Numerical errors generated in simulations of slowly moving shocks

By E. Johnsen and S. K. Lele

1. Motivation and objectives

An accurate treatment of shockwaves is critical in computational fluid dynamics applications, because numerical errors may propagate in the flow field and contaminate the solution. Because of its simplicity and robustness, shock capturing has become over the past decades the most common means for simulating flows with shockwaves in a stable fashion. However, an important drawback of many commonly used shock-capturing schemes is the generation of errors in computations of slowly moving shocks. Such errors may be particularly detrimental in problems in which small fluctuations interact with shockwaves, e.g., in aero-acoustics or turbulence.

This problem was first identified by Woodward & Colella (1984) and studied systematically by Roberts (1990). Depending on the shock speed relative to the grid and on the numerical scheme, errors and oscillations may occur and affect the solution. Jin & Liu (1996) provided an explanation for the occurrence of start-up errors with the Lax-Friedrichs family of solvers, while Arora & Roe (1997) and Karni & Canic (1997) considered other solvers and observed post-shock oscillations. One means of fixing this problem is to introduce additional dissipation (Colella & Woodward 1984; Lin 1995; Karni & Canic 1997; Stiriba & Donat 2003). However, it is not clear how much dissipation is required.

Though such problems related to slowly moving shocks have been known for more than 20 years, the fundamental reason for the occurrence of such errors is not well understood. The goal of the present report is to better understand this so-called slow shock problem, so that one can determine flow regimes for which this issue is likely to be detrimental, and to provide a means to reduce such errors.

2. Numerical methods

For simplicity, the 1-D isothermal Euler equations are considered. If the temperature is assumed constant, the sound speed, \( a \), is constant, and \( dp = a^2 dp \), where \( p \) is the pressure and \( \rho \) is the density. As a result, the Euler equations become:

\[
q_t + f(q)_x = 0, \quad q = \left( \begin{array}{c} \rho \\ \rho u \\ \rho u^2 + \rho a^2 \end{array} \right), \quad f = \left( \begin{array}{c} \rho u \\ \rho u^2 + \rho a^2 \end{array} \right),
\]

(2.1)

where \( u \) is the velocity. The eigenvalues of the flux Jacobian are \( \lambda^{(1)} = u - a \) and \( \lambda^{(2)} = u + a \), with right eigenvectors

\[
r^{(1)} = \left( \begin{array}{c} 1 \\ u - a \end{array} \right), \quad r^{(2)} = \left( \begin{array}{c} 1 \\ u + a \end{array} \right).
\]

(2.2)
A difference between two states (e.g., across a shock) can then be expressed

$$\Delta q = q_R - q_L = \sum_p \alpha_1^{(p)} r_1^{(p)},$$  \hspace{1cm} (2.3)$$

where the wave strengths are then given by:

$$\alpha_1^{(1)} = \frac{\lambda_1^{(2)} \Delta \rho - \Delta (\rho u)}{2a}, \quad \alpha_1^{(2)} = -\frac{\lambda_1^{(1)} \Delta \rho + \Delta (\rho u)}{2a}.$$  \hspace{1cm} (2.4)$$

The characteristic variables are

$$C^\mp = \ln \rho \mp \frac{u}{a}.$$  \hspace{1cm} (2.5)$$

Without loss of generality, consider a right-moving shock \((s > 0)\) with compression (i.e., density gradient) pointing in the negative \(x\)-direction, as shown in Fig. 1. The state ahead of the shock is the right state and that behind the shock is the left state. The isothermal Euler equations can be integrated over a control volume enclosing the shock in a stationary frame of reference to yield the following Rankine-Hugoniot conditions:

$$-s \Delta \rho + \Delta (\rho u) = 0, \quad -s \Delta (\rho u) + \Delta (\rho u^2 + \rho a^2) = 0.$$  \hspace{1cm} (2.6)$$

The velocity behind the shock is given by

$$u_L = u_R + a \sqrt{\frac{\rho_R}{\rho_L}} \left( \frac{\rho_L}{\rho_R} - 1 \right),$$  \hspace{1cm} (2.7)$$

and the shock velocity by

$$s = u_R + a \sqrt{\frac{\rho_L}{\rho_R}}.$$  \hspace{1cm} (2.8)$$

The shock Mach number is defined as

$$M = \sqrt{\frac{\rho_L}{\rho_R}},$$  \hspace{1cm} (2.9)$$

so that, given \(M\) and \(\rho_R\) and choosing \(s\), the initial conditions can be defined:

$$\rho_L = \rho_R M^2, \quad u_R = s - aM, \quad u_L = u_R + a \left( M - \frac{1}{M} \right).$$  \hspace{1cm} (2.10)$$

For the present work, base values of \(M = 2\) and \(\rho_R = 1\) are selected and \(s\) is chosen; the remaining quantities are computed to satisfy Eqs. (2.10). It can then be shown that the eigenvalues depend only on \(s\) and \(M\). The first eigenvalues can be written:

$$\lambda_1^{(1)} = u_L - a = a \left( 1 + \frac{1}{M} \right), \quad \lambda_1^{(2)} = u_R - a = a(M + 1).$$  \hspace{1cm} (2.11)$$

For the present case, \(s\) is small and positive, so that \(\lambda_1^{(1)}\) and \(\lambda_1^{(2)}\) are both negative. The
Numerical errors generated in simulations of slowly moving shocks

For the Lax-Friedrichs scheme, 
\[ \lambda_L^{(2)} = u_L + a = s + a \left( 1 - \frac{1}{M} \right), \quad \lambda_R^{(2)} = u_R + a = s - a(M - 1). \]  
Thus, \( \lambda_L^{(2)} > 0 \) always, but \( \lambda_R^{(2)} \) may be negative if \( u_R < -a \) or \( s < a(M - 1) \).

For simplicity, a finite volume approach is considered with a first-order accurate reconstruction and third-order accurate Total Variation Diminishing Runge-Kutta time-marching (with constant \( \Delta t/\Delta x \)). Four different approximate Riemann solvers are considered: Lax-Friedrichs, Rusanov, HLL and Roe.† The first two schemes are essentially artificial diffusion methods in which an explicit diffusive term is added, while the latter two methods are based on upwinding, for which the sign of the eigenvalues matters. As shown below, all the schemes can be written as a central difference plus a diffusive term, whose artificial viscosity coefficient may vary but which goes to zero as the grid is refined, i.e.,

\[
\rho u_{i+1/2} = \frac{1}{2}(\rho u_i + \rho u_{i+1}) - c_{i+1/2}^{(\rho)}(\rho_{i+1} - \rho_i) - c_{i+1/2}^{(\rho u)}(\rho u_{i+1} - \rho u_i),
\]

\[
(\rho u^2 + \rho a^2)_{i+1/2} = \frac{1}{2} [(\rho u^2 + \rho a^2)_i + (\rho u^2 + \rho a^2)_{i+1}] - m_{i+1/2}^{(\rho)}(\rho_{i+1} - \rho_i) - m_{i+1/2}^{(\rho u)}(\rho u_{i+1} - \rho u_i),
\]

where \( c \) and \( m \) are the artificial viscosity coefficients. The diffusive term can be written more concisely in matrix form:

\[
\begin{pmatrix}
    c_{i+1/2}^{(\rho)} \\
    m_{i+1/2}^{(\rho)}
\end{pmatrix}
\begin{pmatrix}
    \rho_{i+1} - \rho_i \\
    \rho u_{i+1} - \rho u_i
\end{pmatrix}.
\]

The Lax-Friedrichs and Rusanov (or Local Lax-Friedrichs) fluxes are defined as

\[
f_{i+1/2} = \frac{1}{2}(f_i + f_{i+1}) - \frac{\beta_{i+1/2}}{2}(q_{i+1} - q_i).
\]

For the Lax-Friedrichs scheme, \( \beta \) is a global value (i.e., the largest eigenvalue over the whole domain):

\[
\beta_{i+1/2} = \max_{i,p} |\lambda_i^{(p)}|.
\]

For the Rusanov scheme, a local value of \( \beta \) is utilized:

\[
\beta_{i+1/2} = \max_p (|\lambda_i^{(p)}|, |\lambda_{i+1}^{(p)}|).
\]

Thus, the numerical fluxes for Lax-Friedrichs schemes can be written in artificial diffusion form:

\[
(\rho u)_{i+1/2} = \frac{1}{2} [(\rho u)_i + (\rho u)_{i+1}] - \frac{\beta_{i+1/2}}{2}(\rho_{i+1} - \rho_i),
\]

\[
(\rho u^2 + \rho a^2)_{i+1/2} = \frac{1}{2} [(\rho u^2 + \rho a^2)_i + (\rho u^2 + \rho a^2)_{i+1}] - \frac{\beta_{i+1/2}}{2} [(\rho u)_{i+1} - (\rho u)_i].
\]

The artificial viscosity matrix is thus diagonal with equal positive coefficients.

† Listed from most dissipative to least dissipative. For the isothermal equations, Roe and HLL are virtually identical; the proof is not included here for conciseness.
The Harten-Lax-van Leer (HLL) flux is defined as
\[ f_{i+1/2} = \frac{s_{i+1/2}^+ f_i - s_{i+1/2}^- f_{i+1} + s_{i+1/2}^+ s_{i+1/2}^- (q_{i+1} - q_i)}{s_{i+1/2}^+ - s_{i+1/2}^-}, \]  
(2.19)
where the wave speeds are given by
\[ s_{i+1/2}^- = \min(0, \lambda_i^{(1)}), \quad s_{i+1/2}^+ = \max(0, \lambda_i^{(2)}), \]  
(2.20)
The tilde represents Roe-averaged values (see below). The numerical fluxes for the HLL scheme can then be written in artificial diffusion form:
\[ (pu)_{i+1/2} = \frac{\rho u_i + (pu)_{i+1}}{2} - \frac{s_{i+1/2}^+ s_{i+1/2}^-}{s_{i+1/2}^+ - s_{i+1/2}^-} (\rho u_{i+1} - \rho_i) - \frac{1}{2} \left( s_{i+1/2}^+ + s_{i+1/2}^- \right) \left[ (pu)_{i+1} - (pu)_i \right], \]  
(2.21a)
\[ (pu^2 + \rho a^2)_{i+1/2} = \frac{\rho u^2_i + (pu^2 + \rho a^2)_{i+1}}{2} - \frac{\lambda_i^{(1)} + \lambda_i^{(2)}}{2} \frac{s_{i+1/2}^+ s_{i+1/2}^-}{s_{i+1/2}^+ - s_{i+1/2}^-} \left( \rho u_{i+1} - \rho_i \right) \]  
\[ - \left( \frac{\lambda_i^{(1)} + \lambda_i^{(2)}}{2} \frac{s_{i+1/2}^+ s_{i+1/2}^-}{s_{i+1/2}^+ - s_{i+1/2}^-} \right) \left[ (pu)_{i+1} - (pu)_i \right], \]  
(2.21b)
It should be noted that to obtain Eqs. 2.21 the following property of Roe averages (see below) was used: \( \Delta (pu^2) = \Delta ((pu^2)/\rho) = 2\tilde{u} \Delta (pu) - \tilde{u}^2 \Delta \rho. \)
The Roe flux is defined as
\[ f_{i+1/2} = \frac{1}{2} (f_i + f_{i+1}) - \frac{1}{2} \sum_p \tilde{\alpha}_{i+1/2}^p \tilde{\beta}_{i+1/2}^p \tilde{a}_{i+1/2}^p \]  
(2.22)
Here, the tilde denotes Roe-averaged quantities. For the isothermal Euler equations, only the Roe-averaged velocity is required:
\[ \tilde{u}_{i+1/2} = \sqrt{\frac{\rho u_{i+1} + u_{i+1}}{\rho u_{i+1} + 1}}. \]  
(2.23)
The numerical fluxes for the Roe scheme can then be written in artificial diffusion form:
\[ (pu)_{i+1/2} = \rho u_i + (pu)_{i+1} - \frac{\tilde{\lambda}_{i+1/2}^{(1)} + \tilde{\lambda}_{i+1/2}^{(2)}}{4a} [\rho (u)_{i+1} - \rho_i], \]  
(2.24a)
\[ (pu^2 + \rho a^2)_{i+1/2} = \frac{pu^2_i + (pu^2 + \rho a^2)_{i+1}}{2} - \frac{\tilde{\lambda}_{i+1/2}^{(1)} + \tilde{\lambda}_{i+1/2}^{(2)}}{4a} \frac{\tilde{\lambda}_{i+1/2}^{(1)} + \tilde{\lambda}_{i+1/2}^{(2)}}{4a} [\rho (u)_{i+1} - \rho_i], \]  
(2.24b)
In both the Roe and HLL solvers, the artificial diffusion matrix is full and may have negative components.
Numerical errors generated in simulations of slowly moving shocks

3. Results

3.1. Overall behavior

In order to illustrate the types of problems that occur with slowly moving shocks, Fig. 2 shows density and momentum profiles at different times using the Lax-Friedrichs and Roe solvers for a shock speed, \( s = 0.1 \); \( N = 101 \) points are used. The solution exhibits rather different features, depending on the scheme. In the Lax-Friedrichs solution, an undershoot (\( \text{spike} \), as defined by Jin & Liu 1996) develops in the momentum and moves with the shock. This undershoot causes an excess of momentum to propagate downstream of the shock. Though the amplitude of the spike is larger than that of the downstream overshoot of momentum, the momentum deficit and excess are equal, by conservation. In addition, the excess momentum propagating downstreaming causes a mass deficit to propagate as well. The shock then reaches a \( \text{steady-state} \) profile in which the density is smeared and the momentum, though also smeared, exhibits a spike. This scheme is very dissipative, as the shock is smeared over approximately 30 points. In the Roe solution, a large initial spike is generated in the momentum, as well as an overshoot downstream of the shock. Then, further unsteady spikes develop periodically in the momentum profile and propagate with the shock. Oscillations propagating downstream of the shock are constantly generated; the amplitude of these oscillations decreases as they propagate downstream because of the low order of the numerical scheme. Similar features are observed in the density profile, though they are less prominent. There is only one intermediate point in the shock; the \( \text{thickness} \) of the spike and the amplitude of the errors are smaller than in the Lax-Friedrichs case.
Figure 3 shows the profiles of the characteristic variables at different times for the Lax-Friedrichs and Roe solvers. Again, it is clear that the solution exhibits different features, depending on which solver is used. The errors observed in the density and momentum profiles are also present in the characteristic variables. In both solvers, the downstream-propagating errors affect the $C^-$ characteristic variable. Though less obvious, the $C^+$ variable is also affected, but the error is only evident in the smearing of the profile; the $C^+$ variable is of the same characteristic family as the shock.

3.2. Classification of the errors

In the previous section, two main types of errors are observed:

(a) A start-up error in the form of an initial spike in the momentum (with a corresponding downstream-propagating wave), which occurred with both solvers, and

(b) Unsteady post-shock oscillations, which are observed only with the Roe scheme.

First, the formation of the momentum spike is considered, following a traveling wave analysis suggested by Jin & Liu (1996). The viscous isothermal equations are considered:

\[
\frac{\partial \rho}{\partial t} + \frac{\partial (\rho u)}{\partial x} = \epsilon \frac{\partial^2 \rho}{\partial x^2}, \tag{3.1a}
\]

\[
\frac{\partial (\rho u)}{\partial t} + \frac{\partial}{\partial x} \left( \frac{(\rho u)^2}{\rho} + \rho a^2 \right) = \epsilon \frac{\partial^2 (\rho u)}{\partial x^2}. \tag{3.1b}
\]
Numerical errors generated in simulations of slowly moving shocks

Letting \( \xi = x - st \), and with asymptotic states,

\[
\begin{align*}
\rho(x \to -\infty, t) &= \rho_L, & \rho(x \to \infty, t) &= \rho_R, \\
\rho u(x \to -\infty, t) &= \rho u_L, & \rho u(x \to \infty, t) &= \rho u_R, \\
\rho_x(x \to -\infty, t) &= \rho_x(x \to \infty, t) = 0, & \rho u_x(x \to -\infty, t) &= \rho u_x(x \to \infty, t) = 0,
\end{align*}
\]

the viscous isothermal equations then reduce to

\[
\begin{align*}
-s\rho' + (\rho u)' &= \epsilon \rho'', \\
-s(\rho u)' + (\rho u^2 + \rho a^2)' &= \epsilon(\rho u)''.
\end{align*}
\]

Integrating the first equation, the momentum can be computed:

\[
\rho u = \rho u_L + s(\rho - \rho_L) + \epsilon \rho'.
\]

Figure 4. Analysis of momentum spike with the Lax-Friedrichs solver for \( s = 0.1 \).

In general, the diffuse density profile does not change significantly with \( s \). However, when \( s \) is large enough, the term \( s \rho \) overwhelms \( \epsilon \rho' \), so that the spike is no longer prominent. It should be noted that \( \epsilon \) depends on \( s \) through Eq. (2.11), though weakly for small \( s \). Thus, this analysis shows that, though the spike moving with the shock cannot be prevented, the initial downstream-propagating wave can be removed in the Lax-Friedrichs solver by smearing the density and choosing an appropriate momentum profile according to Eq. (3.4).
Next, the occurrence of unsteady post-shock oscillations is investigated. Figure 5 shows momentum profiles generated with the Roe scheme, with \( s = 0.1 \) and \( N = 100 \). For these parameters, the shock takes 0.2 time units to move across one cell, and \( u_L - a = -1.4 \). The generation of the downstream-propagating post-shock oscillations is clearly periodic. In the left plot, it can be seen that the profiles at times 0.03, 0.05 and 0.07 are almost identical and separated by \( \Delta x = 0.02 \); the same holds for the profiles at times 0.04, 0.06 and 0.08. Furthermore, from the right plot, the distance between the peaks (denoted by the arrows) at \( t = 0.30, 0.35 \) and 0.40 is approximately 0.07, so that the velocity at which these oscillations propagate is 0.07/0.05 = 1.4. Thus, this analysis highlights the fact that the period of oscillation shedding is equal to the time it takes for the shock to move across one cell and that these oscillations propagate downstream at a velocity of \( u - a \); in other words, the wavelength of the oscillations is \( \Delta x(u_L - a)/s \).

To further confirm the period of the oscillation shedding, Fig. 6 shows momentum profiles for \( N = 50, 100 \) and 200 at times 1.6, 0.8 and 0.4, respectively. The Roe solver is used with a constant CFL, and \( s = 0.1 \). Though the profiles seem quite different for different \( N \), if the \( N = 100 \) grid is stretched by a factor of 2 and the \( N = 200 \) grid is stretched by a factor of 4, all the profiles collapse onto the \( N = 50 \) curve. This further implies that the amplitude of the oscillations does not decrease with grid refinement for a fixed CFL, though the wavelength of the oscillations does.

As observed previously, the Lax-Friedrichs scheme does not generate such oscillations. The main difference with the Roe scheme is that the coefficients of the artificial viscosity matrix may be non-positive. Figure 7 shows the initial artificial viscosity coefficients for the Roe scheme, as a function of the shock speed. Three states are defined: left, Roe-averaged and right. In the continuity equation, the artificial viscosity coefficients for the momentum are mostly negative, while the coefficients for the density are greater or equal to zero. In the momentum equation, the coefficients are mostly positive, though the density coefficients, at small \( s \), may become negative. Post-shock oscillations are observed in the region between the dashed lines in the figure (discussed in the next section).

### 3.3. Fix and characterization of the errors

In the previous section, it was observed that no post-shock oscillations other than the initial start-up error are generated with the Lax-Friedrichs solver; however, the shock profile diffuses over a wide region. On the other hand, the Roe scheme introduces little dissipation, as the shock is spread over just a few points, but post-shock oscillations are
Numerical errors generated in simulations of slowly moving shocks

In order to prevent oscillations, Karni & Canic (1997) proposed a modified entropy fix using the Roe scheme. The present strategy is to control the artificial viscosity coefficients to avoid excessive dissipation and prevent oscillations. In other words, the goal is to regularize the artificial viscosity term, which is shown in Fig. 8 for the Lax-Friedrichs and Roe solvers, and confine it to a small region.

To provide the most flexibility, the HLL solver is considered, for which:

\[ c^{(\rho)} = \frac{-s^- s^+}{s^+ - s^-} \]
\[ m^{(\rho)} = -\lambda(1) \lambda(2) c^{(\rho)} \]
\[ c^{(\rho u)} = \frac{s^+ + s^-}{s^+ - s^-} \]
\[ m^{(\rho u)} = (\lambda(1) + \lambda(2)) c^{(\rho u)} + c^{(\rho)} \]  

In the case of a slowly moving shock, \( s^+ = 0 \) ahead of the shock, so that \( c^{(\rho)} = 0 \) and \( c^{(\rho u)} = -1/2 \), so that \( m^{(\rho)} > 0 \) and \( m^{(\rho u)} > 0 \). The quantity, \( s^+ = 0 \), is expected to lead to problems, so this quantity is varied in the viscosity coefficients. Figure 9 shows momentum profiles for different global values of \( s^+ \) and compares the optimal global value with a local scheme. A value of \( s^+ = 0 \) corresponds to the original HLL solver for this problem. It can be observed that large values of \( s^+ \) exhibit more dissipative features,
particularly at the spike. Decreasing the value of $s^+$ reduces the dissipation, but the cost for this is the presence of post-shock oscillations. The value of $s^+$ that exhibits the least amount of dissipation, while preventing downstream oscillations, is $s^+ = 0.6 = \lambda_L^{(2)}$. At this stage, it is not clear why this particular value prevents oscillations. In order to reduce the dissipation further, a localized version can be used; such a formulation is necessary when more complex problems are considered. It is noticed that such a fix is required near the shock and only when $\lambda_{j+1}^{(2)} < 0$. Since shocks are typically spread over approximately five cells, the value of $s^+ = \lambda_L^{(2)}$ is set only when:

$$\lambda_j^{(2)} \lambda_{j+1}^{(2)} < 0, \quad j \in [i - 2, i + 2].$$

(3.6)

Though not shown here explicitly, it is noted that a bound is required on the wave speed in the artificial diffusion of density in the continuity equation only. Thus, the fix is required in this field only. This final observation thus raises the question, how much dissipation is required to simulate shocks without generating oscillations?

For practical applications, it is of interest to know the extent of the errors related to slowly moving shocks and the range of $s$ over which they are expected to affect the flow. For simplicity, the errors are computed after the initial transient has occurred. The error
Numerical errors generated in simulations of slowly moving shocks

The error in density is normalized by the jump in the relevant quantity, as follows:

\[ \Delta f = \frac{f_{\text{exact}} - f_{\text{computed}}}{|f_L - f_R|}. \]  

(3.7)

Thus, a positive value implies an undershoot and a negative value implies an overshoot. Figure 10 shows the overshoot in density, the overshoot in momentum downstream of the shock and the undershoot of the spike as a function of the shock speed for \( M = 2 \). The Roe solver, the Roe solver with the fix by Karni & Canic (1997), and the present HLL fix are included. The error in density is not very large in general for this shock strength (up to 1%). The fix proposed by Karni & Canic (1997) works well, and the present fix even better. Though not shown here, the errors increase with increasing \( M \). As an example, the errors with the fix by Karni & Canic (1997) become larger for \( M = 4 \), while with the present fix the error still remains negligible. Similar findings are observed in the momentum overshoot, but in which the error without the fixes can reach 30%. As mentioned in the previous sections, the occurrence of the spike cannot be avoided unless the initial conditions are regularized. For small shock speeds, this error can be extremely large (up to 300%) for all three solvers.

4. Conclusions and future work

In the present report, the errors generated by shocks moving slowly compared to the grid are analyzed. The errors are divided into two classes: start-up errors and post-shock oscillations. It is shown that the traveling wave analysis of Jin & Liu (1996) describes the spike formation in the momentum for the Lax-Friedrichs solver very well. By smearing the density and momentum profiles appropriately, the corresponding downstream-propagating wave can be removed. The post-shock oscillations have been characterized; as a fix for this problem, a lower bound on the wave speed of the HLL solver is set in the artificial viscosity coefficient for density in the continuity equation.

Though not presented here, an extension of the present HLL fix to multi-dimensions is trivial. However, higher-order accurate schemes, such as WENO, exhibit severe post-shock oscillations, even with the Lax-Friedrichs solver. Future work includes providing a fix for such cases and studying interactions between slowly moving shocks and entropy or acoustic waves to understand the effect of the post-shock oscillations on the flow.
field. Finally, this analysis raises the following critical question, how much dissipation is required to simulate shocks accurately?

Acknowledgements

The author is grateful for discussions with Dr. Johan Larsson and Dr. Soshi Kawai, and for comments by Dr. Kazuo Matsuura on the manuscript.

REFERENCES


