom{3}{0} - 1 \binom{3}{1} \right] \right\},$$

(iii) $s = 5$ ($s = 6$)

$$G(\pi) = \frac{2a}{v} \left\{ \Lambda_1 \binom{1}{0} + \Lambda_3 \left[3 \binom{3}{0} - 1 \binom{3}{1} \right] + \Lambda_5 \left[5 \binom{5}{0} - 3 \binom{5}{1} + 1 \binom{5}{2} \right] \right\},$$

etc.

From the above expressions we conjecture that

$$\sum_{k=0}^m (-1)^k (2m + 1 - 2k) \binom{2m + 1}{k} \equiv 0, \quad m = 1, 2, \dots, \tag{18}$$

implying that the only non-zero contribution to the group velocity of the mode $\xi h = \pi$ comes from the first term and that under a general (s)-stage RK scheme

$$G(\pi) = \frac{2a}{v} \Lambda_1 \binom{1}{0} = \frac{2a}{v} \Lambda_1 = a, \tag{19}$$

where the last equality uses the consistency condition $\lambda_1 = 1$, hence $\Lambda_1 = v/2$. Thus the group velocity in (19) always points in the wrong direction, yielding undamped ‘backward’ propagation of saw-tooth wave modes. This is confirmed by simple numerical tests.

NUMERICAL TESTS

Equation (1) is solved numerically with an initial high-frequency wave packet

$$u_0(x) = u(x, 0) = e^{-\sigma x^2} \cos(\pi x).$$

The exact solution at time t is the initial condition shifted to the *left* by a distance at , i.e. $u(x, t) = u_0(x + at)$. In Figure 1 we use the LW scheme (11) with $v = 0.1, 0.05, 0.01$ and 0.005 .

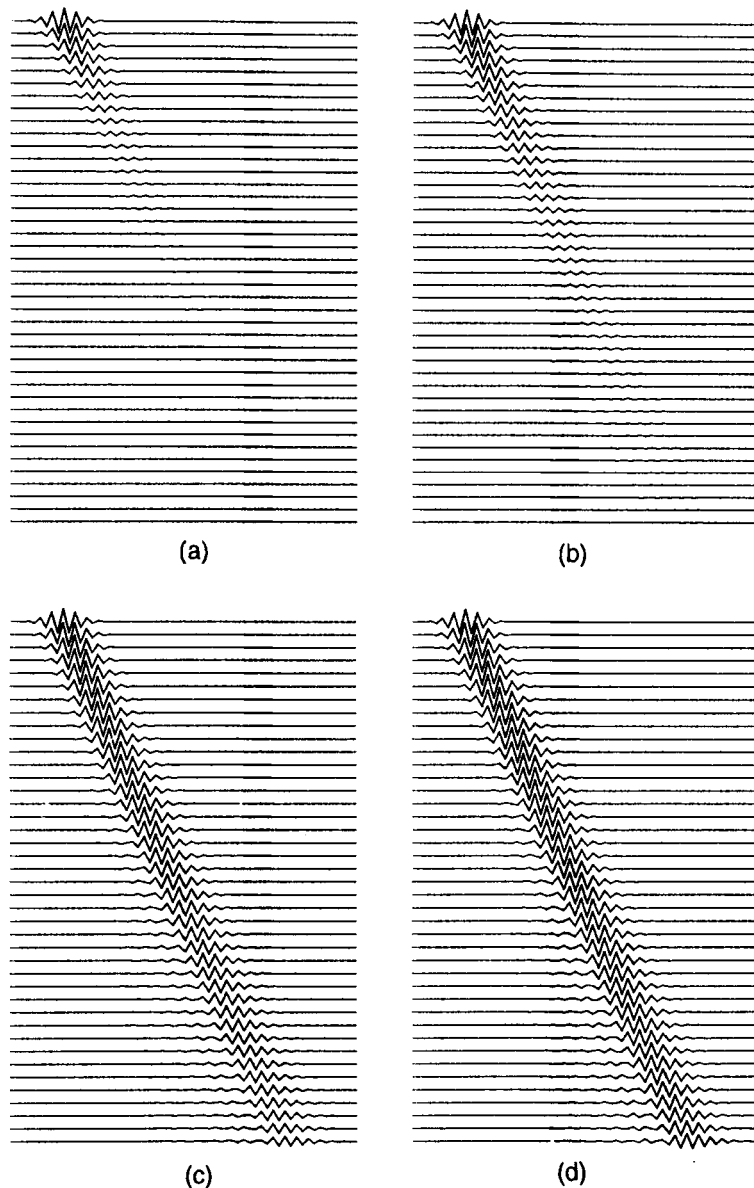


Figure 1. LW scheme (11)—evolution of solution profiles at (A) $\nu = 0.1$, (B) $\nu = 0.05$, (C) $\nu = 0.01$ and (D) $\nu = 0.005$. High-frequency waves propagate 'backwards' and persist longer as $\nu \rightarrow 0$

In all cases the solution is computed until a fixed time T , so the total number of time steps increases as ν decreases. As indicated by the analysis, the initial wave moves in the wrong direction and is lightly damped. It is also noticed that as ν gets smaller, the overall damping decreases and the wave tends to persist. In the limit the wave is undamped. The same test is repeated in Figure 2, this time with the second-order upwind scheme of (12) with $\nu = 0.0005$ and 0.0001 . The wave appears to begin moving in the wrong direction but is strongly damped

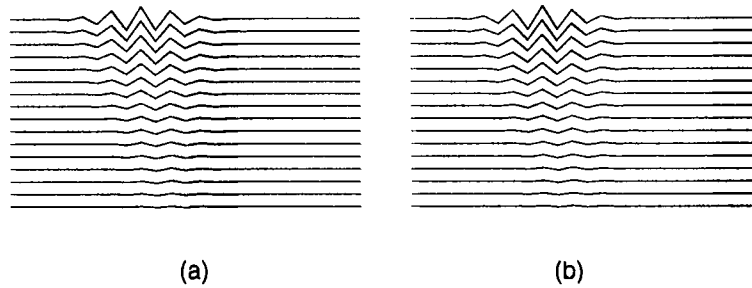


Figure 2. Second-order upwind scheme of (12)—evolution of solution profiles at (A) $\nu = 0.0005$ and (B) $\nu = 0.0001$. As $\nu \rightarrow 0$, the scheme remains dissipative, with waves strongly damped before they succeed to propagate ‘backwards’

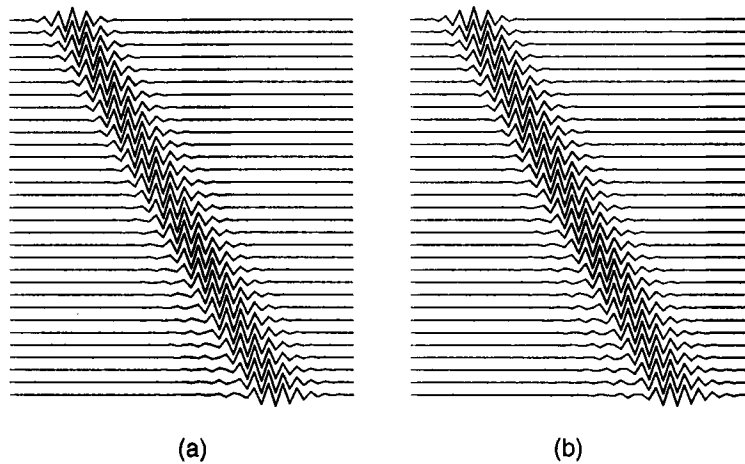


Figure 3. RK scheme (14)—evolution of solution profiles at $\nu = 0.5$; (A) two-stage time integrator and (B) four-stage time integrator with $\lambda_k = k^{-1}$. In both cases waves are not dissipated at all

at a rate e^{-4T} independently of the CFL number ν . In Figure 3 the RK scheme is used with $s = 2$ and 4 and with $\lambda_k = k^{-1}$ at $\nu = 0.5$. The initial wave packet moves ‘backwards’ and is undamped.

CONCLUDING REMARKS

We have considered dissipative numerical approximations to the linear advection equation (1) in the limit of weak dissipation. The analysis was conducted under assumptions typical of far-field wave propagation. We analysed explicit two-level symmetric/upwind schemes of optimal accuracy and symmetric RK schemes. We have shown that when $\nu \rightarrow 0$, the amount of damping per time step of all schemes diminishes and high-frequency spurious solutions are admitted to propagate ‘backwards’. While upwind schemes maintain sufficient dissipation to damp the spurious solutions exponentially in time, the accumulative damping in symmetric schemes tends to unity, yielding practically non-dissipative schemes. This sheds light on why parasitic high-frequency modes constitute a much more severe problem in the symmetric case even when the schemes are dissipative.

REFERENCES

1. A. Iserles, 'Generalized leap-frog methods', *IMA J. Numer. Anal.*, **6**, 381–392 (1986).
2. B. Engquist and S. Osher, 'One-sided difference schemes and transonic flow', *Proc. Natl. Acad. Sci. U.S.A.*, (1980).
3. L. N. Trefethen, 'Group velocity in finite difference schemes', *SIAM Rev.*, **24**, 113–135 (1982).
4. L. N. Trefethen, 'Group velocity interpretation of the stability theory of GKS', *J. Comput. Phys.*, **49**, 199–217 (1983).
5. R. Vichnevetsky, 'Wave propagation analysis of difference schemes for hyperbolic equations: a review', *Int. j. numer. methods fluids*, **7**, 409–452 (1987).
6. S. Karni, 'To the boundary and back—a numerical study', *Int. j. numer. methods fluids*, **13**, 201–216 (1991).
7. A. Iserles, 'Order stars and a saturation theorem for first order hyperbolics', *University of Cambridge DAMPT NA3/81*, 1981.