

On sub-linear convergence for linearly degenerate waves in capturing schemes

J.W. Banks^{a,c,*}, T. Aslam^{b,3}, W.J. Rider^{c,2}

^a Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

^b Shock and Detonation Physics Group, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

^c Computational Sciences Research Institute, Sandia National Laboratory, Albuquerque, NM 87185, USA

Received 19 December 2007; received in revised form 20 March 2008; accepted 3 April 2008

Available online 14 April 2008

Abstract

A common attribute of capturing schemes used to find approximate solutions to the Euler equations is a sub-linear rate of convergence with respect to mesh resolution. Purely nonlinear jumps, such as shock waves produce a first-order convergence rate, but linearly degenerate discontinuous waves, where present, produce sub-linear convergence rates which eventually dominate the global rate of convergence. The classical explanation for this phenomenon investigates the behavior of the exact solution to the numerical method in combination with the finite error terms, often referred to as the modified equation. For a first-order method, the modified equation produces the hyperbolic evolution equation with second-order diffusive terms. In the frame of reference of the traveling wave, the solution of a discontinuous wave consists of a diffusive layer that grows with a rate of $t^{1/2}$, yielding a convergence rate of $1/2$. Self-similar heuristics for higher-order discretizations produce a growth rate for the layer thickness of $\Delta t^{1/(p+1)}$ which yields an estimate for the convergence rate as $p/(p+1)$ where p is the order of the discretization. In this paper we show that this estimated convergence rate can be derived with greater rigor for both dissipative and dispersive forms of the discrete error. In particular, the form of the analytical solution for linear modified equations can be solved exactly. These estimates and forms for the error are confirmed in a variety of demonstrations ranging from simple linear waves to multidimensional solutions of the Euler equations.

© 2008 Elsevier Inc. All rights reserved.

Keywords: Convergence rate; Modified equations; Discontinuity

* Corresponding author. Address: Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA. Tel.: +1 925 667 6727.

E-mail addresses: banks20@llnl.gov, banksj@alum.rpi.edu (J.W. Banks).

¹ This study has been supported by Lawrence Livermore National Laboratory under the auspices of the U.S. Department of Energy through Contract No. DE-AC52-07NA27344.

² This study has been supported by Sandia National Laboratory and performed under the auspices of the U.S. Department of Energy through Contract No. DE-AC04-94AL85000.

³ This study has been supported by Los Alamos National Laboratory and performed under the auspices of the U.S. Department of Energy through Contract No. DE-AC52-06NA25396.

1. Introduction

Simulations of inviscid compressible flows typically result in the computation of an advected discontinuity. Such advected discontinuities are typified by the simple contact discontinuity, but other examples, such as slip lines, also exist. In any case, these discontinuities have the property that the governing equations are linear in a neighborhood of the jump and so they are often referred to as “linearly degenerate” waves [1]. The result of this linearity is that characteristic curves run parallel to each other as well as the discontinuity, and so any error made in the computation of such a wave will persist throughout the simulation for all time. To contrast this, fully nonlinear waves such as shocks, have a natural steepening mechanism which causes the characteristics to enter back into the wave. This can also be viewed as the natural steepening mechanism inherent to shocks. The result is that errors made in the computation of shocks can be absorbed back into the discontinuity and dissipated, thus the idiom that shock waves are “numerical garbage collectors”.

These factors become important when linear jumps are “captured” rather than “tracked”. Certainly for Eulerian schemes [2,3], all waves are captured and so the connection is clear. Lagrangian schemes [4–6] are often used to circumvent the issues associated with captured linear jumps by performing computations in the frame of the fluid, but most interesting simulations require, at the very least, mesh remap which results in the so-called arbitrary-Lagrangian–Eulerian (ALE) schemes [7]. The use of such remap has the potential to cause ALE to schemes suffer the same fate as purely Eulerian schemes although the details may depend on the frequency at which remapping is performed. For many problems then, the convergence rates for ALE methods near linear jumps are limited in a similar way as the convergence rates for Eulerian schemes.

Material interfaces, contact surfaces, and slip lines are all examples of linearly degenerate jumps and are present in most physically interesting simulations. To make matters worse, such jumps can arise spontaneously in the flow. This can happen for example through a simple Riemann problem [8], the interaction of two shocks, or even geometric irregularities interacting with other nonlinear waves [9]. On the other hand, the importance of accurate treatment of all discontinuities, including linear jumps, is paramount in the overall efficacy of a given simulation. The difficulties associated with the computation of discontinuous solutions has been the driver behind such developments as high-resolution methods [10–12] and adaptive mesh refinement [13,14].

Abundant mechanisms to propagate and generate discontinuous solutions, combined with the requisite accuracy requirements for effective simulations, makes the need to have a detailed understanding of the expected convergence behavior of numerical methods near those jumps clear. Of course this applies to linear as well as nonlinear jumps, and this paper focuses on the linear variety. Our focus on linear waves stems from the ironic fact that the linear case is, in some sense, more difficult. Because nonlinear waves can reprocess numerical error while linear waves cannot, the convergence rates for numerical methods near linear jumps are necessarily lower. Thus the convergence character for quantities which jump through linear waves, and potentially for quantities that are coupled to those jumping through linear waves, is ultimately driven by the convergence character near the linear jump and so in this sense a complete understanding for linear jumps is more critical. Furthermore, experience tells us that the convergence character near linear jumps is the primary difficulty. Of course, we do not write off the need to adequately understand behavior near nonlinear jumps, but this is a topic for another paper.

For linear jumps, we posit that the general L_1 order of convergence for linear p th order capturing schemes is $p/(p + 1)$. Recall that capturing schemes are generally non-convergent in the L_∞ norm because at jumps, $O(1)$ errors are always present. This order of convergence is demonstrated rigorously by considering the truncated modified equation (ME) which is solved exactly. From this solution one computes the error and determines the advertised convergence rate. For limited schemes, sometimes called high-resolution schemes or nonlinear schemes, the story is slightly more complex. For non-compressive limiting strategies, such as MinMod, the convergence rate is identical to the unlimited case. In fact the analytic solution to the ME for the second-order MinMod limited scheme is derived and convergence is demonstrated at $2/3$. For compressive limiting strategies, such as double MinMod or SuperBee, the modified equation is shown to contain anti-diffusive modes. Such modes are inherently linearly unstable and the scheme relies on the limiting strategy to maintain overall solution stability. These un-physical anti-diffusive terms provide an artificial steepening mechanism for linear jumps which can result in first-order convergence.

The analysis presented in this paper is in many ways a generalization of what has been presented by many authors before. In [15], Hedstrom presents an analysis of general two level difference schemes and shows how the simple requirement of stability necessitates convergence at order $p/(p+1)$. This pioneering analysis demonstrated the fundamental behavior but failed to give any quantitative description of solution character other than asymptotic convergence rates. Furthermore, the convergence character of compressively limited schemes, which are TVD and hence stable, is not included in this description. This was followed in [16] by Orszag and Jayne where the same convergence character was shown. Soon after, Chin [17,18] derived the exact solution for the ME of a second-order linear scheme as the integral of the Airy function. This valuable addition hinted toward the possibility of solving the ME for arbitrary-order schemes and even the possibility of solving the ME for limited schemes as well. Then during the development of the artificial compression method (ACM) [19], Harten states the fundamental desire to produce difference schemes which can potentially achieve a steady state for linear jumps. Standard schemes are shown to diffuse as a function of time and so he goes on to develop ACM as an alternative. Such schemes are close brethren to modern high-resolution schemes which employ compressive limiting strategies and so can potentially achieve higher rates of convergence. The slow smearing of linear jumps is cited in both [19,20] as the cause of reduced convergence and the rate of smearing is used to devise the same rate $p/(p+1)$. Also in [19] it is shown that the general solution to the ME of a linear scheme for discontinuous data should be representable as a similarity solution. This paper extends the analysis through the next logical step. We provide closed form solutions for the MEs of linear approximations, which were previously lacking. These solutions demonstrate many of the previously known results. We also extend our solutions to limited high-resolution approximations and shows why compressive schemes achieve higher convergence rates. Furthermore, this analysis gives insight into high-resolution schemes showing how second-order limited schemes retain the use of high-order stencils at jumps, and why monotone third-order approximations are not possible.

The rest of this paper is organized as follows. In Section 2 the governing equations and simple numerical schemes are presented. Section 3 presents the essential difficulty and demonstrates sub-linear convergence for the simplest schemes. A detailed analysis of the methods from Section 3 is presented in Section 4. Here the similarity solution argument is presented for the first-order scheme and this forms the basis for much of what follows. Higher-order schemes are analyzed in Section 5 with special emphasis on the exact solution to the ME for linear and nonlinear second-order schemes. Schemes of order higher than two are also investigated and the convergence character shown to hold. A simple example of how this type of convergence character will be exhibited for more complex, two-dimensional Euler flow is presented in Section 6. Some conclusions are given in Section 7. [Appendix A](#) demonstrates how the analyses from this paper extend also to flows containing corners, that is to say flows with discontinuous first derivatives. Finally [Appendix B](#) gives some mathematical detail concerning the derivation of the exact solution of the ME for the MinMod limited scheme.

2. The model problem

We typically seek to understand the nature of solutions to sets of equations such as the Euler equations, or the equations of ideal magneto-hydrodynamics. Furthermore the typical practical setting is one of two or more dimensions further complicating our understanding. It is fortunate that the particular problem under consideration in this work manifests itself for the one-dimensional linear advection equation

$$\frac{\partial}{\partial t} u(x, t) + a \frac{\partial}{\partial x} u(x, t) = 0 \quad (1)$$

with a being the constant flow velocity (without loss of generality take $a > 0$) in just the same way as for the more complex cases. Eq. (1) is a model for the behavior of linearly degenerate waves in the more complex systems of equations. For instance in regions of uniform velocity and pressure in Euler flows, a simple transformation of the equations into the streamline direction reduces the full set of Eq. (1).

To facilitate the discussion, Eqs. (2) and (3) introduce the first-order explicit and implicit upwind schemes respectively

$$v_i^{n+1} = v_i^n - \lambda [v_i^n - v_{i-1}^n] \tag{2}$$

$$v_i^{n+1} = v_i^n - \lambda [v_i^{n+1} - v_{i-1}^{n+1}] \tag{3}$$

where v_i^n is a numerical approximation to $u(x_i, t^n)$ and the so called CFL number is $\lambda = \frac{a\Delta t}{\Delta x}$. Here, the understanding is that the spatial domain $[x_L, x_R]$ has been discretized as $x_i = x_L + i\Delta x$ where $\Delta x = (x_R - x_L)/(N - 1)$ and N is a parameter measuring grid resolution. Similarly, time has been discretized as $t_n = n\Delta t$ with initial conditions $u(x, 0)$ being given at $t = 0$. Numerical stability dictates that λ be less than one for the explicit case (2) but can be arbitrarily large for the implicit case (3). These are the archetypal schemes to which many others reduce near contact discontinuities thus providing the simplest possible context in which to carry out the discussion to follow. It should be noted that although the situation changes slightly for so called “high-resolution” methods, the critical ideas are the same and the ideas presented during the discussion of (2) and (3) lay the groundwork for the rest of this paper.

A more complete understanding of the behavior of the schemes (2) and (3) can be obtained by viewing their respective modified equations, that is to say the equations for which the approximations are exact [2]. After some manipulation, the modified equation for the explicit scheme (2) is found to be

$$\frac{\partial}{\partial t} u(x, t) + a \frac{\partial}{\partial x} u(x, t) - \frac{a\Delta x}{2} (1 - \lambda) \frac{\partial^2}{\partial x^2} u(x, t) + \dots = 0 \tag{4}$$

and for the implicit scheme (3) it is found to be

$$\frac{\partial}{\partial t} u(x, t) + a \frac{\partial}{\partial x} u(x, t) - \frac{a\Delta x}{2} (1 + \lambda) \frac{\partial^2}{\partial x^2} u(x, t) + \dots = 0. \tag{5}$$

Investigation of these modified equations shows that for $0 \leq \lambda \leq 1$ numerical stability is achieved for the explicit scheme because the modification to the actual advection equation (1) represents diffusion (rather than anti-diffusion). For the implicit scheme, $\lambda \geq 0$ is sufficient to guarantee the modification represents diffusion rather than anti-diffusion.

3. The essential difficulty

To illustrate the fundamental difficulty, the two schemes are employed to compute the solution of a simple advecting discontinuity. Of course the reader will immediately notice that discontinuous solutions are suspect as the differential equation (1) is not defined for solutions containing discontinuities. However, reference to many texts, such as [21,1], brings the concept of weak solutions into the picture and for this case the appropriate weak solution is defined through the method of characteristics. Thus if the initial conditions are given as $u(x, 0) = f(x)$ on an infinite domain then the exact solution at any later time is given by $u(x, t) = f(x - at)$. This construction admits discontinuous solutions and the expectation is that the numerical method will converge to such solutions (see [22,1,3] for details). Fig. 1 illustrates the results of a convergence study at various resolutions using both the explicit and implicit first-order upwind methods.

In this computation, the initial conditions are

$$u(x, 0) = \begin{cases} 1.0 & \text{for } x < 0.2 \\ 0.125 & \text{for } x \geq 0.2 \end{cases}$$

and a zero gradient condition is applied at the boundaries. Computations are carried out at four resolutions $\Delta x = 0.05, 0.025, 0.0125,$ and 0.00625 with a fixed CFL $\lambda = 0.8$. The integration is taken to $t_f = 1$ with advection velocity $a = 1$ and the exact solution is overlaid for comparison to the numerical approximations. From this figure it is clear that both numerical methods are in fact converging, but a more quantitative examination shows the flaw. A typical *a priori* expectation is that a first-order numerical method, such as either upwind schemes under consideration, will converge at first-order with errors computed using the L_1 norm. Recall that the convergence rate is the power of two given by the ratio of the error at a given grid spacing to the error at a grid spacing half that of the original. Table 1 shows the errors for these simulations and the corresponding convergence rates for the two schemes at the resolutions of Fig. 1. While it is gratifying to see that the schemes are definitely converging, it may be alarming to see that both schemes are converging at $O(\Delta x^{1/2})$ rather than

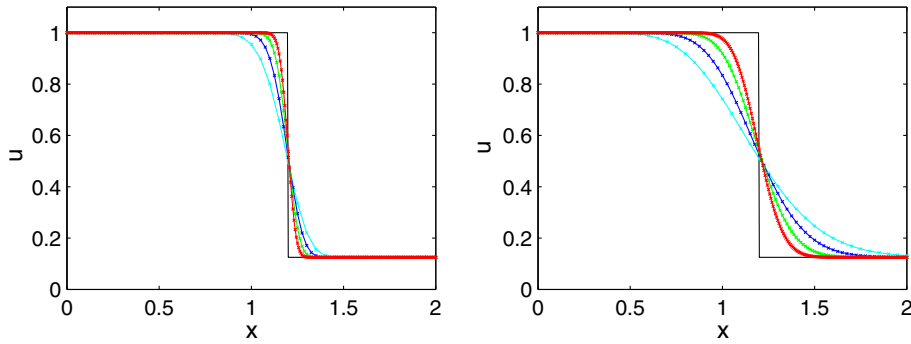


Fig. 1. Convergence of the explicit scheme (left) and implicit scheme (right) for $x \in [0, 2]$, $\Delta x = 2/(N - 1)$ and CFL $\lambda = 0.8$, where $N = 41$ (cyan), 81 (blue), 161 (green) and 321 (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Convergence results for contact wave using the first-order explicit (left) and implicit (right) upwind methods

N	$e_u(N)$ (explicit)	Rate	$e_u(N)$ (implicit)	Rate
41	7.13e-2	–	2.07e-1	–
81	4.99e-2	0.515	1.47e-1	0.494
161	3.51e-2	0.508	1.04e-1	0.499
321	2.46e-2	0.513	7.40e-2	0.491

Shown are L_1 errors at $t = 1$ for grid resolutions determined by N , and the convergence rates at successive resolutions.

the expected $O(\Delta x)$ as $\Delta x \rightarrow 0$. Upon further analysis, this is a byproduct of the fact that the exact solution contains a jump and if the same test is performed for initial conditions defining a smooth transition between states the expected order of convergence, $O(\Delta x)$, will result at some sufficient resolution. Knowledge of this phenomenon is not new and analysis is presented for example in [2,15,16,18–20]. This example simply serves as useful demonstration of this type of behavior.

A natural question to ask is whether this matters for the more complicated equation sets and in multiple dimensions where one is interested in performing computations. All evidence indicates that the answer to this question is yes. For flows governed by the Euler equations one must contend with the formation of discontinuous solutions in finite time as a result of inherent nonlinearities even if the initial data was perfectly smooth. Furthermore, the typical flow one is interested in, contains many such jumps with shocks and contacts interacting in some complicated way. If convergence to a contact is sub-linear and a shock subsequently propagates through it, even if the shock is treated to very high-order it will be infected with low-order errors which force convergence at less than first-order. Ultimately as grids are refined, the smallest convergence rate will dominate the asymptotic rate of convergence for the problem. This state of affairs for linear jumps is, to a large degree, the underlying motivation driving the development of high-order as well as adaptive methods.

4. Analysis of the problem

So what exactly is going on to cause the numerical methods to converge at $O(\Delta x^{1/2})$? This question has been investigated by other authors in a number of ways, but here we seek the answer through a detailed modified equation analysis. If one takes the stance that the modified equation is the PDE which the numerics solve more accurately than the original PDE (presumably a small perturbation to the original problem), then the quest becomes one of understanding the behavior of (4) and (5). Both of these modified equations can be viewed as approximations to the advection–diffusion equation

$$\frac{\partial}{\partial t} \hat{u}(x, t) + a \frac{\partial}{\partial x} \hat{u}(x, t) - v \frac{\partial^2}{\partial x^2} \hat{u}(x, t) = 0, \tag{6}$$

where ν is a viscosity coefficient whose size is dictated by the choice of Δx , the CFL number λ , and the particular scheme (4) or (5). In either case, the same behavior is demonstrated where at fixed λ , a doubling of spatial resolution results in a halving of the numerical value of ν . With the goal of determining an analytic solution for (6), and demonstrating a solution strategy which we will later generalize to high-order and high-resolution approximations, the change of variables

$$\begin{aligned} z &= x - x_0 - at \\ \tau &= t \end{aligned}$$

is introduced which transforms (6) into the familiar heat equation

$$\frac{\partial}{\partial \tau} \hat{u}(z, \tau) - \nu \frac{\partial^2}{\partial z^2} \hat{u}(z, \tau) = 0. \tag{7}$$

On an infinite domain and with discontinuous initial data $u(x, 0) = u_L$ for $x < x_0$ and $u(x, 0) = u_R$ for $x \geq x_0$, one can seek a self similar solution to (7), via the following change of independent variables

$$\zeta(z, \tau) = \frac{z}{\tau^{1/2}}, \tag{8}$$

which reduces the partial differential equation (7) to a second-order ordinary differential equation

$$\zeta \frac{d}{d\zeta} \hat{u}(\zeta) + 2\nu \frac{d^2}{d\zeta^2} \hat{u}(\zeta) = 0. \tag{9}$$

Substituting $s = \frac{d\hat{u}}{d\zeta}$ yields a first-order ODE in the slope

$$\zeta s + 2\nu \frac{ds}{d\zeta} = 0 \tag{10}$$

and upon integration of this separable ODE

$$s = ce^{-\frac{\zeta^2}{4\nu}}, \tag{11}$$

where c is a constant of integration. Integrating (11), setting integration constants to yield the appropriate boundary conditions, and transforming back to the original dependent and independent variables yields the analytic solution to (6) for $t > 0$

$$\hat{u}(x, t; \nu) = \frac{u_L + u_R}{2} + \frac{u_R - u_L}{2} \operatorname{erf}\left(\frac{x - x_0 - at}{\sqrt{4\nu t}}\right), \tag{12}$$

where $\operatorname{erf}(\zeta)$ is the error function

$$\operatorname{erf}(\zeta) = \frac{2}{\pi} \int_0^\zeta e^{-\chi^2} d\chi.$$

Now for a given value $\nu = \nu_1$ and fixed time $t = t_f$, define the L_1 error, e , between (12) and the exact weak solution (13)

$$u(x, t) = \begin{cases} u_l & \text{for } x - x_0 < at \\ u_r & \text{for } x - x_0 \geq at \end{cases} \tag{13}$$

to be

$$e(t_f, \nu_1) = \int_{-\infty}^{\infty} |u(\chi, t_f) - \hat{u}(\chi, t_f)| d\chi. \tag{14}$$

A relatively straightforward simplification shows that

$$e(t_f, \nu_1) = |u_r - u_l| \sqrt{4\nu_1 t_f} \int_0^\infty \operatorname{erfc}(\chi) d\chi, \tag{15}$$

where $\operatorname{erfc}(\zeta) = 1 - \operatorname{erf}(\zeta)$ is the complementary error function. From (15) it is trivial to show that

$$\frac{e(t_f, v_1)}{e(t_f, \frac{v_1}{2})} = \sqrt{2} \tag{16}$$

and thus a halving of the numerical value of v results in a reduction of total error by a factor of $\sqrt{2}$. It should be recognized that although this analysis is carried out on an infinite domain, this is a good approximation to the finite domain where the contact discontinuity is far removed from the computational boundaries. Returning to the original numerical schemes (2) and (3), it is clear that in both cases for fixed λ , $v = O(\Delta x)$ and thus the asymptotic convergence rate is $O(\Delta x^{1/2})$, as shown in Table 1. Fig. 2 shows the numerical solution as computed using the explicit scheme (2) and as computed by the implicit scheme (3). These are overlaid with the analytic solution (12) using the appropriate value for v .

5. Analysis of higher-order schemes

5.1. Second-order linear scheme

The implication of degraded convergence rates for high-resolution schemes has important practical consequences. In fact, high-resolution schemes have become very popular and proved to be quite effective largely because of their ability to mitigate, but not eliminate, this poor behavior. In a high-resolution scheme, a nominally $O(\Delta x^p)$ method switches to a first-order scheme near local maxima and minima in order to maintain monotone profiles. The investigation in this paper shows that this limiting process does not degrade convergence rates for discontinuous linear waves. Rather, we see that the high-resolution schemes attain the same convergence rates as their unlimited linear upwind counterparts.

To understand this, introduce scheme (17) which is a second-order unlimited (linear) approximation in both space and time

$$v_i^{n+1} = v_i^n - \lambda \left[\left(v_i^n + \frac{1}{4}(1 - \lambda)(v_{i+1}^n - v_{i-1}^n) \right) - \left(v_{i-1}^n + \frac{1}{4}(1 - \lambda)(v_i^n - v_{i-2}^n) \right) \right]. \tag{17}$$

This is the second-order unlimited Godunov method, and will serve as a useful template when limited methods are investigated later. The modified equation for this scheme is

$$\frac{\partial}{\partial t} u(x, t) + a \frac{\partial}{\partial x} u(x, t) + \frac{a\Delta x^2}{2} (1 - \lambda) \left(\frac{1}{3}(1 + \lambda) - \frac{1}{2} \right) \frac{\partial^3}{\partial x^3} u(x, t) + \dots = 0. \tag{18}$$

A straightforward convergence study performed in the same way as the previous studies is shown in Table 2. The trend for this scheme appears to follow a convergence rate of $O(\Delta x^{p/(p+1)})$, where $p = 2$. This behavior can be explained through an analysis of (18) above. Again, we will examine the truncated modified equation

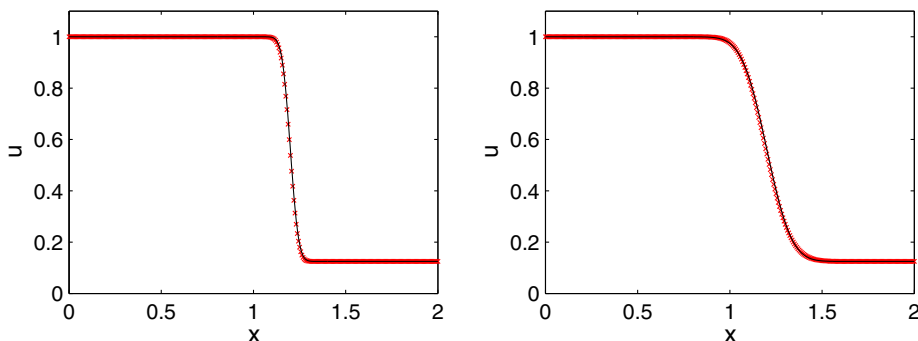


Fig. 2. Red marks show the numerical solution for $N = 321$ as computed using the explicit scheme (left) and by the implicit scheme (right). Overlaid is the analytic solution (12) using the appropriate value for v (black line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2

Convergence for the second-order unlimited scheme with discontinuous initial data approaches the theoretical convergence rate of 2/3

N	$e_u(N)$	Rate
501	6.55e-3	–
1001	4.09e-3	0.679
2001	2.55e-3	0.682
4001	1.61e-3	0.663

$$\frac{\partial}{\partial t} \hat{u}(x, t) + a \frac{\partial}{\partial x} \hat{u}(x, t) + \eta \frac{\partial^3}{\partial x^3} \hat{u}(x, t) = 0, \tag{19}$$

where η is simply

$$\eta = \frac{a\Delta x^2}{2} (1 - \lambda) \left(\frac{1}{3} (1 + \lambda) - \frac{1}{2} \right), \tag{20}$$

as it appears in (18). Note that η changes sign at $\lambda = 1/2$. Eq. (19) is recognized as the linear KdV equation and solutions have been found in other contexts, for example see [21]. As before, applying the change of variables

$$\begin{aligned} z &= x - x_0 - at, \\ \tau &= t, \end{aligned}$$

yields

$$\frac{\partial}{\partial \tau} \hat{u}(z, \tau) + \eta \frac{\partial^3}{\partial z^3} \hat{u}(z, \tau) = 0. \tag{21}$$

For discontinuous initial data $u(x, 0) = u_L$ for $x < x_0$ and $u(x, 0) = u_R$ for $x \geq x_0$, one can seek a self similar solution to (21), via the following change of independent variables:

$$\zeta(z, \tau) = \frac{z}{\tau^{1/3}}. \tag{22}$$

This reduces the partial differential equation (7) to a third-order ordinary differential equation

$$\zeta \frac{d}{d\zeta} \hat{u}(\zeta) - 3\eta \frac{d^3}{d\zeta^3} \hat{u}(\zeta) = 0. \tag{23}$$

Substituting $s = \frac{d\hat{u}}{d\zeta}$ yields a second-order ODE in the slope

$$\zeta s - 3\eta \frac{d^2 s}{d\zeta^2} = 0. \tag{24}$$

The solution to this ODE [23] is given by

$$s = c_1 \text{Ai} \left(\frac{\sqrt[3]{\frac{1}{\eta} \zeta}}{\sqrt[3]{3}} \right) + c_2 \text{Bi} \left(\frac{\sqrt[3]{\frac{1}{\eta} \zeta}}{\sqrt[3]{3}} \right), \tag{25}$$

where Ai and Bi are Airy functions and c_1 and c_2 are integration constants. Because Bi is unbounded as $\zeta \rightarrow \infty$, we must take $c_2 = 0$. Integrating (25), and setting integration constants to yield the appropriate boundary conditions yields (for $1/2 < \lambda < 1$)

$$\hat{u}(\zeta) = \frac{1}{3} (u_L + 2u_R) - \frac{(u_R - u_L) \zeta \left(\sqrt[3]{\frac{1}{\eta} \zeta} \Gamma\left(\frac{2}{3}\right)^2 {}_1F_2\left(\frac{2}{3}, \frac{4}{3}, \frac{5}{3}; \frac{\zeta^3}{27\eta}\right) - 3\Gamma\left(\frac{1}{3}\right) \Gamma\left(\frac{5}{3}\right) {}_1F_2\left(\frac{1}{3}, \frac{2}{3}, \frac{4}{3}; \frac{\zeta^3}{27\eta}\right) \right)}{27 \sqrt[3]{\eta} \Gamma\left(\frac{2}{3}\right) \Gamma\left(\frac{4}{3}\right) \Gamma\left(\frac{5}{3}\right)} \tag{26}$$

where Γ is the Euler Gamma function, and ${}_1F_2$ is a generalized hypergeometric function. The analytic solution to (19) can be written by simply substituting $\zeta = \frac{x-x_0-at}{t^{1/3}}$ into (26). Also note that by using (26), the convergence of the second-order linear scheme is demonstrated to be 2/3. See analysis in Section 5.3 for more details.

Fig. 3 shows a zoom near the true discontinuity, $x = 1.2$ at $t = 1$, for 2 different resolutions. Both the numerical solution, and analytic solution to the modified equation are shown. Note that when $0 < \lambda < 1/2$, the sign of η changes, and expected oscillating dispersive waves will lead the true discontinuity locus, instead of lag as shown in Fig. 3.

5.2. Second-order MinMod scheme

The effect of limiting the solution to maintain monotone profiles (with non-compressive limiters such as MinMod) does not effect this rate of convergence as is shown in Table 3. Here scheme (17) is used but a MinMod limiter is applied to the slopes. Thus the scheme becomes

$$v_i^{n+1} = v_i^n - \lambda \left[\left(v_i^n + \frac{1}{2}(1 - \lambda)\alpha \right) - \left(v_{i-1}^n + \frac{1}{2}(1 - \lambda)\beta \right) \right], \tag{27}$$

where

$$\alpha = \text{MinMod}(v_{i+1}^n - v_i^n, v_i^n - v_{i-1}^n),$$

$$\beta = \text{MinMod}(v_i^n - v_{i-1}^n, v_{i-1}^n - v_{i-2}^n),$$

and

$$\text{MinMod}(b, c) = \begin{cases} b & \text{if } |b| < |c| \text{ and } bc > 0, \\ c & \text{if } |b| \geq |c| \text{ and } bc > 0, \\ 0 & \text{if } bc \leq 0. \end{cases}$$

Note that this is nothing more than a high-resolution Godunov method (see [14,24] for details). As is demonstrated in Table 3, scheme (27) converges at order 2/3 as was the case for the unlimited scheme. This interesting phenomenon shows that for discontinuous data, the sub-linear convergence exhibited by numerical schemes is not a by-product of the limiting procedure, but rather a fundamental flaw in the convergence character of the high-order methods. That is to say that for simple linear jumps, the MinMod scheme has second-order truncation error (uses second-order approximations) everywhere and the use of the limiter never causes a first-order approximation to be used, which would reduce the convergence rate to 1/2.

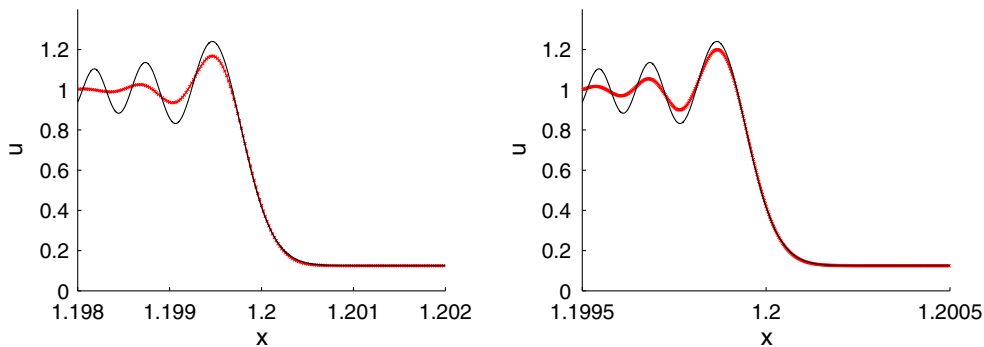


Fig. 3. Comparison of modified equation solution (black line) and numerical discretization (red × marks) for the unlimited scheme (17). Shown are comparisons for $\lambda = 0.8$ for $\Delta x = \frac{1}{50,000}$ (left) and $\Delta x = \frac{1}{400,000}$ (right). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Convergence for the second-order MinMod scheme with discontinuous initial data

N	$e_d(N)$	Rate
501	8.67e-3	–
1001	5.53e-3	0.649
2001	3.51e-3	0.656
4001	2.22e-3	0.661

To better understand the behavior of the second-order MinMod limited scheme, we perform an analysis of its corresponding modified equation (28)

$$\frac{\partial}{\partial t} \hat{u}(x, t) + a \frac{\partial}{\partial x} \hat{u}(x, t) + \frac{a\Delta x^2}{12} (1 - \lambda) \left[(2\lambda - 1) - 3 \operatorname{sign} \left(\frac{\partial}{\partial x} \hat{u}(x, t) \right) \operatorname{sign} \left(\frac{\partial^2}{\partial x^2} \hat{u}(x, t) \right) \right] \frac{\partial^3}{\partial x^3} \hat{u}(x, t) + \dots = 0. \tag{28}$$

Details concerning the derivation of the modified equation for the MinMod scheme can be found in [25] in Chapter 5. The key point to recognize is that the functions used to define the MinMod limiter can be expressed in a form amenable to a modified equation analysis. Specifically,

$$\operatorname{MinMod}(b, c) = \frac{1}{4} (\operatorname{sign}(b) + \operatorname{sign}(c)) (|b + c| - |b - c|),$$

where $|d| = \sqrt{d^2}$ and $\operatorname{sign}(d) = \frac{d}{|d|}$. It can be shown that modified equation (28) can be simplified to the form (19), with η defined as

$$\eta = \begin{cases} -\frac{a\Delta x^2}{6} (1 - \lambda)(2 - \lambda) & \text{if } \frac{\partial \hat{u}}{\partial x} \frac{\partial^2 \hat{u}}{\partial x^2} > 0 \\ \frac{a\Delta x^2}{6} (1 - \lambda^2) & \text{otherwise.} \end{cases} \tag{29}$$

For the most part, the analysis of Section 5.1 carries over for the MinMod scheme. The major difference is that the solution is now broken into two parts depending on the signs of the first and second derivatives. Since the approximate solution is TVD, and thus monotone for this example, the first derivative does not change sign and the second derivative provides the switch in (29). The location of this switch in the similarity solution, call it ξ_s , is found as part of the solution process and in general depends on the left and right values of η and thus the CFL number λ . Note that for $\lambda = 1/2$, the magnitude of η on the left is identical to the magnitude on the right and $\xi_s = 0$. For other values of λ , the solution is written in the general form

$$\hat{u}(\xi) = \begin{cases} C_2 - C_1 \xi \left[\frac{\xi \Gamma(\frac{2}{3}) \sqrt[3]{\frac{1}{\eta_L} F_2 \left(\frac{2, 4, \xi, \xi^3}{3, 3, 3, 27 \eta_L} \right)}}{4\pi \sqrt[3]{3}} - \frac{{}_1F_2 \left(\frac{1, 2, 4, \xi^3}{3, 3, 3, 27 \eta_L} \right)}{\Gamma(\frac{2}{3}) \sqrt[3]{3^2}} \right] & \text{for } \xi < \xi_s, \\ C_4 - C_3 \xi \left[\frac{\xi \Gamma(\frac{2}{3}) \sqrt[3]{\frac{1}{\eta_R} F_2 \left(\frac{2, 4, \xi, \xi^3}{3, 3, 3, 27 \eta_R} \right)}}{4\pi \sqrt[3]{3}} - \frac{{}_1F_2 \left(\frac{1, 2, 4, \xi^3}{3, 3, 3, 27 \eta_R} \right)}{\Gamma(\frac{2}{3}) \sqrt[3]{3^2}} \right] & \text{for } \xi \geq \xi_s. \end{cases} \tag{30}$$

The constants $C_1, C_2, C_3,$ and C_4 are found as functions of ξ_s by enforcing the conditions that $\hat{u}(-\infty) = u_L, \hat{u}(\infty) = u_R,$ continuity of $\hat{u}(\xi)$ at $\xi = \xi_s,$ and continuity of $\frac{d}{d\xi} \hat{u}(\xi)$ at $\xi = \xi_s.$ In order to uniquely define the value for $\xi_s,$ the fact that the numerical scheme is exactly conservative by construction is used and that property is enforced for $\hat{u}(\xi)$ as well. Note that for linear schemes such as those we have investigated thus far, the exact solution of the ME is conservative by construction and so conservation does not pose any additional constraints on the solution. For more details concerning this derivation refer to the appendix in Section B.

For the example under consideration here ($\lambda = 4/5, u_L > u_R$), the values for η are given as

$$\eta_L = -\frac{a\Delta x^2}{6} (1 - \lambda)(2 - \lambda),$$

$$\eta_R = \frac{a\Delta x^2}{6} (1 - \lambda^2),$$

and we demonstrate the solution for two values of $\Delta x.$ The numerical values for the integration constants and ξ_s are shown in Table 4 to five significant figures. Recall $\xi = (x - x_0 - at) / \sqrt[3]{t}$ and so the solution (30) can be written in terms of x and $t.$ Fig. 4 shows a zoom near the true discontinuity, $x = 1.2$ at $t = 1,$ for the two resolutions whose complete solution is presented in Table 4 and Eq. (30). Shown are the exact solution to the modified equation (30) and the numerical approximations. Note the different scales in the two images.

Table 4
Numerical values for integration constants and ξ_s to define the solution in (30)

	C_1	C_2	C_3	C_4	ξ_s
$\Delta x = \frac{1}{50,000}$	-3.9113×10^3	5.2618×10^{-1}	-2.9786×10^3	5.3805×10^{-1}	-7.1958×10^{-5}
$\Delta x = \frac{1}{400,000}$	-1.5645×10^4	5.2618×10^{-1}	-1.1914×10^4	5.3805×10^{-1}	-1.7990×10^{-5}

The values have been found through an iterative procedure and are shown to five significant figures.

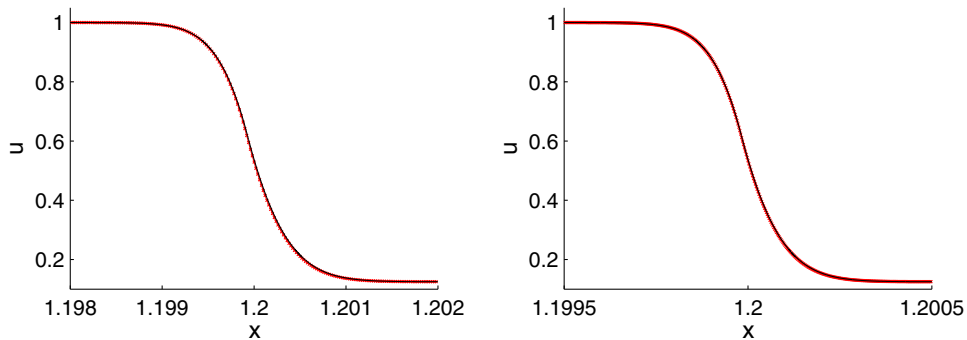


Fig. 4. Comparison of modified equation solution (black line) and numerical discretization (red \times marks) for the MinMod limited scheme (27). Shown are comparisons for $\lambda = 0.8$ for $\Delta x = \frac{1}{50,000}$ (left) and $\Delta x = \frac{1}{400,000}$ (right). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Here it should be noted that the approximate solution obtained using the MinMod scheme (27) is fully second-order accurate except at extrema. Because the solution to the unlimited approximation is monotone on one side of the discontinuity, a fully monotone solution can be pieced together which maintains second-order accuracy everywhere by taking the monotone portion of two separate solutions. Effectively the MinMod limiter chooses between two fully second-order schemes, the Beam-Warming scheme to the left and the Lax-Wendroff scheme to the right, in such a way as to maintain a monotone profile. Unfortunately, the existence of schemes which exhibit this property seems to be a peculiarity of second-order (or lower) approximations. For most higher-order approximations, the solutions of the modified equations are not monotone on either side of a discontinuity. The result is that there is no method which is both monotone and fully p th order accurate for $p > 2$ and arbitrary λ .

As discussed in the introduction, limiting procedures that result in slightly compressive, also called anti-diffusive, schemes, such as Superbee or Ultrabee, are somewhat outside this analysis as their modified equations contains some anti-diffusive portion which serves as an artificial steepening mechanism and allows such schemes to converge at higher rates for discontinuous data. However, because such schemes allow some unstable character into the approximate solution, they rely on limiting to maintain overall solution stability. As such, extreme care must be taken when using such schemes as unphysical results and/or noise in the higher derivatives can occur.

An ill-advised method that has the same fundamental character, but is amenable to modified equation analysis, is the TVD limited downwind method. For $a > 0$, a downwind stencil in a slope limited scheme corresponds to the selection of a slope equal to twice the downwind slope. To make this scheme stable it should conform to bounding via a TVD limiter [26]. This can be accomplished through the selection of a slope defined (as before) by,

$$\alpha = 2\text{MinMod}(v_{i+1}^n - v_i^n, v_i^n - v_{i-1}^n),$$

and

$$\beta = 2\text{MinMod}(v_i^n - v_{i-1}^n, v_{i-1}^n - v_{i-2}^n).$$

The modified equation for this method (31) can easily be computed and immediately produces two important observations: the method is first-order accurate, and any useful computation involves a balance between linear anti-diffusion and both linear and nonlinear dispersion,

$$\begin{aligned} & \frac{\partial}{\partial t} \hat{u}(x, t) + a \frac{\partial}{\partial x} \hat{u}(x, t) + \frac{|a|\Delta x}{2} (1 - \lambda) \frac{\partial^2}{\partial x^2} \hat{u}(x, t) \\ & + \frac{a\Delta x^2}{6} (1 - \lambda) \left[(\lambda - 2) - 3 \text{sign} \left(\frac{\partial}{\partial x} \hat{u}(x, t) \right) \text{sign} \left(\frac{\partial^2}{\partial x^2} \hat{u}(x, t) \right) \right] \frac{\partial^3}{\partial x^3} \hat{u}(x, t) + \dots = 0. \end{aligned} \tag{31}$$

It is this balance that allows the method to achieve a higher convergence rate than the linear or nonlinear second-order methods above. Here, the lower-order anti-dissipative terms never dominate the high-order stabilizing diffusion because the discrete wave profile is never resolved. That is to say that the approximation attains a quasi-steady profile of approximately one to two cell widths. This behavior is different than the standard MinMod scheme where the wave profile grows as a function of time as $t^{1/3}$. The smeared profile is responsible for the observed sub-linear convergence rate. Simple calculations confirm that this limited down-wind method produces a first-order result for discontinuous waves, but at the cost of including controlled unstable stencils in the approximation.

5.3. Higher than second-order schemes

For unlimited schemes with greater than second-order accuracy the story is much the same as before with the understanding that higher-order TVD schemes do not exist. Here we can demonstrate that a convergence rate of $O(\Delta x^{p/(p+1)})$ is expected for the unlimited schemes and non-compressive limiting procedures are not expected to effect this rate. To show the advertised convergence rate notice that a stable p th order schemes will have a modified equation of the form

$$\frac{\partial}{\partial t} u(x, t) + a \frac{\partial}{\partial x} u(x, t) - k \frac{\partial^{p+1}}{\partial x^{p+1}} u(x, t) + \dots = 0, \tag{32}$$

where $k = \tilde{k} \Delta x^p$ and \tilde{k} is some constant (dependent on the CFL λ). In analyzing the behavior of this modified equation it is again useful to perform the change of variables

$$\begin{aligned} z &= x - x_0 - at, \\ \tau &= t, \end{aligned}$$

which after dropping the higher-order terms transforms (32) into the form

$$\frac{\partial}{\partial \tau} \hat{u}(z, \tau) - \hat{k} \frac{\partial^{p+1}}{\partial z^{p+1}} \hat{u}(z, \tau) = 0, \tag{33}$$

where \hat{k} is either plus or minus k depending on the value of p . In a similar manner as before, for

$$\xi = \frac{z^{\frac{p+1}{2}}}{\sqrt{\hat{k}\tau}},$$

there exists a similarity solution

$$\hat{u}(\xi) = \frac{u_L + u_R}{2} + \frac{u_R - u_L}{2} S(\xi),$$

where S is some similarity function whose form is not important but in general will involve hypergeometric functions. A straightforward manipulation shows

$$e(t_f, \hat{k}) = |u_r - u_l| (\hat{k} t_f)^{\frac{1}{p+1}} \frac{2}{p+1} \int_{-\infty}^{\infty} |S(\chi) \chi^{\frac{p-1}{2}}| d\chi. \tag{34}$$

From (32) it is seen that if $\hat{k} = \hat{k}_1$ for a given resolution Δx_1 , then doubling the resolution to $\frac{1}{2} \Delta x_1$ results in $\hat{k} = \hat{k}_2 = \frac{1}{2^p} \hat{k}_1$. At this point all the pieces are in place and it is straight forward to show that

$$\frac{e(t_f, \hat{k}_1)}{e(t_f, \hat{k}_2)} = \left(\frac{p^p \hat{k}_1}{\frac{p^p \hat{k}_1}{2^p}} \right)^{\frac{1}{p+1}} = 2^{\frac{p}{p+1}}.$$

Thus the order of convergence for unlimited stable schemes is established as $O(\Delta x^{p/(p+1)})$ where p is the order of the approximation for smooth flows. Any limiting procedure which does not introduce unstable portions into the modified equation (i.e. a non-compressive limiter) is then bounded by this convergence rate. Detailed analyses of such compressively limited schemes is beyond the scope of this paper. One of the primary advances represented in this analysis is the knowledge that not only does a similarity solution exist, but that the form is known as piecewise generalized hypergeometric functions even for nonlinear schemes.

6. More complex flows

Up until now, the discussion in this paper has centered on understanding the behavior of numerical error produced by shock capturing methods for linear equations with discontinuities, but clearly this is of limited use as linear equations rarely, if ever, accurately model the physical world. In this short section we show some results for the more complex situation of inviscid compressible flows governed by the Euler equations in two space dimensions where sub-linear convergence is demonstrated at a contact surface. Typically, exact weak solutions for the Euler equations are difficult to derive when the flow is more complicated than that arising from a simple Riemann problem in one dimension. For two dimensions then, non-trivial exact weak solutions are very rare.

One exception to this is the so called shock polar solution were two oblique shocks turn the flow on either side of a contact so that the resulting structure is steady in the frame where the shocks and contact intersect. Such flows have been investigated for example in [24,27] where the contact surface separated two different materials. Here a single ideal gas is used and a shock polar solution derived which consists of an upstream slip which is aligned with the computational mesh so that no significant instability growth can arise there, and a post-shock state with no slip. Fig. 5 shows a schematic of such a flow.

The solution represented in Fig. 5 can be derived through a shock polar analysis, as described in [27,24] with reference to [21]. The problems above and below the contact line are solved in isolation using the oblique shock jump conditions and then coupled using the requirements that the velocity and pressure agree in the post-shock region. Such an analysis leads to the shock polar diagram shown in Fig. 6 where the post-shock pressure and streamline deflection angle are parameterized by the shock inclination angle and are plotted for the gas both above and below the contact surface. The eventual shock polar solution is found as the intersection of these two shock polars and is given in Table 5.

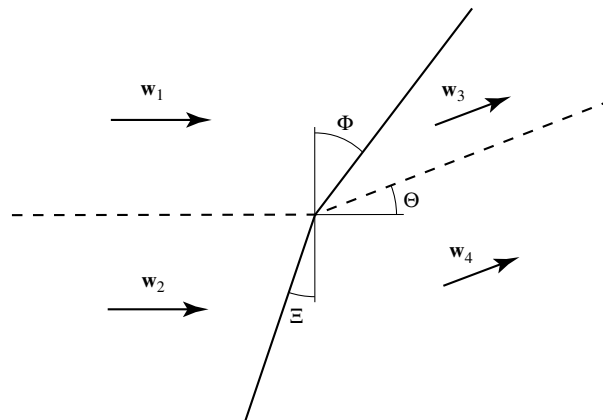


Fig. 5. Oblique shock interaction with a planar material interface. The interface (dashed line) indicates a slip line which is deflected by an angle Θ behind the shock. The oblique shock (solid line) make angles Φ and Ξ relative to the vertical. The flow states are denoted by w_i for $i = 1, 2, 3,$ and 4 .

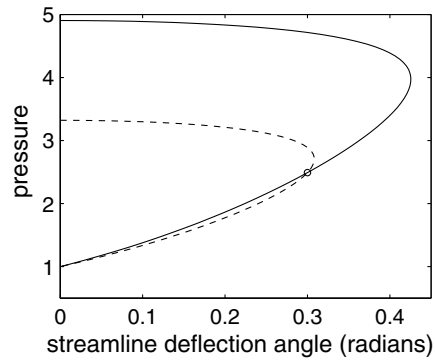


Fig. 6. Shock polars for the flow configuration shown in Fig. 5. The dashed curve shows the downstream pressure p_3 versus flow deflection Θ_3 parameterized by the shock angle Φ for the gas above the dashed contact. The solid curve is p_4 versus Θ_4 parameterized by Ξ for the gas below the dashed contact. The exact solution with states given in Table 5 is given by the point of intersection of these shock polars where $p_3 = p_4 = 2.4916$, $\Theta_3 = \Theta_4 = 0.3$, $\Phi = 0.76125$ and $\Xi = 0.50955$.

Table 5

Primitive variables for the flow configuration shown in Fig. 5

	Region 1	Region 2	Region 3	Region 4
ρ	2	1	3.75651514	1.87825757
u_x	1.24440775	1.54592121	0.816883174	0.816883174
u_y	0	0	0.407359702	0.407359702
p	1	1	2.49155439	2.49155439

The velocities include a constant translation to the left with speed 0.5.

This flow is then investigated numerically by setting the initial condition for the rectangle $\{(x, y) : x \in [-2, 2], y \in [-3, 2]\}$ as the shock polar solution from Table 5 centered about the origin and integrating to a final time of $t = 1$. The solution is then queried along the line $x = -0.25$, $y \in [0, 0.2]$ and the computed solution compared to the exact solution. Note here that this slice through regions 3 and 4 is sufficiently removed from the computational boundaries so that no disturbance generated at the boundaries has yet had opportunity to interact with it. The computational method is a second-order high-resolution Godunov method as discussed in [14]. Errors are computed using the L_1 norm for all primitive quantities and results presented in Table 6 for various resolutions defined by $h = 1/(N - 1)$ with N a measure of the number of grid cells per unit distance.

The results indicate that although the velocities and pressure are tending toward first-order convergence, the density, which jumps across the contact, is not. The analysis presented in this paper indicates a probable cause for this type of sub-linear convergence, thus demonstrating that the phenomenon of sub-linear convergence at contacts is not entirely an academic exercise. Here the same effect is seen for the Euler equations in two space dimensions for a well characterized test case. It is somewhat surprising that not all quantities

Table 6

Convergence results for shock polar along the line $x = -0.25$ for $y \in [0, 0.2]$

N	$e_\rho(N)$	Rate	$e_{u_x}(N)$	Rate	$e_{u_y}(N)$	Rate	$e_p(N)$	Rate
26	6.06e-2	–	9.30e-4	–	5.95e-4	–	1.99e-3	–
51	4.50e-2	0.429	6.72e-4	0.469	2.50e-4	1.25	1.34e-3	0.571
101	2.70e-2	0.737	3.99e-4	0.752	1.42e-4	0.816	6.57e-4	1.03
201	1.64e-2	0.719	2.26e-4	0.820	7.30e-5	0.960	3.23e-4	1.02
401	1.07e-2	0.616	1.26e-4	0.843	3.79e-5	0.946	1.51e-4	1.10
801	6.64e-3	0.688	6.82e-5	0.886	1.88e-5	1.01	7.25e-5	1.06

Here L_1 errors are computed and the tendency is toward first-order convergence for the velocities and pressure but sub-linear convergence at order 2/3 for the density.

converge at sub-linear rates because the flow is highly coupled. In fact this type of de-coupling should not be expected in general. Of course, in most practical simulations, the resulting flow will be quite complex and subsequent interaction with poorly converged interfaces will result in poor convergence for all quantities involved in the interaction even if the convergence rates do de-couple near the contact.

7. Conclusions

We have provided a description concerning the phenomenon of sub-linear convergence rates for numerical methods applied to linear discontinuities. The discussion proceeded through a detailed analysis of the modified equations (ME) of various schemes for the model problem of linear advection. The convergence rate of $1/2$ was derived through the exact solution to the ME with discontinuous data for the explicit and implicit first-order upwind schemes and this rate was verified through a comparison to numerical experiment.

The analysis was then extended to a linear, or unlimited, second-order scheme and the exact solution of the ME again derived. This exact solution was shown to be a generalized hypergeometric function and so the solution was expected to contain oscillations caused by dispersion. Numerical experimentation was carried out for this scheme and not only does the convergence rate match the estimated $2/3$ rate, but the form of the solution agrees with prediction to a very high degree. Second-order limited schemes were then shown to fit into this analysis and the exact solution for the ME from the MinMod limiter was given. This solution was fully second-order, and the TVD and monotonicity constraints of the solution were guaranteed by piecing together two non-oscillatory portions. These two parts correspond roughly to the exact solution of the Beam-Warming ME and the Lax-Wendroff ME. Again the convergence rate is demonstrated to be $2/3$ and this verified through numerical experimentation. Furthermore, the exact solution to the ME and the computed numerical approximation are shown to be in very good agreement. At this point a short discussion of a compressively limited scheme, the double MinMod scheme which corresponds to the TVD limit, was presented. Here it was shown why such a scheme lies somewhat outside this analysis and how steady state may be achieved as a balance between anti-diffusive and dispersive terms of the ME. The inclusion of anti-diffusive terms sheds light as to why such care must be taken when using such compressive schemes.

Linear schemes of arbitrary order were also investigated through this analysis. For a general scheme of order p , the exact solution of the ME with discontinuous data is shown to be a similarity solution involving generalized hypergeometric functions. The convergence rate of $p/(p + 1)$ is derived from these solutions. This approach demonstrates why TVD schemes of order higher than 2 do not exist in that the exact solutions do not contain monotone halves as was the case for $p = 2$.

Finally the results are shown to apply to the Euler equations in two space dimensions. Here an exact solution involving shocks and contacts is derived through a shock polar analysis and computations presented which demonstrate the $2/3$ convergence rate. More complex examples would also demonstrate such a rate, but this is perhaps the simplest non-trivial 2-D example to which the solution is known and which does not suffer from physical instability.

The first appendix shows how this analysis extends to flows containing corners. Such flows are likely to occur near rarefaction corners, and although those waves are strictly nonlinear, the nonlinearity is fairly weak and so the analysis may be important there as well. Finally the second appendix presents some detail concerning the derivation of the exact solution for the MinMod ME.

Appendix A. Convergence at corners (discontinuous first derivatives)

We make one observation concerning continuous flows containing a discontinuous first derivative, a corner. See [20] for an alternate analysis. Such an analysis is important not only for density profiles containing corners, but also for the corners found in rarefaction waves of Euler flows. Although rarefactions are a nonlinear phenomenon, the nonlinearity is quite weak and in practice the convergence of most numerical methods in such regions adheres very closely to that which we would expect from a purely linear wave.

Assume that one is advecting a corner using the first-order upwind method. So that we have it here in front of us, the truncated modified equation for that method is

$$\frac{\partial}{\partial t} \hat{u}(x, t) + a \frac{\partial}{\partial x} \hat{u}(x, t) - v \frac{\partial^2}{\partial x^2} \hat{u}(x, t) = 0.$$

The now familiar analysis uses a change of variables to place us in the reference of the traveling solution

$$\frac{\partial}{\partial \tau} \hat{u}(z, \tau) - v \frac{\partial^2}{\partial z^2} \hat{u}(z, \tau) = 0. \tag{35}$$

A corner is a continuous solution, but the first derivative will jump and so we expect self similar behavior in the first derivative of the computed solution. Taking the derivative $\frac{\partial}{\partial z}$ of (35) yields

$$\frac{\partial}{\partial \tau} \hat{s}(z, \tau) - v \frac{\partial^2}{\partial z^2} \hat{s}(z, \tau) = 0, \tag{36}$$

where

$$\hat{s}(z, \tau) = \frac{\partial}{\partial z} \hat{u}(z, \tau).$$

Obviously, (36) represents familiar ground and the behavior of \hat{s} can be analyzed as we have done before. The difference here is that \hat{s} represents the slope of the solution and so one further integration is required to recover the original solution \hat{u} . At the end of the day, the analytic solution is

$$\hat{u}(z, \tau) = \hat{u}_0 + \frac{1}{2} \left[(\hat{s}_L + \hat{s}_R)z + \frac{e^{-\frac{z^2}{4v\tau}} (\hat{s}_R - \hat{s}_L) \sqrt{4v\tau}}{\sqrt{\pi}} - (\hat{s}_L - \hat{s}_R)z \operatorname{erf} \left(\frac{z}{\sqrt{4v\tau}} \right) \right],$$

where \hat{u}_0 is the solution value at the corner, and \hat{s}_L and \hat{s}_R are the slopes of the solution on the left and right respectively. From such an analytic solution it is trivial to show that the L_∞ error occurs at the corner itself and has the form

$$e_\infty = \frac{(-\hat{s}_L + \hat{s}_R) \sqrt{v\tau}}{\sqrt{\pi}}.$$

Because we are dealing with a first-order method and $v = O(\Delta x)$, the L_∞ convergence rate is $O(\Delta x^{1/2})$. One can also show that the L_1 error is

$$e_1 = |(-\hat{s}_L + \hat{s}_R)v\tau|$$

and so the L_1 convergence rate is $O(\Delta x)$. For a p th order method these results generalize to convergence at $O(\Delta x^{p/(p+1)})$ for the L_∞ norm and $O(\Delta x^{2p/(p+1)})$ for the L_1 norm.

Appendix B. Notes on the exact solution of the ME for the MinMod scheme

We feel it appropriate to say a few words concerning the derivation of the exact solution to the modified equation for the MinMod limited scheme. Recall that we had the general form for the solution

$$\hat{u}(\xi) = \begin{cases} C_2 - C_1 \xi \left[\frac{\xi \Gamma(\frac{2}{3}) \sqrt[3]{\frac{1}{\eta_L}} {}_1F_2 \left(\frac{2, 4, \xi^3}{3, 3, 3; 27\eta_L} \right) - {}_1F_2 \left(\frac{1, 2, 4, \xi^3}{3, 3, 3; 27\eta_L} \right)}{4\pi \sqrt[3]{3}} \right] & \text{for } \xi < \xi_s, \\ C_4 - C_3 \xi \left[\frac{\xi \Gamma(\frac{2}{3}) \sqrt[3]{\frac{1}{\eta_R}} {}_1F_2 \left(\frac{2, 4, \xi^3}{3, 3, 3; 27\eta_R} \right) - {}_1F_2 \left(\frac{1, 2, 4, \xi^3}{3, 3, 3; 27\eta_R} \right)}{4\pi \sqrt[3]{3}} \right] & \text{for } \xi \geq \xi_s. \end{cases}$$

For the left and right boundary conditions we enforce the conditions that $\hat{u}(-\infty) = u_L$ and $\hat{u}(\infty) = u_R$. This produces the two constraints

$$\begin{aligned} C_2 + C_1 \frac{\sqrt[3]{\eta_L}}{\sqrt[3]{3}} &= u_L, \\ C_4 + C_3 \frac{\sqrt[3]{\eta_R}}{\sqrt[3]{3}} &= u_R. \end{aligned} \tag{37}$$

We also require the solution to be continuous at the location of the switch ζ_s and so

$$\begin{aligned}
 C_2 - C_1 \zeta_s & \left[\frac{\zeta_s \Gamma\left(\frac{2}{3}\right) \sqrt[3]{\frac{1}{\eta_L}} {}_1F_2\left(\frac{2}{3}; \frac{4}{3}, \frac{5}{3}; \frac{\zeta_s^3}{27\eta_L}\right)}{4\pi\sqrt[3]{3}} - \frac{{}_1F_2\left(\frac{1}{3}; \frac{2}{3}, \frac{4}{3}; \frac{\zeta_s^3}{27\eta_L}\right)}{\Gamma\left(\frac{2}{3}\right)\sqrt[3]{3^2}} \right] \\
 & = C_4 - C_3 \zeta_s \left[\frac{\zeta_s \Gamma\left(\frac{2}{3}\right) \sqrt[3]{\frac{1}{\eta_R}} {}_1F_2\left(\frac{2}{3}; \frac{4}{3}, \frac{5}{3}; \frac{\zeta_s^3}{27\eta_R}\right)}{4\pi\sqrt[3]{3}} - \frac{{}_1F_2\left(\frac{1}{3}; \frac{2}{3}, \frac{4}{3}; \frac{\zeta_s^3}{27\eta_R}\right)}{\Gamma\left(\frac{2}{3}\right)\sqrt[3]{3^2}} \right].
 \end{aligned}
 \tag{38}$$

We require that the first derivatives match at the location of the switch which gives

$$C_1 \text{Ai}\left(\frac{\sqrt[3]{\frac{1}{\eta_L}\zeta_s}}{\sqrt[3]{3}}\right) = C_3 \text{Ai}\left(\frac{\sqrt[3]{\frac{1}{\eta_R}\zeta_s}}{\sqrt[3]{3}}\right).
 \tag{39}$$

The requirement of continuity of first derivatives is something which is seen in practice and results from the fact that if the first derivatives are not continuous at that point, the MinMod limiter would favor the stencil with the smaller slope. This would then force the location of ζ_s to change until continuity is achieved. Finally, the restriction that the solution of the modified equation agree with the numerics in so far as it is exactly conservative adds the additional constraint

$$\int_{-\infty}^0 (u_L - \hat{u}(\chi)) d\chi = \int_0^{\infty} (\hat{u}(\chi) - u_R) d\chi.
 \tag{40}$$

Notice that these integrals can be found analytically because $\hat{u}(-\infty) = u_L$ and $\hat{u}(\infty) = u_R$. Now Eq. (37)–(40) represent five nonlinear equations for the five unknowns C_1, C_2, C_3, C_4 , and ζ_s and can be solved through an iterative procedure. For values of λ where this process is non-trivial, that is where $\zeta_s \neq 0$, we have used such a numerical approach to determine the root to within a given tolerance although a closed form solution seems possible. One step toward a closed form solution seems to be the fact that ζ_s scales as $O(\Delta x^{2/3})$ while C_1 and C_3 scale as $O(\Delta x^{-2/3})$. This observation also allows the construction of exact solutions for the MinMod ME using Table 4 provided that we restrict ourselves to $\lambda = 4/5$. For $\lambda = 1/2$, this solution process reveals that $\zeta_s = 0$ and the whole solution results in a straightforward way. In these cases $C_1 = C_3, C_2 = C_4$, and a closed form solution is readily realized.

References

[1] P.D. Lax, *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*, SIAM, Philadelphia, 1972.
 [2] R.J. LeVeque, *Numerical Methods for Conservation Laws*, Birkhauser, Basel, 1992.
 [3] J.W. Thomas, *Numerical Partial Differential Equations: Conservation Laws and Elliptic Equations*, Springer-Verlag, New York, 1999.
 [4] M.L. Wilkins, Use of artificial viscosity in multidimensional fluid dynamic calculations, *J. Comput. Phys.* 36 (1980) 281–303.
 [5] E.J. Caramana, M.J. Shashkov, P.P. Whalen, Formulations of artificial viscosity for multi-dimensional shock wave computations, *J. Comput. Phys.* 144 (1998) 70–97.
 [6] E.J. Caramana, D.E. Burton, M.J. Shashkov, P.P. Whalen, The construction of compatible hydrodynamics algorithms utilizing conservation of total energy, *J. Comput. Phys.* 146 (1998) 227–262.
 [7] R. Anderson, N. Elliott, R. Pember, A dynamically adaptive arbitrary Lagrangian–Eulerian method for solution of the Euler equations, *Tech. Rep. Research Report UCRL-JC-151904*, Lawrence Livermore National Laboratory, 2003.
 [8] E.F. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics*, Springer, Berlin, 1999.
 [9] P. Woodward, P. Colella, The numerical simulation of two-dimensional fluid flow with strong shocks, *J. Comput. Phys.* 54 (1984) 115–173.
 [10] J.P. Boris, D.L. Book, Flux-corrected transport. I. SHASTA a fluid transport algorithm that works, *J. Comput. Phys.* 11 (1973) 38–69.
 [11] A. Harten, P.D. Lax, B. van Leer, On upstream differencing and Godunov type schemes for hyperbolic conservation laws, *SIAM Rev.* 25 (1983) 33–61.
 [12] A. Harten, B. Engquist, S. Osher, S. Chakravarthy, Uniformly high order accurate essentially non-oscillatory schemes III, *J. Comput. Phys.* 71 (1987) 231–303.
 [13] M.J. Berger, J. Olinger, Adaptive mesh refinement for hyperbolic partial differential equations, *J. Comput. Phys.* 53 (1984) 484–512.

- [14] W.D. Henshaw, D.W. Schwendeman, An adaptive numerical scheme for high-speed reactive flow on overlapping grids, *J. Comput. Phys.* 191 (2) (2003) 420–447.
- [15] G.W. Hedstrom, The rate of convergence of some difference schemes, *J. SIAM Numer. Anal.* 5 (2) (1968) 363–406.
- [16] S.A. Orszag, L.W. Jayne, Local errors of difference approximations to hyperbolic equations, *J. Comput. Phys.* 14 (1974) 93–103.
- [17] R.C.Y. Chin, Note on the article local error of difference approximations to hyperbolic equations, *J. Comput. Phys.* 16 (1974) 311–313.
- [18] R.C.Y. Chin, Dispersion and gibbs phenomenon associated with difference approximations to initial boundary-value problems to hyperbolic problems, *J. Comput. Phys.* 18 (1975) 233–247.
- [19] A. Harten, The artificial compression method for computation of shocks contact discontinuities. I. single conservation laws, *Commun. Pure Appl. Math.* 30 (5) (1977) 611–638.
- [20] R. Donat, S. Osher, Propagation of error into regions of smoothness for non-linear approximations to hyperbolic equations, *Comput. Methods Appl. Mech.* 80 (1990) 59–64.
- [21] G.B. Whitham, *Linear and Nonlinear Waves*, Wiley Interscience, New York, 1974.
- [22] P. Lax, B. Wendroff, Systems of conservation laws, *Commun. Pure Appl. Math.* 13 (1960) 217–237.
- [23] M. Abramowitz, I.A. Stegun (Eds.), *Handbook of Mathematical Functions: With Formulas Graphs and Mathematical Tables*, Dover, 1965.
- [24] J.W. Banks, D.W. Schwendeman, A.K. Kapila, W.D. Henshaw, A high-resolution Godunov method for compressible multi-material flow on overlapping grids, *J. Comput. Phys.* 223 (2007) 262–297.
- [25] F.F. Grinstein, L. Margolin, W.J. Rider (Eds.), *Implicit Large Eddy Simulation: Computing Turbulent Fluid Flow*, Cambridge University Press, New York, 2007.
- [26] P.K. Sweby, High resolution schemes using flux limiter for hyperbolic conservation laws, *SIAM J. Numer. Anal.* 21 (1984) 995–1011.
- [27] T.D. Aslam, J.B. Bdzil, Numerical and theoretical investigations on detonation-inert confinement interactions, in: *The Twelfth Symposium (International) on Detonation*, 2002, pp. 483–488.