Local oscillations in finite difference solutions of hyperbolic conservation laws

4 authors, including:

Jiequan Li
Institute of Applied Physics and Computational Mathematics
81 PUBLICATIONS   1,437 CITATIONS

SEE PROFILE

Huazhong Tang
Peking University
99 PUBLICATIONS   1,580 CITATIONS

SEE PROFILE

Gerald Warnecke
Otto-von-Guericke-Universität Magdeburg
181 PUBLICATIONS   3,054 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Two phase flows with phase transition   View project

Project   High order methods   View project

LOCAL OSCILLATIONS IN FINITE DIFFERENCE SOLUTIONS OF HYPERBOLIC CONSERVATION LAWS

JIEQUAN LI, HUAZHONG TANG, GERALD WARNECKE, AND LUMEI ZHANG

ABSTRACT. It was generally expected that monotone schemes are oscillation-free for hyperbolic conservation laws. However, recently local oscillations were observed and usually understood to be caused by relative phase errors. In order to further explain this, we first investigate the discretization of initial data that trigger the chequerboard mode, the highest frequency mode. Then we proceed to use the discrete Fourier analysis and the modified equation analysis to distinguish the dissipative and dispersive effects of numerical schemes for low frequency and high frequency modes, respectively. It is shown that the relative phase error is of order $\mathcal{O}(1)$ for the high frequency modes $u_j^n = \lambda_k^n e^{i\xi j}$, $\xi \approx \pi$, but of order $\mathcal{O}(\xi^2)$ for low frequency modes ($\xi \approx 0$). In order to avoid numerical oscillations, the relative phase errors should be offset by numerical dissipation of at least the same order. Numerical damping, i.e. the zero order term in the corresponding modified equation, is important to dissipate the oscillations caused by the relative phase errors of high frequency modes. This is in contrast to the role of numerical viscosity, the second order term, which is the lowest order term usually present to suppress the relative phase errors of low frequency modes.

1. INTRODUCTION

It is expected that monotone schemes give stable approximations to the scalar hyperbolic conservation law

$$(1.1) \qquad u_t + f(u)_x = 0, \quad x \in \mathbb{R}, \ \ t > 0.$$

The resulting solution satisfies the TVD, i.e. total variation diminishing property, the maximum principle and the entropy inequality as well as monotonicity preserving; see [6]. The monotonicity preserving property only implies that if the initial data is monotone, then the solution should have the same property for all time $t > 0$. Theoretically it does not exclude the onset of oscillations at local extrema; see the discussion in [2]. It was a common understanding and observation in practice that the numerical viscosity of such schemes may sufficiently offset relative phase errors and suppress oscillations completely. However, to the contrary, local oscillations were observed in [1, 2, 5, 9] and analyzed in [1, 2] by checking the formation of new local extrema in solutions. The purpose of this paper is to

analyze and understand this seemingly paradoxical phenomenon by applying the method of discretized Fourier analysis and the modified equation analysis, which will help us to further understand local oscillations caused individually by low and high frequency modes.

We limit our attention to 3-point finite difference schemes for (1.1) in a viscosity form [8]

$$(1.2) \quad u_j^{n+1} = u_j^n - \frac{\nu}{2}\big(f(u_{j+1}^n) - f(u_{j-1}^n)\big) + \frac{Q_{j+1/2}^n}{2}(u_{j+1}^n - u_j^n) - \frac{Q_{j-1/2}^n}{2}(u_j^n - u_{j-1}^n),$$

where the *mesh ratio* $\nu = \tau/h$ is assumed to be a constant, τ and h are step sizes in time and space, respectively, u_j^n denotes an approximation of $u(jh, n\tau)$, the terms $Q_{j+1/2} \in]0, 1]$ are the *coefficients of numerical viscosity*. If $Q_{j+1/2}^n = 1, \nu|a_{j+1/2}^n|$, or $(\nu a_{j+1/2}^n)^2$, (1.2) is, namely, the Lax-Friedrichs (LxF) scheme, the Courant-Isaacson-Rees (CIR) scheme, or a version of the Lax-Wendroff (LW) scheme [4], respectively, where

$$a_{j+1/2}^n = \begin{cases} \frac{f(u_{j+1}^n) - f(u_j^n)}{u_{j+1}^n - u_j^n}, & u_{j+1}^n \neq u_j^n, \\ f'(u_j^n), & u_{j+1}^n = u_j^n. \end{cases}$$

The above CIR scheme (see Tadmor [8, page 371]) is a generalization of the simplest upwind scheme for linear advection equation ($f(u) = au$ and $Q_{j+1/2}^n \equiv |a\nu|$) to the case of nonlinear flux functions. In case the coefficients $Q_{j\pm1/2}$ do not depend on the numerical solution u_j^n the condition $\nu|f'(u_{j\pm1}^n)| \leq Q_{j\pm1/2}^n \leq 1$ implies that the scheme (1.2) is monotone. The LxF scheme is monotone. For the linear advection equation $f(u) = au$ the upwind scheme is also monotone. On the other hand, the LW scheme is nonmonotone. If $Q_{j+1/2} = q \in]0, 1[$ is constant with the time step restriction $\max\{\nu|f'(u)|\} \leq q$, the scheme (1.2) is usually called a generalized LxF scheme and it is also monotone under this condition. The special case $q = 1/2$ is the modified LxF scheme; see Tadmor [8]. This scheme turns out to be a special case in the analysis below; see also [2].

Numerical oscillations caused by the LW scheme are quite well understood; see [7, page 100]. Assume that $f(u) = au$, i.e. $a_{j+1/2}^n = a$ and $Q_{j+1/2}^n \equiv q$ are constant, for all $j \in \mathbb{Z}$. The relative phase error of low frequency modes $u_j^n = \lambda_k^n e^{i2\pi kjh} = \lambda_k^n e^{i\xi j}$, for kh or scaled wave number $\xi = 2\pi kh$ small and $i = \sqrt{-1}$, is of order $\mathcal{O}(kh) = \mathcal{O}(\xi)$, and cannot be offset by the small dissipation of order $\mathcal{O}((kh)^2) = \mathcal{O}(\xi^2)$. As is well known, the LW scheme is second order accurate both in time and space. In contrast, for first order accurate monotone schemes, the numerical viscosity is enough to control numerical oscillations caused by the relative phase errors of low frequency modes because the dissipation, i.e. amplitude error, has a magnitude of order higher than that of the relative phase errors. However, for high frequency modes, the situation is quite different. We take the highest frequency mode, *a chequerboard mode*, as the initial data

$$(1.3) \quad u_j^0 = (-1)^j,$$

and catch a glimpse of the resolution of high frequency modes by (1.2). The solution of the 3-point scheme (1.2) for constant q (i.e., the generalized LxF scheme) at $t = n\tau$ is easily seen to be

$$(1.4) \quad u_j^n = (1 - 2q)^n (-1)^j.$$

This shows that the solution of (1.2) at time $t = n\tau$ for $n \in \mathbb{N}$ is still a chequerboard mode, except for the modified LxF scheme with $q = 1/2$. The amplitude of the solution is diminishing if $0 < q < 1$ but keeps invariant for the LxF scheme with $q = 1$ or for the unstable central scheme with $q = 0$. For $q = 1/2$ it is wiped out immediately. The typical chequerboard mode usually contaminates the solution and causes unwanted oscillations in computations. The above rough analysis motivates us to investigate the interrelation among oscillations, relative phase errors and numerical dissipations, i.e. numerical viscosity and numerical damping, systematically.

We show that numerical oscillations in the second order accurate LW solution are caused by the relative phase error of low frequency modes; while local observable oscillations in first order accurate monotone schemes are caused by high frequency modes, particularly of the form (1.3). Numerical viscosity is not sufficient to offset the oscillations by the high frequency modes. In nonlinear cases the nonlinearity of the flux function $f(u)$ does not have any effect on this matter.

Due to the role of the frequency of the Fourier modes in the oscillations, we consider the discretization of general initial data

(1.5) $$u(x, 0) = u_0(x),$$

in different ways, which possibly contain the chequerboard mode (1.3). It becomes obvious from the analysis of the discretizations of a single square signal with an odd or even number of grid points: The former contains the highest frequency chequerboard mode (1.3) and the resulting solution displays the oscillatory phenomenon; but the latter displays exactly the expected nonoscillatory behavior of a monotone scheme; cf. [1, 2].

In order to understand the resolution of Fourier modes of different frequencies by the scheme (1.2), we proceed to use the discrete Fourier analysis to find that the relative phase error of high frequency modes is of order $\mathcal{O}(1)$ and causes severe oscillations unless there exists a strong numerical dissipation to suppress these errors of the high frequency modes. To further understand this observation, we will respectively consider the smooth solution $(U^s)_j^n$ corresponding to low frequency modes and the oscillatory solution $(U^0)_j^n$ related to high frequency modes, and check their respective modified equations. As in [7], for the smooth part, the modified equation has the familiar form

$$\partial_t U^s + a \partial_x U^s = \alpha_1^s(q, a, h)\tau \partial_x^2 U^s + \alpha_2^s(q, a, h)\tau^2 \partial_x^3 U^s + \cdots .$$

While for high frequency modes, the modified equation is of the form

$$\partial_t U^o + a \partial_x U^o = \frac{\ln|2q - 1|}{\tau} U^o + \alpha_1^o(q, a, h)\tau \partial_x^2 U^o + \alpha_2^o(q, a, h)\tau^2 \partial_x^3 U^o + \cdots ,$$

where $q \neq 1/2$. Here α_j^s, α_j^o, $j = 1, 2, \ldots$, are all uniformly bounded functions. The zero order term $\frac{\ln(|2q-1|)}{\tau} U^o$ is a numerical damping and has the order $\mathcal{O}(1)$. This term displays a stronger dissipative effect, compared with the numerical viscosity $\alpha_1^o(q, a, h)\tau \partial_x^2 U^o$. Here it exerts dominant effects of dissipation on the high frequency modes. For the LxF scheme ($q = 1$) the damping term vanishes and numerical viscosity, the second order term, $\alpha_1^o(q, a, h)\tau \partial_x^2 U^o$ takes effect of dissipation of order $\mathcal{O}(\xi)$, thus it is not sufficient to offset the relative phase error of order $\mathcal{O}(1)$.

Thus we can clarify the oscillations in the solutions: *The oscillations are caused by large phase errors and they are not offset by the numerical dissipation of the same*

order. For low frequency modes, the relative phase errors are offset by the numerical viscosity of lower order. For high frequency modes, the error is of order $\mathcal{O}(1)$, and thus the numerical viscosity is not enough to dampen the resulting oscillations and numerical damping is important in this aspect.

We organize this paper into seven sections. In Section 2 we present several examples to display oscillations. In Section 3 we take two different ways to discretize initial data and find chequerboard modes in the presence of initial discrete values. Then the generalized LxF scheme is used to glimpse at the resolution of Fourier modes in Section 4. The main analysis is made in Section 5 to investigate the dissipative and dispersive mechanisms by using linear advective equations. In Section 6 we investigate the nonlinear version and obtain the same conclusion. In Section 7 we summarize the conclusion with more discussions.

2. Local oscillations in the solutions by generalized LxF schemes

The fact that the second order LW scheme produces oscillations is well known. In this section we present several examples to display local oscillations in the solution of (1.1) by the first order generalized LxF schemes, as observed in [1, 2, 9], although the schemes are monotone and TVD under a certain restriction. From these examples we see that the different ways of initial discretization lead to distinct solution behaviors. This motivates the analysis of the discretization of initial data in the next section.

Example 1. Linear advection equation. Consider the advection equation $u_t + u_x = 0$ over the region $x \in [0,1]$ by using the LxF scheme. We take the grid points $M = 50$, $\nu = \tau/h = 0.8$, and use periodic boundary condition just for simplicity. We first look at the impulsive initial data $u_j^0 = 1$ for $j = M/2$, and $u_j^0 \equiv 0$ otherwise. The numerical solution is shown in Figure 2.1(a) to display clear oscillations. Note that the total variation keeps invariant 2. Then we investigate the case that distributed square pulse initial data $u_j^0 = 1$ for $j = M/2$, $M/2+1$ and $u_j^0 \equiv 0$ are taken otherwise. The numerical solution displays exactly the opposite behavior. No oscillation is present; see Figure 2.1(b). However, the total variation is decreasing to be 0.3398 at time $t = 1$.

Example 2. Burgers equation. In order to see if the above oscillatory phenomenon is strongly influenced by the nonlinearity, we use the inviscid Burgers equation $u_t + (u^2/2)_x = 0$ as a nonlinear example. The same initial data are taken as for the above linear case. Similar numerical results are displayed in Figure 2.2. Therefore, the oscillations are not connected to the nonlinearity, just as we pointed out for chequerboard mode in the introduction. This result is in sharp contrast to the fact that the nonlinearity introduces an additional effect of dissipation at discontinuities, such as the step function data, that can be observed when comparing solutions to linear advection and the Burgers equation.

These two examples were presented in [1, 2]. The next example is for a nonlinear system.

Example 3. Compressible Euler equations. This example is for nonlinear systems. We consider the system of compressible Euler equations

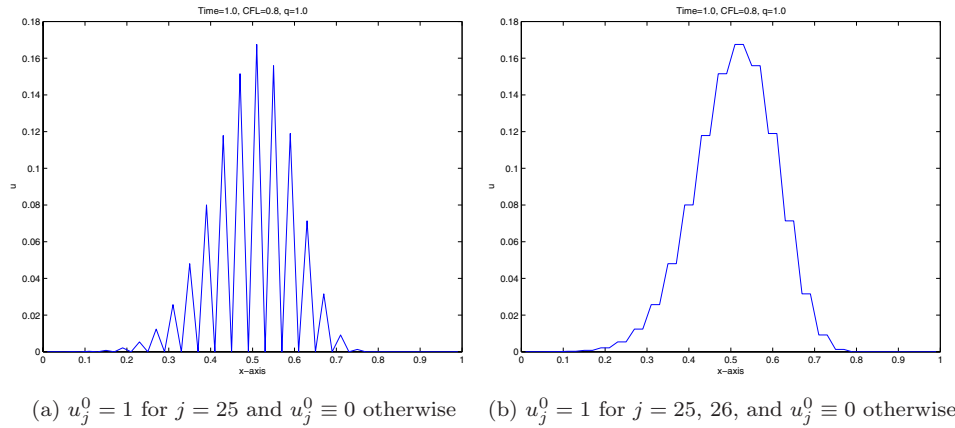$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} = 0,$$

(a) $u_j^0 = 1$ for $j = 25$ and $u_j^0 \equiv 0$ otherwise   (b) $u_j^0 = 1$ for $j = 25$, 26, and $u_j^0 \equiv 0$ otherwise

FIGURE 2.1. Numerical results for the advection equation $u_t + u_x = 0$, 50 grid points are used, $\nu = 0.8$.

$u_j^0 = 1$ for $j = 25$ and $u_j^0 \equiv 0$ otherwise $u_j^0 = 1$ for $k = 25$, 26, and $u_j^0 \equiv 0$ otherwise
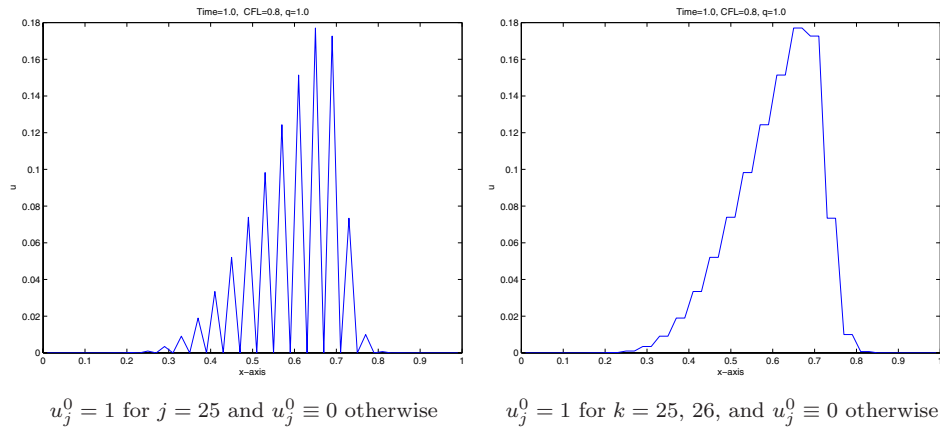
FIGURE 2.2. Numerical results for the Burgers equation $u_t + (u^2/2)_x = 0$, 50 grid points are used.

with $U = (\rho, \rho u, E)^\top$ and $F(U) = (\rho u, \rho u^2 + p, u(E + p))^\top$, where ρ is the density, u the velocity, p the pressure, and $E = \frac{1}{2}\rho u^2 + \rho e$ the total energy with the internal energy e. We take the equation of state $p = (\gamma - 1)\rho e$ for polytropic gases. The index $\gamma = 1.4$ is taken for air. Similar to the case of a single equation, we still use odd and even points to discretize a square-shaped signal initial data. In Figure 2.3, we see the obvious oscillations by the LxF scheme $q = 1$. However, if we use $q = 0.9$ in the scheme, the solution shown in Figure 2.4 is nonoscillatory. On the other hand, if we express the middle initial state with two points, i.e. an even number, the results are displayed in Figure 2.5. No oscillation is present.

We observe that the presence of oscillations is not only related to the ways in which the initial data are approximated, but also to the numerical viscosity coefficient q. For the LxF scheme, the local oscillations are very strong, although the total variation property is not violated due to a strong decay of numerical solution amplitude.
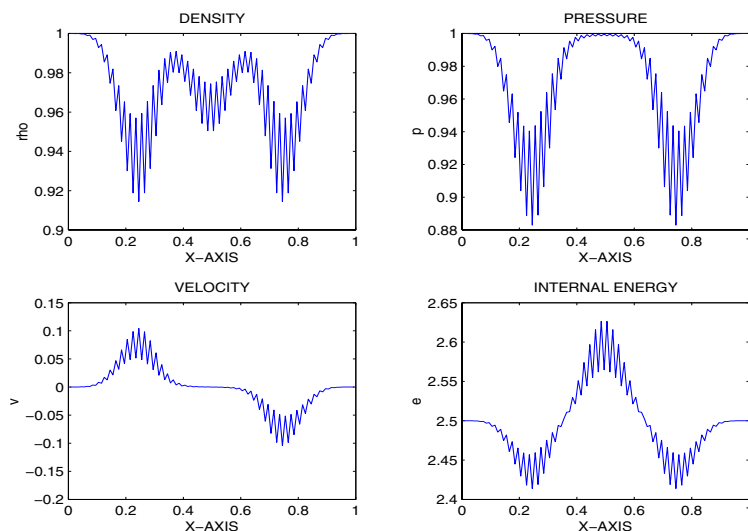
FIGURE 2.3. Numerical results for the Euler equations at time $t = 0.25$, 100 grid points are used, $CFL = 0.6$, and $q = 1$. The initial data are $(\rho_j^0, u_j^0, p_j^0) = (0.125, 0, 0.1)$ for $j = 49, 50, 51$, and $(\rho_j^0, u_j^0, p_j^0) = (1, 0, 1)$ otherwise. The Courant number is taken as usual: $CFL = \max_j\{c_j^n + |u_j^n|\}\nu$ for each n, c_j^n is the local sound speed.
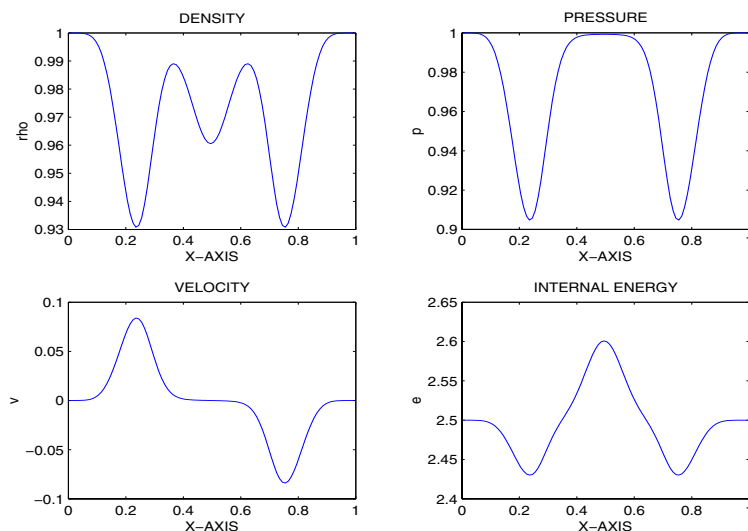

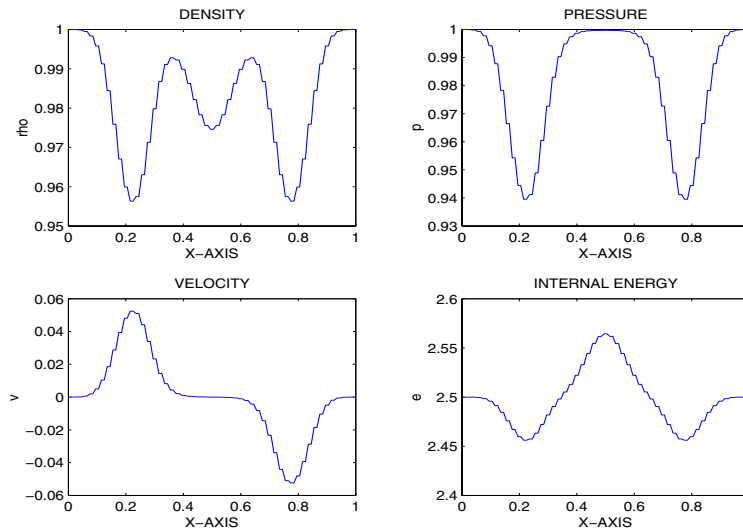FIGURE 2.4. Same as Figure 2.3 except with $q = 0.9$.

FIGURE 2.5.   Same as Figure 2.3 except with the initial data of $(\rho_j^0, u_j^0, p_j^0) = (0.125, 0, 0.1)$ for $j = 50, 51$ and $(\rho_j^0, u_j^0, p_j^0) = (1, 0, 1)$ otherwise.

3. Chequerboard modes in the initial discretization

As observed in the last section and also in [1], the numerical solutions display very distinct behaviors if the initial data are discretized in different ways. It turns out that chequerboard modes are present and affect the solutions if the initial data contains a square signal and are discretized with an odd number of grid points. This motivates us to discuss the discretization of initial data

$$(3.1) \qquad\qquad u(x,0) = u_0(x), \quad x \in [0,1],$$

with M grid points and $h = 1/M$. For simplicity, we assume that M is even, and $u_0(0) = u_0(1)$. The numerical solution value at the grid point x_j is denoted by u_j^0. We express this grid point value u_j^0 by using the usual discrete Fourier sums, as in [10, page 120]. We use the *scaled wave number* $\xi = 2\pi k h$ and obtain

$$(3.2) \qquad u_j^0 = \sum_{k=-M/2+1}^{M/2} c_k^0 e^{i\xi j}, \quad i^2 = -1, \quad j = 0, 1, \ldots, M-1,$$

where the coefficients c_k^0 are, in turn, expressed as

$$(3.3) \qquad c_k^0 = \frac{1}{M} \sum_{j=0}^{M-1} u_j^0 e^{-i\xi j}, \qquad k = -M/2 + 1, \ldots, M/2.$$

We pay special attention to the particular case that

$$c_k^0 = \begin{cases} 1, & \text{if } k = M/2, \\ 0, & \text{otherwise,} \end{cases}$$

i.e. the initial data are taken to be just the highest Fourier mode which is a single chequerboard mode oscillation

$$u_j^0 = e^{i2\pi \frac{M}{2} jh} = e^{i\pi j} = (-1)^j.$$

3.1. **Single square signal case.** We start with the simple initial data of a square signal, i.e. an initial function of the following type:

$$(3.4) \qquad u(x,0) = \begin{cases} 1, & 0 < x^{(1)} < x < x^{(2)} < 1, \\ 0, & \text{otherwise.} \end{cases}$$

We use the following two methods to approximate the step function (3.4) as a grid function: One uses an odd number of grid points to take the value one of the square signal and the other uses the next smaller even number of grid points. They could be seen as approximations to some given fixed interval with end points not represented on the mesh, which we do not explicitly specify here. We use the discrete Fourier sum to clarify the difference.

(i) **Discretization with an <u>odd</u> number of grid points.** Take $j_1, j_2 \in \mathbb{N}$ such that $j_1 + j_2$ is an even number. We set $x^{(1)} = (\frac{M}{2} - j_1)h$ and $x^{(2)} = (\frac{M}{2} + j_2)h$. We first discretize the square signal (3.4) with $p := j_1 + j_2 + 1$ nodes, i.e. an odd number of grid points, such that

$$(3.5) \qquad u_j^0 = \begin{cases} 1, & \text{if } j = M/2 - j_1, \ldots, M/2 + j_2, \\ 0, & \text{otherwise.} \end{cases}$$

Substituting them into (3.3), we obtain by simple calculation,

$$(3.6) \qquad c_k^0 = h \sum_{j=0}^{M-1} u_j^0 e^{-i\xi j} = \begin{cases} \frac{(-1)^k e^{i\xi j_1}(1 - e^{-i\xi p})}{M(1 - e^{-i\xi})}, & \text{for } k \neq 0, \\ ph, & \text{for } k = 0. \end{cases}$$

We give special attention to the term

$$c_{M/2}^0 = (-1)^{j_1 + M/2} h,$$

since M is even and p is odd. Hence the initial data (3.5) can be expressed in the form

$$(3.7) \qquad u_j^0 = (-1)^{j + j_1 + M/2} h + ph + \sum_{k \neq 0, M/2} \frac{(-1)^k e^{i\xi(j + j_1)}(1 - e^{-i\xi p})}{M(1 - e^{-i\xi})}.$$

(ii) **Discretization with an <u>even</u> number of grid points.** Rather than (i) above, we use $p := j_1 + j_2$ even number of grid points to express the square signal in (3.4) as follows:

$$u_j^0 = \begin{cases} 1, & \text{if } j = M/2 - j_1 + 1, \cdots, M/2 + j_2, \\ 0, & \text{otherwise.} \end{cases}$$

Then we substitute these initial data into (3.3) to obtain

$$c_k^0 = \begin{cases} \frac{(-1)^k e^{i\xi(j_1 - 1)}[1 - e^{-i\xi(p-1)}]}{M(1 - e^{-i\xi})}, & \text{for } k \neq 0, \\ (p - 1)h, & \text{for } k = 0, \end{cases}$$

and $c_{M/2}^0 = 0$. The initial data can be written by using the discrete Fourier sums as

$$(3.8) \qquad u_j^0 = 0 \times (-1)^j + (p-1)h + \sum_{k \neq 0, M/2} \frac{(-1)^k e^{i\xi(j+j_1-1)}[1 - e^{-i\xi(p-1)}]}{M(1 - e^{-i\xi})}.$$

Comparing (3.7) with (3.8), we observe an essential difference lies in the fact that a checkerboard mode $(-1)^{k+j_1+M/2}$ is present in (3.7), but it is filtered out in (3.8). This is closely related to the oscillatory phenomenon observed in [1, 2, 9].

3.2. **Step function initial data case.** For more general piecewise constant initial data (3.1), there are analogous discrete Fourier sum expressions. We divide the computational domain $[0, 1]$ into L subintervals I_l ($l = 1, 2, \ldots, L$), $\bigcup_{l=1}^L I_l = [0, 1]$, the number of the discrete points of a subinterval I_l is M_l, $M_1 + M_2 + \cdots + M_L = M$, and the initial data (3.1) are expressed as

$$(3.9) \qquad u_j^0 = \sum_{l=1}^L \bar{U}_0^l \cdot \chi_l(j),$$

where $\bar{U}_0^l$ are constants, and $\chi_l(j)$ is the characteristic function on I_l,

$$\chi_l(j) = \begin{cases} 1, & \text{if } x_j \in I_l, \\ 0, & \text{otherwise.} \end{cases}$$

Note that (3.9) can be regarded as the superposition of several single square signals of the form (3.4). Then we express (3.9) as a discrete Fourier sum of the form (3.2) with c_k^0. For $k \neq 0, M/2$, we have

$$c_k^0 = \frac{1}{M} \sum_{l=1}^L \bar{U}_0^l \left[\sum_{j=0}^{M-1} \chi_l(j) e^{-i\xi j} \right]$$

$$= \frac{1}{M} \left[\bar{U}_0^1 \sum_{j=0}^{M_1 - 1} e^{-i\xi j} + \bar{U}_0^2 \sum_{j=M_1}^{M_1 + M_2 - 1} e^{-i\xi j} + \cdots + \bar{U}_0^L \sum_{j=p_{L-1}}^{M-1} e^{-i\xi j} \right]$$

$$(3.10) \qquad = \frac{1}{M(1 - e^{-i\xi})} \sum_{l=1}^L \bar{U}_0^l (e^{i\xi M_l} - 1) e^{-i\xi p_l},$$

where $p_l = M_1 + \cdots + M_l$; and for $k = 0, M/2$, we have

$$c_0^0 = \frac{1}{M}(\bar{U}_0^1 M_1 + \bar{U}_0^2 M_2 + \cdots + \bar{U}_0^L M_L),$$

$$c_{M/2}^0 = \frac{1}{M} \sum_{l=1}^L \bar{U}_0^l \left[\sum_{j=0}^{M-1} \chi_l(j)(-1)^j \right].$$

Thus, the initial data are expressed as

$$u_j^0 = \frac{1}{M} \sum_{l=1}^L \bar{U}_0^l \left[\sum_{m=0}^{M-1} \chi_l(m)(-1)^m \right] (-1)^j + \frac{1}{M}(\bar{U}_0^1 M_1 + \bar{U}_0^2 M_2 + \cdots + \bar{U}_0^L M_L)$$

$$(3.11) \quad + \frac{1}{M} \sum_{k \neq 0, M/2} \frac{1}{(1 - e^{-i\xi})} \sum_{l=1}^L \bar{U}_0^l (e^{i\xi M_l} - 1) e^{-i\xi p_l}.$$

Similar to the case of a single square signal, it depends on $c_{M/2}^0$ whether there is a chequerboard mode in the discrete initial data. Therefore, we have three cases here also.

(i) If the number M_l of the discrete points in each I_l is odd, $c_{M/2}^0$ is

$$c_{M/2}^0 = \frac{1}{M} \sum_{l=1}^{L} \bar{U}_0^l \left[\sum_{j=0}^{M-1} \chi_l(j)(-1)^j \right] = \frac{1}{M} \sum_{l=1}^{L} \bar{U}_0^l (-1)^{l+1}.$$

(ii) If the number M_l of the discrete points in each I_l is even, $c_{M/2}^0$ vanishes,

$$c_{M/2}^0 = \frac{1}{M} \sum_{l=1}^{L} \bar{U}_0^l \left[\sum_{j=0}^{M-1} \chi_l(j)(-1)^j \right] = 0.$$

(iii) If the number of the discrete points in some I_l is odd while in the others it is even, the summation in the even case is zero and in the odd case is 1 or -1. That is $c_{M/2}^0 = \frac{1}{M} \sum_{l=1}^{L} \bar{U}_0^l \phi(l)$, where

$$\phi(l) = \begin{cases} 0, & \text{if } I_l \text{ is in the even case,} \\ 1 \text{ or } (-1), & \text{if } I_l \text{ is in the odd case.} \end{cases}$$

This case is just the superposition of (i) and (ii) above.

It is observed that there is no chequerboard mode for case (ii). For case (i), this summation may be zero when the factors cancel. However, since this summation is taken in the global sense and the chequerboard mode exists in each subinterval, the solution may still contain oscillations due to the finite propagation speed property of the scheme. To some extent, this is analogous to the fact that the scheme (1.2) is TVD under a certain restriction, but that local oscillations are still observed. We summarize all of the above analysis in the following proposition.

Proposition 3.1. *Suppose that the initial data (3.1) are given as, or approximated by, a step function. We can discretize them as the superposition of several single square signals. For each square signal we have two different types of discretizations. If they are discretized with an odd number of grid points, the chequerboard, i.e. highest frequency, mode is present. In contrast, if they are discretized with an even number of grid points, this mode is suppressed.*

4. A GLIMPSE OF CHEQUERBOARD MODE PROPAGATION

In this section we simply look at the (generalized) LxF scheme for the linear advection equation, $f(u) = au$,

$$(4.1) \qquad u_j^{n+1} = u_j^n - \frac{\nu a}{2}(u_{j+1}^n - u_{j-1}^n) + \frac{q}{2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n),$$

and catch a glimpse of the resolution of high frequency modes, where q under the condition $|\nu a| \leq q \leq 1$ is a parameter for monotone schemes. If $q = 1$, then (4.1) is the LxF scheme; taking $q = |\nu a|$ it is the upwind scheme; and for $q = a^2 \nu^2$ being smaller than the monotonicity range it is the nonmonotone LW scheme.

As usual for stability analysis, the solution to (4.1) is expressed analogously to (3.2) in the standard form of a discrete Fourier sum using $\xi = 2\pi kh$:

$$(4.2) \qquad u_j^n = \sum_{k=-M/2+1}^{M/2} c_k^n e^{i\xi j}.$$

The coefficients c_k^n are obtained successively and expressed as

$$(4.3) \qquad c_k^n = \left(1 + q(\cos\xi - 1) - i\nu a \sin\xi\right)^n c_k^0.$$

In correspondence with the two kinds of discretization of a single square signal in Section 3, the Fourier coefficients c_j^0 have different expressions, and the solutions are expressed, respectively, as follows.

(i) Odd discretization case. With the initial data (3.7), the solution of (4.1) is

$$(4.4) \qquad u_j^n = \frac{1}{M}(1-2q)^n(-1)^{j+j_1+M/2} + \sum_{k=-M/2+1}^{M/2-1} c_k^n e^{i\xi j}.$$

(ii) Even discretization case. With the initial data (3.8), we have

$$(4.5) \qquad u_j^n = 0 \times (1-2q)^n(-1)^j + \sum_{k=-M/2+1}^{M/2-1} c_k^n e^{i\xi j}.$$

We emphasize that the Fourier coefficients c_k^n have different expressions in correspondence with the odd and even number of nodes taken for the discretization of the signal in the initial data. By comparing (4.4) and (4.5), we see that in the odd case the chequerboard mode does not vanish if it exists initially, unless we have the modified LxF scheme $q = 1/2$, although it decays with a rate of $|2q-1|$ at each time step; see Figure 4.1. Therefore, a proper discretization of initial data (3.1) would be important to suppress these oscillations in the numerical solution of (1.1). This may not be feasible in real flow applications.

Remark 4.1. (1) For the LxF scheme with $q = 1$ the solution $u_j^n = (-1)^{n+j}$ oscillates between 1 and -1 alternately if the chequerboard mode initial data are taken. The large numerical dissipation does not have any effect.

(2) In case $q \neq 1$, i.e. $|1-2q| < 1$, the chequerboard mode is damped out quickly. In particular, for the modified LxF scheme $q = 1/2$ the chequerboard mode is eliminated and has no influence on the solution at all. We will analyze this in Section 5.

Remark 4.2 (Remark on Tang and Warnecke's paper). In [9] a $(2K + 1)$-point scheme of the following type was investigated

$$u_j^{n+1} = u_j^n - \frac{\nu}{2K}\left(f(u_{j+K}^n) - f(u_{j-K}^n)\right) + \frac{\tau\alpha}{2Kh}(u_{j+K}^n - 2u_j^n + u_{j-K}^n),$$

where $\alpha = \max_u |f'(u)|$ and K is a positive integer. This scheme with $K = 1$ is considered as a generalized Lax-Friedrichs scheme; especially in the linear case when $f(u) = au$ this is the upwind scheme. Using a similar analysis as above, we obtain the following solution if we take the square pulse initial data,

$$u_j^n = \left\{1 + \frac{\alpha\nu}{K}\left[(-1)^K - 1\right]\right\}^n (-1)^j + \sum_{k=-M/2+1}^{M/2-1} c_k^n e^{i\xi j},$$
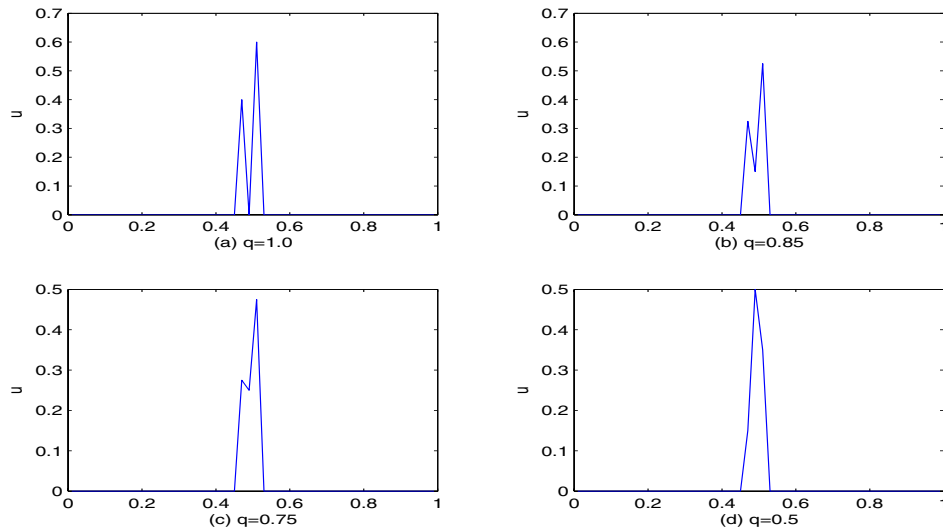
FIGURE 4.1. The decay of oscillations as the parameter q decreases: The initial data are the same as for Figure 2.1(a), the CFL number is 0.2 and only one time step is taken.

where $c_k^n = [1 + \frac{\alpha\nu}{K}(\cos(\xi K) - 1) - ia\nu \sin(\xi K)]c_j^0$. There is a chequerboard mode in the solution and naturally the solution displays such oscillations.

5. Analysis of numerical dissipation and phase error

In this section we attempt to analyze the numerical dissipation and phase error mechanisms of (4.1), particularly on high frequency modes and explain the phenomenon of oscillations caused by high frequency modes. We apply discrete Fourier analysis and the method of modified equation analysis. Both of the methods give complimentary results, which are consistent. We show that as $0 \leq q \leq 1$ is away from $1/2$, the damping on high frequency modes becomes weak. In particular, there is no damping effect in the LxF scheme (i.e. $q = 1$) and the unstable central scheme (i.e. $q = 0$). We note that in [2] the dependence of oscillatory properties on the numerical viscosity was also investigated.

5.1. **Discrete Fourier analysis.** We use the method of discrete Fourier analysis to discuss the dissipation and phase error mechanism of the generalized LxF scheme (4.1). Remember that it is monotone if $0 < |\nu a| < q \leq 1$. We are particularly concerned with the phase accuracy of Fourier modes. Denote a Fourier mode using the scaled wave number $\xi = 2\pi kh$ by $e^{i\xi}$. Then using it as initial data for a linear finite difference scheme results after n time steps in the solution

$$(5.1) \qquad u_k^n = \lambda_k^n e^{i\xi} = (\lambda(k))^n e^{i\xi}, \quad i^2 = -1,$$

where λ_k^n is the amplitude. The modulus of the ratio

$$\lambda(k) = \lambda_k^{n+1}/\lambda_k^n$$

is the *amplitude* of the mode for one time step. For the scheme (4.1) we have with $\nu = \tau/h$ in particular

$$(5.2) \qquad\qquad \lambda(k) = 1 + q(\cos\xi - 1) - i\nu a \sin\xi$$

and

$$
\begin{aligned}
(5.3) \qquad |\lambda(k)|^2 &= \left(1 + q(\cos\xi - 1)\right)^2 + (\nu a)^2 \sin^2\xi \\
&= 1 + 4(a^2\nu^2 - q)\sin^2(\xi/2) + 4(q^2 - a^2\nu^2)\sin^4(\xi/2).
\end{aligned}
$$

For the upwind scheme with $q = \nu|a|$ the last term on the right-hand side vanishes, whereas for the LW scheme with $q = \nu^2 a^2$ the second term vanishes. Also, we see that for the schemes of type (4.1), under the conditions $0 < \nu^2 a^2 \le q \le 1$ (see Figure 5.1), we have from (5.3) the estimate

$$(5.4) \quad |\lambda(k)|^2 = 1 + 4(a^2\nu^2 - q)\left(\sin^2(\xi/2) - \sin^4(\xi/2)\right) + 4q(q-1)\sin^4(\xi/2) \le 1.$$

Thus these conditions imply that the schemes are linearly stable. Conversely, assume that $|\lambda(k)|^2 \le 1$, then from (5.3) we have $(a^2\nu^2 - q) + (q^2 - a^2\nu^2)\sin^2(\xi/2) \le 0$, which further implies that $(a^2\nu^2 - q) + [q^2 - a^2\nu^2 + q(q-1)]\sin^2(\xi/2) \le 0$ must hold for $\xi \in [0, \pi]$. Now for $\xi = 0$ we obtain $a^2\nu^2 \le q$. For $\xi = \pi$ this gives $q(q-1) \le 0$ or since we have $0 < a^2\nu^2 < q$ we obtain $q \le 1$. Therefore, these conditions are necessary and sufficient for stability.
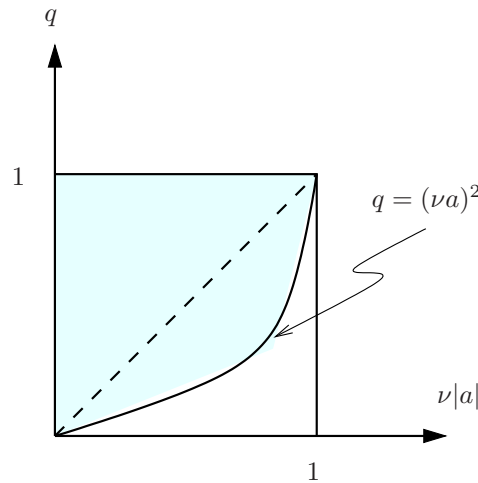

FIGURE 5.1. Range of stability of parameter q over CFL number νa.

The exact solution of the Fourier mode $e^{i\xi}$ for $x = h$ after one time step τ is $e^{i(\xi - 2\pi ak\tau)} = e^{-i2\pi ak\tau}e^{i\xi} = \lambda_{exact}(k)e^{i\xi}$. The exact amplitude $\lambda_{exact}(k)$ has modulus 1. We see from (5.3) that the *amplification error*, i.e. the error in amplitude modulus, is of order $\mathcal{O}(\xi)$ for the monotone schemes and order $\mathcal{O}(\xi^2)$ for the LW scheme. If the modulus of $\lambda(k)$ is less than one, the effect of the multiplication of a solution component with $\lambda(k)$ is called *numerical dissipation* and then the amplification error is called *dissipation error*. If the modulus is larger than 1, this leads to the amplification of the Fourier mode, i.e. instability of any solution

containing it. Further, comparing the exponents of $\lambda(k)$ and $\lambda_{exact}(k)$ there is a *phase error* $\arg\lambda(k) - (-2\pi a k\tau)$. The *relative phase error* is then defined as

$$E_p(k) := \frac{\arg\lambda(k)}{-2\pi a k\tau} - 1 = -\frac{\arg\lambda(k)}{\nu a\xi} - 1.$$

A mode is a low frequency mode if $\xi\approx 0$ and a high frequency mode if $\xi\approx\pi$. We first look at the low frequency modes

$$(U^s)_j^n := \lambda_k^n e^{i\xi j}, \quad \xi\approx 0.$$

For $k = \xi = 0$ we have $\lambda(k) = 1$. From (5.3) we obtain

$$(5.5) \qquad \frac{d(|\lambda(k)|^2)}{dq} = 2(1 + q(\cos\xi - 1))(\cos\xi - 1) < 0,$$

for fixed $\xi \in]0, \pi/2]$. This implies that the dissipation becomes weaker as q decreases. The LxF scheme with $q = 1$ has the largest numerical dissipation for low frequency modes.

The phase of the low frequency modes is approximated by Taylor expansion at $\xi = 0$:

$$(5.6) \quad \arg\lambda = \arctan\left(\frac{-\nu a\sin\xi}{1 + q[\cos\xi - 1]}\right) \approx -\nu a\xi\left(1 + \frac{3q - 1 - 2\nu^2 a^2}{6}\xi^2 + \cdots\right).$$

This phase has a relative error $E_p(k)$ of order $\mathcal{O}(\xi^2)$. For the LW scheme, this phase error causes oscillations, which cannot be suppressed by the weaker dissipation of order $\mathcal{O}(\xi^2)$, compared to the dissipation error $\mathcal{O}(\xi)$ of the upwind scheme.

For high frequency modes (5.1), $\xi\approx\pi$, the situation is very different. We introduce the decomposition $\xi = \pi + \xi'$, i.e. $\xi' = 2\pi k'h$ with $kh = 1/2 + k'h$, and thus $\xi'\approx 0$. We write the modes in the form

$$(5.7) \qquad (U^h)_j^n = \lambda_k^n e^{i\xi j} = \lambda_k^n e^{i(\pi + \xi')j} = (-1)^{j+n}\lambda_{k'}^n e^{i\xi' j},$$

with $\lambda_{k'}^n = (-1)^{j+n}e^{i\pi j}\lambda_k^n$ and set

$$(U^o)_j^n := \lambda_{k'}^n e^{i\xi' j}.$$

The factor $(U^o)_j^n$ can be regarded as a perturbation amplitude of the chequerboard modes $(e^{i\pi})^{j+n} = (-1)^{j+n}$. The dissipation (amplitude error) depends only on $\lambda_{k'}^n$. Then substituting $(U^h)_j^n$ into (4.1) yields

$$\lambda' := \lambda_{k'}^{n+1}/\lambda_{k'}^n = -1 + q(1 + \cos\xi') - i\nu a\sin\xi'.$$

Therefore, we have

$$(5.8) \qquad \begin{aligned} |\lambda'|^2 &= (1 - q(1 + \cos\xi'))^2 + \nu^2 a^2\sin^2\xi' \\ &= 1 + 4(a^2\nu^2 - q)\cos^2(\xi'/2) + 4(q^2 - a^2\nu^2)\cos^4(\xi'/2). \end{aligned}$$

This is consistent with a shift of π in the variable ξ in (5.3). Regarding the high frequency modes, for all schemes the amplitude error is $\mathcal{O}(1)$. At $\xi' = 0$ we have

$$(5.9) \qquad |\lambda'|^2 = 1 - 4q(1 - q),$$

so we have the lowest amplitude error for $q = 1$ or near zero, the highest for the modified LxF scheme with $q = 1/2$. Obviously, for small ξ', $|\lambda'|^2$ is an increasing function of q if $q > 1/2$ because

$$(5.10) \qquad d(|\lambda'|^2)/dq = -2[1 - q(1 + \cos(\xi'))](1 + \cos\xi') > 0.$$

That is, (4.1) becomes much more dissipative for high frequency modes as the parameter q decreases, which is in sharp contrast to the situation for low frequency modes; see (5.5).

Note that for the LW scheme with $q = \nu|a|$ and the upwind scheme with $q = \nu|a|$ the situation depends on the choice of the CFL number $\nu|a|$ or equivalently the step size for the computation; see Figure 5.1. We disregard the singular case of CFL number 1 in which the schemes reproduce the exact solution for linear fluxes. But for CFL numbers near one they are not very dissipative for the high frequency modes, whereas this is a strong dissipation for a CFL number giving $q = 1/2$.

Furthermore, let us look at the relative phase error. We compute

$$\arg \lambda' = \tan^{-1}\left(\frac{-\nu a \sin \xi'}{-1 + q(1 + \cos \xi')}\right)$$

(5.11)
$$= \frac{-\nu a \xi'}{2q - 1} - \frac{\nu a}{3(2q-1)^2}\left[\frac{q+1}{2} - \frac{\nu^2 a^2}{2q-1}\right]\xi'^3 + \mathcal{O}(\xi'^5).$$

Then for the high frequency modes $(U^h)_j^n$, we have by recalling that $\xi = \pi + \xi'$,

$$(U^h)_j^n = (-1)^{j+n}\lambda_k'^n e^{i\xi' j}$$
$$= |\lambda'|^n e^{in(-\pi + \arg \lambda')} \cdot e^{ij(\pi + \xi')}$$
$$= |\lambda'|^n e^{i(j\xi - 2\pi k a n\tau)} \cdot e^{in(-\pi + \arg \lambda' + \nu a \xi)}.$$

Therefore, the relative phase error of high frequency modes at each time step is using (5.11)

$$E_p(k) = -\frac{-\pi + \arg \lambda' + \nu a \xi}{\nu a \xi}$$
$$= -\frac{\pi(1 - \nu a)}{\nu a \xi} + \frac{2(q-1)\nu a \xi'}{(2q-1)\nu a \xi}$$
(5.12)
$$- \frac{1}{3(2q-1)^2\xi}\left[\frac{q+1}{2} - \frac{\nu^2 a^2}{2q-1}\right]\xi'^3 + \mathcal{O}(\xi'^5).$$

We note that $\xi \approx \pi$. Therefore, the relative phase error has $\mathcal{O}(1)$. This error is huge, and strong numerical dissipation is needed to suppress it.

We summarize the above Fourier analysis in the following proposition.

Proposition 5.1. *We distinguish low and high frequency Fourier modes $u_j^n = \lambda_k^n e^{ij\xi}$, $\xi = 2\pi kh$, and they behave differently.*

(i) For the low frequency modes ($\xi \sim 0$), the relative phase error is of order $\mathcal{O}(\xi^2)$ (see (5.6)), and the amplitude error (dissipation) becomes smaller as the parameter q decreases; see (5.5). The order of amplitude error is $\mathcal{O}(\xi)$ for the monotone schemes and $\mathcal{O}(\xi^2)$ for the LW scheme.

(ii) For the high frequency modes ($\xi \sim \pi$), the relative phase error is of order $\mathcal{O}(1)$ (see (5.12)), the amplitude error becomes larger as the parameter q is closer to $1/2$; see (5.9).

5.2. **Modified equation analysis.** As we know, the amplitude error and relative phase error of the Fourier modes have a correspondence with dissipation and phase error mechanisms displayed by related partial differential equations. Hence we use the method of modified equation analysis to further investigate the mechanisms of dissipation and phase error of (1.2). Particularly, we want to see how the dissipation

offsets the large phase error of high frequency modes. This method of modified equation analysis was originally introduced for low frequency modes [11]. Here it is especially used for high frequency modes, as its usefulness was clearly shown in [7]. We begin in this section with the linear case, and still use notation $\nu = \tau/h$. The nonlinear case is left to the next section.

As in [7, page 173], we will consider a smooth solution $(U^s)_j^n$ and an oscillatory solution $(U^h)_j^n$, respectively. The oscillatory solution $(U^h)_j^n$ is written as

$$(5.13) \qquad (U^h)_j^n = (-1)^{j+n}(U^o)_j^n,$$

where $(U^o)_j^n$ is viewed as the perturbation amplitude of the chequerboard mode.

The smooth solution $(U^s)_j^n$ satisfies (4.1), i.e.

$$(5.14) \quad (U^s)_j^{n+1} = (U^s)_j^n - \frac{\nu a}{2}((U^s)_{j+1}^n - (U^s)_{j-1}^n) + \frac{q}{2}((U^s)_{j+1}^n - 2(U^s)_j^n + (U^s)_{j-1}^n).$$

Then we derive a modified equation for this solution, and the notation $\widetilde{U}^s$ corresponds to the associated exact solution

$$(5.15)$$
$$\partial_t \widetilde{U}^s + a\partial_x \widetilde{U}^s = \frac{1}{2\tau}\left(qh^2 - a^2\tau^2\right)\partial_x^2 \widetilde{U}^s + a\left(-\frac{h^2}{6} + \frac{1}{2}qh^2 - \frac{1}{3}a^2\tau^2\right)\partial_x^3 \widetilde{U}^s + \cdots.$$

The process of how to derive this equation can be found in [7, page 169]. It is evident that the numerical viscosity of (4.1) becomes stronger for low frequency modes as q is larger, and vice versa. In particular, for the LW scheme with $q = \nu^2 a^2$ the dissipation mechanism comes from the fourth order term and therefore is quite weak. This is consistent with the fact observed by the Fourier analysis that the dissipation effect becomes weaker as the viscosity coefficient q decreases, provided that the scheme is stable.

However, the numerical dissipation of (1.2) is very different for the high frequency modes. Let us discuss the perturbation $(U^o)_j^n$ of the oscillatory solution (5.13). Substituting (5.13) into (4.1) yields

$$(5.16) \qquad \frac{(U^o)_j^{n+1} - (U^o)_j^n}{\tau} = \frac{q-2}{\tau}(U^o)_j^n - \frac{\nu a}{2\tau}[(U^o)_{j+1}^n - (U^o)_{j-1}^n]$$
$$+ \frac{q}{2\tau}[(U^o)_{j+1}^n + (U^o)_{j-1}^n].$$

Compared to (5.14) for the low frequency modes, (5.16) contains an extra term $\frac{q-2}{\tau}(U^o)_j^n$, which plays the key role of damping on high frequency modes. We use the notation $\tilde{U}^o(jh, n\tau)$ to express $(U^o)_j^n$ inserted into (5.16) and apply the standard approach (see [7, page 173]). That is, taking the standard Taylor expansion yields

$$(5.17) \qquad \mathcal{D}_{+t}\widetilde{U}^o + a\partial_x\widetilde{U}^o = \frac{2(q-1)}{\tau}\widetilde{U}^o + \frac{a^2 q\tau}{2\nu^2}\partial_x^2\widetilde{U}^o - \frac{a^3\tau^2}{6\nu^2}\partial_x^3\widetilde{U}^o + \cdots,$$

where D_{+t} is a forward difference operator in time and can be expressed as $D_{+t} = (e^{\tau\partial_t} - 1)/\tau$. Note that in (5.16) the term $\frac{q-2}{\tau}(\widetilde{U}^o)_j^n$ is unusual compared to classical modified equation analysis. Next, we write (5.17) as

$$(e^{\tau\partial_t} - 1)\widetilde{U}^o = \beta\widetilde{U}^o - a\tau\partial_x\widetilde{U}^o + \frac{a^2 qh^2}{2}\partial_x^2\widetilde{U}^o - \frac{a^3\tau h^2}{6}\partial_x^3\widetilde{U}^o + \cdots$$

where $\beta = 2(q-1)$. Note the following basic facts, namely the formal operator expansion

$$\tau \partial_t = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(e^{\tau \partial_t} - 1)^m}{m},$$

and the well-known power series

$$\frac{1}{(1+z)^2} = \sum_{m=0}^{\infty} (-1)^m (m+1) z^m, \qquad \frac{1}{(1+z)^3} = \sum_{n=0}^{\infty} (-1)^m \frac{(m+1)(m+2)}{2} z^m.$$

Let $C_m^{\ell} = \frac{m!}{(m-\ell)! \ell!}$ denote the binomial coefficients for $\ell \leq m$. For $z \in]-1, 1[$ and $0 < q \leq 1$ with $q \neq 1/2$ we obtain, by ignoring terms of orders higher than three, that

$$\tau \partial_t \widetilde{U}^o = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \beta^m \widetilde{U}^o + \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} C_m^1 \beta^{m-1} (-a\tau \partial_x) \widetilde{U}^o$$

$$+ \left\{ \sum_{m=2}^{\infty} \frac{(-1)^{m+1}}{m} C_m^2 \beta^{m-2} a^2 \tau^2 \partial_x^2 + \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} C_m^1 \beta^{m-1} \frac{qh^2}{2} \partial_x^2 \right\} \widetilde{U}^o$$

$$+ \left\{ \sum_{m=3}^{\infty} \frac{(-1)^{m+1}}{m} C_m^3 \beta^{m-3} (-a\tau \partial_x)^3 \right.$$

$$+ \sum_{m=2}^{\infty} \frac{(-1)^{m+1}}{m} C_m^1 C_{m-1}^1 \beta^{m-2} (-a\tau \partial_x) \cdot \frac{qh^2}{2} \partial_x^2$$

$$\left. + \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} C_m^1 \beta^{m-1} \left(-\frac{a\tau h^2}{6} \partial_x^3 \right) \right\} \widetilde{U}^o + \cdots$$

$$= \ln |\beta + 1| \widetilde{U}^o - \frac{a\tau}{1+\beta} \partial_x \widetilde{U}^o + \left\{ -\frac{a^2 \tau^2}{2(1+\beta)^2} + \frac{qh^2}{2(1+\beta)} \right\} \partial_x^2 \widetilde{U}^o$$

$$+ \left\{ \frac{-a^3 \tau^3}{3(1+\beta)^3} + \frac{a\tau}{(1+\beta)^2} \frac{qh^2}{2} - \frac{1}{1+\beta} \cdot \frac{a\tau h^2}{6} \right\} \partial_x^3 \widetilde{U}^o + \cdots$$

$$= \ln |2q - 1| \widetilde{U}^o - \frac{a\tau}{2q-1} \partial_x \widetilde{U}^o + \frac{1}{2} \frac{[q(2q-1)h^2 - a^2 \tau^2]}{(2q-1)^2} \partial_x^2 \widetilde{U}^o$$

$$+ \frac{a\tau}{6} \frac{[(q+1)(2q-1)h^2 - 2a^2 \tau^2]}{(2q-1)^3} \partial_x^3 \widetilde{U}^o + \cdots.$$

Thus we derive the modified equation for the oscillatory part

$$\partial_t \widetilde{U}^o + \frac{a}{2q-1} \partial_x \widetilde{U}^o$$

$$= \frac{\ln |2q - 1|}{\tau} \widetilde{U}^o + \frac{h^2}{2\tau} \frac{[q(2q-1) - \nu^2 a^2]}{(2q-1)^2} \partial_x^2 \widetilde{U}^o$$

$$+ \frac{ah^2}{6} \frac{[(q+1)(2q-1) - 2\nu^2 a^2]}{(2q-1)^3} \partial_x^3 \widetilde{U}^o + \cdots.$$

We introduce for $q \neq 1/2$ a rescaling that becomes singular for $q = 1/2$, namely we set $x' = x(2q - 1)$ and omit the use of a primed variable. This gives,

$$\partial_t \widetilde{U}^o + a\partial_x \widetilde{U}^o$$

(5.18)
$$= \frac{\ln|2q-1|}{\tau} \widetilde{U}^o + \frac{h^2}{2\tau}[q(2q-1) - \nu^2 a^2]\partial_x^2 \widetilde{U}^o$$

$$+ \frac{ah^2}{6}[(q+1)(2q-1) - 2\nu^2 a^2]\partial_x^3 \widetilde{U}^o + \cdots .$$

Unlike the modified equation (5.15) for the low frequency modes the numerical dissipation comes from two terms: zero order term $\frac{\ln|2q-1|}{\tau}\widetilde{U}^o$ and the second order term $\frac{a^2\tau}{2\nu^2 a^2}[q(2q-1) - \nu^2 a^2]\partial_x^2 \widetilde{U}^o$. The former exerts more dominant dissipation than the latter, which can be explained from the following well-known fact that linear source terms decay exponentially in time, whereas second order diffusive terms as in the heat equation have a much lower algebraic decay. Consider

(5.19)
$$\begin{cases} v_t = \alpha v + \mu v_{xx}, \quad \alpha < 0, \quad \mu > 0, \\ v(x, 0) = v_0(x). \end{cases}$$

The solution expression is

(5.20)
$$v(x, t) = \frac{e^{\alpha t}}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} v_0(y) e^{-\frac{(x-y)^2}{4\mu t}} \, dy.$$

From this solution we clearly see the decay property in time.

We call the zero order term in (5.18) $\frac{\ln|2q-1|}{\tau}\widetilde{U}^o$ a *numerical damping* term and the second order term $\frac{h^2}{2\tau}[q(2q-1) - \nu^2 a^2]\partial_x^2 \widetilde{U}^o$ a *numerical viscosity*. They play distinct dissipation roles in controlling the amplitude of high frequency modes. Furthermore, two more remarks are in order.

(i) For the LxF scheme with $q = 1$ the modified equation for the perturbation part $\widetilde{U}^o$ is

(5.21) $$\partial_t \widetilde{U}^o + a\partial_x \widetilde{U}^o = \frac{a^2\tau}{2\nu^2 a^2}[1 - \nu^2 a^2]\partial_x^2 \widetilde{U}^o + \frac{a^3\tau^2}{3\nu^2 a^2}[1 - \nu^2 a^2]\partial_x^3 \widetilde{U}^o + \cdots .$$

Although this part is dissipated through the numerical viscosity term if $|\nu a| < 1$, this dissipation is still weak in comparison with the numerical damping term of the form $\frac{\ln(|2q-1|)}{\tau}\widetilde{U}^o$. In particular, $\widetilde{U}^0 \equiv 1$ is a solution of (5.21) if the constant unity initial data are given, which implies that the chequerboard mode is not perturbed and therefore not damped at all. This explains why the oscillations in the LxF scheme are observed. This was already highlighted above through the discrete Fourier analysis.

(ii) As $0 < q < 1$, the strong damping term $\frac{\ln(|2q-1|)}{\tau}\widetilde{U}^o$ suppresses the oscillations well no matter how the viscosity term behaves. In particular, if $q = 1/2$, the oscillation is damped out immediately, by noting that

(5.22)
$$\lim_{q \to 1/2+0} \ln(|2q - 1|) = -\infty.$$

So the damping becomes infinite for $q = 1/2$. This is consistent with two previous observations. The first is that in the solution (1.4) to the chequerboard initial data (1.3) the amplitude vanishes in this case. The second is the computational evidence shown in Figure 4.1(d).

The high frequency modes are dissipated very quickly if $0 < q < 1$, unlike the LxF scheme with $q = 1$ or the unstable central scheme with $q = 0$, even through there is a chequerboard mode initially when we have $\widetilde{U}^0 = 1$ as initial data for (5.18). The solution decays algebraically, i.e.,

$$(5.23) \qquad |2q - 1|^{\frac{1}{\tau}} \widetilde{U^o} = \mathcal{O}(1).$$

Therefore, the chequerboard mode is damped algebraically. This explains why the oscillations are less visible for $0 < q < 1$ than those in the standard LxF scheme for $q = 1$ or the unstable central scheme for $q = 0$.

6. Modified equation analysis for nonlinear cases

Now we consider monotone schemes for nonlinear cases of (1.1) and restrict to the 3-point conservative schemes. Thus the schemes take the form

$$(6.1) \qquad u_j^{n+1} = u_j^n - \frac{\tau}{h}\big(\hat{f}(u_{j+1}^n, u_j^n) - \hat{f}(u_j^n, u_{j-1}^n)\big),$$

where the numerical flux $\hat{f}(v, w)$ satisfies the consistent condition, $\hat{f}(u, u) = f(u)$. In [3], it was shown that for smooth solutions, the modified equation of (6.1) is

$$(6.2) \qquad \partial_t u + f(u)_x = \tau\big(d(u, \tau/h)u_x\big)_x + \mathcal{O}(\tau^2),$$

where we have the diffusion coefficient $d(u, \tau/h) \geq 0$. For simplicity, we continue to discuss the generalized LxF schemes

$$(6.3) \qquad u_j^{n+1} = u_j^n - \frac{\tau}{2h}\big(f(u_{j+1}^n) - f(u_{j-1}^n)\big) + \frac{q}{2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n).$$

These schemes are monotone if $\nu \cdot \max_j\{|f'(u_j^n)|\} \leq q \leq 1$.

In order to understand the resolution of high frequency modes by (6.3), we use the same approach as in Section 5. Only consider high frequency data

$$u_j^n = (-1)^{j+n}(U^o)_j^n,$$

and investigate the evolution by (6.3). Then the oscillatory part $(U^o)_j^n$ satisfies,

$$(U^o)_j^{n+1} = (U^o)_j^n - \frac{(-1)^{j+n+1}\tau}{2h}\big[f\big((-1)^{j+n+1}(U^o)_{j+1}^n\big) - f\big((-1)^{j+n-1}(U^o)_{j-1}^n\big)\big]$$

$$(6.4) \qquad + \frac{q}{2}[(U^o)_{j+1}^n - 2(U^o)_j^n + (U^o)_{j-1}^n] + 2(q-1)(U^o)_j^n.$$

Note that $(-1)^{j+n+1}(U^o)_j^n = (-1)^{j+n-1}(U^o)_j^n$ and so

$$f\big((-1)^{j+n+1}(U^o)_{j+1}^n\big) - f\big((-1)^{j+n-1}(U^o)_{j-1}^n\big) = f\big((-1)^{j+n+1}(U^o)_{j+1}^n\big)$$

$$(6.5) \qquad\qquad - f\big((-1)^{j+n+1}(U^o)_j^n\big) + f\big((-1)^{j+n-1}(U^o)_j^n\big)$$

$$\qquad\qquad - f\big((-1)^{j+n-1}(U^o)_{j-1}^n\big).$$

Then we obtain for U^o, by taking the Taylor expansion at $(-1)^{j+n+1}(U^o)_j^n$ (resp. $(-1)^{j+n-1}(U^o)_j^n$),

$$(6.6) \qquad (e^{\tau\partial_t} - 1)\widetilde{U}^o = 2(q-1)\widetilde{U}^o - \tau f'(-\widetilde{U}^o)\partial_x\widetilde{U}^o + \frac{qh^2}{2}\partial_{xx}\widetilde{U}^o + \mathcal{O}(h^3),$$

or

$$(6.7) \qquad (\partial_t + \frac{\tau}{2}\partial_t^2 + \cdots)\widetilde{U}^o = 2(q-1)\widetilde{U}^o - f'(-\widetilde{U}^o)\partial_x\widetilde{U}^o + \frac{qh^2}{2\tau}\partial_{xx}\widetilde{U}^o + \mathcal{O}(h^2).$$

We only concentrate on the dissipation effect of the scheme for high frequency modes and therefore the terms of orders higher than two will be ignored. Thus we compute formally

$$
\begin{aligned}
\tau\partial_t\widetilde{U}^o &= \sum_{m=1}^{\infty}\frac{(-1)^{m+1}}{m}\left[2(q-1)-\tau f'(-\widetilde{U}^o)\partial_x+\frac{qh^2}{2}\partial_{xx}\right]^m\widetilde{U}^o+\mathcal{O}(h^3) \\
&= \sum_{m=1}^{\infty}\frac{(-1)^{m+1}}{m}(2(q-1))^m\widetilde{U}^o \\
&\quad -\sum_{m=1}^{\infty}\frac{(-1)^{m+1}}{m}C_m^1(2(q-1))^{m-1}\tau f'(-\widetilde{U}^o)\partial_x\widetilde{U}^o \\
&\quad +\sum_{m=2}^{\infty}\frac{(-1)^{m+1}}{m}C_m^2(2(q-1))^{m-2}[-\tau f'(-\widetilde{U}^o)\partial_x]^2\widetilde{U}^o \\
&\quad +\sum_{m=1}^{\infty}\frac{(-1)^{m+1}}{m}C_m^1(2(q-1))^{m-1}\cdot\frac{qh^2}{2}\partial_{xx}\widetilde{U}^o+\mathcal{O}(h^3) \\
&= \ln|2q-1|\widetilde{U}^o-\sum_{m=0}^{\infty}(-1)^m[2(q-1)]^m\cdot\tau f'(-\widetilde{U}^o)\partial_x\widetilde{U}^o \\
&\quad -\frac{1}{2}\sum_{m=0}^{\infty}(-1)^m(m+1)[2(q-1)]^m\cdot\tau^2 f'(-\widetilde{U}^o)\partial_x(f'(-\widetilde{U}^o)\partial_x\widetilde{U}^o) \\
&\quad +\sum_{m=0}^{\infty}(-1)^m[2(q-1)]^m\cdot\frac{qh^2}{2}\partial_{xx}\widetilde{U}^o+\mathcal{O}(h^3) \\
&= \ln|2q-1|\,\widetilde{U}^o-\frac{\tau}{2q-1}f'(-\widetilde{U}^o)\partial_x\widetilde{U}^o \\
&\quad -\frac{\tau^2 f'(\widetilde{U}^o)}{2(2q-1)^2}[-f''(-\widetilde{U}^o)(\partial_x\widetilde{U}^o)^2+f'(-\widetilde{U}^o)\partial_x^2\widetilde{U}^o] \\
(6.8) &\quad +\frac{qh^2}{2(2q-1)}\partial_x^2\widetilde{U}^o+\mathcal{O}(h^3).
\end{aligned}
$$

Then the modified equation of (6.3) for U^o is

$$
\begin{aligned}
\widetilde{U}_t^o+\frac{f'(-\widetilde{U}^o)}{2q-1}\widetilde{U}_x^o &= \frac{\ln(|2q-1|)}{\tau}\widetilde{U}^o+\frac{\tau}{2(\tau/h)^2}\frac{[q(2q-1)-\frac{\tau^2}{h^2}(f'(-\widetilde{U}^o))^2]}{(2q-1)^2}\widetilde{U}_{xx}^o \\
(6.9) &\quad +\frac{\tau f'(-\widetilde{U}^o)f''(-\widetilde{U}^o)}{2(2q-1)^2}(\widetilde{U}_x^o)^2+\mathcal{O}(h^2).
\end{aligned}
$$

Compared to the linear case we discussed in the last section, there is a numerical damping term to control the amplitude of the high frequency modes.

Remark 6.1. In the manipulation of (6.8) we formally use (6.6) to calculate the term $f'(-\widetilde{U}^o)\partial_x$ which finally leads to the modified equation (6.9). Indeed, in order to avoid this formalism, we can use (6.7) to make differentiation successively, instead of (6.6).

7. Discussion and conclusion

In the present paper, we are discussing the local oscillations in the particular generalized LxF scheme. This discussion has general implications into more extensive (monotone) schemes and multidimensional cases. Indeed, the generalized LxF scheme is monotone and thus TVD under a certain restriction. The TVD property is proposed to describe the global property of solutions of (1.1). The phenomenon of oscillations we are investigating is *local* and does not contradict this *global* TVD property. As far as hyperbolic problems are concerned, local properties should be given more attention because of the finite propagation of waves.

We have individually analyzed the resolution of the low and high frequency modes $u_j^n = \lambda_k^n e^{ij\xi}$, $\xi = 2\pi kh$ in numerical solutions. Our approach is the discrete Fourier analysis and the modified equation analysis, which are applied to investigating the numerical dissipative and dispersive mechanisms as well as relative phase errors. We summarize as follows.

1. Relative phase error. For the low frequency modes, the error is of order $\mathcal{O}(\xi^2)$, while for high frequency modes the error is of order $\mathcal{O}(1)$ after each time step, which is generally independent of the parameter q.

2. Numerical dissipation. For the low frequency modes, the dissipation is usually of order $\mathcal{O}(\xi)$ for the scheme (4.1), which closely depends on the parameter q. As $q = \nu^2 a^2$, (4.1) becomes the LW scheme and it has the amplitude error $\mathcal{O}(\xi^2)$. For high frequency modes, the scheme usually has the numerical damping of order $\mathcal{O}(1)$ that becomes stronger as q is closer to $1/2$, unless it vanishes for the limit case ($q = 1$ or 0), in which the amplitude is dissipated via the numerical viscosity of second order.

Thus we conclude that *the relative phase errors should be at least offset by the numerical dissipation of the same order. Otherwise the oscillation could be caused.* In the second order accurate LW scheme the oscillations are caused by the relative phase error of low frequency modes, while in the first order LxF schemes, oscillations are caused by the relative phase error of high frequency modes. In order to control the oscillations by high frequency modes, the strong numerical damping (zero order term) is necessary to add.

The presence of high frequency modes results from the initial/boundary conditions. In Section 3, we show that the discretization may produce the chequerboard modes $(-1)^{j+n}$. As discussed in [7], the classical box scheme has serious difficulty in controlling such modes, and a weighted time-average technique is used to cure it. Such a technique is indeed used to introduce an *artificial numerical damping*, without sacrificing the accuracy. Compared to the LxF scheme ($q = 1$), the scheme (4.1) ($0 < q < 1$) introduces the numerical damping as well, which is stronger as q is closer to $1/2$. Then (4.1) becomes the modified LxF scheme [8]. Hence, in order to control the oscillations caused by high frequency modes, *the numerical damping plays an important role.*

Acknowledgement

References

[1] M. Breuss, The correct use of the Lax-Friedrichs method, *M2AN Math. Model. Numer. Anal.* 38(2004), 519–540. MR2075758 (2006e:65137)

[2] M. Breuss, An analysis of the influence of data extrema on some first and second order central approximations of hyperbolic conservation laws, *M2AN Math. Model. Numer. Anal.*, 39(2005), 965–993. MR2178569 (2007f:35183)

[3] A. Harten, J. M. Hyman and P. D. Lax, On finite-difference approximations and entropy conditions for shocks, *Comm. Pure Appl. Math.*, 29(1976), 297–322. MR0413526 (54:1640)

[4] P. Lax and B. Wendroff, Systems of conservation laws, *Comm. Pure Appl. Math.*, 13(1960), 217–237. MR0120774 (22:11523)

[5] P.D. LeFloch and J.-G. Liu, Generalized monotone schemes, discrete and paths of extrema, and discrete entropy conditions, *Math. Comp.*, 68(1999), 1025-1055. MR1627801 (99i:65092)

[6] R. LeVeque, *Numerical Methods for Conservation Laws*, Lectures in Mathematics ETH Zurich, 2nd ed. Basel, Birkhäuser Verlag, 1992. MR1153252 (92m:65106)

[7] K.W. Morton and D.F. Mayers, *Numerical Solution of Partial Differential Equations*, 2nd edition, Cambridge University Press, 2005. MR2153063 (2006a:65003)

[8] E. Tadmor, Numerical viscosity and the entropy condition for conservative difference schemes. *Math. Comp.*, 43(1984), 369–381. MR758189 (86g:65163)

[9] H.-Z. Tang and G. Warnecke, A note on $(2K + 1)$-point conservative monotone schemes, *M2AN Math. Model. Numer. Anal.*, 38(2004), 345–357. MR2069150 (2005f:65111)

[10] J.W. Thomas, *Numerical Partial Differential Equations: Finite Difference Methods*, Springer-Verlag, 1995. MR1367964 (97a:65001)

[11] R. F. Warming and B. J. Hyett, The modified equation approach to the stability and accuracy of finite difference methods, *J. Comput. Phys.*, 14(1974), 159–179. MR0339526 (49:4284)

School of Mathematics, Capital Normal University, Beijing 100037, People's Republic of China
E-mail address: jiequan@mail.cnu.edu.cn

LMAM, School of Mathematical Sciences, Peking University, Beijing 100871, People's Republic of China
E-mail address: hztang@math.pku.edu.cn

Institut für Analysis und Numerik, Otto-von-Guericke-Universität, PSF 4120, 39016 Magdeburg, F.R. Germany
E-mail address: warnecke@ovgu.de

The high school attached to the Central University for Nationalities, Beijing 100081, People's Republic of China
E-mail address: zhanglumei99@163.com