

AN ANALYSIS OF THE INFLUENCE OF DATA EXTREMA ON SOME FIRST AND SECOND ORDER CENTRAL APPROXIMATIONS OF HYPERBOLIC CONSERVATION LAWS

MICHAEL BREUSS¹

Abstract. We discuss the occurrence of oscillations when using central schemes of the Lax-Friedrichs type (LFt), Rusanov's method and the staggered and non-staggered second order Nessyahu-Tadmor (NT) schemes. Although these schemes are monotone or TVD, respectively, oscillations may be introduced at local data extrema. The dependence of oscillatory properties on the numerical viscosity coefficient is investigated rigorously for the LFt schemes, illuminating also the properties of Rusanov's method. It turns out, that schemes with a large viscosity coefficient are prone to oscillations at data extrema. For all LFt schemes except for the classical Lax-Friedrichs method, occurring oscillations are damped in the course of a computation. This damping effect also holds for Rusanov's method. Concerning the NT schemes, the non-staggered version may yield oscillatory results, while it can be shown rigorously that the staggered NT scheme does not produce oscillations when using the classical minmod-limiter under a restriction on the time step size. Note that this restriction is not the same as the condition ensuring the TVD property. Numerical investigations of one-dimensional scalar problems and of the system of shallow water equations in two dimensions with respect to the phenomenon complete the paper.

Mathematics Subject Classification. 35L65, 65M06, 65M12.

Received: November 4, 2004. Revised: March 25, 2005.

INTRODUCTION

The phenomenon we discuss within this paper is concerned with oscillations appearing at local data extrema in the numerical simulation of hyperbolic conservation laws when using central methods. The aim of this paper is on the one hand the detailed investigation of the influence of the coefficients of numerical viscosity defining especially the Lax-Friedrichs type (LFt) central schemes on the occurrence of this phenomenon. The properties found for these methods also shed light on the corresponding properties of Rusanov's method. On the other hand, we clarify the role of the formulation of methods using staggered or non-staggered grids, respectively, with respect to the subject of our investigation. This is especially important for the construction of higher-order central schemes for which the Nessyahu-Tadmor (NT) schemes we investigate are the classical prototype.

Keywords and phrases. Conservation laws, numerical methods, finite difference methods, central schemes.

¹ Technische Universität Braunschweig, Computational Mathematics, Pockelsstraße 14, 38106 Braunschweig, Germany.

This work was done during a research stay at Mathématiques Appliquées de Bordeaux, Université Bordeaux 1, 351, Cours de la Libération, 33405 Talence Cedex, France. The author would like to thank the Deutsche Forschungsgemeinschaft (DFG) for supporting this stay under grant No. BR 2245/1-1.

© EDP Sciences, SMAI 2005

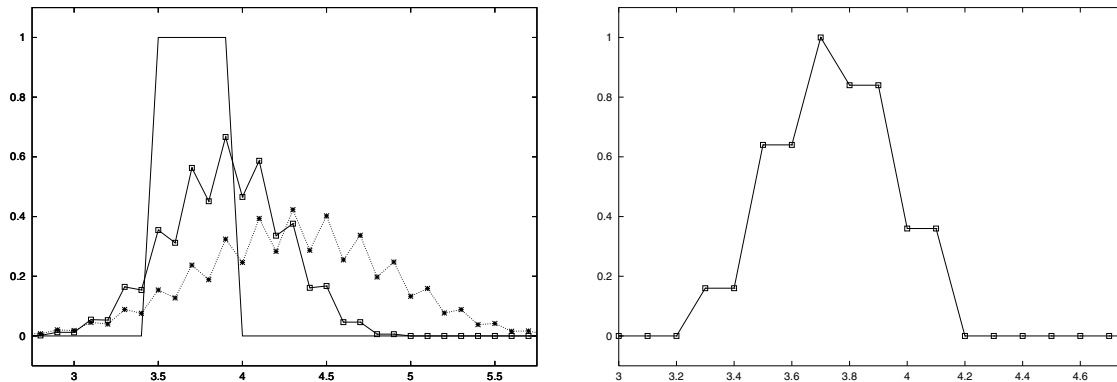


FIGURE 1. This figure shows (a) on the left hand side the initial condition (line) together with the numerical solutions obtained using the Lax-Friedrichs scheme after 10 (line with boxes) and 30 (dotted line with stars) time steps, respectively. On the right hand side (b), we zoom on the situation directly after the first time step using the initial condition displayed in (a). A strict local maximum is clearly observable.

Central schemes as those discussed within this paper play an important role as tools for the numerical approximation of conservation laws. They are usually regarded as being very robust as well as simple to implement in comparison to the Godunov type schemes since no particular knowledge about the structure of solutions of Riemann problems is required for the algorithmical definition of the methods. Also, no computation of the Jacobian of the fluxes is usually incorporated within these methods, yielding quite efficient algorithms. For useful descriptions of central schemes, see *e.g.* [6, 13–15] and the references therein.

All of the central schemes we consider in this paper – LfT schemes, Rusanov’s method and the NT schemes – are TVD under a CFL condition and thus monotonicity preserving. The monotonicity preserving property means, that a monotone data sequence is mapped to a new monotone data sequence by the method; however, concerning *e.g.* the situation at a local extremum in 1-D with monotone data profiles to the left and right of it, this property yields no assertion about the behaviour of the method exactly at the intermediate point between the monotone parts of the data. Generally, such data extrema may arise in any simulation where the data assume a non-monotone profile, and especially if the initial data are not restricted to a Riemann problem. For example, at some point during a simulation only a peak may be left of an extremum formerly consisting of several discretization points. However, as is pointed out in [1] for the case of the Lax-Friedrichs method, such a peak can cause severe oscillations in the numerical solution. A notably simple but devastating example can be given by discretizing an initial square signal transported *via* the linear advection equation $u_t + u_x = 0$. Under a usual CFL condition, the solution obtained by using the Lax-Friedrichs method shows after a few time steps the features given in Figure 1a. The investigations in [1] prove, that the data extremum displayed in Figure 1b which occurs in the course of the computation is indeed the reason for the oscillations. The question arises – for which we seek to give an answer within this paper – to what extent this behaviour carries over to other central schemes.

With respect to the monotonicity property of the LfT schemes and Rusanov’s method, we note that the monotonicity of a scheme does not prevent the occurrence of new extrema as is pointed out in [10], see also the pioneering work of Harten [5] on TVD schemes in this context: there, monotone schemes with the property that the number of extrema is preserved are classified as generalized monotone schemes. It was shown in [10] that the LfT schemes are not generalized monotone schemes for coefficients of numerical viscosity $q \in (2/3, 1]$, however, the authors showed this by mentioning one trivial example (they set the flux function to zero) without a further analysis of the underlying phenomenon and without considering more general situations for central schemes: their work is mainly dedicated to higher order reconstructions for use within the MUSCL approach.

Oscillations in monotone schemes are also reported in [17], taking into account schemes with $(2k + 1)$ -point stencils.

In [1], a detailed analysis of oscillatory states within the numerical solutions obtained by using the classical Lax-Friedrichs method is given. The investigations in [1] seemed to hint that the decoupling of nodes which takes place within the classical Lax-Friedrichs scheme is responsible for the behaviour of it at local extrema. Within this paper, we show that the occurrence of oscillatory states is a more general feature of central schemes, even when the odd-even decoupling of nodes is not an issue within the definition of the method. On the other hand, we also show that the coupling effect of nodes within a scheme is advantageous in order to prevent or damp oscillations. Let us emphasize that this property is important for a useful definition of higher order central schemes. The investigations within this paper show, that one needs to be very careful when constructing a central scheme, especially on a non-staggered grid. One may understand some parts of our investigations as a justification of the approach presented in [7]: there, non-staggered central schemes are derived by integration using staggered grids followed by an averaging projection onto the initial non-staggered cells.

We show rigorously for Lax-Friedrichs type (LFt) methods, that the properties of the corresponding operator strongly depend on the choice of the viscosity coefficients; surprisingly, schemes with relatively large viscosity are prone to oscillations. In [1] it is shown that the classical Lax-Friedrichs method [9] may conserve the local variation at data extrema exactly, explaining the occurrence of oscillations as well as their stability once they appear. In this paper, we prove that the Lax-Friedrichs scheme is the only LFt scheme with this property. LFt schemes with lesser viscosity coefficients may also produce oscillations, as remarked in [10] for the trivial example of a zero flux; however, the variation of such oscillations generally decreases in time. Moreover, we show that all LFt schemes with a coefficient of numerical viscosity q with $q \leq 1/2$ do not generate new extrema under fairly general conditions. Note that the LFt methods play an important role as relaxed schemes within relaxation methods [8], and thus they not only represent the most simple generalizations of the classical Lax-Friedrichs scheme but are of independent interest. The results obtained for the LFt schemes can be used to understand corresponding properties of Rusanov's method [12]. We then proceed with a discussion of the staggered and of the non-staggered Nessyahu-Tadmor (NT) schemes given in [15]. The consideration of these schemes is important since many higher-order central schemes were derived based on the developments in [15]. It turns out that the staggered grid version is much more robust concerning the investigated effect than the non-staggered scheme. We show rigorously, that the non-staggered scheme does not yield oscillatory states under a restriction on the time step size. Notably, this restriction is more severe than the restriction found in [15] ensuring the TVD property; however, in parallel to [15] the restriction has the meaning of a sufficient condition and not of a necessary one. Various numerical tests confirm the theoretical assertions.

The paper is organized according to the outlined content. Within the first section we study the mentioned schemes from a theoretical point of view. The second section is devoted to numerical tests, and we finish the paper with conclusive remarks.

1. THEORETICAL INVESTIGATIONS

As indicated, we investigate within this section the LFt schemes in detail, while Rusanov's method is briefly discussed using the obtained results. After that, the NT schemes are the subject of our interest, using a slightly different approach as for the investigations before. All discussions within this section are concerned with scalar 1-D problems.

1.1. LFt schemes

The LFt schemes we study are of the form

$$U_j^{n+1} = U_j^n + \frac{q}{2} (U_{j+1}^n - 2U_j^n + U_{j-1}^n) - \frac{1}{2}\lambda [f(U_{j+1}^n) - f(U_{j-1}^n)]. \quad (1)$$

In (1), q with $0 < q \leq 1$ denotes the coefficient of numerical viscosity, see *e.g.* [3] for further explanations concerning this notion. By λ we denote as usual the ratio of grid parameters $\Delta t/\Delta x$. For simplicity, we omit the superscript n denoting the time level $n\Delta t$ when appropriate, and in that case we refer to U_j^{n+1} as U_j^{new} .

Obviously, for $q = 1$ we retrieve the classical Lax-Friedrichs scheme. Another important choice is $q = 1/2$ yielding the modified Lax-Friedrichs scheme [3, 16]. Setting $q = 0$ would result in a well-known unstable central scheme, and thus this value is excluded completely within our investigations.

In order to guarantee the stability of the LFT methods, one has to impose the validity of a numerical viscosity dependent CFL condition. It is useful to introduce this condition using the CFL number $CFL \in (0, 1)$, where we exclude $CFL = 1$ since it involves the additional consideration of trivial special cases within the technical details without giving a substantial contribution. Taking into account the relevant finite data range \mathcal{I} , the CFL condition then reads

$$\lambda |f'(u)| < CFL \cdot q \quad \forall u \in \mathcal{I}. \tag{2}$$

Note that the validity of the CFL condition (2) also implies the validity of the monotonicity property of the LFT schemes.

Before we speak about oscillations, we want to define clearly what is meant. Thus, we give in the same fashion as in [1] the following definition:

Definition 1.1. Let a data sequence $\{U_l, U_m, U_r\}$ be given. We define the state U_m^{new} obtained by application of the scheme under consideration as an oscillation if one of the two assertions holds:

$$\begin{aligned} &U_l \leq U_m \geq U_r \text{ and } U_l^{new} > U_m^{new} < U_r^{new} \text{ is true,} \\ \text{or } &U_l \geq U_m \leq U_r \text{ and } U_l^{new} < U_m^{new} > U_r^{new} \text{ is true.} \end{aligned}$$

It is useful to consider, for the beginning, a single data maximum, *i.e.*, the given data

$$\dots, U_{ll}, U_l, U_m, U_r, U_{rr}, \dots$$

feature the structure

$$\dots = U_{ll} = U_l, U_l < U_m > U_r \text{ and } U_r = U_{rr} = \dots \tag{3}$$

Following the Definition 1.1 these data lead to an oscillation if

$$U_l^{new} - U_m^{new} > 0 \quad \text{and} \quad U_r^{new} - U_m^{new} > 0. \tag{4}$$

For the situation (3), we now compute in a straightforward fashion the values corresponding to these conditions for the LFT schemes (1):

$$\begin{aligned} U_l^{new} - U_m^{new} &= \frac{q}{2} (U_l + U_m) - \frac{\lambda}{2} (f(U_m) - f(U_l)) + (1 - q) U_l \\ &\quad - \frac{q}{2} (U_l + U_r) + \frac{\lambda}{2} (f(U_r) - f(U_l)) - (1 - q) U_m \\ &= \frac{q}{2} (U_m - U_r) - \frac{\lambda}{2} (f(U_m) - f(U_r)) + (1 - q) (U_l - U_m), \end{aligned} \tag{5}$$

$$\begin{aligned} U_r^{new} - U_m^{new} &= \frac{q}{2} (U_m + U_r) - \frac{\lambda}{2} (f(U_r) - f(U_m)) + (1 - q) U_r \\ &\quad - \frac{q}{2} (U_l + U_r) + \frac{\lambda}{2} (f(U_r) - f(U_l)) - (1 - q) U_m \\ &= \frac{q}{2} (U_m - U_l) + \frac{\lambda}{2} (f(U_m) - f(U_l)) + (1 - q) (U_r - U_m). \end{aligned} \tag{6}$$

Thus, we can derive by (4)–(6) after some further elementary calculations the oscillation conditions

$$q - \lambda \frac{f(U_m) - f(U_r)}{U_m - U_r} > 2(1 - q) \frac{U_m - U_l}{U_m - U_r} \tag{7}$$

and

$$q + \lambda \frac{f(U_m) - f(U_l)}{U_m - U_l} > 2(1 - q) \frac{U_m - U_r}{U_m - U_l}. \tag{8}$$

Remark 1.2.

- (1) The left hand sides of (7) and (8) are positive by the CFL condition (2).
- (2) The data ratios on the right hand sides of (7) and (8) are always positive in the situation discussed since U_m denotes a local extremum.
- (3) For $q = 1$, *i.e.*, for the classical Lax-Friedrichs scheme, the oscillation conditions (7) and (8) are always fulfilled, independently of the actual time step size, the specific flux function and the data ratio. The reason is that $q = 1$ yields the right hand sides of (7) and (8) as being zero, and by the CFL condition (2) the left hand sides are always positive.
- (4) Considering the right hand sides of (7) and (8), we observe that the value $q = 1/2$ corresponding to the modified Lax-Friedrichs scheme seems to be an important value because for $q > 1/2$ the right hand sides become smaller than the occurring data ratios, while for $q < 1/2$ the value of the data ratios is amplified.

Heuristically, this means that the modified Lax-Friedrichs scheme may be the optimal compromise avoiding oscillations on the one hand and enabling a useful time step size on the other hand, see (2).

Note that the third remark slightly extends the properties of the classical Lax-Friedrichs scheme found in [1] since there the monotonicity of the flux is assumed.

There are some further qualitative points to mention which indicate which behaviour we may expect in numerical computations. For the corresponding discussion we keep a fixed spatial mesh width Δx , in corresponding computations we always use $\Delta x = 0.1$.

For instance, let us assume $f'(\cdot) > 0$ for all values within the relevant data range, and let us discuss the situation

$$\frac{U_m - U_r}{U_m - U_l} > 1, \text{ i.e., } \frac{U_m - U_l}{U_m - U_r} < 1. \tag{9}$$

These conditions are equivalent to $U_l > U_r$. By (9), the right hand side in (7) tends to be small, while the right hand side in (8) tends to be large. Taking into account $f'(\cdot) > 0$ on the corresponding left hand sides, this means that the oscillation conditions (7) and (8) may be fulfilled in that case for relatively moderate values for q and Δt , respectively.

Elaborating on these thoughts *via* a simple example, we solve numerically the linear advection equation

$$u_t + u_x = 0, \tag{10}$$

choosing $q = 0.8$ and $\lambda = 0.4$ together with the data

$$U_l = 0.6, U_m = 1 \text{ and } U_r = 0.2 \tag{11}$$

for a situation as in (3). The initial condition as well as the numerical solution after one time step are shown in Figure 2. The oscillation at $x = 0$ is clearly visible.

If the data value structure is the opposite of (9), the right hand side of (7) tends to be large while the right hand side in (8) tends to be small. In this situation, the oscillatory condition (7) is harder to fulfill than (8), however, it may especially be fulfilled for relatively *large viscosities* q and *small time step sizes* Δt . We also find, that the same type of choice for these parameters generally leads to oscillations in the case just discussed

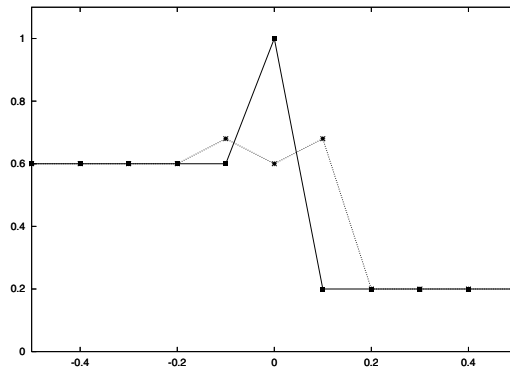


FIGURE 2. This figure shows the arising oscillation for $q = 0.8$, $\lambda = 0.4$ and the data in (11). We display the initial data (thick line) together with the numerical solution after one time step (dotted line).

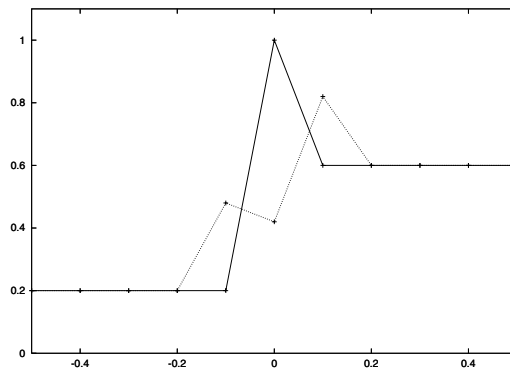


FIGURE 3. This figure shows the arising oscillation for $q = 0.9$, $\lambda = 0.2$ and the data in (12). Again we show the initial data (thick line) together with the numerical solution after one time step (dotted line).

before. This is in contrast to the usual intuition that much numerical viscosity and small time step sizes yield very stable, non-oscillating numerical solutions!

For a numerical experiment concerned with the latter case, we solve again the linear advection equation (10), this time with $q = 0.9$ and $\lambda = 0.2$, choosing the data just the other way round as in (11), *i.e.*,

$$U_l = 0.2, U_m = 1 \text{ and } U_r = 0.6. \tag{12}$$

The numerical result after one time step is displayed in Figure 3.

A further indication, that the stated thoughts are based on solid ground is given when considering $U_l \equiv U_r \equiv U$ in (7) and (8). Then both these conditions can be summarized *via*

$$\pm \lambda \frac{f(U_m) - f(U)}{U_m - U} > 2 - 3q.$$

Independently of the actual behaviour of f , at least one of the values on the left hand side is non-positive, and thus this inequality can only be fulfilled with a negative right hand side, yielding $q > 2/3$. Also, the

absolute value of the left hand side will be small for small Δt , while keeping Δx fixed. Thus, in the mentioned combination of large numerical viscosity and small time step sizes the oscillation conditions tend to be fulfilled.

We can make these heuristic considerations precise by proving the following assertion.

Lemma 1.3. *Consider the LFT schemes (1) under the fulfilled CFL condition (2) for an arbitrarily chosen but fixed Lipschitz continuous monotone flux f . For any given numerical viscosity $q \in (2/3, 1)$, one can find a time step size $\Delta\tau$ and data values U_l, U_m and U_r in a structure of data values as in (3), such that the scheme creates new extrema.*

Proof. For the beginning of the proof, let us note that $q = 1$ is not taken into account here since this choice corresponds to the Lax-Friedrichs methods dealt with in Theorem 1.4, see also [1] for more details concerning this scheme.

Concerning the monotonicity assumption, let us assume $f'(\cdot) \geq 0$ over the relevant range of data. The case $f'(\cdot) \leq 0$ can be treated analogously.

Because of $f'(\cdot) \geq 0$, a sufficient condition so that the oscillation condition (8) is fulfilled is

$$q > 2(1 - q) \frac{U_m - U_r}{U_m - U_l},$$

yielding directly the relation

$$\frac{q}{2(1 - q)} > \frac{U_m - U_r}{U_m - U_l}. \tag{13}$$

The terms on the left hand side of the other oscillation condition (7) can be estimated using (2) and a value $U^* \in \mathcal{I}$ by

$$q - \lambda \frac{f(U_m) - f(U_r)}{U_m - U_r} = q - \lambda f'(U^*) > q - CFL \cdot q.$$

Thus, we obtain after a few simple manipulations a sufficient condition for the validity of (7) as

$$CFL < 1 - \frac{U_m - U_l}{U_m - U_r} \frac{2(1 - q)}{q}. \tag{14}$$

Because of $CFL > 0$, we can extract from (14) the condition

$$\frac{U_m - U_l}{U_m - U_r} \frac{2(1 - q)}{q} < 1. \tag{15}$$

The relation (15) is then a sufficient condition to find a time step size $\Delta\tau$ such that an oscillation arises. (If (15) is fulfilled, then the right hand side in (14) takes on a value $\delta > 0$. We may set then *e.g.* $CFL := (1/2)\delta$, showing that (15) is indeed a sufficient condition as stated.) The inequality (15) is equivalent to

$$\frac{q}{2(1 - q)} > \frac{U_m - U_l}{U_m - U_r}. \tag{16}$$

Comparing (13) and (16), we realize, that the values on the right hand sides are, respectively, the reciprocal of each other. Putting the sufficient conditions for oscillatory states from those equations together, we obtain

$$\frac{q}{2(1 - q)} > \max \left(\frac{U_m - U_r}{U_m - U_l}, \frac{U_m - U_l}{U_m - U_r} \right). \tag{17}$$

Let us note again, that in (17) the maximum is taken from two positive values since U_m denotes a local extremum. Furthermore, since the arguments of the maximum function are reciprocal to each other, obviously

$$\max \left(\frac{U_m - U_r}{U_m - U_l}, \frac{U_m - U_l}{U_m - U_r} \right) \geq 1$$

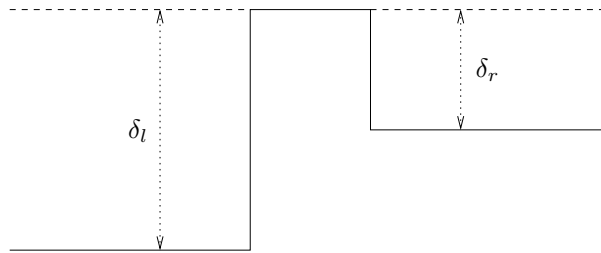


FIGURE 4. The figure shows a local maximum surrounded by constant states. The level of the maximum is represented by a dashed line. The computation of the total variation especially incorporates summing up the data differences – denoted here δ_l and δ_r – left and right of a local extremum.

holds. Since we are free to choose the values U_l, U_m and U_r in order to create an oscillation, we obtain the sought sufficient condition on the coefficient of numerical viscosity q for the possible occurrence of oscillations from (17):

$$\frac{q}{2(1-q)} > 1, \text{ i.e., } q > \frac{2}{3}. \quad \square$$

The meaning of the Lemma 1.3 is that LfT schemes with coefficients of numerical viscosity $q \in (2/3, 1)$ are not generalized monotone schemes in the sense of [10], even when considering slightly more general situations and not only $f \equiv 0$ as it was done in [10]: there, generalized monotone schemes are defined by being monotone as well as satisfying the number of extrema diminishing property. Moreover, we have shown by considering $q < 1$ that oscillations may appear independently from an odd-even-decoupling of nodes as it is the case within the Lax-Friedrichs scheme.

We can also prove the following result illuminating the nature of oscillations occurring.

Theorem 1.4. *For general Lipschitz continuous fluxes, the classical Lax-Friedrichs scheme (i.e., $q = 1$) is the only LfT scheme which conserves the local variation of the data at an arising oscillation exactly.*

Proof. For the proof we employ a variation of a procedure developed in [1]. We compute in a straightforward fashion the variation of the data in the case of an oscillation arising at a local maximum peak since this will be sufficient to prove the assertion of the theorem.

By Figure 4 we illustrate the procedure of the computation of the variation of the data directly at the local maximum, just before an oscillation takes place. The dashed line indicates the height of the local maximum.

This dashed line is taken over into the Figure 5 which shows the situation as the oscillation appears after the next time step when using data as displayed in Figure 4. The idea behind the computations within the proof is illustrated: we compute the quantities indicated by $\delta_i, i = 1, 2$, within Figure 5 and we seek to obtain their equality, respectively, in order to show the conservation of variation property.

Therefore, we again investigate a data sequence $\{U_{ll}, U_l, U_m, U_r, U_{rr}\}$ with $U_{ll} = U_l, U_l < U_m > U_r$ and $U_{rr} = U_r$ as in (3). For the computations which follow, we use the general form of LfT schemes (1) with an up to now unspecified coefficient of numerical viscosity q .

As observable by the Figures 4 and 5, for the exact conservation of the local variation we have to verify

$$U_m - U_l^{new} = U_r^{new} - U_m^{new} \quad \text{and} \quad U_m - U_r^{new} = U_l^{new} - U_m^{new}. \quad (18)$$

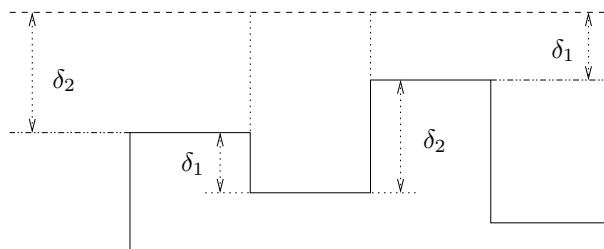


FIGURE 5. The figure sketches the principle concerning the conservation of the variation at an oscillation. We observe which lengths can be identified when considering the differences of the new data with respect to the level of the former local maximum indicated by the dashed line.

We now compute in a straightforward fashion these differences:

$$\begin{aligned} U_m - U_l^{new} &= U_m - \frac{q}{2}(U_m - U_l) + \frac{\lambda}{2}(f(U_m) - f(U_l)) \\ &= \left(1 - \frac{q}{2}\right)U_m - \left(1 - \frac{q}{2}\right)U_l + \frac{\lambda}{2}(f(U_m) - f(U_l)), \end{aligned} \tag{19}$$

$$\begin{aligned} U_m - U_r^{new} &= U_m - \frac{1}{2}(U_m - U_r) - \frac{\lambda}{2}(f(U_m) - f(U_r)) \\ &= \left(1 - \frac{q}{2}\right)U_m - \left(1 - \frac{q}{2}\right)U_r - \frac{\lambda}{2}(f(U_m) - f(U_r)), \end{aligned} \tag{20}$$

and from (5) and (6) we recall

$$U_l^{new} - U_m^{new} = \frac{q}{2}(U_m - U_r) - \frac{\lambda}{2}(f(U_m) - f(U_r)) + (1 - q)(U_l - U_m), \tag{21}$$

$$U_r^{new} - U_m^{new} = \frac{1}{2}(U_m - U_l) + \frac{\lambda}{2}(f(U_m) - f(U_l)) + (1 - q)(U_r - U_m). \tag{22}$$

As indicated, we obtain from (18)–(22) the equalities

$$\begin{aligned} \left(1 - \frac{q}{2}\right)U_m - \left(1 - \frac{q}{2}\right)U_l + \frac{\lambda}{2}(f(U_m) - f(U_l)) &= \frac{q}{2}(U_m - U_l) + \frac{\lambda}{2}(f(U_m) - f(U_l)) + (1 - q)(U_r - U_m) \\ &\Leftrightarrow \\ (1 - q)[U_r + U_l] &= (1 - q)[2U_m] \end{aligned} \tag{23}$$

and

$$\begin{aligned} \left(1 - \frac{q}{2}\right)U_m - \left(1 - \frac{q}{2}\right)U_r - \frac{\lambda}{2}(f(U_m) - f(U_r)) &= \frac{q}{2}(U_m - U_r) - \frac{\lambda}{2}(f(U_m) - f(U_r)) + (1 - q)(U_l - U_m) \\ &\Leftrightarrow \\ (1 - q)[U_r + U_l] &= (1 - q)[2U_m]. \end{aligned} \tag{24}$$

Obviously, (23) and (24) are identical. Moreover, since U_m is a strict local extremum, we cannot obtain

$$U_r + U_l = 2U_m,$$

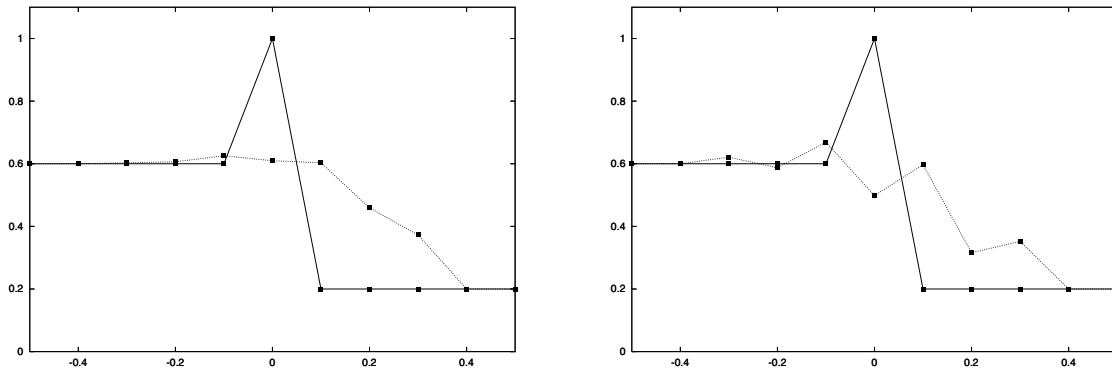


FIGURE 6. The figure contains the numerical solutions after three time steps together with the initial data employing (a) to the left $q = 0.8$ and $\lambda = 0.4$ and (b) to the right $q = 0.95$ and $\lambda = 0.2$. The initial data are marked by the thick lines.

and thus the only way the equalities (23) and (24) may hold is by $q = 1$, yielding the Lax-Friedrichs scheme as being in general the only LfT type scheme which conserves the local variation exactly. \square

Since the classical Lax-Friedrichs scheme conserves the local variation exactly at an arisen oscillation, oscillations once arisen do not simply vanish by application of this scheme. Thus, the Theorem 1.4 yields a very important consequence: even if a LfT scheme creates oscillations, one can hope when using a different scheme than the classical Lax-Friedrichs scheme, that these oscillations tend to vanish during the course of a computation simply due to the application of the method! This consideration is also important with respect to Rusanov’s method since there the coefficients of numerical viscosity are in general not equal to one.

We explore the evolution of oscillations by an example. We use again the data and parameters documented via (11) and Figure 2. Within Figure 6a, we observe that the oscillation shown in Figure 2 has nearly vanished after two subsequent time steps. However, this behaviour depends of course on the data and the computational parameters in use: for instance, choosing the same data but $q = 0.95$ and $\lambda = 0.2$, we observe in Figure 6b – which is analogous to Figure 6a – that the oscillation has spread.

By the snapshots displayed in Figure 7 taken for the same parameters as in the latter case, we observe in more detail the corresponding temporal evolution using once more the data from (11). Besides the recognizable spreading of oscillations, we also observe by the plot of the temporal evolution of the total variation shown in Figure 7d, that the total variation of the data decreases in time, as expected.

Note that we did not up to now take into account other profiles than the constant ones to the left and right from an extremum. We also note that we did not exclude other LfT schemes with viscosity coefficients less than $2/3$ from being prone to oscillations. Since these points are quite important, we give the following assertion.

Theorem 1.5. *Let a general Lipschitz continuous flux function f be given. Moreover, let us assume the validity of the following assumptions on the data value structure:*

- (1) *The data values are piecewise composed of non-strictly monotone profiles.*
- (2) *If there exist strict local extrema then these are isolated, i.e., if m denotes the index of a strict local extremum, there is no index in $\{m - 1, m + 1\}$ denoting another strict local extremum.*

Then the modified Lax-Friedrichs scheme ($q = 1/2$) as well as the LfT schemes with $q \in (0, 1/2)$ are generalized monotone schemes with respect to the next time step under the CFL condition (2).

Proof. Since the LfT schemes are monotone and thus monotonicity preserving under the CFL condition (2), all we need to discuss is the situation at extrema.

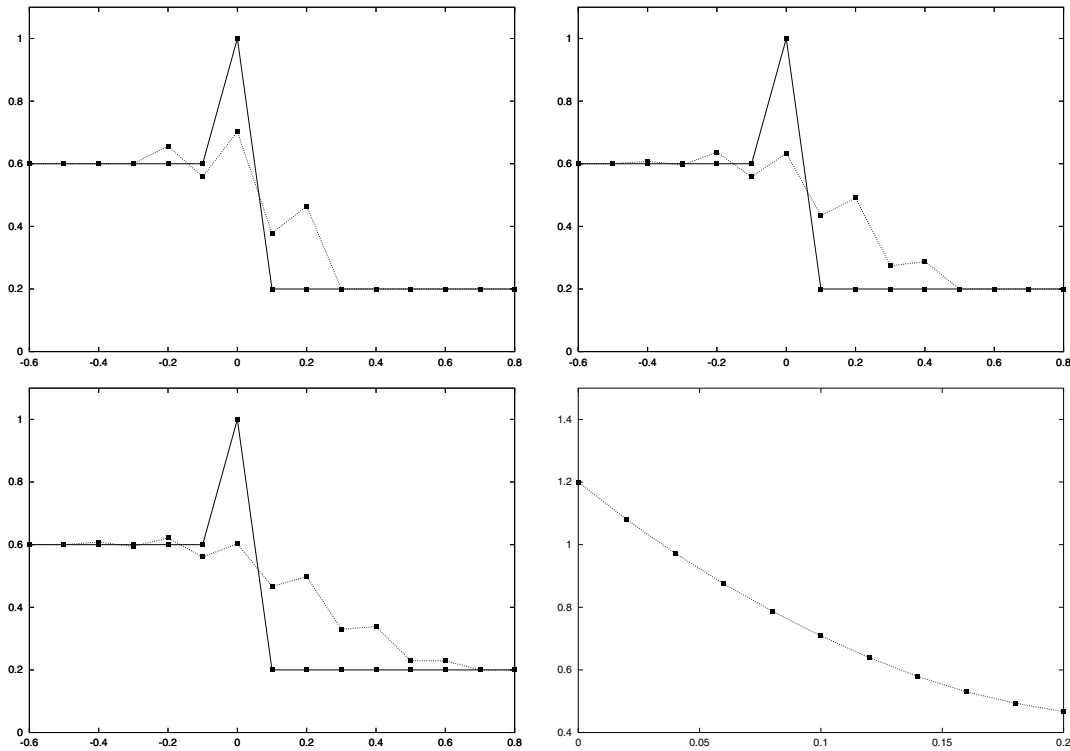


FIGURE 7. From left to right and top to bottom: the figure displays the evolution of the initial data after (a) 2, (b) 4 and (c) 6 time steps, together with (d) a plot of the evolution of the total variation (0.4 being the minimum here since this is exactly the size of the jump in the data from 0.6 to 0.2). The computational parameters are $q = 0.95$ and $\lambda = 0.2$.

The situation modeling the data structure at extrema is given by the data sequence

$$\{U_l, U_l, U_m, U_r, U_{rr}\},$$

where U_m denotes (without a restriction on the generality of our approach) a local maximum and where the other data are suitably arranged. We exclude by the assumptions of the theorem *a priori* the situations

$$U_l > U_l < U_m \quad \text{or} \quad U_m > U_r < U_{rr}$$

since in this case we would have strict local extrema at the indices $\{l, m\}$ or $\{m, r\}$, respectively.

For the oscillation conditions (4) we obtain by straightforward computation using (1) the following coupled conditions:

$$q(U_m - U_r) - \lambda[f(U_m) - f(U_r)] > q(U_l - U_l) + \lambda[f(U_l) - f(U_l)] + 2(1 - q)(U_m - U_l), \quad (25)$$

$$q(U_m - U_l) + \lambda[f(U_m) - f(U_l)] > q(U_r - U_{rr}) - \lambda[f(U_r) - f(U_{rr})] + 2(1 - q)(U_m - U_r). \quad (26)$$

Concerning the situation at non-strict extrema, *i.e.*, $U_l \leq U_m \geq U_r$, the oscillation conditions (25) and (26) cannot be fulfilled simultaneously. In the case $U_m = U_r$, the left hand side of (25) becomes zero, as well as we have on the right hand side because of $U_m \geq U_l$ a contribution

$$2(1 - q)(U_m - U_l) \geq 0,$$

while the other term on the right hand side can be estimated in the fashion

$$\begin{aligned} q(U_r - U_{rr}) - \lambda[f(U_r) - f(U_{rr})] &= q(U_r - U_{rr}) - \lambda \frac{f(U_r) - f(U_{rr})}{U_r - U_{rr}} (U_r - U_{rr}) \\ &= q(U_r - U_{rr}) - \lambda f'(U_{rr,r}^*) (U_r - U_{rr}), \quad U_{rr,r}^* \in (U_{rr}, U_r) \\ &= (q - \lambda f'(U_{rr,r}^*)) (U_r - U_{rr}) \\ &> 0 \quad \text{by (2)}. \end{aligned}$$

Thus, for $U_m = U_r$ the oscillation condition (25) automatically fails. The case $U_m = U_l$ can be treated analogously, leading to the failure of the condition (26).

Therefore, we only need to discuss strict local extrema with the structure of data values

$$U_l \leq U_l < U_m > U_r \geq U_{rr}.$$

For such data we have to show, that the two coupled inequalities (25) and (26) cannot be fulfilled simultaneously for $q \leq 1/2$.

Using the same kind of simple manipulations and analogous notions as above, we can derive for the corresponding terms on the right hand sides of (25) and (26) the relations

$$q(U_k - U_{kk}) \pm \lambda[f(U_k) - f(U_{kk})] = \underbrace{(q \pm \lambda f'(U_{kk,k}^*))}_{>0 \text{ by (2)}} (U_k - U_{kk}) \geq 0, \quad k \in \{l, r\}. \tag{27}$$

Employing essentially the same procedure, we also obtain for $k \in \{l, r\}$ the estimates

$$q(U_m - U_k) \pm \lambda[f(U_m) - f(U_k)] = (q \pm \lambda f'(U_{k,m}^*)) (U_m - U_k) < 2q(U_m - U_k). \tag{28}$$

By the derived relations (27) and (28), (25) and (26) yield the following necessary conditions for the occurrence of oscillations:

$$2q(U_m - U_r) > 2(1 - q)(U_m - U_l), \tag{29}$$

$$2q(U_m - U_l) > 2(1 - q)(U_m - U_r). \tag{30}$$

Evidently, the conditions (29) and (30) can only be fulfilled for

$$2q > 2(1 - q), \quad \text{i.e., for } q > \frac{1}{2}. \quad \square$$

The importance of Theorem 1.5 is obvious, since one can now be sure to have a reliable method when using *e.g.* the modified Lax-Friedrichs scheme (at least in the scalar case in 1-D).

One could ask if it is possible to extend under more specific circumstances the assertion of Theorem 1.5 with respect to coefficients of numerical viscosity $q > 1/2$. For example, let us assume $f'(\cdot) \geq 0$, so that instead of using the relations (27) which were derived for a general Lipschitz continuous flux, we now take into account some more information. By estimating

$$q + \lambda f'(\cdot) < 2q \quad \text{and} \quad q - \lambda f'(\cdot) < q \quad \text{since } f'(\cdot) \geq 0,$$

we can derive in the same fashion as within the proof of Theorem 1.5 above the set of necessary conditions

$$q(U_m - U_r) > 2(1 - q)(U_m - U_l), \tag{31}$$

$$2q(U_m - U_l) > 2(1 - q)(U_m - U_r). \tag{32}$$

Because of

$$\max \left(\frac{U_m - U_r}{U_m - U_l}, \frac{U_m - U_l}{U_m - U_r} \right) \geq 1,$$

and since we may choose any special set of data values $\{U_l, U_m, U_r\}$, we may write down the necessary conditions (31) and (32) in the form

$$\frac{q}{2(1-q)} > 1 \quad \text{and} \quad \frac{2q}{2(1-q)} > 1, \tag{33}$$

respectively, and we see that the latter inequality is less restrictive than the first one. Especially, if

$$\frac{q}{2(1-q)} > 1$$

which is equivalent to $q > 2/3$ is fulfilled, the validity of the second inequality in (33) follows. We may now specifically impose $U_r < U_l$, *i.e.*,

$$\frac{U_m - U_r}{U_m - U_l} =: a \geq 1,$$

leading by (31) and (32) to the conditions

$$\frac{q}{2(1-q)} > \frac{1}{a} \quad \text{and} \quad \frac{q}{2(1-q)} > \frac{a}{2}.$$

Summarizing these expressions, we obtain the condition

$$\frac{q}{2(1-q)} > \max \left\{ \frac{1}{a}, \frac{a}{2} \right\}. \tag{34}$$

Since the quantity $q/[2(1-q)]$ grows monotonously in q for $q \in (0, 1)$, the task remains to minimize the expression on the right hand side of (34) over $a \geq 1$ in order to find the lowest value of q for which the necessary conditions for the occurrence of oscillations are fulfilled. Obviously, this minimum is assumed at $a = \sqrt{2}$. Using this value within the expressions (34), we obtain the relations

$$q > \frac{\sqrt{2}}{1 + \sqrt{2}}, \quad \text{or, equivalently,} \quad q > \frac{2}{2 + \sqrt{2}},$$

yielding approximately $q > 0.5858$.

1.2. Rusanov's Method

Rusanov's method reads

$$U_j^{n+1} = U_j^n - \frac{1}{2}\lambda \left[f(U_{j+1}^n) - f(U_{j-1}^n) \right] + \frac{\lambda}{2} a_{j+1/2} (U_{j+1}^n - U_j^n) - \frac{\lambda}{2} a_{j-1/2} (U_j^n - U_{j-1}^n), \tag{35}$$

where the parameters controlling the numerical viscosity $a_{j\pm 1/2}$ at the cell interfaces are chosen taking into account the local signal velocities, *i.e.*,

$$a_{j-1/2} = \max_{U \in \text{conv}(U_{j-1}^n, U_j^n)} |f'(U)|,$$

where $\text{conv}(\cdot, \cdot)$ denotes the convex hull of the arguments.

Obviously, the factors $\lambda a_{j\pm 1/2}$ take on the role of the viscosity coefficient q , and for $a_{j+1/2} = a_{j-1/2}$ which especially holds in the linear case, one can identify directly

$$\lambda a_{j\pm 1/2} \equiv q \tag{36}$$

and write Rusanov’s method as a LFt scheme. Let us also note that, at an extremum with index k , it is likely that $a_{k-1/2} = a_{k+1/2}$ holds.

By (36), we see that Rusanov’s method does in general not even locally degenerate to the classical Lax-Friedrichs scheme since this would be equivalent to the situation

$$\lambda a_{j\pm 1/2} = 1. \tag{37}$$

The fact that (37) does not occur at all (not even for one of the signs \pm) in the context discussed within this paper is due to our notion of the CFL condition (2), *i.e.*, especially it is due to the CFL number $CFL \in (0, 1)$ avoiding a ratio of mesh parameters λ large enough so that (37) could hold in special cases.

Employing exactly the same procedure as used up to now, the oscillation conditions (4) can be written down in the form

$$\begin{aligned} \lambda a_{m+1/2} (U_m - U_r) - \lambda [f(U_m) - f(U_r)] \\ > \lambda a_{l-1/2} (U_l - U_u) + \lambda [f(U_l) - f(U_u)] + 2(1 - \lambda a_{m-1/2}) (U_m - U_l) \end{aligned} \tag{38}$$

and

$$\begin{aligned} \lambda a_{m-1/2} (U_m - U_l) + \lambda [f(U_m) - f(U_l)] \\ > \lambda a_{r+1/2} (U_r - U_{rr}) - \lambda [f(U_r) - f(U_{rr})] + 2(1 - \lambda a_{m+1/2}) (U_m - U_r). \end{aligned} \tag{39}$$

Note the similarity of (38) and (39) with (25) and (26), respectively.

The assertions derived for the LFt schemes can obviously be employed as a means of necessary or sufficient conditions for the occurrence or avoidance of oscillations. Let us summarize the may be most important property of Rusanov’s method in this context in the framework of a corollary.

Corollary 1.6. *If we choose λ in a manner such that $\lambda a_{j-1/2} \leq 1/2$ holds for all j , then Rusanov’s method is a generalized monotone scheme within the same framework as used in Theorem 1.5.*

Following these guiding considerations, we can easily find an oscillatory state computed by using Rusanov’s method. The example we use is defined by solving Burgers equation

$$u_t + \left(\frac{1}{2}u^2\right) = 0$$

using the parameter $\lambda = 0.8$ and for simplicity the data

$$\dots = U_{ll} = U_l = 0.8, U_m = 1 \text{ and } 0.4 = U_r = U_{rr} = \dots$$

The result of the computation after one time step is shown in Figure 8. Of course, after the discussion of the LFt schemes, it is not surprising that any oscillation will vanish again after some time steps.

1.3. The Nessyahu-Tadmor schemes

We now consider the staggered as well as the non-staggered form of the NT scheme which can both be found in [15]. For the purpose of the definition of the methods, let us introduce the notion for predicted values

$$U_k^{1/2} = U_k - \frac{1}{2}\lambda f'_k \tag{40}$$

where f'_k indicates the second order reconstruction of the corresponding quantity using also a limiter. Within this paper, we use

$$f'_k := \text{minmod}(\Delta f_{k+1/2}, \Delta f_{k-1/2}) := \text{minmod}(f_{k+1} - f_k, f_k - f_{k-1}) \tag{41}$$

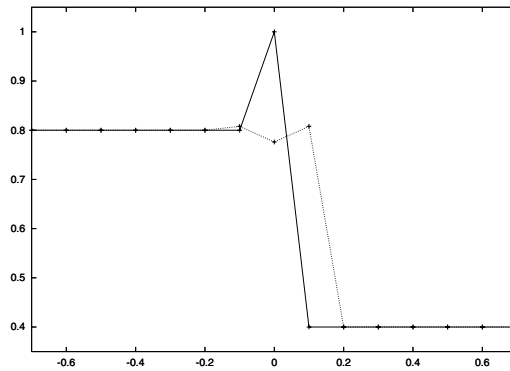


FIGURE 8. The figure contains the initial data (thick line) together with the numerical solution (dotted line) using Rusanov's method after one time step.

where minmod denotes the usual minmod-function, see *e.g.* [3, 12] for general discussions. The quantity U'_k is defined analogously.

Using the predicted values from (40), we obtain the corrector steps of the *staggered NT scheme*

$$U_{j+1/2}^{new} = \frac{1}{2} (U_j + U_{j+1}) + \frac{1}{8} (U'_j - U'_{j+1}) - \lambda [f(U_{j+1}^{1/2}) - f(U_j^{1/2})] \tag{42}$$

and of the *non-staggered NT scheme*

$$U_j^{new} = \frac{1}{2} (U_{j-1} + U_{j+1}) + \frac{1}{4} (U'_{j-1} - U'_{j+1}) - \frac{\lambda}{2} [f(U_{j+1}^{1/2}) - f(U_{j-1}^{1/2})], \tag{43}$$

respectively. As can easily be seen, the non-staggered form reduces in the simple scenarios

$$\dots = U_l = U_l < U_m > U_r = U_{rr} = \dots$$

to the classical LF scheme, indicating already an oscillatory property in more general cases.

It is not helpful to write down analogously to the procedure before the oscillation conditions for the non-staggered NT scheme. Instead, let us sketch an approach by which one may elaborate on the oscillatory properties of the scheme. For this purpose, let us write down the correction step of the non-staggered scheme (with additional terms which add up to zero) in the form

$$U_j^{new} = \frac{1}{2} (U_{j-1} + U_{j+1}) - \frac{\lambda}{2} [f(U_{j+1}) - f(U_{j-1})] + \frac{1}{4} (U'_{j-1} - U'_{j+1}) + \frac{\lambda}{2} [f(U_{j+1}) - f(U_{j+1}^{1/2})] + \frac{\lambda}{2} [f(U_{j-1}^{1/2}) - f(U_{j-1})]. \tag{44}$$

Thus we have written the non-staggered NT scheme as the LF scheme plus a correction. Splitting the additional terms within (44) into the parts

$$\frac{1}{4} U'_{j-1} + \frac{\lambda}{2} [f(U_{j-1}^{1/2}) - f(U_{j-1})] \quad \text{and} \quad -\frac{1}{4} U'_{j+1} + \frac{\lambda}{2} [f(U_{j+1}) - f(U_{j+1}^{1/2})],$$

it is a simple task to rewrite the scheme (40), (43) for instance for the case of a local maximum

$$\dots \leq U_l < U_l < U_m > U_r \geq U_{rr} \geq \dots$$

into the form

$$U_m^{new} = \frac{1}{2} (\tilde{U}_l + \tilde{U}_r) - \frac{\lambda}{2} [f(\tilde{U}_l) - f(\tilde{U}_r)] \quad \text{with} \quad \tilde{U}_l \leq U_m, \tilde{U}_r \leq U_m.$$

Depending on the particular flux function in use and on the given data (and also on the relations of \tilde{U}_{kk} to \tilde{U}_k , $k \in \{l, r\}$), an oscillation may appear. Because of all these various influences, it is not very instructive to discuss the procedure for a general framework; rather than that, let us point out that the possibility to obtain very easily such an expression *together* with the bounds $\tilde{U}_l \leq U_m$ and $\tilde{U}_r \leq U_m$ are an alarming sign.

The procedure mentioned strongly indicates that the non-staggered NT scheme may give oscillatory states, and referring to [1] we may also argue that oscillations once they appear do not simply vanish by application of the scheme (40), (43). Let us refer to the section about numerical tests for a numerical verification of the validity of these heuristic impressions.

It is interesting to elaborate on the staggered NT scheme (40), (42) in the context discussed in this paper.

Let us investigate first the local Lax-Friedrichs method

$$U_{j+1/2}^{new} = \frac{1}{2} (U_j + U_{j+1}) - \lambda [f(U_{j+1}) - f(U_j)]. \tag{45}$$

This method is of interest since the staggered NT scheme (40), (42) recovers the formula (45) for $U'_k \equiv 0 \equiv f'_k$, $k \in \{j, j+1\}$.

It is useful to consider an arbitrary data tuple $\{U_l, U_r\}$. Let f be a general Lipschitz continuous flux with Lipschitz constant L , and let us recall that the scheme is monotone under the CFL condition

$$\lambda L \leq \frac{1}{2}. \tag{46}$$

Then the application of (45) at this data tuple yields

$$\frac{1}{2} (U_l + U_r) - \underbrace{\lambda [f(U_r) - f(U_l)]}_{\geq -\lambda L |U_r - U_l|} \stackrel{(46)}{\geq} \frac{1}{2} (U_l + U_r) - \frac{1}{2} |U_r - U_l| = \begin{cases} U_l : U_l \leq U_r, \\ U_r : U_l \geq U_r. \end{cases} \tag{47}$$

Analogously, we find

$$\frac{1}{2} (U_l + U_r) - \underbrace{\lambda [f(U_r) - f(U_l)]}_{\leq \lambda L |U_r - U_l|} \stackrel{(46)}{\leq} \frac{1}{2} (U_l + U_r) + \frac{1}{2} |U_r - U_l| = \begin{cases} U_l : U_l \geq U_r, \\ U_r : U_l \leq U_r. \end{cases} \tag{48}$$

Thus, by (47) and (48) the new value $U_{(l+r)/2}$ is always in the convex hull of U_l and U_r . Obviously, no new extremum can be generated by the discussed method because of this property!

Let us now turn to the staggered NT method (40), (42). We want to prove two assertions, one under the assumption that the flux is monotone, while the other assertion is more general. It turns out that it is instructive to begin with the more specialized case since there the meaning of the details is more obvious. Hence, we proceed by stating the following

Theorem 1.7. *Consider the same assumption on the data structure as within Theorem 1.5. Let us also assume $f'(\cdot) \geq 0$. Then the staggered Nessyahu-Tadmor scheme (40), (42) using the minmod-limiter does not create new extrema at data extrema under the condition*

$$\lambda \max_{U \in \mathcal{I}} f'(U) \leq 0.375, \tag{49}$$

where \mathcal{I} is the given range of data values.

Proof. Without a restriction on the generality of the approach, let us restrict our attention to the case of a local maximum; the case of a local minimum can be treated analogously.

We consider the data

$$\{U_{ll}, U_l, U_r, U_{rr}\}$$

with a possible maximum either in U_l or in U_r . This means, we have to distinguish the two possible situations of a local maximum

$$(a) U_{ll} < U_l > U_r \geq U_{rr} \quad \text{and} \quad (b) U_{ll} \leq U_l < U_r > U_{rr}. \tag{50}$$

Note that all possible situations of interest are covered *via* (50). We discuss only case (a) since the proceeding concerning case (b) is essentially the same.

Because of our assumption $f'(\cdot) \geq 0$ we know that

$$\text{minmod}(f(U_l) - f(U_{ll}), f(U_r) - f(U_l)) = 0 \tag{51}$$

holds. Using this with (40), we obtain after adding terms equal to zero the update formula

$$U_{(l+r)/2}^{new} = \underbrace{\frac{1}{2}(U_l + U_r)}_A - \underbrace{\lambda[f(U_r) - f(U_l)]}_B - \underbrace{\frac{1}{8}U_r' + \lambda[f(U_r) - f(U_r^{1/2})]}_C. \tag{52}$$

Part A locates the new value exactly at the average of U_l and U_r . The task is now to estimate the deviation from this average introduced by B and C. For the estimations, we employ a number $\theta_a \in (0, 1]$ with $\lambda L \leq \theta_a$. Then we obtain by $f'(\cdot) \geq 0$ for B

$$-\lambda[f(U_r) - f(U_l)] = \lambda[f(U_l) - f(U_r)] \begin{cases} \leq \theta_a (U_l - U_r), \\ \geq 0. \end{cases} \tag{53}$$

Similarly, because of

$$\text{minmod}(f(U_r) - f(U_l), f(U_{rr}) - f(U_l)) \leq 0,$$

we have

$$U_r^{1/2} = U_r - \frac{\lambda}{2} f_r' \begin{cases} \geq U_r, \\ \leq U_r - \frac{\lambda}{2} [f(U_r) - f(U_l)] \leq U_r + \frac{\theta_a}{2} (U_l - U_r). \end{cases} \tag{54}$$

Employing (54), we obtain for C:

$$\underbrace{-\frac{1}{8}U_r'}_{\geq 0} + \lambda \underbrace{[f(U_r) - f(U_r^{1/2})]}_{\leq 0} \begin{cases} \leq \frac{1}{8}(U_l - U_r), \\ \geq -\lambda\theta_a (U_r^{1/2} - U_r) \geq -\frac{\lambda^2\theta_a^2}{2}(U_l - U_r). \end{cases} \tag{55}$$

By (53), (54) and (55), the conditions for

$$U_{(l+r)/2}^{new} \in \text{conv}(U_l, U_r)$$

are

$$\frac{1}{2}(U_l + U_r) + \theta_a (U_l - U_r) + \frac{1}{8}(U_l - U_r) \leq U_l \tag{56}$$

and

$$\frac{1}{2}(U_l + U_r) - \frac{\lambda^2\theta_a^2}{2}(U_l - U_r) \geq U_r. \tag{57}$$

A simple computation yields the equivalence of (56), (57) with $\theta_a \leq 3/8 = 0.375$. The analogous procedure for case (b) results in $\theta_b \leq +\sqrt{2} - 1 \approx 0.4142$. Taking the minimum of θ_a and θ_b yields the assertion. \square

Note that the assertion of the Theorem 1.7 can be applied in a local fashion at an extremum if the TVD property is ensured for the remaining parts of the data.

As indicated before, we now give a more general statement.

Theorem 1.8. *Consider a general Lipschitz continuous flux f . Then the staggered Nessyahu-Tadmor scheme (40), (42) using the minmod-limiter does not create new extrema under the condition*

$$\lambda \max_{U \in \mathcal{I}} |f'(U)| \leq +\sqrt{\frac{1}{2}} - \frac{1}{2} \approx 0.2071, \tag{58}$$

where \mathcal{I} is the given range of data values.

Proof. We use the same notations as in the proof of Theorem 1.7. Let us consider a general data tuple $\{U_l, U_r\}$. The proceeding is again to estimate the deviation from the average $(U_l + U_r)/2$.

In the general case, the same procedure which lead to (52) results in

$$U_{(l+r)/2}^{new} = \frac{1}{2}(U_l + U_r) \underbrace{-\lambda [f(U_r) - f(U_l)]}_A + \underbrace{\frac{1}{8}(U'_l - U'_r)}_B + \underbrace{\lambda [f(U_r) - f(U_r^{1/2})]}_C \underbrace{-\lambda [f(U_l) - f(U_l^{1/2})]}_D. \tag{59}$$

We now estimate the parts A – D using again the Lipschitz constant L with respect to f . We obtain

$$A \leq \lambda |f(U_r) - f(U_l)| \leq \lambda L |U_l - U_r| \leq \theta |U_l - U_r|,$$

$$B \leq \frac{1}{8} (|U'_l| + |U'_r|) \leq \frac{1}{8} (2|U_l - U_r|) = \frac{1}{4} |U_l - U_r|.$$

The procedure for the estimations of C and D can be summarized using $k \in \{l, r\}$ via

$$\left| \pm \lambda [f(U_k) - f(U_k^{1/2})] \right| \leq \theta |U_k - U_k^{1/2}| \leq \frac{\theta^2}{2} |U_l - U_r|,$$

compare (54), (55). Note that we have – corresponding to C and D – two times the latter contribution within the resulting condition

$$\theta |U_l - U_r| + \frac{1}{4} |U_l - U_r| + \theta^2 |U_l - U_r| \leq \frac{1}{2} |U_l - U_r|. \tag{60}$$

The condition (60) ensures that the deviation from the average of U_l and U_r is not large enough to leave $\text{conv}(U_l, U_r)$. The largest value for θ which satisfies (60) is obviously obtained by solving the quadratic equation

$$\theta^2 + \theta = \frac{1}{4},$$

yielding the assertion with (58). □

Note that the number θ found here is lower than the corresponding value ensuring the TVD property mentioned in [15], Corollary 3.3. However, within Theorem 1.8 we deal with a sufficient condition in order to avoid oscillations; in practice one may employ a higher number restricting the time step size without observing any oscillations.

Note also that we heavily relied on the properties of the classical minmod-limiter in order to estimate the derivatives occurring. When using a limiter which enables a larger variation within the local data reconstruction, a sufficient condition for the avoidance of oscillations will feature a smaller time step size.

2. NUMERICAL TESTS

In this section, we discuss the results of two kinds of numerical experiments.

One experiment is concerned with an one-dimensional scalar test case which was already found useful in [1], *i.e.*, the approximation of a N-wave solution of Burgers equation. We compare the numerical solutions obtained by all the schemes discussed. Rusanov's method and the NT schemes are used only in this test case since these schemes are inherently one-dimensional in the formulations subject to the theoretical investigations in this paper; a two-dimensional extension of these schemes is not trivial.

The second test case is devoted to the 2-dimensional shallow water equations. There we are considering a radial dam break problem given in a similar fashion in [12]. Here, we investigate if the numerical phenomenon discussed is also recognizable when approximating systems of equations in more than one dimension since this is clearly most relevant for practical purposes. For the tests, we use a straightforward 2-D extension of the LFT schemes presented in (1).

2.1. Burgers equation

The test problem is concerned with Burgers equation

$$u_t + \left(\frac{1}{2} u^2 \right)_x = 0$$

supplemented by the initial condition

$$u_0(x) = \begin{cases} \frac{x}{2} & : -2 \leq x \leq 2 \\ 0 & : \text{else.} \end{cases}$$

The exact solution is the N-wave

$$N(x, t) = \begin{cases} \frac{x}{t+2} & : \sqrt{2(t+2)} \leq x \leq \sqrt{2(t+2)} \\ 0 & : \text{else.} \end{cases}$$

Since every solution of Burgers equation decays to an N-wave, this example is of fundamental importance, see *e.g.* [2, 12] for discussions.

Within the figures corresponding to this test problem, we show snapshots of the exact as well as well as of the numerical solutions obtained by all the schemes discussed. The numerical initial conditions are always the same, see Figure 9, while we employ different parameter values λ .

We begin our discussion with Rusanov's method. In Figure 10 we observe an example for an evolution. Especially, we have taken $\lambda = 0.98$ in order to show small oscillations after the first time step exactly at the former extrema: this is due to the fact that the viscosity parameters computed within the scheme are near 1 for the data at the extrema. Since these coefficients are computed anew for the second step and since they are then considerably smaller at the oscillation, this oscillation vanishes immediately. This figure shows from left to right and top to bottom: the numerical solution after (a) one, (b) two and (c) 20 time steps using $\lambda = 0.98$, and for comparison of the latter result also (d) the numerical solution after 40 time steps using $\lambda = 0.49$. In (c) and (d), the snapshots of the numerical solutions (dotted lines) are supplemented by the corresponding exact solutions (lines). Note that we also show by plot (d) in that figure that there is not much difference between the solutions obtained by the choices $\lambda = 0.98$ or $\lambda = 0.49$ (the latter choice yielding no oscillations at all) in the long run, respectively.

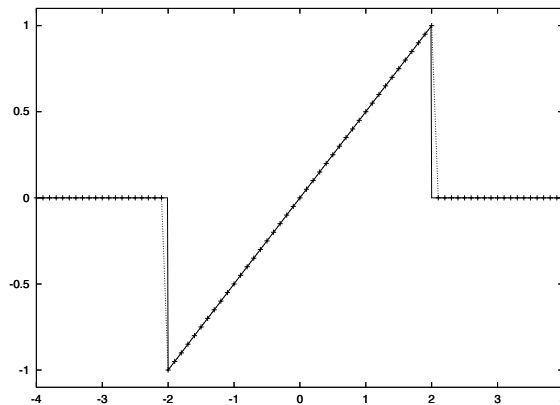


FIGURE 9. The figure contains a plot of the initial data (line) together with the numerical initial condition (line with crosses) used for all the tests.

The Figure 11 shows results obtained by using the traditional LF scheme, *i.e.*, $q = 1$. It shows from left to right and top to bottom: the numerical solution after (a) one, (b) 5 and (c) 25 time steps using $\lambda = 0.8$, and also (d) the numerical solution after 100 time steps using $\lambda = 0.2$. As within Figure 10, the snapshots of exact solutions are always drawn with lines and the numerical solutions are depicted by dotted lines. Obviously, the local extrema within the initial data yield numerical solutions with severe oscillations. Note especially, that Figure 11(d) shows that much numerical viscosity as it is present within the Lax-Friedrichs method for small time step sizes Δt does not give better results.

By the Figures 12 to 15 we study the effects of lowering the viscosity parameter q within the LFt schemes. These figures are completely analogous to each other, and we always use $\lambda = 0.4$. We observe in every case from left to right and top to bottom: the numerical solution after (a) one, (b) two, (c) 5 and (d) 25 timesteps, in the latter case supplemented in the usual style by the corresponding exact solution. We observe that the occurring oscillations decay very rapidly when q stays away from $q = 1$. For comparison, we also give the results obtained by the modified Lax-Friedrichs scheme ($q = 1/2$). The presented experiments show, that the decaying effect predicted by the theoretical investigations seems to be very influential in practical computations. In fact, even for small deviations of the viscosity coefficient from $q = 1$ we observe a strong damping effect, emphasizing the special role of the Lax-Friedrichs scheme. However, there are some points to mention concerning this: we used a relatively large time step size, and the linear distribution of values in the middle of the N-wave also has a regularizing effect. Moreover, our test example was chosen in a fashion so that the phenomenon of interest could appear already at the first time step. In order to avoid (or explain) the discussed phenomenon during the course of other computations, the theoretical discussion given before remains important.

We then turn to the NT schemes. By the Figures 16 and 17, we compare the results obtained by the staggered and the non-staggered method, respectively. These are from left to right and top to bottom: the numerical solution after (a) 2, (b) 10, (c) 20 and (d) 50 time steps. In (c) and (d), the snapshots of the numerical solutions (dotted lines) are again supplemented by the corresponding exact solutions (lines). The parameter λ is chosen in both cases as $\lambda = 0.3$ (satisfying the sufficient non-oscillatory condition we derived for the staggered scheme locally since f is monotone over the range of critical data near the extrema). We observe the stability and high-resolution property of the staggered scheme we expected, while there are severe oscillations in the case of the non-staggered scheme, similarly to those observed in Figure 11 for the classical Lax-Friedrichs method.

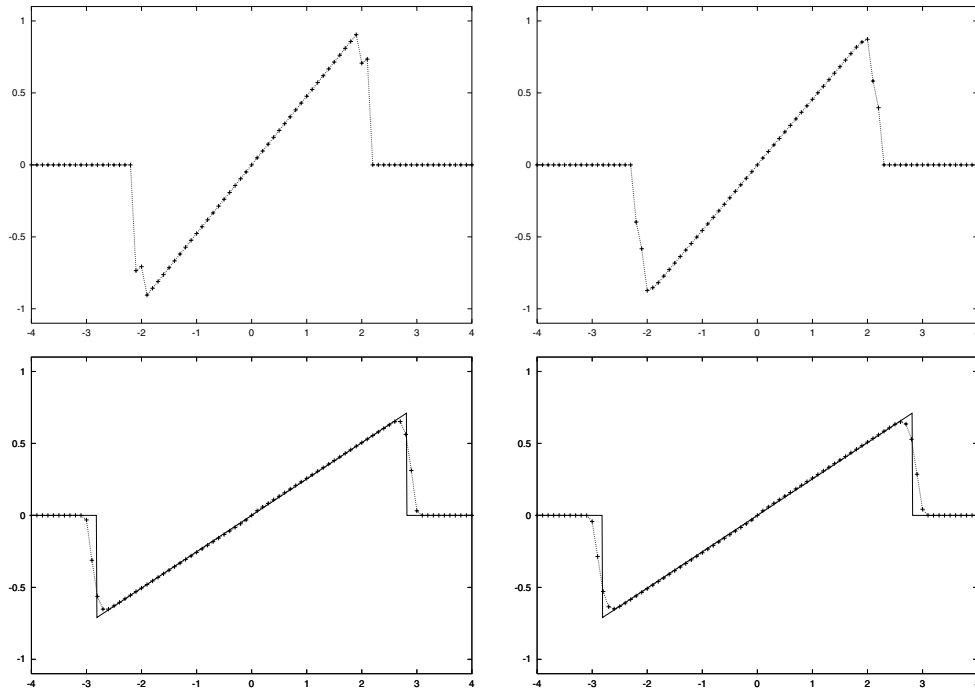


FIGURE 10. Rusanov's method.

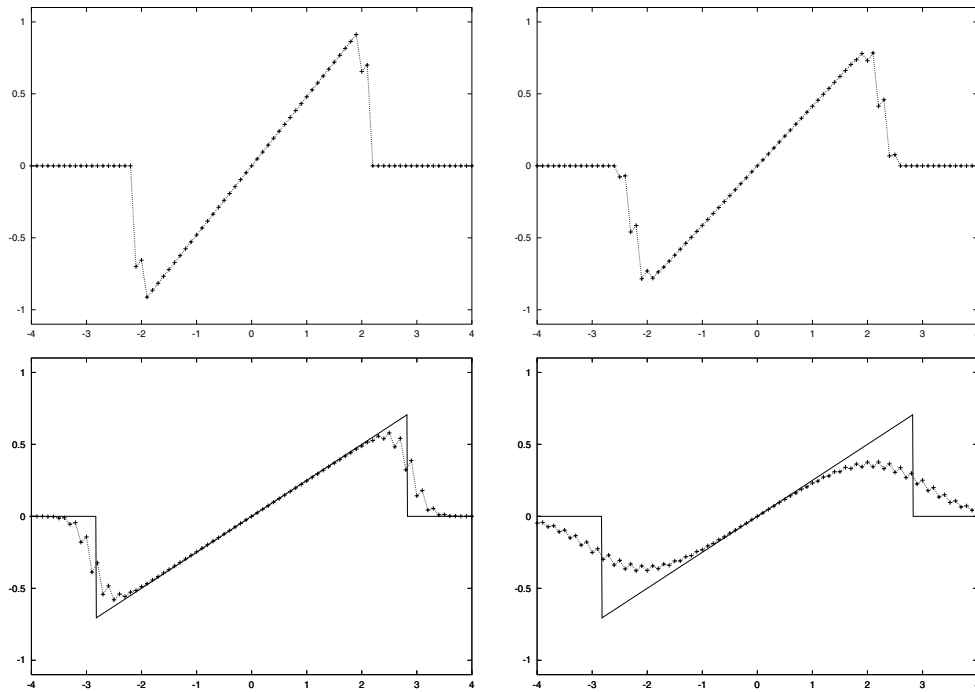


FIGURE 11. The Lax-Friedrichs scheme.

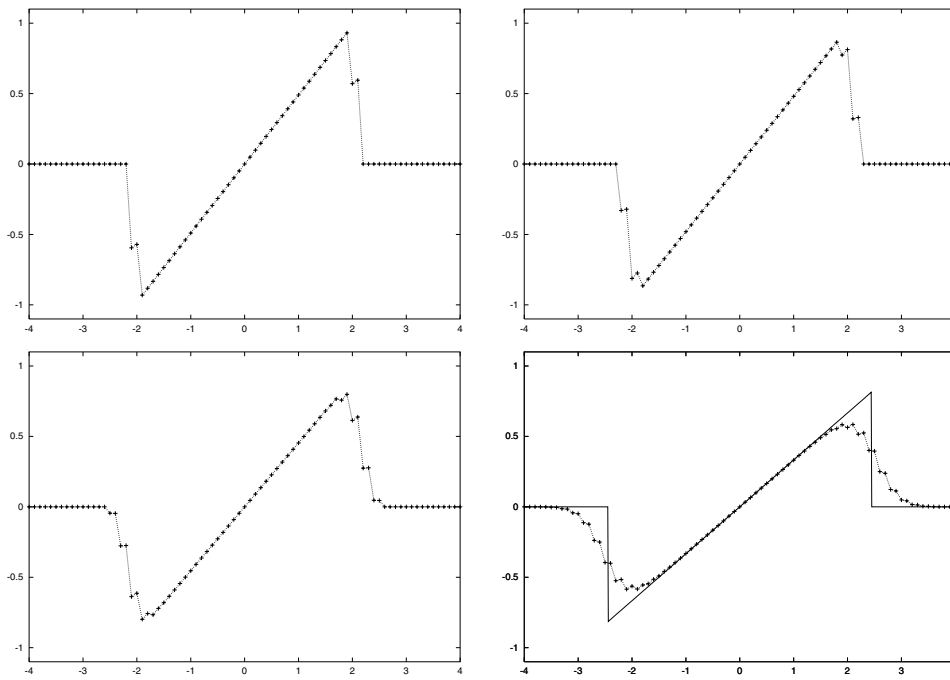


FIGURE 12. LfT scheme with $q = 0.99$.

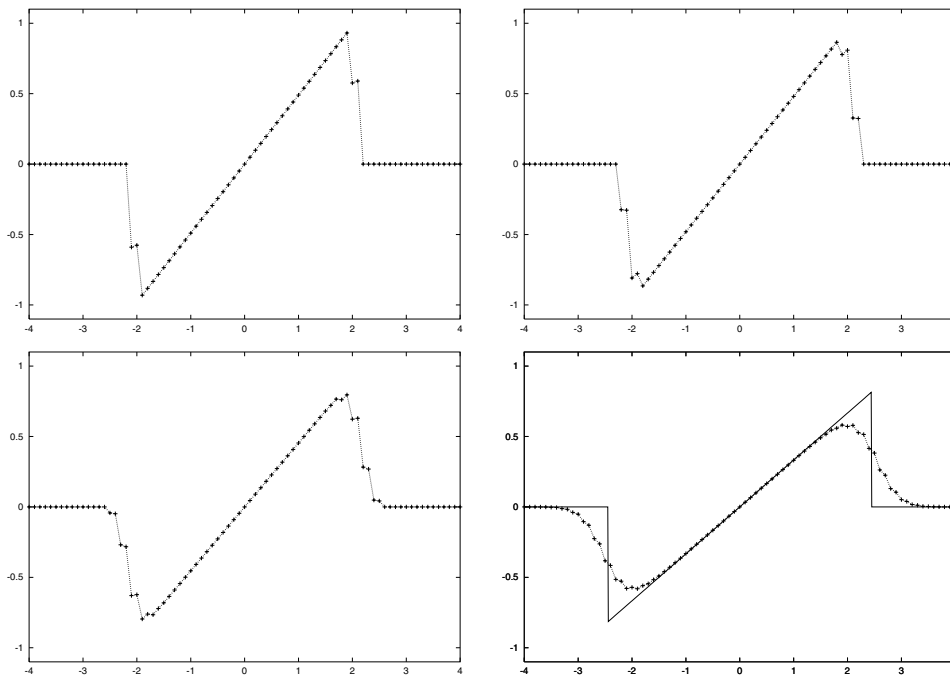


FIGURE 13. LfT scheme with $q = 0.98$.

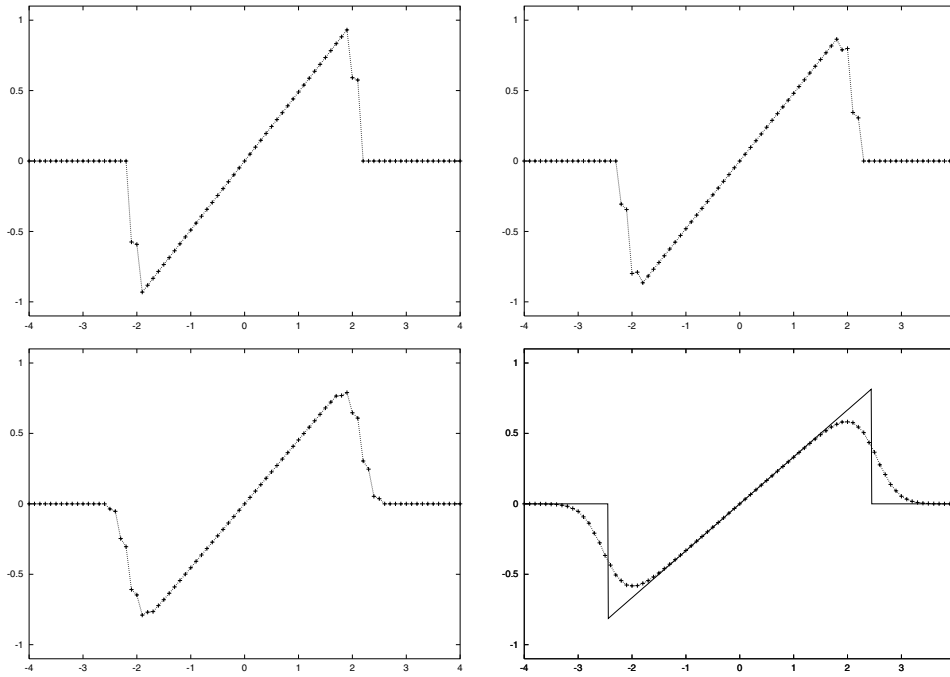


FIGURE 14. Lf scheme with $q = 0.95$.

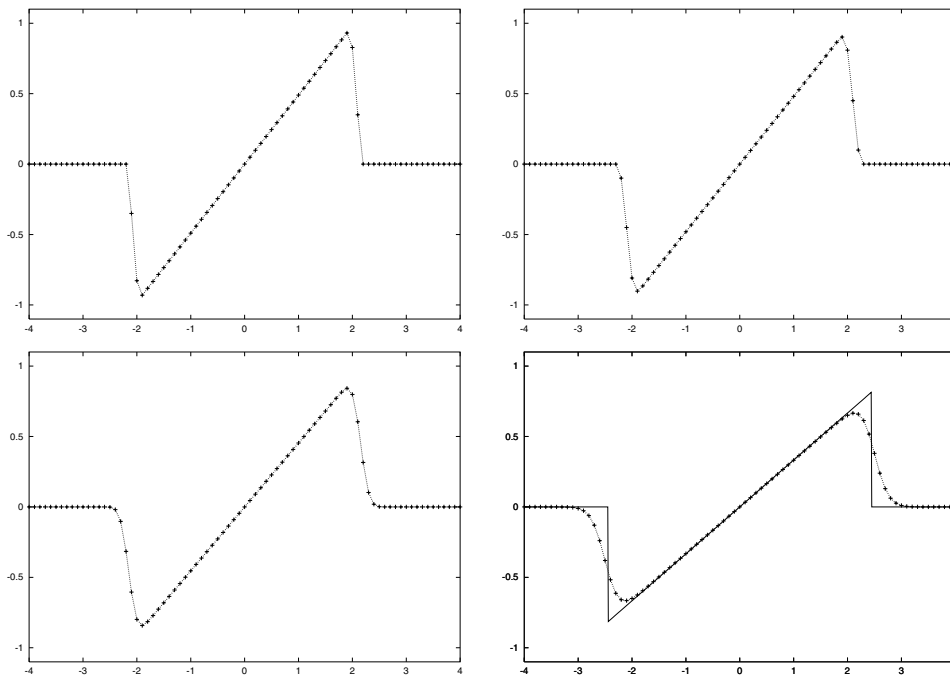


FIGURE 15. The modified Lax-Friedrichs scheme, *i.e.*, $q = 0.5$.

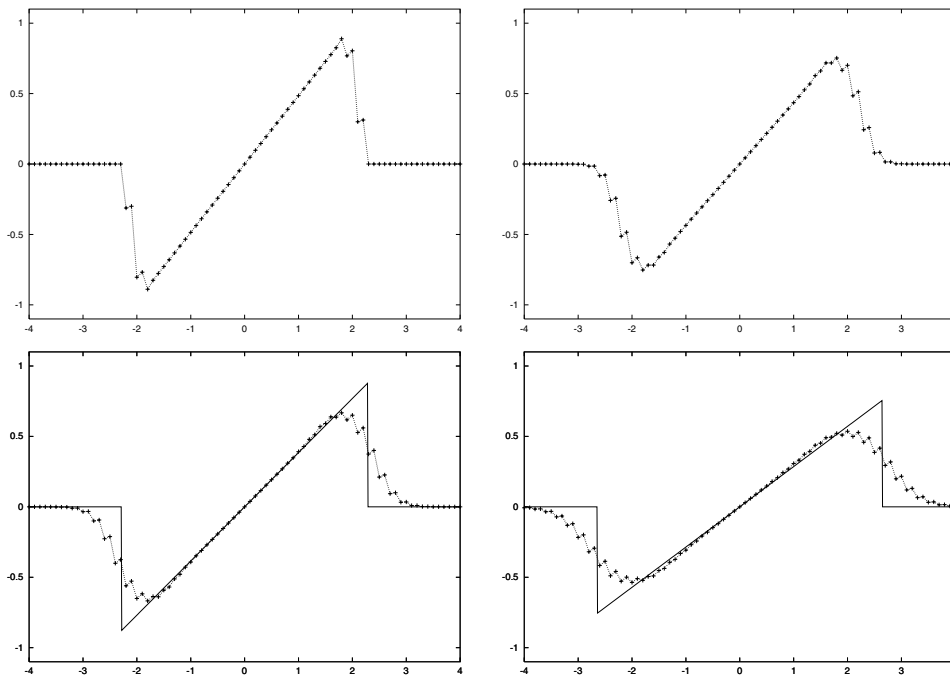


FIGURE 16. The non-staggered NT scheme.

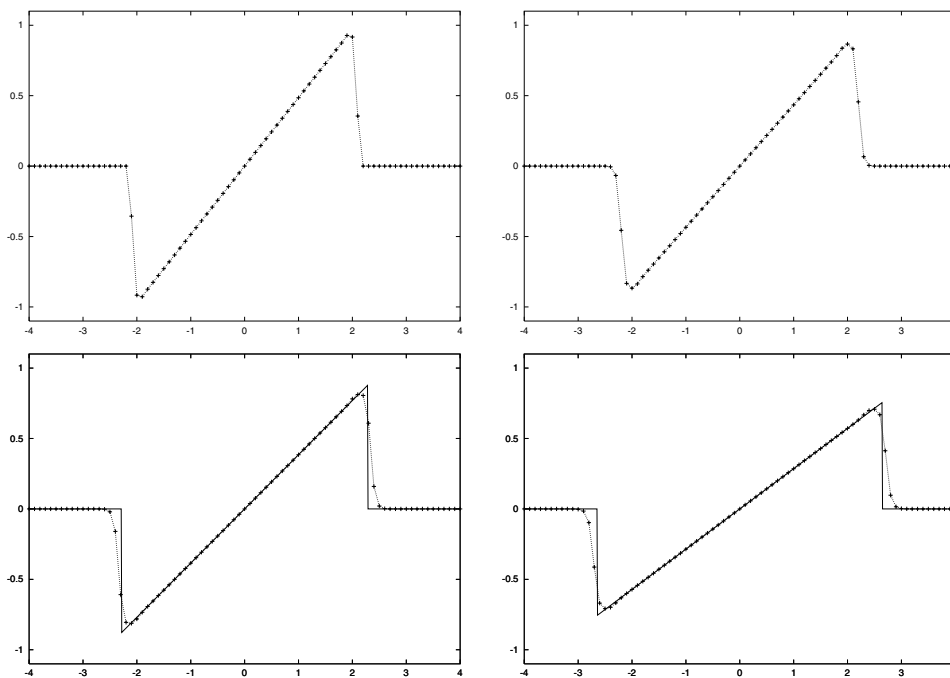


FIGURE 17. The staggered NT scheme.

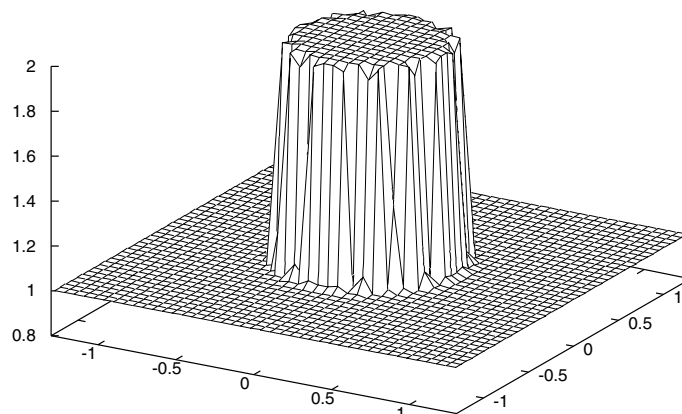


FIGURE 18. The figure shows a plot of the initial data used for the dam break problem.

2.2. The Shallow Water equations

For the numerical computations in this paragraph we use the two-dimensional form of the LFT schemes which are a generalization of the two-dimensional form of the LF scheme as it can be found in [4], *i.e.*, we deal with

$$\begin{aligned}
 U_{j,k}^{n+1} = & U_{j,k}^n + \frac{q}{4} (U_{j+1,k}^n - 2U_{j,k}^n + U_{j-1,k}^n) - \frac{\lambda_x}{2} [f(U_{j+1,k}^n) - f(U_{j-1,k}^n)] \\
 & + \frac{q}{4} (U_{j,k+1}^n - 2U_{j,k}^n + U_{j,k-1}^n) - \frac{\lambda_y}{2} [g(U_{j,k+1}^n) - g(U_{j,k-1}^n)].
 \end{aligned} \tag{61}$$

The scheme (61) is then applied componentwise on the shallow water system of equations without bottom topography

$$\begin{cases}
 h_t + (hu)_x + (hv)_y = 0, \\
 (hu)_t + (hu^2 + \frac{1}{2}gh^2)_x + (huv)_y = 0, \\
 (hv)_t + (huv)_x + (hv^2 + \frac{1}{2}gh^2)_y = 0,
 \end{cases} \tag{62}$$

where g is the gravitational constant, h is the depth and (u, v) the velocity vector, so that (hu) and (hv) are the momenta in the two directions; see also *e.g.* [12] and the references therein.

The test problem is a radial dam break problem with the initial depth $h = 2$ inside a circular dam of radius $r = 1/2$ and $h = 1$ outside, see Figure 18, while $(u, v) = (0, 0)$ initially. Let us note that we do not use the radial symmetry in order to reduce the dimension of the problem since we want to observe if our phenomenon shows a dependence on the orientation of the grid in the 2-D case.

For the relevant parameters for the numerical tests we always choose $\Delta x = \Delta y = 0.04$ and $\Delta t = 0.004$ (yielding $\lambda = 0.1$) together with a grid of 125×125 mesh points. We investigate the state obtained after 20 timesteps, *i.e.*, at $t = 0.08$.

Similarly to the experiments before, we first discuss the classical Lax-Friedrichs choice $q = 1$, and then we modify the coefficient of numerical viscosity. It is instructive to discuss the solutions obtained for the depth h , and thus we restrict our attention to h .

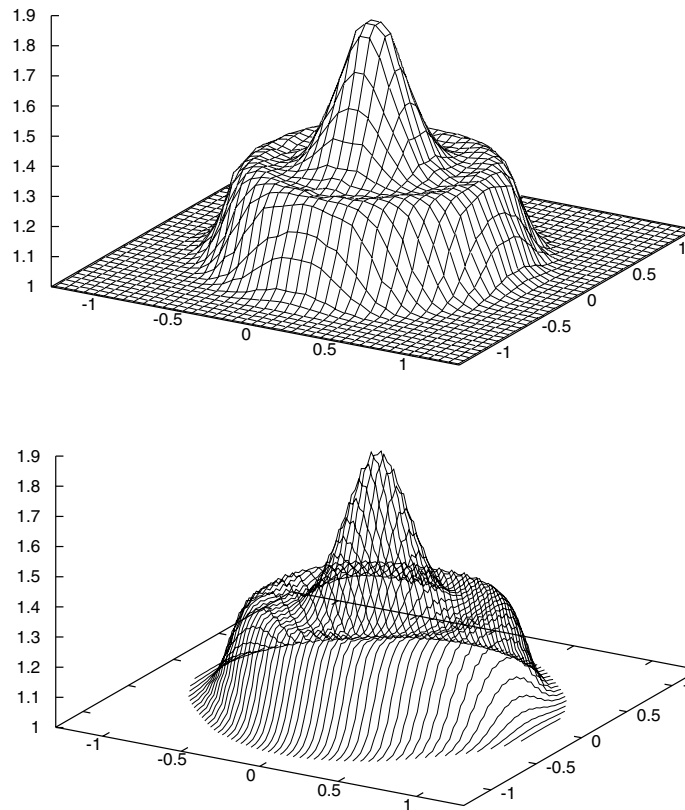


FIGURE 19. The figure shows the numerical solution at $t = 0.08$ using two different display styles.

At $t = 0$, the dam breaks, and at the state of our interest at $t = 0.08$ we especially observe a radial shock front followed by a rarefaction. Two plots of the complete numerical approximation (using $q = 1$ in (61)) of this state are given in Figure 19. At the top, we observe nothing unusual because of the style used for creating this figure; the same data displayed below in a different fashion reveal many wiggles spread over the whole computational domain.

Zooming into several parts of the solution, see Figure 20, we clearly observe oscillations spreading over the remaining part of the initial pillar of water and over the rarefaction following the shock.

Note, that the latter oscillations are not observable when looking only along the grid diagonal, see Figure 21: however, while the plot of the values along the diagonal does not feature an oscillation, the top figure clearly shows that there are oscillations (in y -direction only) present over the rarefaction, also in the diagonal part. Displays of the data along several lines confirm especially the symmetry of the effect exactly in the directions of both the x - and y -axis, thus we only show one plot of this. Within the Figure 21, we also show the effect of using different coefficients of numerical viscosity, *i.e.*, here we display the plots of values along the x -axis for $q = 0.9$ and $q = 0.6$, respectively. As in the scalar case discussed before, a choice $q < 1$ immediately introduces a dissipation effect with respect to oscillations, confirming the results of the experiments presented in the previous section.

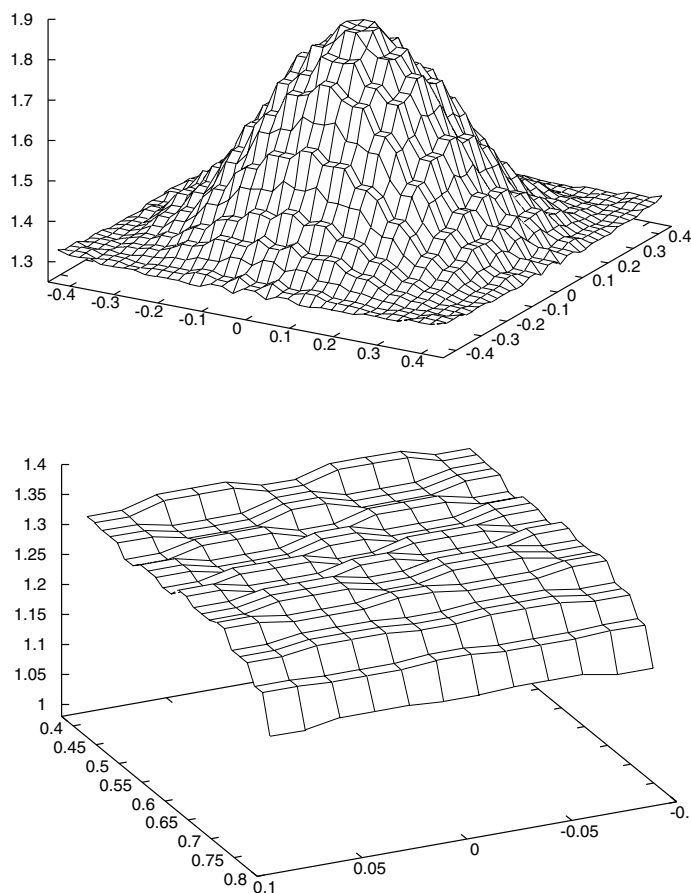


FIGURE 20. Zooms of parts of the solution shown in the total in Figure 19: at the top the remaining part of the initial water pillar, at the bottom a part of the rarefaction following the shock.

3. CONCLUSIVE REMARKS

We have analyzed in detail the influence of the viscosity coefficient on the occurrence and the decay of oscillations arising at data extrema for several central schemes.

Especially the numerical tests indicate, that any other of the central first order schemes than the Lax-Friedrichs scheme strongly damps oscillations if these arise. This adds an important information about the actual behaviour of the investigated schemes and confirms the developed theoretical statements.

Another important property of the investigated central schemes is concerned with the staggered grid approach: It turns out, that the use of a staggered grid is not only a construction step as in [15], it is also to an significant extent responsible for the stability of the scheme. This is of importance for a construction of higher-order non-staggered central schemes which does not follow the mentioned recipe presented in [7]. The construction of such schemes is surely an issue for future developments because non-staggered grids are advantageous if complex geometries and boundary conditions come into play.

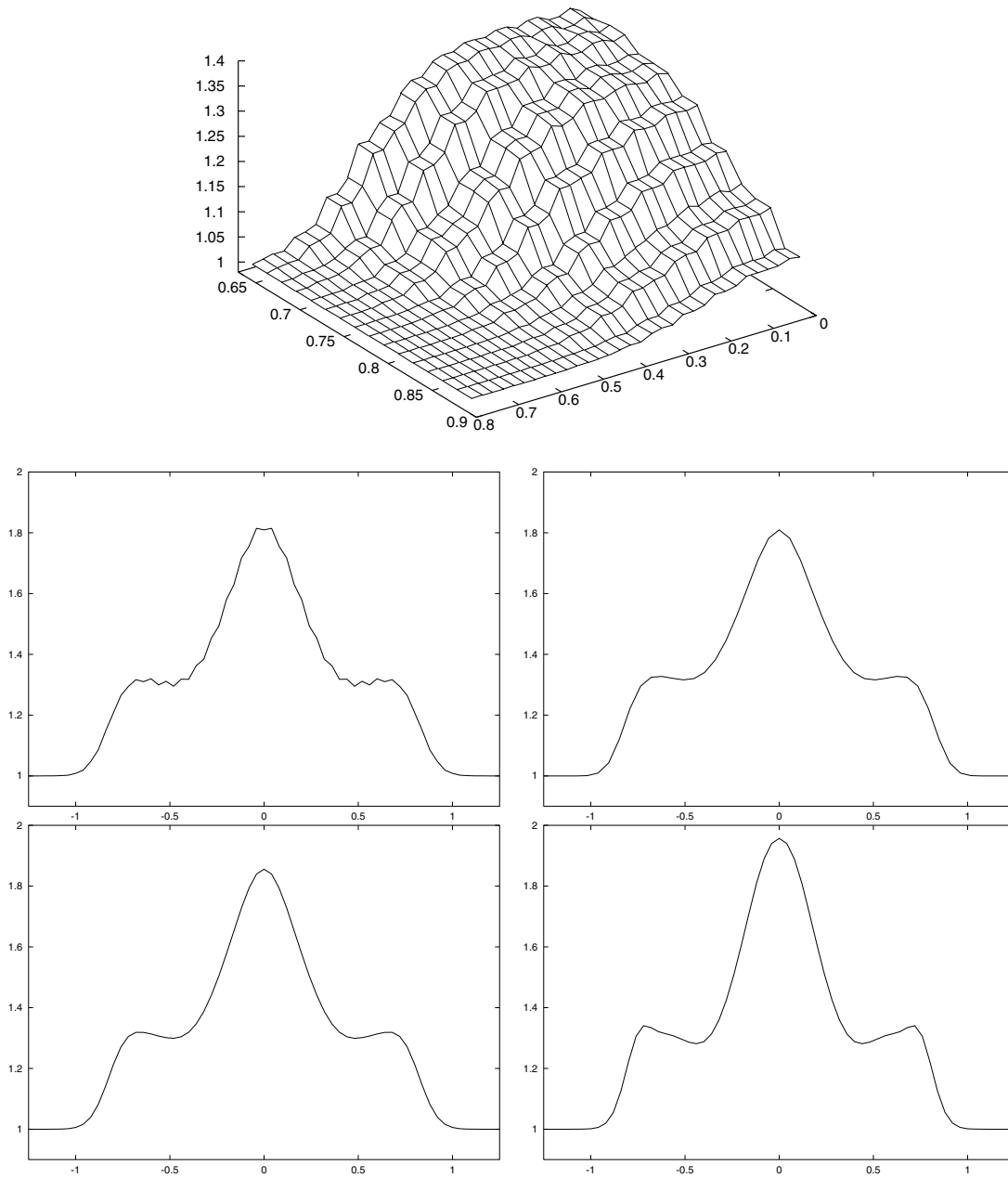


FIGURE 21. This figure shows at the top a zoom on the diagonal part of the solution obtained for $q = 1$. The bottom quartet of figures is arranged from left to right and top to bottom as: (a) plot of the values along the x - or y -axis, (b) plot of the values along a diagonal, (c) plot of the values along the x -axis for $q = 0.9$ and (d) the same for $q = 0.6$.

REFERENCES

- [1] M. Breuß, The correct use of the Lax-Friedrichs method. *ESAIM: M2AN* **38** (2004) 519–540.
- [2] L. Evans, *Partial Differential Equations*. American Mathematical Society (1998).
- [3] E. Godlewski and P.-A. Raviart, *Hyperbolic systems of conservation laws*. Edition Marketing (1991).
- [4] E. Godlewski and P.-A. Raviart, *Numerical Approximation of Hyperbolic Systems of Conservation Laws*. Springer Verlag, New York (1996).
- [5] A. Harten, On a class of high order resolution total variation stable finite difference schemes. *SIAM J. Numer. Anal.* **21** (1984) 1–23.
- [6] G.-S. Jiang and E. Tadmor, Non-oscillatory central schemes for multidimensional hyperbolic conservation laws. *SIAM J. Sci. Comput.* **19** (1998) 1892–1917.
- [7] G.-S. Jiang, D. Levy, C.T. Lin, S. Osher and E. Tadmor, High-resolution nonoscillatory central schemes with nonstaggered grids for hyperbolic conservation laws. *SIAM J. Numer. Anal.* **35** (1998) 2147–2168.
- [8] S. Jin and Z. Xin, The relaxation scheme for systems of conservation laws in arbitrary space dimension. *Comm. Pure Appl. Math.* **45** (1995) 235–276.
- [9] P.D. Lax, Weak solutions of nonlinear hyperbolic equations and their numerical approximation. *Comm. Pure Appl. Math.* **7** (1954) 159–193.
- [10] P.G. Lefloch and J.-G. Liu, Generalized monotone schemes, discrete paths of extrema, and discrete entropy conditions. *Math. Comp.* **68** (1999) 1025–1055.
- [11] R.J. Leveque, *Numerical Methods for Conservation Laws*. Birkhäuser Verlag, 2nd Edition (1992).
- [12] R.J. Leveque, *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press (2002).
- [13] D. Levy, G. Puppo and G. Russo, Compact central WENO schemes for multidimensional conservation laws. *SIAM J. Sci. Comput.* **22** (2000) 656–672.
- [14] X.D. Liu and E. Tadmor, Third order nonoscillatory central schemes for hyperbolic conservation laws. *Numer. Math.* **79** (1998) 397–425.
- [15] H. Nessyahu and E. Tadmor, Non-oscillatory central differencing for hyperbolic conservation laws. *J. Comput. Phys.* **87** (1990) 408–436.
- [16] E. Tadmor, Numerical viscosity and the entropy condition for conservative difference schemes. *Math. Comp.* **68** (1984) 1025–1055.
- [17] H. Tang and G. Warnecke, A note on $(2k + 1)$ -point conservative monotone schemes. *ESAIM: M2AN* **38** (2004) 345–358.