The University of Reading
School of Mathematical and Physical Sciences

A Comparative Study of Computational Methods in Cosmic Gas Dynamics Continued

George Fitton

August 2010

This dissertation is a joint MSc in the Departments of Mathematics & Meteorology and is submitted in partial fulfilment of the requirements for the degree of Master of Science

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged

Signature _____

Abstract

Based on the 1981 paper "A Comparative Study of Computational Methods in Cosmic Gas Dynamics" by van Albada, van Leer and Roberts, Jr. [23] we begin our investigation which continues their search to find a reliable computational method for cosmic flow problems. Our first aim was to re-analyse their work in order to confirm their results and conclusions and more importantly to better understand the problem at hand. This was done with the aid of much more powerful processing capabilities enabling plot animation and therefore much more detailed observations of the temporal progression of the solution. This led to the conclusion that the transient waves that occur in the system are much stronger than were previously thought. Our second aim was to add to the comparison of methods: the explicit approximate Riemann solvers, the Roe flux-splitting method and the HLL scheme, an additional flux-corrected transport limiter and the inclusion of flux-limiters to Roe's method. The result of the investigation is an increase in the number of numerical methods available to solve the representative astrophysical flow problem and a better understanding of how the solutions the methods produce, progress over time.

Acknowledgments

I would like to thank Pete Sweby for his unceasing patience and understanding. I would also like to thank my friends and family for their support. Most importantly I would like to thank my girlfriend for her most excellent brew making skills.

Table of Contents

List of Tables

List of Figures

1 Introduction

There are many theoretical investigations where astrophysical modelling leads to systems of conservation laws similar to the Euler equations for the density of matter in space. Shu *et al*. [19] and Roberts [14] described such a system for a spiral galaxy consisting of two alternating arms of high and low density separated by discontinuities that are propagating shock waves. It is the location and rendition of these shock waves that has a significant impact on the overall solution, and so an accurate computation of the discontinuity in the flow field is of great importance. The results presented in this study are specific to the computation of shocks in astrophysical flow however the numerical treatment of shocks is valuable in many other areas. For example; The Euler equations are used extensively in aerodynamics in modelling the flow of air around an airfoil. These are typically three dimensional problems, although 2D and even 1D problems can result in a discontinuity in the pressure. Unsteady problems also arise in modelling wing flutter, or the flow patterns around rotating helicopter blades or the blades of a turbine. At high speeds these problems also involve the generation of shock waves. The scales involved in all of these problems varies considerably however the treatment of the shocks remains the same.

In van Albada *et al.*'s paper [23] a search is undertaken to find a reliable computational method for cosmic gas flow. In order to find a reliable method a comparison of a sample of carefully selected methods is done using a representative problem. The representative problem is a simplified version of the Shu and Roberts spiral galaxy system described above. In [14] Roberts uses the one-dimensional equations for steady-state gas that included a forcing term due to the spiral field of the stars to demonstrate the ease with which mild stellar spiral structures can induce shocks. Woodward [26] then showed that simplified, time-dependent versions of Roberts' equations could model the evolution of the flow. It is these simplified, time-dependent equations of Woodward and his set of parameter values that make up van Albada *et al.*'s representative test problem. Since our investigation is an extension of [23] it seemed only natural for us to use the same test problem. There are several distinguishing characteristics of this problem that make it highly suitable for the investigation: the major role of the source terms, the development of strong shocks and the significant rotational effects of the system. For a method to model these characteristics requires reliability and accuracy. Such a demanding test allows for a comprehensive comparison of the virtues and failings of all our sample methods. The details of the test problem are described in §2.

The comparison in [23] consisted of five algorithms: three commonly used in astrophysics, two of which are first-order accurate, a second-order accurate central differencing scheme and a second-order accurate upwind-differencing scheme. A detailed description of the methods is outlined in

§3. In order to compare our methods with those of van Albada *et al*. we first needed to recreate their results. Herein lies our first objective. This task however, proved more challenging than first anticipated as the transients in the system were much stronger than had previously been described in [23]. The result of this was that the testing procedure used in [23], where each method was run for a fixed time (1200$\pm$150 steps), would produce some solutions that had reached steady-state and some that had not. This meant the comparison of values would be unfair and therefore unreliable. The solution to this problem was simply to make sure each scheme had run for long enough to have reached steady-state. This was done by introducing a tolerance based on the root mean squared error (RMSE). Each scheme was then run until it was within the defined tolerance (see §3.1 and §4). This proved a much more versatile and reliable way to test and compare each of the methods.

Our second objective was to compare and test van Albada *et al*.'s list of methods with our sample of additional methods that consisted of two first-order approximate Riemann solvers, an additional flux-corrected transport method and the addition of flux limiters to one of the first-order approximate Riemann solvers. In §4 and §5 we analyse the results the additional methods produce and compare them to the results produced from the methods used in [23], keeping in mind the purpose of this all is to answer the question posed by van Albada *et al*. in the first place "What reliable, accurate, efficient and easy-to-program method should be used for this calculation?".

2 Basic Equations

Shu *et al*. [19] describes how the perturbations in the flow of interstellar gas resulting from a steady forcing F from spiral gravitational fields result in the nonlinear response of the density. It is this response, that if strong enough to exceed a critical value, forms the shock that is the basis of our test problem. In this section we present a form of Roberts' simplified equations that produce this very shock.

The fundamental isothermal equations of motion for gas flow may be written by an observer in an inertial frame of reference as:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{q}) = 0, \tag{2.1}$$

$$\frac{\partial \mathbf{q}}{\partial t} + \mathbf{q} \cdot \nabla \mathbf{q} = -\frac{c^2}{\rho} - \nabla \Phi, \tag{2.2}$$

where t denotes time, ρ is the gas density, $\mathbf{q}$ is the velocity, c the (constant) sound speed and Φ is the gravitational potential. The isothermal assumption is used since interstellar gas cools by radiative processes on a much shorter time-scale than that of any dynamical processes.

In the absence of the spiral forcing the gas flow is at a base state of motion. An equilibrium state of purely circular gas flow where the total smoothed central gravitational force field is exactly balanced by the inertial force associated with the rotation of the disk as a whole. For purely circular flow we have angular velocity $\Omega(\omega)$ at radius ω. A steady spiral field with small pitch angle α is assumed to rotate rigidly with pattern speed Ω_p. A convenient coordinate system is one which rotates at this speed and is aligned with the equipotential contours of the spiral. The coordinates parallel and perpendicular to the equipotential contours are denoted by ξ and η, respectively. The velocity components in this coordinate system are written as

$$\begin{aligned} v &= q_\xi, \\ u &= q_\eta. \end{aligned} \tag{2.3}$$

If we assume that the spiral has a pitch angle $\alpha \ll 1$, the equilibrium velocities are approximately

$$\begin{aligned} v_0 &= \omega(\Omega - \Omega_p), \\ u_0 &= \alpha\omega(\Omega - \Omega_p). \end{aligned} \tag{2.4}$$

In this approximation derivatives with respect to η (normal to the spiral arms) are retained, but derivatives with respect to ξ (along the spiral arms) are discarded.

For a two-armed spiral the resulting equations can be written as the system of conservation laws

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}(\mathbf{u})}{\partial \eta} = \mathbf{s}(\mathbf{u}), \tag{2.5}$$

where the vector of conserved quantities is

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \end{pmatrix}, \tag{2.6}$$

the vector of fluxes is

$$\mathbf{f} = \begin{pmatrix} \rho u \\ \rho(u^2 + c^2) \\ \rho u v \end{pmatrix}, \tag{2.7}$$

and the vector of source terms is

$$\mathbf{s} = \begin{pmatrix} 0 \\ 2\Omega(v - v_0)\rho + \dfrac{2}{\alpha\omega}\rho A \sin\hat{\eta} \\ -\dfrac{\kappa^2}{2\Omega}(u - u_0)\rho \end{pmatrix}. \tag{2.8}$$

The spiral phase $\hat{\eta}$ is defined by

$$\hat{\eta} = \frac{2\eta}{\alpha\omega}, \tag{2.9}$$

and the epicyclic frequency κ by

$$\kappa^2 = \frac{2\Omega}{r}\frac{d}{dr}(\omega^2\Omega). \tag{2.10}$$

In this approximation the flow is periodic; in terms of the spiral phase the periodicity condition reads

$$\mathbf{u}(\hat{\eta}, t) = \mathbf{u}(\hat{\eta} + 2\pi, t). \tag{2.11}$$

The driving term $(2/\alpha\omega)\rho A \sin\hat{\eta}$ comes from the assumed gravitational field of the stellar component.

For the test problem we adopt parameters thought to be appropriate for the neighbourhood of the Sun in our own galaxy: $\Omega = 25$ km s^{-1}/kpc, $\kappa = 31.3$ km s^{-1}/kpc, $\Omega_p = 13.5$ km s^{-1}/kpc, $c = 8.56$ km s^{-1}/kpc, $\omega = 10$ kpc and $\alpha = \sin(6.7^\circ) \approx 0.11667$. For the amplitude A we choose $A = 72.92$ (km s^{-1})^2, which makes the amplitude of the spiral force F, 2.0% of the equilibrium for $\omega\Omega^2$.

In steady state Eq. (2.5) becomes

$$\frac{du}{d\eta} = \frac{u}{u^2 + c^2}\left[2\Omega(v - v_0) + \frac{2A}{\alpha\omega}\sin\frac{2\eta}{\alpha\omega}\right], \tag{2.12}$$

$$\frac{dv}{d\eta} = -\frac{\kappa^2}{2\Omega}\frac{u - u_0}{u}, \tag{2.13}$$

A procedure for solving the steady state Eqs. (2.12) and (2.13) plus the periodicity condition (2.11) is described in Roberts [14]. Insufficient details are given to be able to recreate the solution however a very fine grid approximation (6400 zones) to the exact solution is used as a reference solution to compare our results with. The scheme used for this was the best performing scheme from [23], the MacCormack method (pre-shock and decompression smooth zones). Some of the issues that occur with such an approximation is a bias towards the MacCormack method run at a low resolution and the detrimental features reducing the accuracy of the method will also be prevalent in the reference solution.

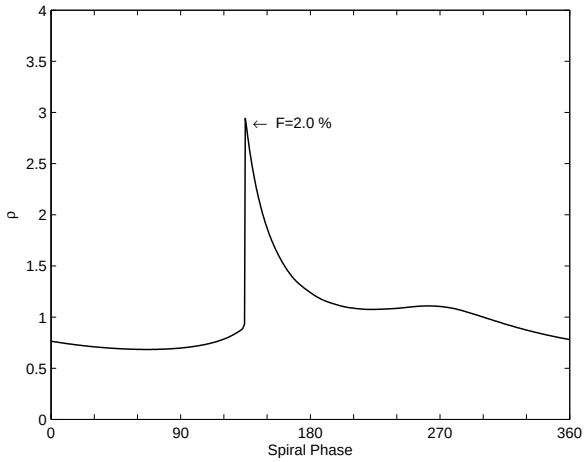

Figure 1: High resolution plot of the variation in gas density, ρ, as a function of the phase angle $\hat{\eta}$ with the parameters set to those defined in [26] and §2

Some of the noteworthy features visible in figure 1 are the shock at spiral phase $131.68°$, the rapid decompression after the shock and the secondary structure near a spiral phase of $270°$ which is caused by resonance effects. It is the time-dependent version of this problem that challenges a numerical method to cope with the shock, while also resolving the remaining smooth structure of the flow. In the next section we discuss how we transfer our continuous model and equations into discrete counterparts making them suitable for numerical evaluation.

3 Methods

We divide the spatial region $(0, \pi\alpha\omega)$ into N zones centred at the grid points equal to $(i - \frac{1}{2})\Delta\eta$, where $i = 1, 2, ..., N$ and $\Delta\eta = \pi\alpha\omega/N$ (or $\Delta\hat{\eta} = 2\pi/N$), and advance the approximate solution from time t_n to time t_{n+1} (where $t_{n+1} = t_n + \Delta t$) by means of a decretised version of Eq. (2.5). The approximate value of $\mathbf{u}$ at the point (η_i, t_n) is denoted by $\mathbf{u}_i^n$, $\mathbf{f}_i^n$ and $\mathbf{s}_i^n$, and are defined as $\mathbf{f}(\mathbf{u}_i^n, \eta_i)$ and $\mathbf{s}(\mathbf{u}_i^n, \eta_i)$ or, as explained in §3.8.2, $\mathbf{s}(\mathbf{u}_i^n, \eta_i, \Delta t)$. In all methods discussed here, $\mathbf{u}_i^n$ approximates the average value of the solution over zone i. A subscript $i + \frac{1}{2}$ denotes an interpolated numerical value at the zone boundary $\eta = i\Delta\eta$, or a finite difference across this boundary; a superscript $n + \frac{1}{2}$ denotes an approximate value $t_{n+1/2} = (n + \frac{1}{2})\Delta t$.

Our sample of numerical methods[1] includes

1. The Beam scheme (B) [16].
2. Second-order flux-splitting method (FS2) [24].
3. MacCormack's [12] method (MC2).
4. Flux Corrected Transport (FCT) methods
   - i) MacCormack's method
     - ◇ SHASTA [2]
     - ◇ Zalesak* [28]
   - ii) Lax-Wendroff [9] method (LW)
     - ◇ SHASTA
     - ◇ Zalesak*
5. Approxiate Riemann Solvers
   - i) Roe's [15] method* (R*)
     - ◇ Flux-limiters* [17]
       - – minmod
       - – superbee
       - – van Leer
     - ◇ Source term decomposed* [11]
   - ii) Harten, Lax & va Leer* (HLL*) [6]

An asterisk (*) denotes the method is an addition to the original sample used in [23].

[1] Godunov's method (G) could not be attempted as the referenced ICASE report [5] in [23], that includes the exact Riemann solution, is no longer available and there was insufficient time to reconstruct an exact solution ourselves.

Of our sample, the B, R* and HLL* methods are of first order accuracy meaning they approximate Eq. (2.5) with an error $O(\Delta\eta)$. All of the first-order methods and FS2 are based on upwind-differencing where the approximation of $\partial\mathbf{f}/\partial\eta$ makes a distinction between the positive and negative contributions of the wave motion. The methods MC2 and LW are of second-order accuracy meaning they approximate Eq. (2.5) with an error $O(\Delta\eta^2)$ and use central differencing which does not make the above distinction. Note the addition of flux-corrected limiters and flux limiters makes first-order methods second-order accurate.

Eq. (2.5) can be written in the form of the difference equations

$$\frac{\mathbf{u}_i^{n+1} - \mathbf{u}_i^n}{\Delta t} + \frac{\mathbf{h}_{i+1/2}^v - \mathbf{h}_{i-1/2}^v}{\Delta\eta} = \mathbf{s}_i^{n+1/2}, \tag{3.1}$$

with $v = n$ for the first order methods and $v = n + \frac{1}{2}$ for the higher order methods. The higher order methods, with the exception of Roe's method with flux limiters, are two-step algorithms where time centring is achieved using first-order accurate first steps at $t_{n+1/2}$ and t_{n+1}.

Since Eq. (3.1) has a numerical flux vector

$$\mathbf{h}_{i+1/2}^v \equiv \mathbf{h}(\mathbf{u}_{i-k+1}^n, ..., \mathbf{u}_{i+k}^n), \tag{3.2}$$

that is consistent with the analytic flux in the sense

$$\mathbf{h}(\mathbf{u}, ..., \mathbf{u}) = \mathbf{f}(\mathbf{u}), \tag{3.3}$$

the scheme is said to be in *conservation form* (see Appendix §6.1). Any scheme that can be written in this form is said to be *conservative*. The benefits of having a conservative scheme for a homogeneous equation is that, by the Lax-Wendroff theorem, if the numerical solution of a conservative scheme converges, it converges toward a weak solution of the conservation law and thus most importantly has the exact shock speed. For a non-homogeneous equation the situation is slightly more complicated but the conservation property is still very important. With this in mind it is then the particular choice of $\mathbf{h}$ that distinguishes one scheme from another. All of the methods are usually stable under the Courant-Friedrichs-Lewy condition, which says that the largest radial wave or signal speed in a cell must not exceed the numerical signal speed $\Delta\eta/\Delta t$. In the next section we describe the test conditions under which our methods will be implemented.

3.1 Test Setup

We attempted to keep the testing as close to that used in [23] as possible. Van Albada *et al*. uses Roberts' exact solution to the steady-state equations (2.12) and (2.13) as a reference solution however as previously stated we could not obtain an exact solution so a very fine grid approximation is used. Since no exact time-dependent solution exists we are restricted to testing the methods on the accuracy of the steady-state they produce. In order to determine whether a scheme has reached steady-state we would typically test for $|(\mathbf{u}^{n+1} - \mathbf{u}_{\text{ref}}) - (\mathbf{u}^n - \mathbf{u}_{\text{ref}})| = 0$ where $\mathbf{u}_{\text{ref}}$ is the reference solution as defined earlier. However, as the solutions do not converge to an absolute value and instead reach a steady oscillation about the reference solution, the conventional test returns a false positive as it passes through the reference solution. An alternative approach was therefore necessary. The final computed solutions of van Albada *et al*'s first tests are displayed as the root mean squared error (RMSE) as a percentage of the equilibrium values excluding 8 points straddling the shock, including 5 points in the decompression region. This is done to mainly indicate the accuracy in the smooth part of the solution. In keeping with van Albada *et al*'s methods it is the RMSE as a percentage of the equilibrium value ρu that will be used to test if the solution is within a certain margin of error σ. The reason for doing this is that for a steady-state solution we expect a uniform ρu distribution. Figure 3 highlights the disrupting effects of the transient waves in the ρu distribution for one of the methods used in [23]. Transient waves are mentioned in van Albada *et al*. as they destroy the stationary solution for the Phoenical SHASTA FCT method but are not mentioned to cause problems anywhere else.

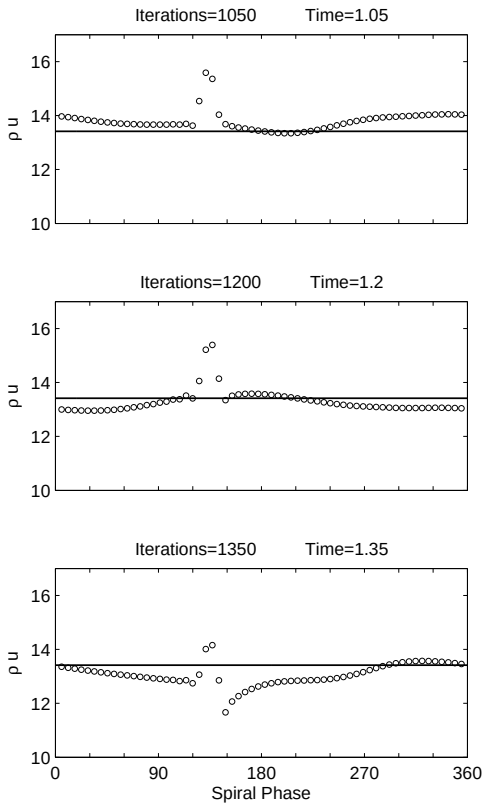

Figure 2: A MC2 time evolution plot of ρu for $t_n =$ 1050, 1200 and 1350 as used in [23].

Although the transients aren't mentioned as a problem in the other methods an alternative more accurate way to approximate the source term is suggested (for the direct integration of the source

term see §3.8.1) for maintaining stability in runs over 2,000 time-steps and it is this that has proved the most effective method for reducing the transients. In [23] the use of a direct integration method in the source term evaluation is used for stability in longer time runs only. We used it in all of our schemes except for the decomposition of the source term in R*. This was done to reduce the time it took to achieve steady-state. In general the time taken to reach steady-state was halved when using the more accurate source term evaluation. The results from using a more accurate source term evaluation are shown in §4 along side the specific details of the testing procedure.

Although we are testing the methods once they have reached steady-state, evaluating the methods on the time required for them to reach the steady-state starting, say, from uniform initial values $(\rho, u, v)_i^0 = (1, u_0, v_0)$, is unfair because the better methods will perform the worst because the transients will persist until they are damped out by the numerical viscosity because of the periodicity of the flow, which is highest for the least accurate schemes. As in [23] the fairest test to use is the reference solution itself as the initial-value distribution, and compare how well the various methods preserve it. However as later shown even when starting from the reference solution strong transient waves occur and it is some time before they are damped out. This constitutes the first test performed. Five of the additional methods that performed well in this test were applied to the problem with uniform initial values, mainly to determine their "robustness". This constitutes the second test.

The results presented are based on a computational grid of 64 zones; All of the methods use a constant time-step corresponding to a global Courant number of 0.5. In this span the fastest moving signals can traverse the computational domain about 10 times.

Even though conclusions about the accuracy with which the schemes deal with transient phenomena should not be given because there is no time-dependent exact solution to compare it with, we still present the number of steps required to reach our defined limit σ and also an elapsed time e. The elapsed time although dependent on the efficiency of the writing of the program and also the hardware of the computer system gives some idea of the computational cost in running each method and its subroutines.

All methods presented are compiled and run using Matlab R2009b numerical computing software on an Intel Core 2 Duo processor with 2 GB 1067 MHz DDR3 RAM.

3.2 The Beam Scheme (B)

There are essentially two approaches for identifying upwind directions, namely the Godunov approach and the flux-vector splitting approach. These two approaches are often referred to as the Riemann approach and the Boltzmann approach. The beam scheme, used mainly in astrophysics, is an example of the Boltzmann approach, where mass and momentum are transported by pseudo-particles with a velocity distribution $f(w)$. The velocity distribution is the sum of the number of delta functions (the beams); for present calculations we used three beams:

$$f(w) = \frac{1}{6}\rho\delta(w - [u - c\sqrt{3}]) + \frac{2}{3}\rho\delta(w - u) + \frac{1}{6}\rho\delta(w - [u + c\sqrt{3}]) \qquad (3.4)$$

although the middle beam is not really needed for this isothermal problem. Assuming that the velocity distribution is uniform and constant in each cell during the time-step, we can compute the positive direction, $\mathbf{f}^+(\mathbf{u})$, and in the negative direction, $\mathbf{f}^-(\mathbf{u})$:

$$\mathbf{f}^+(\mathbf{u}) = \begin{cases} \mathbf{f}(\mathbf{u}) = \begin{pmatrix} \rho u \\ \rho(u^2 + c^2) \\ \rho uv \end{pmatrix} & u \geq c\sqrt{3} \\ \mathbf{f}(\mathbf{u}) = \begin{pmatrix} \frac{1}{6}\rho(5u + c\sqrt{3}) \\ \frac{1}{6}\rho[4u^2 + (u + c\sqrt{3})^2] \\ \frac{1}{6}\rho v(5u + c\sqrt{3}) \end{pmatrix} & 0 \leq u < c\sqrt{3} \\ \mathbf{f}(\mathbf{u}) = \begin{pmatrix} \frac{1}{6}\rho(u + c\sqrt{3}) \\ \frac{1}{6}\rho(u + c\sqrt{3})^2 \\ \frac{1}{6}\rho v(u + c\sqrt{3}) \end{pmatrix} & -c\sqrt{3} < u < 0 \\ 0 & u \leq -c\sqrt{3} \end{cases} \qquad (3.5.1)$$

and $\mathbf{f}^-(\mathbf{u})$ is obtained from

$$\mathbf{f}^-(\mathbf{u}) + \mathbf{f}^+(\mathbf{u}) = \mathbf{f}(\mathbf{u}). \qquad (3.5.2)$$

The net flux across the cell interface at $\eta_{i+1/2}$, to be used in the scheme (3.1), is

$$\mathbf{h}^n_{i+1/2} = \mathbf{f}^+(\mathbf{u}^n_i) + \mathbf{f}^-(\mathbf{u}^n_{i+1}). \qquad (3.6)$$

That Eq. (3.6) leads to upwind differencing becomes clear when we write down the central difference of $\mathbf{h}$ needed in scheme (3.1):

$$\mathbf{h}^n_{i+1/2} - \mathbf{h}^n_{i-1/2} = \mathbf{f}^+(\mathbf{u}^n_i) - \mathbf{f}^+(\mathbf{u}^n_{i-1}) + \mathbf{f}^-(\mathbf{u}^n_{i+1}) - \mathbf{f}^-(\mathbf{u}^n_i). \qquad (3.7)$$

The results obtained with the beam scheme are displayed in figures 6 and 7.

3.3 Second-order Flux-splitting Method (FS2)

We can change any first-order upwind-differencing method into a second-order method by advancing the cell boundary values in the numerical flux vector and source terms to the intermediate time level $t_{n+1/2}$. Doing so allows the interaction between the cells to be fully ignored.

We choose $\mathbf{w}$ to be a vector of (not necessarily conserved) quantities describing the state of the gas, in particular

$$\mathbf{w} = \begin{pmatrix} \rho \\ u \\ v \end{pmatrix} = \begin{pmatrix} \rho \\ \\ \mathbf{q} \end{pmatrix}. \tag{3.8}$$

We then assume that the initial values for $\mathbf{q}$ form a piecewise linear distribution

$$\mathbf{w}^n(\eta) = \mathbf{w}_i^n + (\eta - \eta_i)\frac{(\delta\mathbf{w})_i^n}{\Delta\eta} \quad \eta_{i-1/2} < \eta < \eta_{i+1/2} \tag{3.9.1}$$

with

$$(\delta\mathbf{q})_i^n = c \cdot \text{ave}\left(\frac{\mathbf{q}_{i+1}^n - \mathbf{q}_i^n}{c}, \frac{\mathbf{q}_i^n - \mathbf{q}_{i-1}^n}{c}\right) \tag{3.9.2}$$

and

$$(\delta\rho)_i^n = \rho_i^n \cdot \text{ave}\left(2\frac{(\rho_{i+1}^n - \rho_i^n)}{(\rho_{i+1}^n + \rho_i^n)}, 2\frac{(\rho_i^n - \rho_{i-1}^n)}{(\rho_i^n + \rho_{i-1}^n)}\right) \tag{3.9.3}$$

where ave (a, b) is an averaging procedure[2] to be specified later. The formulation in Eq. (3.9.3) guarantees positivity for ρ when substituted in Eq. (3.9.1). Thus we have

$$\left(\frac{\partial\mathbf{w}}{\partial\eta}\right)_i^n = \frac{(\delta\mathbf{w})_i^n}{\Delta\eta}, \tag{3.10}$$

allowing us to calculate $(\partial\mathbf{w}/\partial t)_i^n$ from the appropriate modification (see Appendix §6.2) of Eq. (2.5). The cell averages are now advanced to $t_{n+1/2}$ and boundary values are calculated (the source terms have already been advanced, by Eq. (3.44) see §3.9.2):

$$\mathbf{w}_i^{n+1/2} = \mathbf{w}_i^n + \frac{\Delta t}{2}\left(\frac{\partial\mathbf{w}}{\partial\eta}\right)_i^n, \tag{3.11.1}$$

$$\mathbf{w}_{(i\pm1/2)\mp}^{n+1/2} = \mathbf{w}_i^{n+1/2} \pm \frac{\Delta\eta}{2}(\delta\mathbf{w})_i^n, \tag{3.11.2}$$

$$\mathbf{u}_{(i\pm1/2)\mp}^{n+1/2} = \mathbf{u}(\mathbf{w}_{(i\pm1/2)\mp}^{n+1/2}). \tag{3.11.3}$$

The time-centred fluxes at cell boundary $i \pm \frac{1}{2}$ can now be computed from $\mathbf{u}_{(i\pm1/2)-}^{n+1/2}$ and $\mathbf{u}_{(i\pm1/2)+}^{n+1/2}$ by any upwind-biased numerical flux formula.

[2]We have made the assumption that $b = 2(\rho_i^n - \rho_{i-1}^n)/(\rho_i^n + \rho_{i-1}^n)$ and not $b = 2(\rho_i^n - \rho_{i-1}^n)/(\rho_i^n + \rho_{i+1}^n)$ as stated in **[?]** as we believe it to be a typo. Both were tested and the difference is negligible.

Here we use a formula, due to van Leer [25], based on flux-vector splitting and therefore related to the flux in B; however, no particular velocity distribution is used in its derivation. The forward and backward fluxes of mass and momentum are defined according to

$$\mathbf{f}^{+}(\mathbf{u}) = \begin{cases} \mathbf{f}(\mathbf{u}) & u \geq c \\ \begin{pmatrix} \frac{\rho}{4c}(u+c)^2 \\ \frac{\rho}{2}(u+c)^2 \\ \frac{\rho}{4c}(u+c)^2 v \end{pmatrix} & |u| < c \\ 0 & u \leq -c \end{cases} \tag{3.12}$$

and (3.5.2), while **h** again is given by (3.6). The split flux (3.5.1) is smoother than (3.5), having continuous first derivatives. Furthermore, the reduced mass flux in (3.12) relative to (3.5.1) results in a reduced numerical diffusion.

In FS2 we then use

$$\mathbf{h}_{i+1/2}^{n+1/2} = \mathbf{f}^{+}(\mathbf{u}_{(i+1/2)-}^{n+1/2}) + \mathbf{f}^{-}(\mathbf{u}_{(i+1/2)+}^{n+1/2}). \tag{3.13}$$

The function ave (a, b) is chosen such that it tends to $\frac{1}{2}(a+b)$ if a and b are subsequent finite differences of a smooth solution, but tends to the smallest value where the solution is not smooth (see van Leer, [22]). We specifically choose

$$\text{ave}\,(a,b) = \frac{(b^2+\epsilon^2)a + (a^2+\epsilon^2)b}{a^2+b^2+2\epsilon^2}, \tag{3.14}$$

where ϵ^2 is a small non-vanishing bias of order $O((\Delta\eta)^3)$. This type of averaging prevents central differencing across a discontinuity in the solution or in its first derivative, which would lead to numerical oscillations. The bias prevents the undesirable clipping of smooth extremum but otherwise has negligible influence. In the actual computations we used $\epsilon^2 = 0.008$, but the results are not very sensitive to the precise value of ϵ^2. The results obtained with the second-order flux splitting methods are displayed in figures 8 and 9.

3.4 MacCormack Method (MC2)

The MC2 method is a variation of the two-step Lax-Wendroff [9] (LW) method (see Appendix §6.3 for formulation of method) but is much simpler in its application. The application of MC2 proceeds in two steps; a predictor step which is followed by a corrector step. Typically the predictor step uses a forward difference and the corrector step uses a backward difference. This can be changed to a backward difference predictor and a forward difference corrector either way there is a slight bias in the solution due to the one-sided differences. If desired, this balance can be reduced by alternating the direction of the predictor and corrector spatial differences from one time level to the next. The alternating predictor corrector approach is taken in [23] however through testing and comparison of the different variations we found the best performing MC2 method (see Appendix §6.6) was the forward backward difference MacCormack method (MC2FB).

The forward predictor step determines provisional values at t_{n+1} (source term is advanced by Eq. (3.44))

$$\bar{\mathbf{u}}_i^{n+1} = \mathbf{u}_i^n - \frac{\Delta t}{\Delta \eta}(\mathbf{f}_{i+1}^n - \mathbf{f}_i^n), \tag{3.15}$$

followed by a backward corrector step which determines the final values at t_{n+1}

$$\mathbf{u}_i^{n+1} = \frac{1}{2}\left[\mathbf{u}_i^n + \bar{\mathbf{u}}_i^{n+1} - \frac{\Delta t}{\Delta \eta}(\bar{\mathbf{f}}_i^n - \bar{\mathbf{f}}_{i-1}^n)\right]. \tag{3.16}$$

The corrector corresponds to inserting

$$\mathbf{h}_{i+1/2}^{n+1/2} \equiv \frac{1}{2}(\mathbf{f}_{i+1}^n + \bar{\mathbf{f}}_i^{n+1}) \tag{3.17}$$

into (3.1). This method is formally second order accurate in both space and time.

Although MacCormack's method is slightly dissipative, an explicit smoothing term had to be added in order to control nonlinear instabilities in the test problem. This was implemented by adding the term

$$\mathbf{d}_i^n = b\frac{\Delta t}{\Delta \eta}[\nu_{i+1/2}^n(\mathbf{u}_{i+1}^n - \mathbf{u}_i^n) - [\nu_{i-1/2}^n(\mathbf{u}_i^n - \mathbf{u}_{i-1}^n)] \tag{3.18}$$

to the right-hand side of Eq. (3.16). The coefficient b is an adjustable constant of order unity. For our testing we set $b = 1.0$.

$$\nu_{i+1/2}^n = |\mathbf{u}_{i+1}^n - \mathbf{u}_i^n|, \tag{3.19}$$

for the artificial-viscosity coefficient was made on the basis of Liebovich's [10] experience with MC2 in two-dimensional calculations. The results obtained with MC2 are displayed in figures 10 and 11.

3.5 Flux-corrected Transport methods (FCT)

It is well know that higher order schemes (order 2 and above) suffer from dispersive ripples when approximating steep gradients which is particularly relevant in our case. And that lower order schemes or higher order schemes with a zeroth order diffusive term added produce no ripples but suffer from excessive numerical diffusion. FCT is a technique developed by Boris and Book [1] which embodies the best of both of the above worlds. Formally the procedure is as follows:

1. Update the solution using a low order scheme with numerical flux $\mathbf{h}^L$. Where $\mathbf{u}_i^{td}$ is the "transported and diffused" solution. One of the simplest ways of doing this is by adding to the scheme of Lax-Wendroff (see Appendix §6.3 Eq. (6.15)) or MacCormack (3.16) the strong artificial diffusion term

$$\mathbf{d}_i^n = \frac{1}{8}(\mathbf{u}_{i+1}^n - 2\mathbf{u}_i^n + \mathbf{u}_{i-1}^n), \tag{3.20}$$

2. Compute an anti-diffusive flux as the difference between a high order numerical flux $\mathbf{h}^H$ (MC2 or LW without the diffusive term) and the low order flux used in stage one

$$\mathbf{h}_{i+1/2}^A = \mathbf{h}_{i+1/2}^H - \mathbf{h}_{i+1/2}^L \tag{3.21}$$

3. Limit the anti-diffusive flux in such a manner that no new extrema are introduced in stage four

$$\mathbf{h}_{i+1/2}^C = C_{i+1/2}\mathbf{h}_{i+1/2}^A \qquad 0 \leq C \leq 1 \tag{3.22}$$

4. Correct the solution using $\mathbf{h}_{i+1/2}^C$

$$\mathbf{u}_i^{n+1} = \mathbf{u}_i^{td} - \frac{1}{8}(\mathbf{h}_{i+1/2}^C - \mathbf{h}_{i-1/2}^C) \tag{3.23}$$

In their original paper, Boris and Book proposed the following limitation on the anti-diffusive flux for their SHASTA [2] (Sharp and Smooth Transport Algorithm)

$$\mathbf{h}_{i+1/2} = \min\{8|\mathbf{h}_{i-1/2}|, |\mathbf{h}_{i+1/2}|, 8|\mathbf{h}_{i+3/2}|\}\text{sgn}(\mathbf{h}_{i+1/2}) \tag{3.24}$$

Zalesak [27] proposed an alternative flux limiting algorithm

$$C_{i+1/2} = \begin{Bmatrix} \min(1, Q_i^-/P_i^-) & \text{if } P_i^- > 0 \\ 0 & \text{if } P_i^- = 0 \end{Bmatrix} \tag{3.25}$$

where , Q_i^- and P_i^- are defined in the Appendix §6.4. The results obtained with the FCT methods are displayed in figures 12 to 19.

3.6 Approximate Riemann Solvers

Whilst for some scalar conservation laws the Riemann problem is easily solved, this is not the case for non-linear systems of conservation laws where the k^{th} wave may be a rarefaction, shock or contact discontinuity.

Up to now the only methods employed to solve this problem have been non-upwind biased schemes based on the exact solution of the Riemann problem (G). The solution must therefore be constructed in phase space to link the two end states by a path of valid waves. Here an iterative procedure is often required which, since this must be used at every cell boundary at every time step, will make it the most computationally expensive task of the whole method. To simplify the process and reduce this overhead approximate Riemann solvers, which do not employ iteration, are often used. We wish to use a method that employs upwind bias on each characteristic field. This can be achieved either by approximating the Riemann states and applying the physical flux, or by approximating the numerical flux directly. We look at the latter, and outline the distinguishing features of some of the approximate Riemann solvers used.

3.6.1 Roe's Method (R*)

Perhaps the simplest approximate Riemann solver is due to Roe. The solution is based on solving a localised Riemann problem to calculate the flux at the interface of each grid zone. The basic premise of this problem is that changes in a flow can be transmitted only through waves, and only at some given speeds, which represent the eigenvalues of the governing non-linear equation system. For a three equation system there are three wave speeds corresponding to the speed the gas is travelling and the speed of sound relative to the gas speed in the upstream and downstream directions. Since the solution to the equation set changes only across one of these waves, the solution to the system in any point in space and time can be represented by the summation of the states to the extreme left or right of the space, plus or minus one or more of the state changes across the waves. To start it useful to remember the system of conservation laws given in §2 by Eq. (2.5).

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}(\mathbf{u})}{\partial \eta} = \mathbf{s}(\mathbf{u}).$$

In quasi-linear form

$$\frac{\partial \mathbf{u}}{\partial t} + A(\mathbf{u})\frac{\partial \mathbf{u}}{\partial \eta} = \mathbf{s}, \tag{3.26}$$

where $A(\mathbf{u})$ is the Jacobian matrix $J_{\mathbf{f}} = \partial \mathbf{f}/\partial \mathbf{u}$.

Roe linearises this form of the equations in each interval (η_i, η_{i+1}) by replacing the Jacobian by interval-wise constant matrices $\tilde{A}(\mathbf{u}_i, \mathbf{u}_{i+1})$ which satisfy his *Property U*, i.e. for any two adjacent states $\mathbf{u}_L$ and $\mathbf{u}_R$:

- $\tilde{A}(\mathbf{u}_L, \mathbf{u}_R)$ is diagonalisable with real eigenvalues (hyperbolicity);
- $\tilde{A}(\mathbf{u}_L, \mathbf{u}_R) \to A(\mathbf{u})$ as $\mathbf{u}_L$ and $\mathbf{u}_R \to \mathbf{u}$ (consistency)
- $\mathbf{f}(\mathbf{u}_L - \mathbf{u}_R) = \tilde{A}(\mathbf{u}_L, \mathbf{u}_R)(\mathbf{u}_L - \mathbf{u}_R)$ (conservation)

The first two conditions are readily satisfied if $\tilde{A}$ is taken to be the Jacobian evaluated at an averaged state, i.e. $\tilde{A}(\mathbf{u}_L, \mathbf{u}_R) = A(\bar{u})$. Once $\tilde{A}$ has been obtained, it is diagonalised which results in a set of decoupled linear advection equations in each interval.

The flux differences $\mathbf{f}_R - \mathbf{f}_L$ in each interval are then decomposed onto local eigenvectors

$$\Delta\mathbf{f} = \mathbf{f}_R - \mathbf{f}_L = \sum_{k=1}^{3} \tilde{\alpha}^{(k)} \lambda^{(k)} \mathbf{x}^{(k)} \tag{3.27}$$

where $\tilde{\alpha}^{(k)}$, $\tilde{\lambda}^{(k)}$ and $\tilde{\mathbf{x}}^{(k)}$ are the coefficient for Δu, the eigenvalue and eigenvector, respectively, corresponding to the k^{th} characteristic field of $\tilde{A}$.

Roe's original scheme updated the solution by upwinding and directly adding these flux difference components.

It may be placed in the framework of intercell fluxes by integration around the half cell $(x_{i+1/2}, x_{i+1}) \times (t^n, t^{n+1})$ (similar to the derivation of Godunov's scheme [4]) resulting in the flux

$$\mathbf{h}_{i+1/2} = \frac{1}{2}(\mathbf{f}_{i+1} + \mathbf{f}_i) - \frac{1}{2} \sum_{k=1}^{3} \tilde{\alpha}^{(k)} |\lambda|^{(k)} \mathbf{x}^{(k)}. \tag{3.28}$$

In this formulation $\tilde{A}$ can be seen to be identified with the cell interfaces (see Appendix §6.5 for term evaluation).

Because the resulting individual approximate Riemann problems are linear, their solutions contain only discontinuities and not expansion fans. For this reason Roe's original method is not entropy satisfying, although a number of entropy fixes have since been proposed. The results obtained with the R* are displayed in figures 20 and 21.

3.6.2 HLL Method (HLL*)

For the purpose of computing a Godunov flux, Harten, Lax and van Leer [6] presented a novel approach for solving the Riemann problem approximately. The resulting Riemann solvers have become known as HLL Riemann solvers. In this approach an *approximation for the intercell numerical flux is obtained directly*. The central idea is to assume, for the solution, a wave configuration that consists of *two waves* separating three constant states.

If $s^R_{i+1/2}$ and $s^L_{i+1/2}$ are upper and lower bounds, respectively, for the largest and smallest signal velocities resulting from the solution of the Riemann problem centred at $\eta_{i+1/2}$, then the intermediate state is obtained from conservation to be

$$\mathbf{u}^{\text{HLL}}_{i+1/2} = \frac{s^R_{i+1/2}\mathbf{u}_i - s^L_{i+1/2}\mathbf{u}_i}{s^R_{i+1/2} - s^L_{i+1/2}} - \frac{\mathbf{f}_{i+1} - \mathbf{f}_i}{s^R_{i+1/2} - s^L_{i+1/2}}. \tag{3.29}$$

Integrating this solution substituted into the conservation law over the half cell $(\eta_{i+1/2}, \eta_i) \times (t^n, t^{n+1})$ results in the HLL flux. Note the line of interest is the vertical line $\eta/t = 0$ in the Riemann problem as it is this which gives the flux

$$\mathbf{f}^{\text{HLL}}_{i+1/2} = \begin{cases} \mathbf{f}_i & \frac{\eta}{t} \le s^L_{i+1/2} \\ \mathbf{f}^{\text{HLL}}_{i+1/2} & s^L_{i+1/2} \le \frac{\eta}{t} \le s^R_{i+1/2} \\ \mathbf{f}_{i+1} & s^L_{i+1/2} \le \frac{\eta}{t} \end{cases} \tag{3.30}$$

where

$$\mathbf{f}^{\text{HLL}}_{i+1/2} = \frac{s^R_{i+1/2}\mathbf{f}_i - s^L_{i+1/2}\mathbf{f}_{i+1} + s^L_{i+1/2}s^R_{i+1/2}(\mathbf{u}_{i+1} - \mathbf{u}_i)}{s^R_{i+1/2} - s^L_{i+1/2}} \tag{3.31}$$

It now remains to specify the upper and lower bounds $s^R_{i+1/2}$ and $s^L_{i+1/2}$

1. One possibility is to evaluate them directly

$$s^L_{i+1/2} = \frac{u_i + u_{i+1}}{2} - c, \quad s^R_{i+1/2} = \frac{u_i + u_{i+1}}{2} + c \tag{3.32}$$

2. or, as suggested by Davis to use the maximum eigenvalue evaluated at the right state and the minimum eigenvalue evaluated at the left state respectively.

$$s^L_{i+1/2} = \min(u_L - c, u_R - c), \quad s^R_{i+1/2} = \max(u_L + c, u_R + c) \tag{3.33}$$

3. Another alternative, suggested by Davis and also Einfeldt is to use Roe's averaged eigenvalues as estimates.

$$s^L_{i+1/2} = \tilde{u}_{i+1/2} - c, \quad s^R_{i+1/2} = \tilde{u}_{i+1} + c \tag{3.34}$$

The results obtained with the HLL schemes are displayed in figures 28 to 35.

3.7 Flux Limiters (FL*)

Flux limiters are a class of schemes similar to the Flux-Corrected Transport methods which encompass the Total Variation Diminishing (TVD) algorithms of several authors. There are two main differences however between the approach adopted here and that of Boris and Book (and later Zalesak). Firstly the FCT algorithm was essentially a two-step procedure, whereas here we adopt a single-step approach; and secondly the FCT limiter was constricted by unity whilst we allow a more generous upper limit.

In order to apply the flux limiters to Roe's method we must first write the scheme in *flux form*.

$$\mathbf{h}_{i+1/2} = \frac{1}{2}(\mathbf{f}_{i+1} + \mathbf{f}_i) - \frac{1}{2}\sum_{k=1}^{3}(1 - \phi_i(1 - |\nu|))\tilde{\alpha}^{(k)}|\lambda|^{(k)}\mathbf{x}^{(k)} \tag{3.35}$$

where $\nu = (\Delta t/\Delta\eta)\lambda^{(k)}$. The *flux limiter* ϕ is a function of the ratio of consecutive gradients of the solution, i.e.

$$\phi_i = \phi(r_i) \tag{3.36}$$

where

$$r_i = \frac{\delta u^n_{i-1/2}}{\delta u^n_{i+1/2}} \tag{3.37}$$

The Flux Limiter framework encompasses the work of several authors, including:

1. Roe's *minmod* flux limiter which adaptively switches between Lax-Wendroff and Warming and Beam
$$\phi_{mm}(r) = \max(0, \min(1, r)) \tag{3.38}$$
2. Roe's compressive *superbee* limiter
$$\phi_{sb}(r) = \max(0, \min(2r, 1)), \min(r, 2)) \tag{3.39}$$
3. Van Leer's limiter which switches between non-conservative forms of Lax-Wendroff and Warming and Beam
$$\phi_{vL} = \frac{r + |r|}{1 + |r|} \tag{3.40}$$

All of the above limiters have the property

$$\frac{\phi(r)}{r} = \phi\left(\frac{1}{r}\right) \tag{3.41}$$

which is a symmetry property ensuring that the top of a discontinuity is treated symmetrically to the bottom of a discontinuity. Not all limiters have this property and neither do the FCT limiters mentioned in §3.5. The results obtained with the addition of flux limiters to R* are displayed in figures 22 to 27.

3.8 Source Terms

In Albada *et al.*'s study [23] a choice is made between two different approaches to evaluating the source term of Eq. (2.5). For runs less than 2,000 time-steps a point-wise method is used and for runs greater than 2,000 time-steps a direct integration method is used for greater stability. The point-wise method is only applied to MC2 and LW and so time centring is achieved in a two step procedure. As previously mentioned for all of our methods we chose to use the direct integration method for faster convergence however both methods will be presented here as well as a third method that decomposes the source term using the same technique as used in the Appendix §6.5.

3.8.1 Point-wise method

If we take our discretised version of Eq. (2.5) with source term

$$\mathbf{u}_i^{n+1} = \mathbf{u}_i^n - \frac{\Delta t}{\Delta \eta}(\mathbf{h}_{i+1/2}^{n+1/2} - \mathbf{h}_{i-1/2}^{n+1/2}) + \Delta t \mathbf{s}_i^{n+1/2}. \tag{3.42}$$

Using a two-step predictor corrector algorithm corresponds to inserting

$$\mathbf{s}_{i+1/2}^{n+1/2} \equiv \frac{1}{2}(\mathbf{s}_{i+1}^n + \bar{\mathbf{s}}_i^{n+1}) \tag{3.43}$$

where $\bar{\mathbf{s}}_i^{n+1}$ is the predicted value at time level t_{n+1}. The source term is then evaluated at each point in the spatial region, hence point-wise.

3.8.2 Direct Integration Method

We start by integrating the following equation

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{s}, \tag{3.44}$$

over a half time-step, then continuing by integrating

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}}{\partial \eta} = 0, \tag{3.45}$$

over a full time step, and finishing off with integrating Eq. (3.44) over a half time-step again. The advantage of this method is that Eq. (3.44) can be accurately integrated however this may not always be possible.

For the current problem the analytic solution of (3.44) is

$$\begin{aligned}
&\rho(t_n+\tau,\eta_i) = \rho_i^n,\\
&u(t_n+\tau,\eta_i) - u_0 = (u_i^n - u_0)\cos(\kappa\tau)\\
&\qquad\qquad + \frac{2\Omega}{\kappa}\left(v_i^n - v_0 + \frac{A}{\alpha\omega\Omega}\sin\hat{\eta}\right)\sin(\kappa\tau),\\
&v(t_n+\tau,\eta_i) - v_0 = -\frac{A}{\alpha\omega\Omega}\sin\hat{\eta}\\
&\qquad\qquad + \left(v_i^n - v_0 + \frac{A}{\alpha\omega\Omega}\sin\hat{\eta}\right)\cos(\kappa\tau) - \frac{\kappa}{2\Omega}(u_i^n - u_0)\sin(\kappa\tau)
\end{aligned} \tag{3.46}$$

As stated in [23] the cause of the instability that arises is the same that would make a linear first-order algorithm for integrating (3.44) unstable: instead of choosing $(u(t_n+\tau), v(t_n+\tau))$ on the ellispe given by (3.46), the linearised version will put it on the tangent to that ellipse, thus always leading to an amplification of the disturbance.

3.8.3 Decomposition

A possible way to treat non-homogeneous terms is to approximate the r.h.s term of Eq. (2.5), **s**, by:

$$\tilde{\mathbf{s}} = \frac{1}{2}[\mathbf{s}_i + \mathbf{s}_{i+1}] \tag{3.47}$$

We now decompose the term into wave components through the β terms as is done in equations (6.28 through to 6.61):

$$\tilde{\beta}^{(1)} = \Delta\eta\frac{\tilde{\mathbf{s}}_3 - \tilde{v}}{\tilde{\lambda}^{(1)}\tilde{\mathbf{s}}_1} \quad \tilde{\beta}^{(2)} = \Delta\eta\frac{\tilde{\mathbf{s}}_2 - (\tilde{u}-c)\tilde{\mathbf{s}}_1}{2c\tilde{\lambda}^{(2)}} \quad \tilde{\beta}^{(3)} = \Delta\eta\frac{-\tilde{\mathbf{s}}_2 + (\tilde{u}+c)\tilde{\mathbf{s}}_1}{2c\tilde{\lambda}^{(3)}} \tag{3.48}$$

The full scheme now reads

$$\mathbf{u}_i^{n+1} = \mathbf{u}_i^n - \frac{\Delta t}{\Delta\eta}(\mathbf{f}_{i+1/2}^n - \mathbf{f}_{i-1/2}^n) + \frac{\Delta t}{2}(\tilde{\mathbf{s}}_{i+1/2}^n + \tilde{\mathbf{s}}_{i-1/2}^n), \tag{3.49}$$

where

$$\mathbf{f}_{i+1/2}^n = \frac{1}{2}(\mathbf{f}_{i+1} + \mathbf{f}_i) - \frac{1}{2}\sum_{k=1}^{3}(\tilde{\alpha}^{(k)} + \tilde{\beta}^{(k)})|\lambda|^{(k)}\mathbf{x}^{(k)} \tag{3.50}$$

The results obtained with the decomposition of the source term are displayed in figures 28 and 29. The next section presents the results obtained from applying the methods to the representative problem.

4 Results

The first method attempted was that of MacCormack. To maintain consistency with van Albada *et al*. the solution was plotted at $t_n = 1200 \pm 150$. The results from which were far from acceptable due to the large variation in ρ, u and v over such a short period of time. By animating the temporal progression of our solution we found the disruption of the convergence was because of strong transients. This is difficult to see from figure 3 and in general would not be picked up without animation. A better way to illustrate it is with the distribution of ρu as shown in §3.1. It is important to recall that all of the results that have been plotted in figures 4-5 are based on the RMSE of the smooth values straddling the shock. The reason for this is that no scheme is expected to perfectly model the shock and the area of interest is the accuracy with which the pre-shock and decompression zones are modelled. It is therefore important to continue to analyse the data in this way to maintain consistency within both sets of data (ours and Albada *et al*.). Figure 4 gives a much clearer picture of what is happening. The scheme is converging to a steady oscillation that is within 1% of the reference solution. Notice the marked improvement resulting from the use of the direct integration of the source term. This increase in the rate convergence was also seen when applied to LW. Although the graphs give a better understanding of what is happening it was our wish to use this information to form a more reliable testing method.

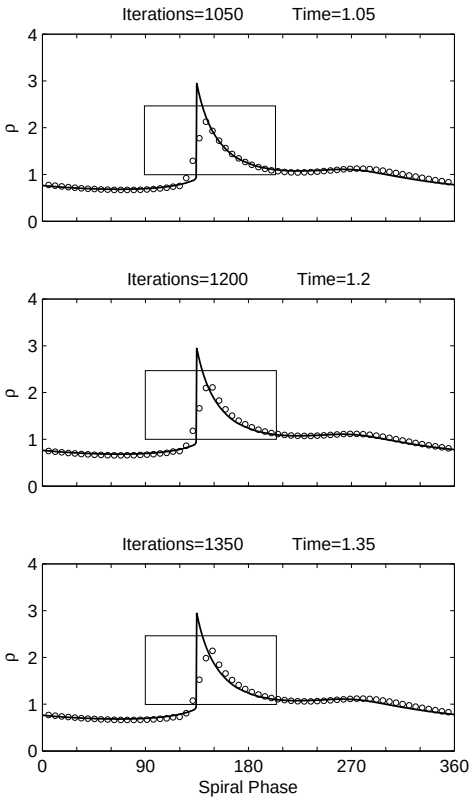

Figure 3: A MC2 time evolution plot of ρ for $t_n = 1050$, 1200 and 1350 as used in [23],

Using a simple peak finding algorithm (see figure 5) it was possible to include a realtime bound, σ, on the accuracy of the scheme. Using this bound meant that when run, each scheme's maximum RMSE of ρu was known to be within that bounded value. Once reached the values of ρ, u and v would then be returned and output to file along with the number of time-steps n and the elapsed time e.

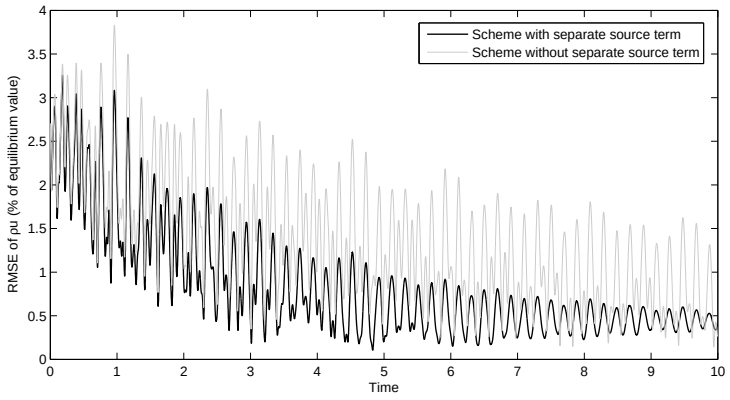

Figure 4: Plot comparing the use of the direct integration method for source term evaluation.

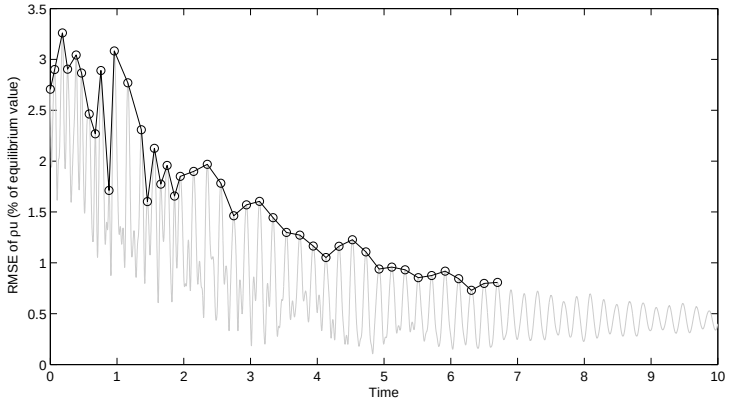

Figure 5: Plot highlighting the capabilities of the peak finding algorithm.

Since the RMSE values the methods produce are mainly an indication of the accuracy in the smooth part of the solution the analysis of the shock will come from visual inspection of the plots. With this in mind we should ask the, question what kinds of solution features cause numerical problems?

To begin with, many numerical methods have difficulties when a wave speed equals zero. In one dimension, wave speeds equal zero at *sonic points*. For our particular problem the sonic point is located at a spiral phase $\hat{\eta} = 155.53°$. In other words $\mathbf{f}'(\mathbf{u}^*) = A(\mathbf{u}^*) = 0$, where $\mathbf{u}^*$ is the solution at the sonic point. Sonic points usually signal a change in the wind direction. Unless specific steps are taken, many numerical methods produce significant errors near sonic points. Besides sonic points, shock waves, contact discontinuities, expansion fans, and other non-smooth features are major stumbling blocks for our numerical methods. Typical symptoms include oscillations, overshoots, and a smearing that spreads the discontinuity over a region of several cells.

4.1 Test One

4.1.1 The Beam Scheme

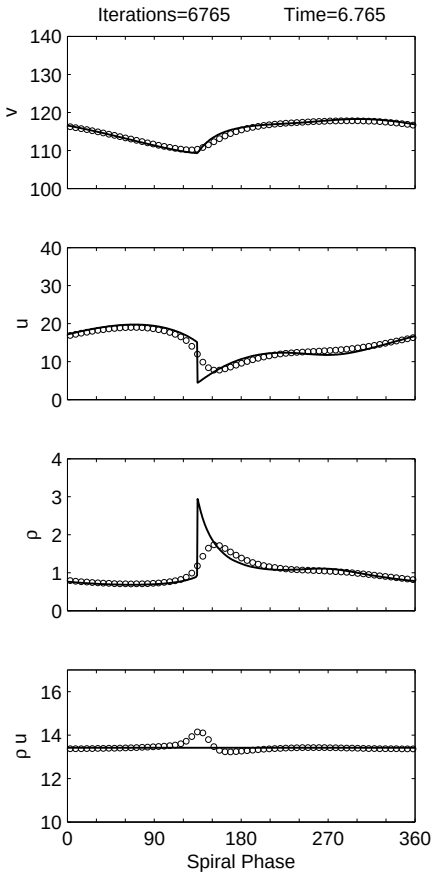

Figure 6: Numerical results (circles) obtained with the beam scheme (B), starting from the reference solution (line) and run until the RMSE of ρu is within 1%.

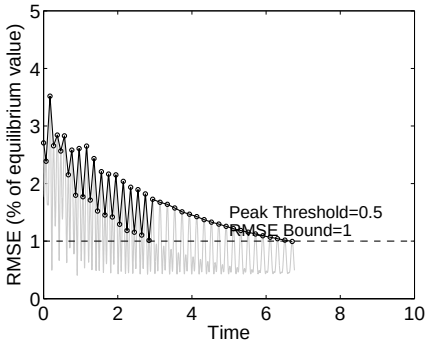

Figure 7: The beam scheme time series plot of the RMSE of ρu.

From figure 6 we can see there is strong numerical diffusion occurring when it encounters the shock which results in the density maximum being severely underestimated. The consequence of the maximum density being under-shot is barely subsonic values of u. This related response is evidence the solutions are coupled.

Note the displacement of shocks in the downstream direction is a typical problem encountered by upwind-differencing methods. Notice how the zone immediately before the shock is not influenced by the downstream subsonic region.

The non-uniform distribution of ρu across the discontinuity is further evidence the shock it is not being modelled accurately. Only the most general features of the smooth region are represented. This result coincides with that of van Albada *et al*.

4.1.2 Second-order flux-splitting method

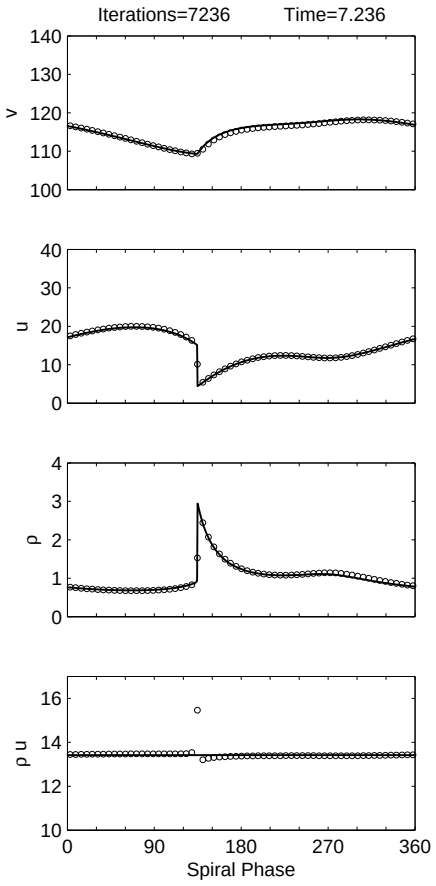

Figure 8: Results of the second-order flux-splitting method (FS2) run until the RMSE of ρu is within 1%.

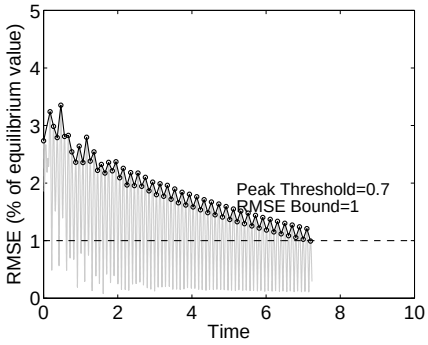

Figure 9: The second-order flux-splitting method (FS2) time series plot of the RMSE of ρu.

FS2 produces an oscillation-free, sharp rendition of the shock showing the efficiency of applying dissipation through the averaging function Eq. (3.14) is well worth the additional programming effort.

Notice how only one point in the plot of ρu struggles with the approximation of the discontinuity when compared to the previous values of the beam scheme. The accuracy with which the shock is modelled is evidence that flux-splitting is a very good approach to take when modelling shocks. This result coincides with that of van Albada *et al*.

4.1.3 MacCormack's method (FB)

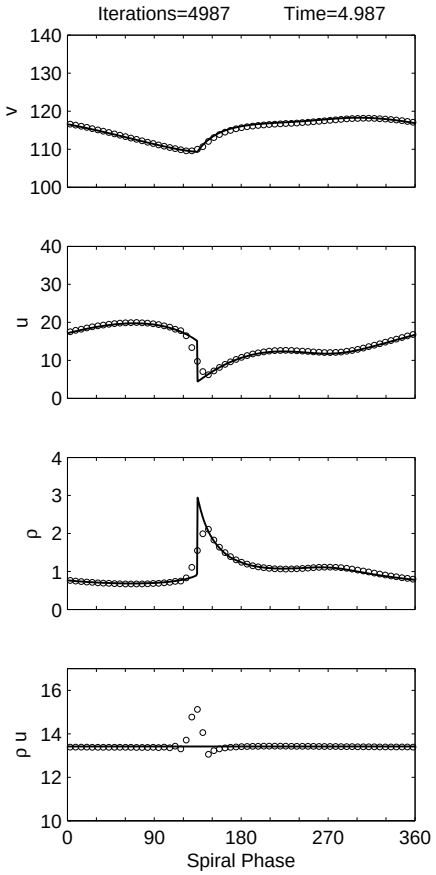

Figure 10: Results of the MacCormack method (MC2) run until the RMSE of ρu is within 1%.

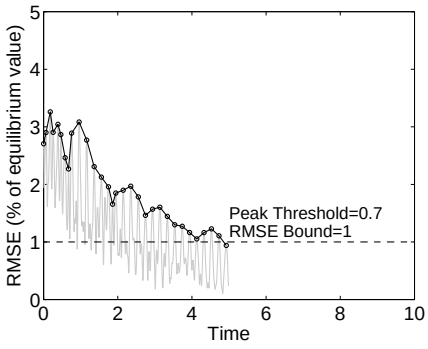

Figure 11: The MacCormack method (MC2) time series plot of the RMSE of ρu.

Visually our results confirm those presented in [23]. The MacCormack method produces a reasonably narrow shock that is in the right place. MC2's representation of the smooth region either side of the shock is the best out of all of the methods however, as previously explained this is expected since the reference solution is calculated using MC2 and the testing is therefore biased towards the results produced by MC2.

Van Albada *et al*. describes oscillations occurring before the shock in both the density ρ and the velocity component u. From figure 9 these observations are not obvious, however it is clear the pre-shock oscillations are still occurring as they are visible in the plot of ρu. This suggests that the solutions presented in [23] are still converging to a steady oscillation.

One thing that was noticed with the animation of the solution was that the transients that occur in the temporal progression propagate only in the forward direction. This may be the reason the forward backward predictor corrector sequence performed better.

4.1.4 SHASTA
Flux Corrected Transport (LW)

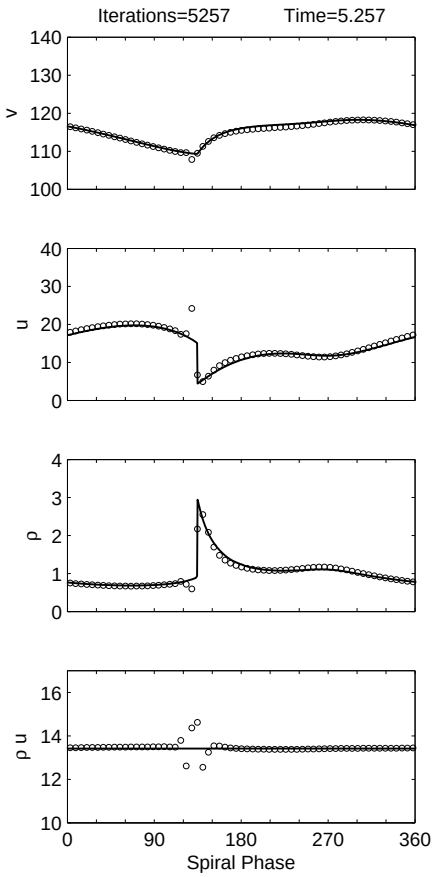

Figure 12: Results of the SHASTA flux-corrected transport (LW) method (FCTLWS) run until the RMSE of ρu is within 1%.

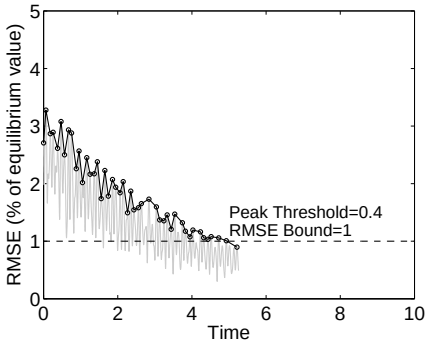

Figure 13: The SHASTA flux-corrected transport (LW) method (FCTLWS) time series plot of the RMSE of ρu.

Similar to the previous methods the results from the SHASTA flux-corrected transport method confirm those in [23]. Large oscillations before the shock are translated across the rest of the system resulting in an unsatisfactory solution. While this method is instantly discarded in [23] the next figure shows there are potential benefits to using flux-correcting limiters in our methods.

4.1.5 SHASTA
Flux Corrected Transport (MC2)

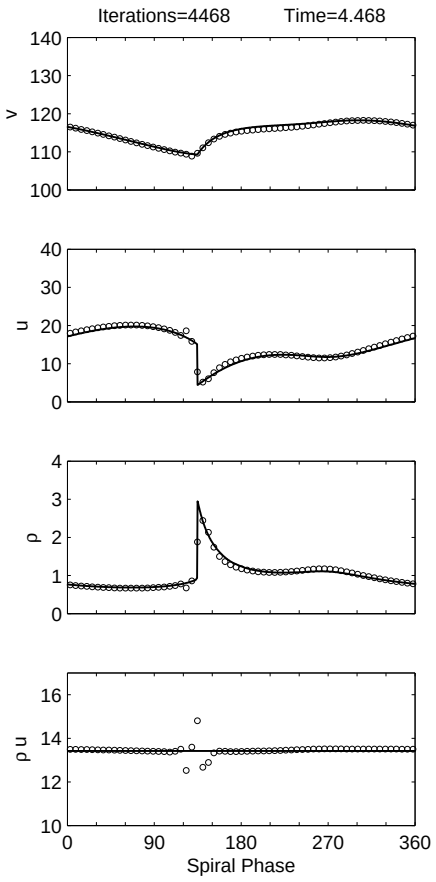

Figure 14: Results of the SHASTA flux-corrected transport (MC2) method (FCTMC2S) run until the RMSE of ρu is within 1%.

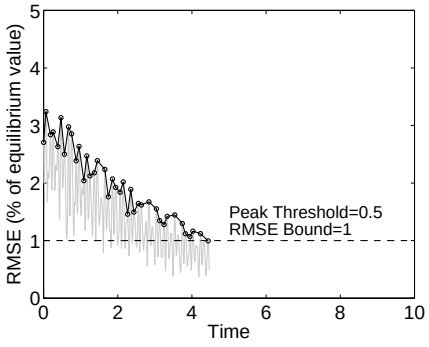

Figure 15: The SHASTA flux-corrected transport (MC2) method (FCTMC2S) time series plot of the RMSE of ρu.

Since we had already written a program for the MacCormack method the addition of a flux-correcting limiter was straight forward. Figure 14 shows a fair rendition of the smooth region and also a sharp, narrow, well positioned shock, however, we must keep in mind that the performance in the smooth region is once again biased due to the use of MC2 in the reference solution.

Although oscillations still occur they are damped. This can also be gauged on the amount of disturbance in the plot of ρu which is slightly less than in LW.

4.1.6 Zalesak
Flux Corrected Transport (LW)

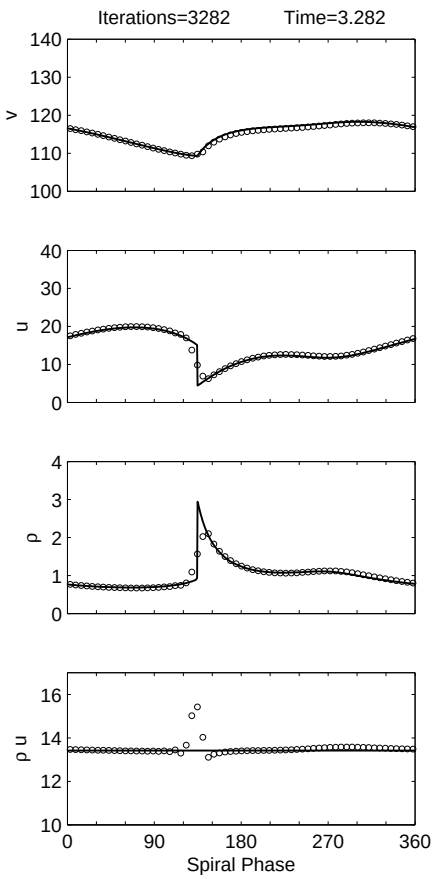

Figure 16: Results of the Zalesak flux-corrected transport (LW) method (FCTLWZ) run until the RMSE of ρu is within 1%.

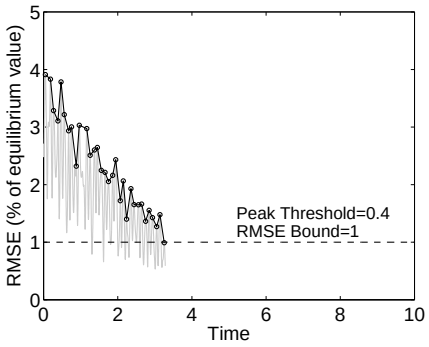

Figure 17: The Zalesak flux-corrected transport (LW) method (FCTLWZ) time series plot of the RMSE of ρu.

The Zalesak Flux-Corrected Transport method produces almost identical results to MC2 although not quite as accurate as shown in table 1. The addition of the strong diffusive term is anti-diffused by the flux-correcting method but not enough to the point where the density maximum matches that of the reference solution and it is therefore under-shot. The corresponding u values are therefore also underestimated.

Considering that the complexity of Zalesak's algorithm requires more programming effort and the result achieved is slightly worse than MC2 the Zalesak algorithm should be deemed unsuitable for this problem however, since the solution is so close to MC2's, the rate of convergence should be included as a contributing factor. The number of steps it takes to achieve steady-state is nearly half the number of steps it takes for MC2 to converge to steady-state. This result however is inconsistent with the rest of our findings.

4.1.7 Zalesak
Flux Corrected Transport (MC2)

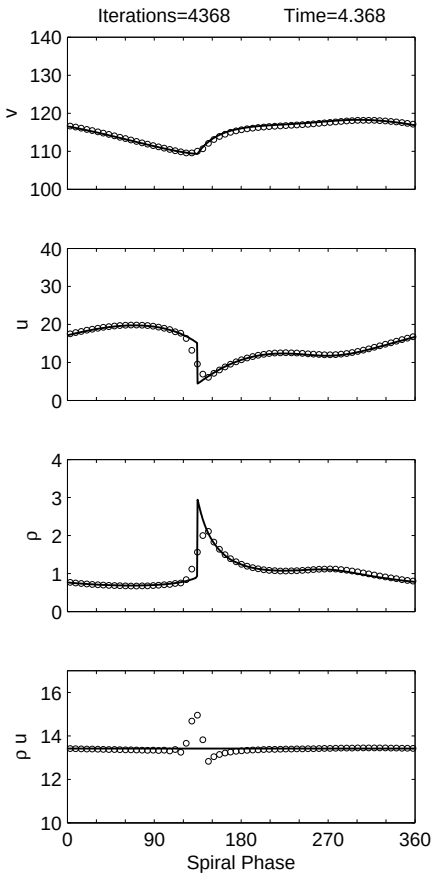

Figure 18: Results of the Zalesak flux-corrected transport (MC2) method (FCTMC2Z) run until the RMSE of ρu is within 1%.

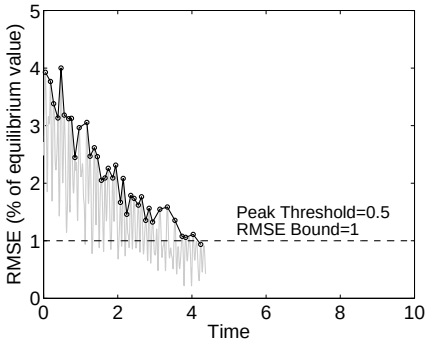

Figure 19: The Zalesak flux-corrected transport (MC2) method (FCTMC2Z) time series plot of the RMSE of ρu.

The results from the application of the Zalesak FCT method to MC2 returns near identical results to its application to LW however it does not benefit from an increased rate of convergence. It is therefore discarded as an appropriate method for the problem.

4.1.8 Roe's method

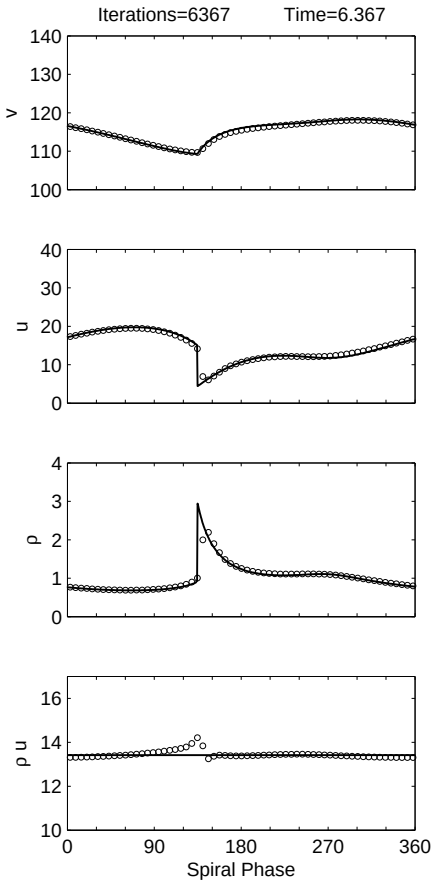

Figure 20: Results of Roe's method (R*) run until the RMSE of ρu is within 1%.

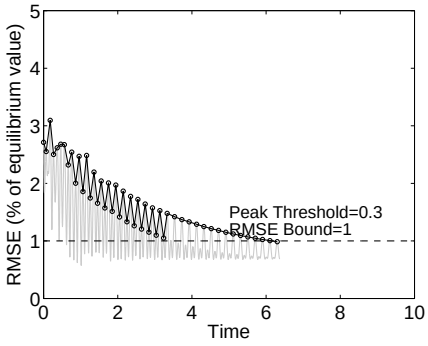

Figure 21: Roe's method (R*) time series plot of the RMSE of ρu.

Roe's method is by far the best performing first-order method. The density peak, although slightly displaced downstream, is sharper and much closer to the reference solution than any of the other first-order methods. There is a slight under-shoot of the peak density so the shock rendition is not of the same quality of FS2 but is much better than MC2.

Although the shock produced is sharper the representation of the smooth region is quite poor. It is for this reason that R* does not perform well in terms of RMSE even with the addition flux limiters.

One thing to note is the distribution of ρu still has some visible oscillations albeit very small. This implies that although the RMSE (as a percentage of the equilibrium value) is within 1 percent it is yet to reach steady state.

4.1.9 Roe's Method
Minmod flux limiter

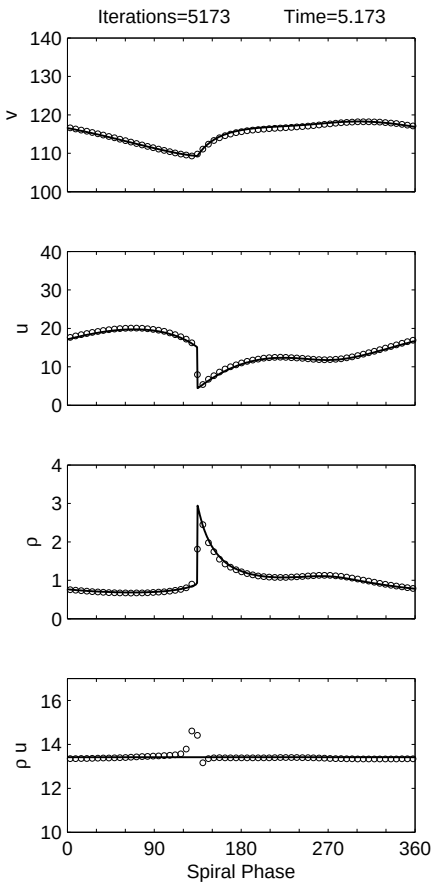

Figure 22: Results of Roe's method (R*) with minmod flux limiter applied (RFLM*) run until the RMSE of ρu is within 1%.

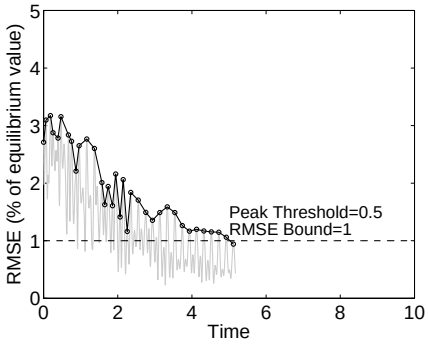

Figure 23: Roe's method (R*) with minmod flux limiter applied (RFLM*) time series plot of the RMSE of ρu.

Roe's method with the minmod flux limiter is one of the best performing second-order methods and the best performing flux limiter for this problem. The shock is sharp and narrow, and both the decompression and subsonic regions are well represented.

Although the smooth zones features are more prominent they are still not of the same quality as MC2 and as previously this hinders the overall result of the method.

4.1.10 Roe's Method
Superbee flux limiter

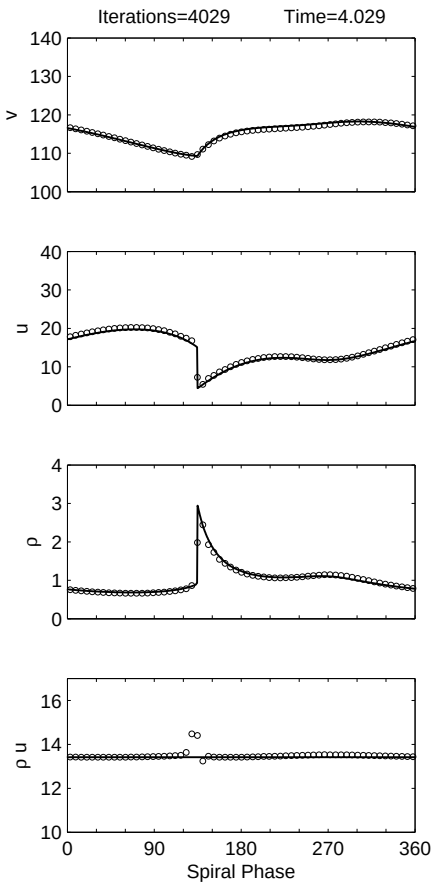

Figure 24: Results of Roe's method (R*) with superbee flux limiter applied (RFLS*) run until the RMSE of ρu is within 1%.

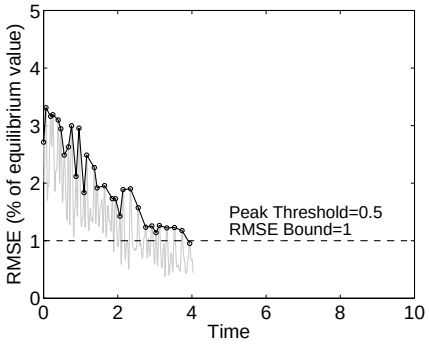

Figure 25: Roe's method (R*) with superbee flux limiter applied (RFLS*) time series plot of the RMSE of ρu.

The addition of the superbee flux limiter to Roe's method produces very similar results to the minmod flux limiter but cannot compete in smooth zone representation. The limiter over-shoots the smooth zone just before the shock.

Since each of the limiters corresponds to a boundary of the TVD region of the scheme it is clear from these results that the accuracy of the scheme depends on which boundary is chosen. Since the superbee limiter is at the top most boundary the amount it allows for variation is the largest out of the limiters. This results in the largest over-shoot in the smooth region. This may also give it the best approximation to the maximum density value but this is not visible at this grid resolution.

4.1.11 Roe's Method
Van Leer flux limiter

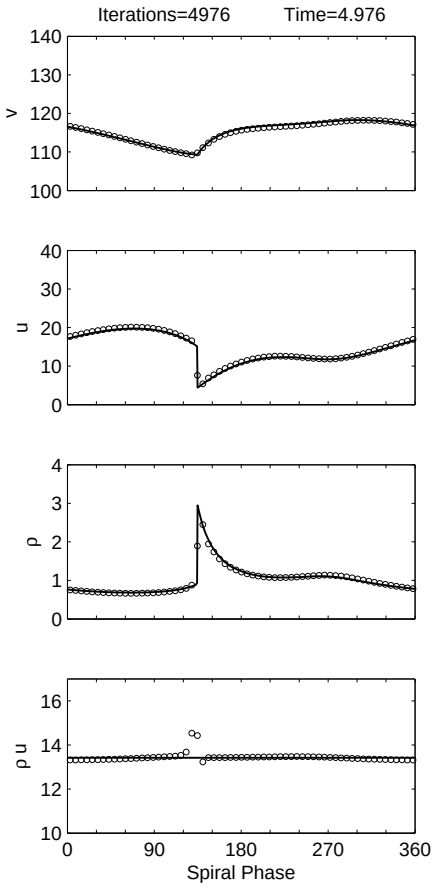

Figure 26: Results of Roe's method (R*) with van Leer limiter applied (RFLV*) run until the RMSE of ρu is within 1%.

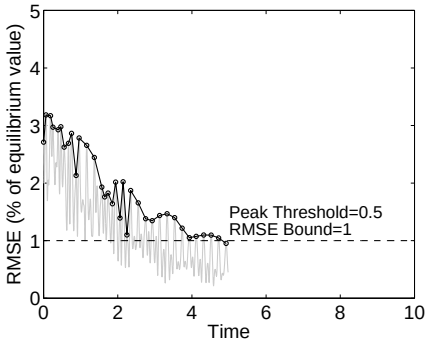

Figure 27: Roe's method (R*) with van Leer flux limiter applied (RFLV*) time series plot of the RMSE of ρu.

Similar to the addition of the superbee flux limiter the addition of the van Leer flux limiter to Roe's method produces accurate results but overshoots the smooth zone just before the shock.

This further confirms our observations regarding the correlation between the TVD region and Roe's representation of the smooth zone. The van Leer flux limiter is approximately central in the TVD region and the results shown here are central in their approximation to the smooth zone compared to the minmod flux limiter. Since this may not be the case for all problems the choice of limiter must be seen as a disadvantage as it is additional implementation that is needed just as the diffusion term coefficient must be chosen for MC2.

4.1.12 Roe's Method
Source Term Decomposed

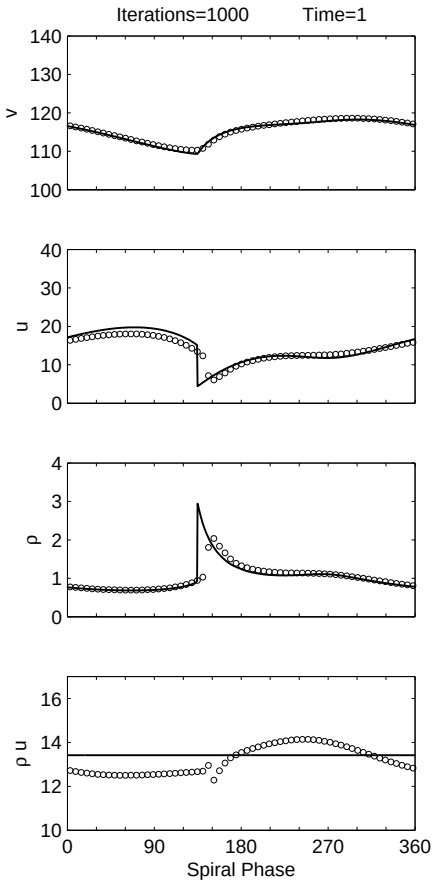

Figure 28: Results of Roe's method (R) with the source term decomposed (RSD) run until the RMSE of ρu is within 1%.

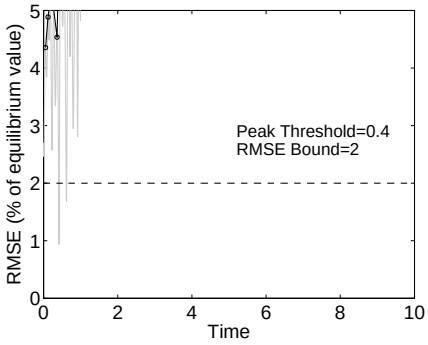

Figure 29: Roe's method (R) with the source term decomposed (RSD) time series plot of the RMSE of ρu.

With the source term decomposed, the shock propagates downstream very quickly (within 100 steps). Strong transient waves then begin to oscillate and as illustrated by figure 29 intensify until the solution is destroyed. It is our belief that this is due to the implementation of the code, which could not be completely debugged due to time constraints, not the method.

This is unfortunate because the result expected should have hopefully allowed the source terms to also cope with the shock which would have resulted in the distribution of ρu being completely uniform.

4.1.13 HLL Scheme Algorithm 1

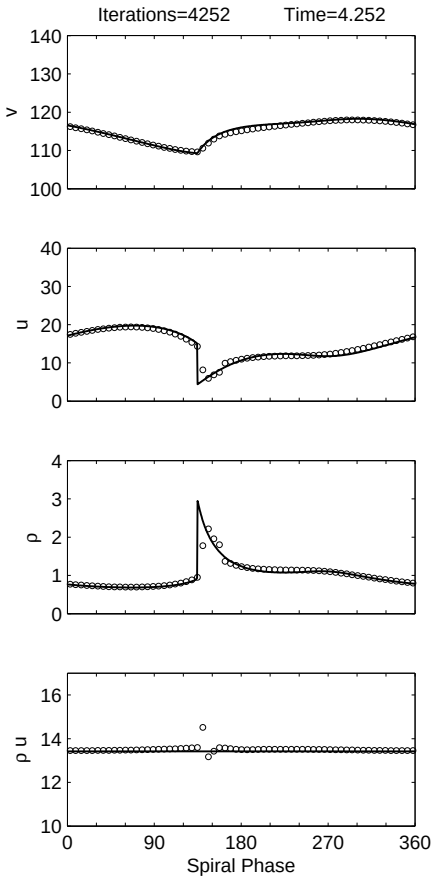

Figure 30: Results of the Harten, Lax & van Leer scheme (HLL) with algorithm 1 (HLL1) run until the RMSE of ρu is within 1%.

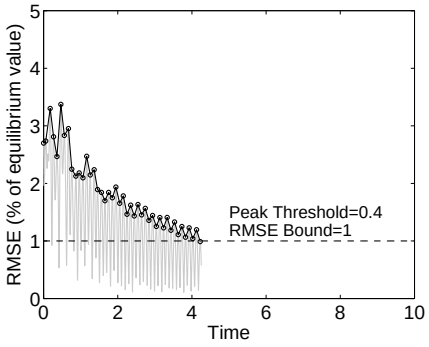

Figure 31: Harten, Lax & van Leer scheme (HLL) with algorithm 1 (HLL1) time series plot of the RMSE of ρu.

For the HLL scheme the shock rendition is narrow but also slightly displaced downstream. The scheme encounters difficulties when when it approximates the solution across the sonic point. The downstream subsonic region cannot numerically influence the upstream supersonic region by sound waves moving upstream, while the numerical diffusion across those waves vanishes in the sonic point (wave speeds are zero at sonic point). The solutions on either side therefore are not strongly coupled. This is also the case in van Albada *et al*'s Godunov method. It is typical for schemes to produce problematic solutions about sonic points.

The overall smooth region representation either side of the shock is poor. The peaks and troughs of the smooth features are under and over approximated respectively. As with B only the general smooth features are represented. The HLL scheme combined with using algorithm 1 to find the shock speed is therefore unsuited for this type of problem.

Although the schemes performance is poor in general the uniformity of the distribution of the ρu values is very good. It is almost on par with the results produced by FS2. This suggests that a methods overall performance is not correlated to the instability of the distribution but rather the accuracy with which the shock is rendered.

4.1.14 HLL Scheme Algorithm 2

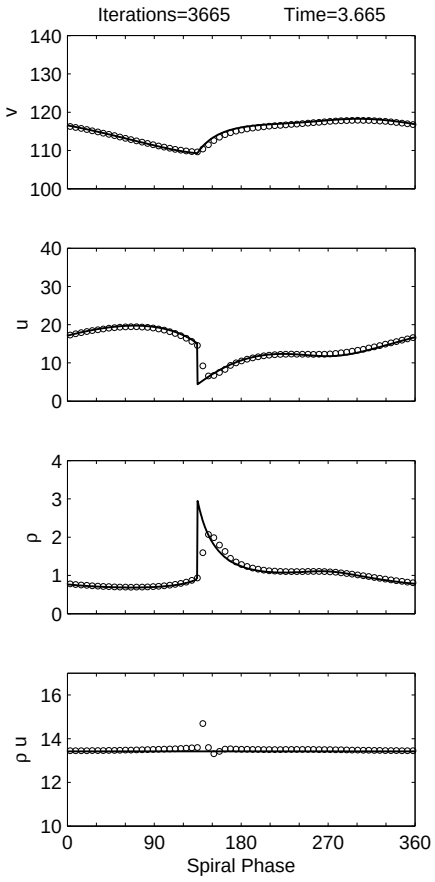

Figure 32: Results of the Harten, Lax & van Leer scheme (HLL) with algorithm 2 (HLL2) run until the RMSE of ρu is within 1%.

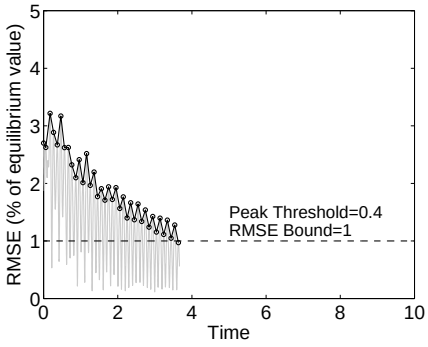

Figure 33: Harten, Lax & van Leer scheme (HLL) with algorithm 2 (HLL2) time series plot of the RMSE of ρu.

The use of algorithm 2 to calculate shock speeds resolves the issue occurring at the sonic point. The shock is closer fitting as a result however, the peak density is underestimated and also slightly displaced as with algorithm 1.

The solution, similar to algorithm 1, provides a poor representation of the smooth region of the flow either side of the shock, which implies that it is the HLL method that cannot cope with modelling the smooth flow as a pose to the type of algorithm used for wave speed calculation.

4.1.15 HLL Scheme Algorithm 3

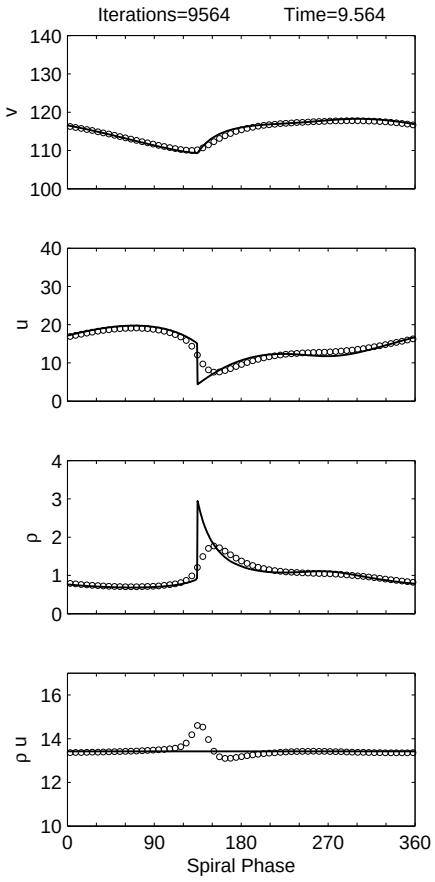

Figure 34: Results of the Harten, Lax & van Leer scheme (HLL) with algorithm 3 (HLL3) run until the RMSE of ρu is within 1%.

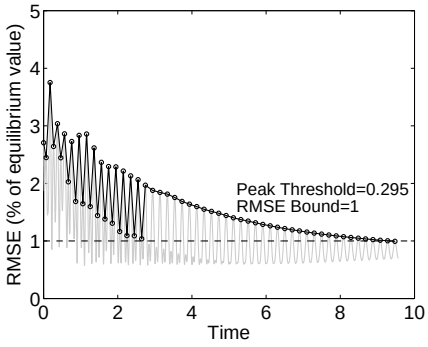

Figure 35: Harten, Lax & van Leer scheme (HLL) with algorithm 3 (HLL3) time series plot of the RMSE of ρu.

The results produced by the HLL scheme using algorithm 3 to calculate the wave speeds are very similar to those of method B. The use of Roe's averaged variables to calculate wave speeds results in strong numerical diffusion which causes the density maximum to be under estimated and also displaced downstream several zones.

4.1.16 Numerical Results

Scheme	RMSE (% of equilibrium value) ρ	u	v	$\sum$	ρu	n	t_n	e	σ
B	5.03	5.71	0.23	10.97	0.61	2,397	2.40	39	2%
	5.19	5.56	0.27	11.02	0.49	6,765	6.77	133	1%
FS2	1.10	2.68	0.19	3.97	1.24	2,283	2.28	78	2%
	1.14	2.16	0.24	3.54	0.78	7,236	7.24	213	1%
MC2	1.35	1.84	0.15	3.34	1.23	1,791	1.79	15	2%
	1.48	2.27	0.17	3.92	0.23	4,987	4.99	50	1%
FCTLWS	4.14	4.14	0.18	8.46	1.22	1,600	1.60	14	2%
	3.93	4.41	0.30	8.64	0.49	5,257	5.26	56	1%
FCTMC2S	4.00	3.93	0.24	8.17	1.26	1,776	1.78	15	2%
	3.90	4.37	0.32	8.59	0.48	4,468	4.47	43	1%
FCTLWZ*	2.88	2.10	0.36	5.34	1.29	2,088	2.09	39	2%
	3.03	1.79	0.35	5.17	0.59	3,282	3.28	74	1%
FCTMC2Z*	3.29	1.88	0.38	5.55	1.13	2,095	2.10	64	2%
	2.79	1.24	0.29	4.32	0.43	4,368	4.37	105	1%
R*	1.64	2.89	0.31	4.84	1.35	1,087	1.09	40	2%
	1.68	2.87	0.19	4.74	0.69	6,367	6.37	261	1%
RFLM*	2.12	3.78	0.30	6.20	1.10	1,679	1.68	65	2%
	2.02	2.74	0.17	4.93	0.43	5,173	5.17	210	1%
RFLS*	3.02	2.64	0.34	6.99	1.33	1,114	1.11	43	2%
	2.76	3.59	0.25	6.60	0.44	4,029	4.03	167	1%
RFLV*	2.16	2.57	0.13	4.86	1.42	1,597	1.60	62	2%
	2.23	2.92	0.19	5.34	0.45	4,976	4.98	208	1%
HLL1*	2.96	4.12	0.33	7.41	1.48	1,478	1.48	27	2%
	3.19	4.13	0.32	7.64	0.57	4,252	4.25	84	1%
HLL2*	1.75	3.12	0.30	5.17	1.54	1,281	1.28	23	2%
	1.77	2.91	0.28	4.96	0.56	3,665	3.67	73	1%
HLL3*	4.63	5.23	0.22	10.08	1.06	2,787	2.79	52	2%
	4.75	5.08	0.27	10.10	0.70	9,564	9.56	217	1%

Table 1: Test 1. RMSEs in the smooth region of the numerical solution, obtained with 14 schemes after the RMSEs are within 2 and 1 percent of the equilibrium values with a Courant number of 0.5, using a high resolution MC2 solution as the initial-value distribution. The results correspond to the figures 5-35 after n steps at $t_n = n\Delta t$, where $\Delta t = 10^{-3}$ and e is the elapsed time as described in the previous section.

4.2 Test Two

Scheme	RMSE (% of equilibrium value) ρ	u	v	$\sum$	ρu	n	t_n	e	σ
FS2	2.10	3.68	0.19	5.97	1.36	4,102	2.28	78	2%
	2.14	1.66	0.24	4.04	0.63	6,922	7.24	213	1%
MC2	1.48	1.79	0.20	3.97	1.51	4,066	4.07	44	2%
	1.60	2.24	0.19	4.03	0.52	6,708	6.71	94	1%
R*	1.95	5.00	0.31	7.26	1.67	1,281	1.28	53	2%
	1.65	2.79	0.27	4.08	0.69	3,006	3.01	131	1%
RFLM*	2.23	2.65	0.28	5.16	1.64	3,416	3.42	120	2%
	2.14	2.31	0.21	4.66	0.61	5,468	5.47	201	1%
RFLV*	2.27	3.13	0.24	5.64	1.56	4,162	4.16	148	2%
	2.22	2.85	0.25	5.32	0.70	5,571	5.57	204	1%

Table 2: Test 2. RMSEs in the smooth region of the numerical solution, obtained with 14 schemes after the RMSEs are within 2 and 1 percent of the equilibrium values with a Courant number of 0.5, using a uniform initial-value distribution. The results correspond to the figures 5-34 after n steps at $t_n = n\Delta t$, where $\Delta t = 10^{-3}$ and e is the elapsed time as described in the previous section.

Scheme	RMSE (% of equilibrium value) ρ	u	v	$\sum$	ρu	n	t_n	σ
FS2	2.10	3.68	0.19	5.97	1.36	4,102	2.28	2%
	2.14	1.66	0.24	4.04	0.63	6,922	7.24	1%
FS2 [23]	2.40	4.80	0.20	7.40	7.00	2,400	2.39	n/a [1]
MC2	1.48	1.79	0.20	3.97	1.51	4,066	4.07	2%
	1.60	2.24	0.19	4.03	0.52	6,708	6.71	1%
MC2 [23]	3.10	3.90	0.19	7.19	7.00	2,400	2.39	n/a

Table 3: Comparison of the results taken from [23]'s test 2 and the results in table 1 for FS2 and MC2.

[1] n/a is an abbreviation for not available

5 Conclusion

Our first objective was to recreate the results produced in van Albada *et al*. and confirm their conclusions in order for us to make our own comparisons and conclusions with the updated list of methods. In order decide whether or not our conclusions agree it is appropriate to first define those of van Albada *et al*. based on their first test.

1. The best second-order method (FS2) outperforms the best first-order method (G) by a huge margin, on a grid with a resolution comparable to what is feasible in two-dimensional calculations.

2. The success of the second-order method FS2 derives from the simple and effective averaging procedure that replaces central differencing.

3. The second-order central differencing scheme (MC2), while accurate in rendering the smooth parts of a solution, cannot compete with the second-order upwind-differencing method in shock representation.

In response to this we can say the following three statements in confidence based on the results of our first test

1. The best second-order method (FS2, since we are taking into account the shock rendition also) no longer outperforms the best first-order method (R*) by a huge margin, on a grid with a resolution comparable to what is feasible in two-dimensional calculations. The difference is now within 1 percent compared to the previous 7 percent.

2. The success of the second-order method still derives from the simple and effective averaging procedure that replaces central differencing.

3. The second-order central differencing scheme (MC2), while sufficiently accurate in rendering the smooth parts of a solution, cannot compete with the second-order upwind-differencing method in shock representation still holds.

Our second test reinforces all three of the above conclusions and also shows that the stronger transients that result from an initial distribution do not dampen the solutions the methods produced in the first test. While firm conclusions about the accuracy of these schemes applied to transient phenomena should not be given, since the solutions obtained cannot be calibrated with a time-dependent exact solution, we can say in confidence that using the RMSEs to test for a steady-state solution results in a more reliable comparison than using a fixed time run comparison. Van Albada *et al*. states the RMSEs "do not do not differ much from scheme to scheme" after 2,400 time-steps. Our results suggest that even without the use of the exact solution as an initial-value distribution (test 2), which eliminates the error in using a high resolution approximation to the exact, a marked difference is visible between the accuracy of shorter and longer runs. A reason for this may be the precision and accuracy that the floating point calculations were made with are much lower than those used in our testing. One final interesting observation is that as mentioned in our test setup, the

better scheme would would fair worse since the numerical viscosity would be is most for less accurate schemes. The underlying trend of the results confirmed.

We have two major concerns about our model and testing procedure. The first being the use of an approximation as reference solution. An exact solution is essential. Although the results we obtained are informative and representative of the methods there is clear bias towards the scheme that was used for the reference solution (MC2). This therefore also implies that there is a bias against the remaining methods. Since this bias is the same for all of the methods the results presented are still useful however it is necessary to take this factor into account when assessing the values of the methods that use MC2 in tables 1 and 2. Our second concern is that in order to determine the oscillations peak values our peak finding algorithm would have to read through a time series of values and return a peak after a defined threshold. This procedure lead to a value being returned beyond the maximum. This compromises the reliability of the comparisons and should be taken into account in the future.

We attempted the implementation of some implicit schemes but again due to time constraints could not include them in our study (see [13]). Some suggestions for future direction are the implementation of the methods presented implicitly and also an investigation into some more modern methods such as the MPDATA scheme. For the FCT methods used here and in [23] the evaluation of the anti-diffused term is done by taking the difference between a high order method (MC2 or LW) and a low order method (MC2 and LW with a diffusing term). An alternative approach to this would be to use R* for the lower-order method or some first-order alternative and MC2, LW, FS2 *etc*. as the higher order method or some alternative. In order to summarise our results it seems appropriate to ask the same question we started with "What reliable, accurate, efficient and easy-to-program method should be used for this calculation?"

On the basis of our present test results we can make three suggestions

- ◇ For flow problems involving shocks of importance we recommend the second-order upwind-differencing method FS2.
- ◇ For problems involving shocks where the smoother regions are of more importance we recommend the MacCormack method MC2.
- ◇ For efficiency and a balance between smooth region and shock rendition we recommend Roe's method (R*). Note to improve the accuracy of both smooth regions and shocks the addition of the flux limiter to Roe's method via the flux form requires very little additional programming effort. The choice of limiter can be related to the gradient of the pre-shock and decompression zones. For higher gradients upper TVD bound limiters (superbee) and lower gradients either no limiter or lower bound limiters (minmod).

6 Appendix

6.1 Conservation Form

If we sum the conservative form over our computational domain it is easily seen that the numerical flux differences 'telescope' (remembering the source term has already been advance by Eq. (3.44) and therefore it is a homogeneous case we are considering)

$$\begin{aligned}\sum_i^N \mathbf{u}_i^{n+1} &= \sum_i^N \mathbf{u}_i^n - \frac{\Delta t}{\Delta \eta}\sum_i^N (\mathbf{h}_{i+1/2}^v - \mathbf{h}_{i-1/2}^v) + \Delta t \sum_i^N \mathbf{s}_i^n \\ &= \sum_i^N \mathbf{u}_i^n - \frac{\Delta t}{\Delta \eta}\sum_i^N (... + (\mathbf{h}_{1/2}^v - \mathbf{h}_{-1/2}^v) + ...) + \Delta t \sum_i^N \mathbf{s}_i^n \\ &= \sum_i^N \mathbf{u}_i^n - \frac{\Delta t}{\Delta \eta}(\mathbf{h}_{k+1}^v - \mathbf{h}_k^v) + \Delta t \sum_i^N \mathbf{s}_i^n\end{aligned}$$

Rearranging gives,

$$\frac{\sum_i^N \mathbf{u}_i^{n+1}\Delta\eta - \sum_i^N \mathbf{u}_i^n \Delta\eta}{\Delta t} - \Delta\eta \sum_i^N \mathbf{s}_i^n = -(\mathbf{h}_{k+1}^v - \mathbf{h}_k^v) \tag{6.1}$$

and thus the numerical solution mimics the analytic conservation property

$$\frac{d}{dt}\int_{\eta_k}^{\eta_{k+1}} \mathbf{u}(\eta,t)d\eta - \int_{\eta_k}^{\eta_{k+1}} \mathbf{s}(\eta,t)d\eta = -(\mathbf{f}(\mathbf{u}_{k+1}) - \mathbf{f}(\mathbf{u}_k)) \tag{6.2}$$

6.2 Modification of Equation Eq. (2.5)

The modification of Eq. (2.5) in order to obtain $(\partial \mathbf{w}/\partial t)_i^n$ is not explicitly described in [23]. We felt it necessary to include for completeness.

We start with Eq. (2.5) in component form

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \eta}(\rho u) = \mathbf{s}_1, \tag{6.3}$$

$$\frac{\partial}{\partial t}(\rho u) + \frac{\partial}{\partial \eta}(\rho(u^2 + c^2)) = \mathbf{s}_2, \tag{6.4}$$

$$\frac{\partial}{\partial t}(\rho v) + \frac{\partial}{\partial \eta}(\rho u v) = \mathbf{s}_3. \tag{6.5}$$

We then manipulate Eqs. (6.3, 6.4 and 6.5) such that

$$\frac{\partial \rho}{\partial t} = \frac{\partial \mathbf{w}_1}{\partial t} = -\frac{\partial}{\partial \eta}(\rho u) + \mathbf{s}_1 = (-\rho)\frac{\partial u}{\partial \eta} + (-u)\frac{\partial \rho}{\partial \eta} + \mathbf{s}_1 \tag{6.6}$$

$$(6.4) - u \cdot (6.3) = \rho\frac{\partial u}{\partial t} + c^2\frac{\partial \rho}{\partial \eta} + \frac{\partial}{\partial \eta}(\rho u^2) - u\frac{\partial}{\partial \eta}(\rho u) + \mathbf{s}_2 \tag{6.7}$$

$$= \rho\frac{\partial u}{\partial t} + c^2\frac{\partial \rho}{\partial \eta} + \frac{\partial}{\partial \eta}(\rho u^2) - \left[\frac{\partial}{\partial \eta}(\rho u^2) - (\rho u)\frac{\partial u}{\partial \eta}\right] + \mathbf{s}_2 \tag{6.8}$$

$$= \rho\frac{\partial u}{\partial t} + c^2\frac{\partial \rho}{\partial \eta} + (\rho u)\frac{\partial u}{\partial \eta} + \mathbf{s}_2 \tag{6.9}$$

Rearranging Eq. (6.9)

$$\frac{\partial u}{\partial t} = \frac{\partial \mathbf{w}_2}{\partial t} = -\frac{c^2}{\rho}\frac{\partial \rho}{\partial \eta} - u\frac{\partial u}{\partial \eta} + \mathbf{s}_2 = -c^2\frac{\partial}{\partial \eta}(\ln \rho) - u\frac{\partial u}{\partial \eta} + \mathbf{s}_2 \tag{6.10}$$

Finally

$$(6.5) - v \cdot (6.3) = \rho\frac{\partial v}{\partial t} + \frac{\partial}{\partial \eta}(\rho u v) - v\frac{\partial}{\partial \eta}(\rho u) + \mathbf{s}_3 \tag{6.11}$$

$$= \rho\frac{\partial v}{\partial t} + (u\rho)\frac{\partial}{\partial \eta}v + \mathbf{s}_3 \tag{6.12}$$

Rearranging Eq. (6.12)

$$\frac{\partial v}{\partial t} = \frac{\partial \mathbf{w}_3}{\partial t} = -u\frac{\partial v}{\partial \eta} + \mathbf{s}_3 \tag{6.13}$$

6.3 Lax-Wendroff Scheme

To apply the Lax-Wendroff scheme to a non-linear system it is first written in two-step predictor corrector form (remembering source terms have already been advanced by Eq. (3.44)):

First a Lax-Friedrich predictor

$$\bar{\mathbf{u}}_{i+1/2}^{n+1/2} = \frac{1}{2}\left[(\mathbf{u}_{i+1}^n + \mathbf{u}_i^n) - \frac{\Delta t}{\Delta \eta}(\mathbf{f}_{i+1}^n - \mathbf{f}_i^n)\right], \tag{6.14}$$

followed by a second stage Leapfrog corrector

$$\mathbf{u}_i^{n+1} = \mathbf{u}_i^n - \frac{\Delta t}{\Delta \eta}(\bar{\mathbf{f}}_{i+1/2}^{n+1/2} - \mathbf{f}_{i-1/2}^{n+1/2}), \tag{6.15}$$

6.4 Zalesak Flux-Corrected Transport Algorithm

We seek to limit the anti-diffusive flux $\mathbf{h}_{i+1/2}^A$ such that

$$\mathbf{h}_{i+1/2}^C = C_{i+1/2}\mathbf{h}_{i+1/2}^A, \quad 0 \le C_{i+1/2} \le 1 \tag{6.16}$$

and such that $\mathbf{h}_{i+1/2}^C$ acting in concert with $\mathbf{h}_{i-1/2}^C$ will not allow

$$\mathbf{u}_i^{n+1} = \mathbf{u}_i^{td} - \frac{1}{8}(\mathbf{h}_{i+1/2}^C - \mathbf{h}_{i-1/2}^C) \tag{6.17}$$

to exceed some maximum value $\mathbf{u}_i^{\max}$ nor fall below some minimum value $\mathbf{u}_i^{\min}$.

We define three quantities:

$$P_i^+ = \text{the sum of all anti-diffusive fluxes } \textit{into} \text{ grid point } i \quad (6.18)$$

$$= \max(0, \mathbf{h}_{i-1/2}^A) - \min(0, \mathbf{h}_{i+1/2}^A) \quad (6.19)$$

$$Q_i^+ = \frac{1}{8}(\mathbf{u}_i^{\max} - \mathbf{u}_i^{td}) \quad (6.20)$$

$$R_i^+ = \begin{Bmatrix} \min(1, Q_i^+/P_i^+) & \text{if } P_i^+ > 0 \\ 0 & \text{if } P_i^+ = 0 \end{Bmatrix}. \quad (6.21)$$

Similarly we define three corresponding quantities:

$$P_i^- = \text{the sum of anti-diffusive fluxes } \textit{away from} \text{ grid point } i \quad (6.22)$$

$$= \max(0, \mathbf{h}_{i+1/2}^A) - \min(0, \mathbf{h}_{i-1/2}^A) \quad (6.23)$$

$$Q_i^- = \frac{1}{8}(\mathbf{u}_i^{td} - \mathbf{u}_i^{\min}) \quad (6.24)$$

$$R_i^- = \begin{Bmatrix} \min(1, Q_i^-/P_i^-) & \text{if } P_i^- > 0 \\ 0 & \text{if } P_i^- = 0 \end{Bmatrix}. \quad (6.25)$$

We come now to determine the quantities $\mathbf{u}_i^{\max}$ and $\mathbf{u}_i^{\min}$. We choose a sfae choice although other alternatives are available in [27]

$$\mathbf{u}_i^{\max} = \max(\mathbf{u}_{i-1}^{td}, \mathbf{u}_i^{td}, \mathbf{u}_{i+1}^{td}) \quad (6.26)$$

$$\mathbf{u}_i^{\min} = \min(\mathbf{u}_{i-1}^{td}, \mathbf{u}_i^{td}, \mathbf{u}_{i+1}^{td}) \quad (6.27)$$

6.5 Roe Decomposition

Remembering that we are looking to find the state averaged variables for the numerical flux vector as defined by Eq. (3.28)

$$\mathbf{h}_{i+1/2} = \frac{1}{2}(\mathbf{f}_{i+1} + \mathbf{f}_i) - \frac{1}{2}\sum_{k=1}^{3} \tilde{\alpha}_{i+1/2}^{(k)} |\tilde{\lambda}_{i+1/2}^{(k)}| \tilde{\mathbf{x}}_{i+1/2}^{(k)},$$

where $\tilde{\alpha}^{(k)}$, $\tilde{\lambda}^{(k)}$ and $\tilde{\mathbf{x}}^{(k)}$ are the coefficient for Δu, the eigenvalue and eigenvector, respectively, corresponding to the k^{th} characteristic field of $\tilde{A}$.

6.5.1 Eigenvalues

In order to calculate the flux vector as defined by Eq. (3.28) we need to find the eigenvalues of our system.

Let

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \end{pmatrix} = \begin{pmatrix} \rho \\ m \\ n \end{pmatrix}, \tag{6.28}$$

and

$$\mathbf{f} = \begin{pmatrix} \rho u \\ \rho(u^2 + c^2) \\ \rho uv \end{pmatrix} = \frac{1}{\rho} \begin{pmatrix} m\rho \\ m + \rho^2 c^2 \\ mn \end{pmatrix}, \tag{6.29}$$

where $\partial\mathbf{f}/\partial\mathbf{u}$ is the Jacobian

$$J_{\mathbf{f}}(\rho, m, n) = \begin{pmatrix} \frac{\partial x_1}{\partial \rho} & \frac{\partial x_1}{\partial m} & \frac{\partial x_1}{\partial n} \\ \frac{\partial x_2}{\partial \rho} & \frac{\partial x_2}{\partial m} & \frac{\partial x_2}{\partial n} \\ \frac{\partial x_3}{\partial \rho} & \frac{\partial x_3}{\partial m} & \frac{\partial x_3}{\partial n} \end{pmatrix} \tag{6.30}$$

now

$$A(\mathbf{u}) = \frac{\partial \mathbf{f}}{\partial \mathbf{u}} = \frac{1}{\rho^2} \begin{pmatrix} 0 & 1 & 0 \\ -m^2 + \rho^2 c^2 & 2m\rho & 0 \\ -mn & n\rho & m\rho \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ c^2 - u^2 & 2u & 0 \\ -uv & v & u \end{pmatrix} \tag{6.31}$$

The eigenvalues of A are then given by

$$\det(A - \lambda I) = 0 \tag{6.32}$$

where λ are the eigenvalues and I is the identity matrix

$$\begin{vmatrix} -\lambda & 1 & 0 \\ c^2 - u^2 & 2u - \lambda & 0 \\ -uv & v & u - \lambda \end{vmatrix} = (\lambda - u)(\lambda - (u + c))(\lambda - (u - c)) = 0 \tag{6.33}$$

The eigenvalues are: $\lambda^{(1)} = u$, $\lambda^{(2)} = u + c$ and $\lambda^{(3)} = u - c$

6.5.2 Eigenvectors

Formally, if A is a linear transformation, a non-null vector $\mathbf{x}$ is an eigenvector of A if there is a scalar λ such that

$$A\mathbf{x}^{(\mathbf{k})} = \lambda^{(\mathbf{k})}\mathbf{x}^{(\mathbf{k})} \tag{6.34}$$

where $\mathbf{x} = (x, y, z)^{\mathrm{T}}$. For $\lambda^{(1)} = u$

$$\begin{pmatrix} 0 & 1 & 0 \\ c^2 - u^2 & 2u & 0 \\ -uv & v & u \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = u \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{6.35}$$

after some manipulation $\mathbf{x}^{(1)} = (0, 0, 1)^{\mathrm{T}}$. For $\lambda^{(2)} = u + c$

$$\begin{pmatrix} 0 & 1 & 0 \\ c^2 - u^2 & 2u & 0 \\ -uv & v & u \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = u + c \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{6.36}$$

after some manipulation $\mathbf{x}^{(2)} = (1, u + c, v)^{\mathrm{T}}$. For $\lambda^{(3)} = u - c$

$$\begin{pmatrix} 0 & 1 & 0 \\ c^2 - u^2 & 2u & 0 \\ -uv & v & u \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = u - c \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{6.37}$$

after some manipulation $\mathbf{x}^{(3)} = (1, u - c, v)^{\mathrm{T}}$

We define X to be a matrix with these eigenvectors as columns

$$X = \begin{pmatrix} 0 & 1 & 1 \\ 0 & u + c & u - c \\ 1 & v & v \end{pmatrix} \tag{6.38}$$

Note we can now easily diagonalise A with $X^{-1}AX = \Lambda$ or $A = X\Lambda X^{-1}$ where Λ is a diagonal matrix with $\lambda^{(k)}$ on the diagonal.

6.5.3 The Parameter Vector

Suppose that the components of the vectors **u** and **f** may be written as quadratic functions of the components of another vector say Z. We shall call this the *parameter vector*

$$\mathbf{z} = \sqrt{\rho} \begin{pmatrix} 1 \\ u \\ v \end{pmatrix}, \tag{6.39}$$

whence

$$\mathbf{u} = \begin{pmatrix} (z_1)^2 \\ z_1 z_2 \\ z_1 z_3 \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} z_1 z_2 \\ (z_2)^2 + c^2 (z_1)^2 \\ z_2 z_3 \end{pmatrix}, \tag{6.40}$$

Using the identity

$$\Delta(pq) = \bar{p}\Delta q + \bar{q}\Delta p \tag{6.41}$$

we may write

$$\Delta\mathbf{u} = \tilde{B}\Delta\mathbf{z}, \quad \Delta\mathbf{f} = \tilde{C}\Delta\mathbf{z} \tag{6.42}$$

where $\Delta(\cdot) = \cdot_R - \cdot_L$ and $\tilde{B}$ and $\tilde{C}$ are constant matrices.

For our isothermal system then

$$\Delta\mathbf{u} = \begin{pmatrix} \Delta(z_1)^2 \\ \Delta z_1 z_2 \\ \Delta z_1 z_3 \end{pmatrix} = \begin{pmatrix} 2\bar{z}_1 & 0 & 0 \\ \bar{z}_2 & \bar{z}_1 & 0 \\ \bar{z}_3 & 0 & \bar{z}_1 \end{pmatrix} \cdot \begin{pmatrix} \Delta z_1 \\ \Delta z_2 \\ \Delta z_3 \end{pmatrix} \tag{6.43}$$

and

$$\Delta\mathbf{f} = \begin{pmatrix} \Delta(z_1 z_2) \\ \Delta((z_2)^2 + c^2(z_1)^2) \\ \Delta(z_2 z_3) \end{pmatrix} = \begin{pmatrix} \bar{z}_2 & \bar{z}_1 & 0 \\ 2c^2\bar{z}_1 & 2\bar{z}_2 & 0 \\ 0 & \bar{z}_3 & \bar{z}_2 \end{pmatrix} \cdot \begin{pmatrix} \Delta z_1 \\ \Delta z_2 \\ \Delta z_3 \end{pmatrix} \tag{6.44}$$

i.e.

$$\tilde{B} = \begin{pmatrix} 2\bar{z}_1 & 0 & 0 \\ \bar{z}_2 & \bar{z}_1 & 0 \\ \bar{z}_3 & 0 & \bar{z}_1 \end{pmatrix}, \quad \tilde{C} = \begin{pmatrix} \bar{z}_2 & \bar{z}_1 & 0 \\ 2c^2\bar{z}_1 & 2\bar{z}_2 & 0 \\ 0 & \bar{z}_3 & \bar{z}_2 \end{pmatrix}, \tag{6.45}$$

Thus we have

$$\Delta\mathbf{f} = \tilde{C}\Delta\mathbf{z} = \tilde{C}(\tilde{B}^{-1}\Delta\mathbf{u}) \tag{6.46}$$

and when compared with

$$\Delta\mathbf{f} = \tilde{A}(u_L, u_R)\Delta\mathbf{u} \tag{6.47}$$

gives a choice of

$$\tilde{A}(u_L, u_R) = \tilde{C}\tilde{B}^{-1} \tag{6.48}$$

where

$$\tilde{B}^{-1} = \frac{1}{2\bar{z}_1^2}\begin{pmatrix} \bar{z}_1 & 0 & 0 \\ -\bar{z}_2 & 2\bar{z}_1 & 0 \\ -\bar{z}_3 & 0 & 2\bar{z}_1 \end{pmatrix}, \tag{6.49}$$

and so

$$\tilde{A}(u_L, u_R) = \begin{pmatrix} 0 & 1 & 0 \\ (c^2 - \dfrac{\bar{z}_2^2}{\bar{z}_1^2}) & \dfrac{2\bar{z}_2}{\bar{z}_1} & 0 \\ 0 & \dfrac{\bar{z}_3}{\bar{z}_1} & \dfrac{\bar{z}_2}{\bar{z}_1} \end{pmatrix}, \tag{6.50}$$

remembering for our system

$$A(\mathbf{u}) = \begin{pmatrix} 0 & 1 & 0 \\ c^2 - u^2 & 2u & 0 \\ -uv & v & u \end{pmatrix}, \tag{6.51}$$

and comparing with $\tilde{A}$ gives us the Roe averaged variables

$$\tilde{u} = \frac{\bar{z}_2}{\bar{z}_1} = \frac{\sqrt{\rho_R}u_R + \sqrt{\rho_L}u_L}{\sqrt{\rho_R} + \sqrt{\rho_L}} \tag{6.52}$$

$$\tilde{v} = \frac{\bar{z}_2}{\bar{z}_1} = \frac{\sqrt{\rho_R}v_R + \sqrt{\rho_L}v_L}{\sqrt{\rho_R} + \sqrt{\rho_L}} \tag{6.53}$$

Note that for this example $\tilde{\rho}$ is arbitrary and indeed not required.

6.5.4 Coefficient Evaluation

The coefficients $\tilde{\alpha}^{(k)}$ are given by

$$\tilde{\alpha}^{(k)}\tilde{X} = \Delta\mathbf{u}. \tag{6.54}$$

We are therefore solving the following system

$$\begin{pmatrix} \alpha^{(1)} \\ \alpha^{(2)} \\ \alpha^{(3)} \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 0 & \tilde{u} + c & \tilde{u} - c \\ 1 & \tilde{v} & \tilde{v} \end{pmatrix} = \Delta \begin{pmatrix} \rho \\ \rho u \\ \rho v \end{pmatrix}, \tag{6.55}$$

by the simple manipulation of

$$\alpha^{(2)} + \alpha^{(3)} = \Delta\rho, \tag{6.56}$$

$$\alpha^{(2)}(\tilde{u} + c) + \alpha^{(3)}(\tilde{u} - c) = \Delta(\rho u), \tag{6.57}$$

$$\alpha^{(1)} + \tilde{v}\alpha^{(2)} + \tilde{v}\alpha^{(3)} = \Delta(\rho v), \tag{6.58}$$

to obtain the coefficients

$$(6.58) - \tilde{v} \cdot (6.56) \rightarrow \alpha^{(1)} = \Delta(\rho v) - \tilde{v}\Delta\rho, \tag{6.59}$$

$$(6.57) - (\tilde{u} - c) \cdot (6.56) \rightarrow \alpha^{(2)} = \frac{\Delta(\rho u) - (\tilde{u} - c)\Delta\rho}{2c}, \tag{6.60}$$

$$(6.57) - (\tilde{u} + c) \cdot (6.56) \rightarrow \alpha^{(3)} = \frac{-\Delta(\rho u) + (\tilde{u} + c)\Delta\rho}{2c}. \tag{6.61}$$

6.6 MacCormack Method Comparison Table

The table consists of the results for

- ◇ MacCormack Backward Predictor Forward Corrector (MC2BF)
- ◇ MacCormack Forward Predictor Backward Corrector (MC2FB)
- ◇ MacCormack Time Alternating (MC2TA)

	RMSE (% of equilibrium value)								
Scheme	ρ	u	v	$\sum$	ρu	n	t_n	e	σ
MC2BF	1.36	1.90	0.15	3.41	1.22	1,792	1.79	14	2%
	1.50	2.36	0.17	4.03	0.22	4,993	4.99	53.56	1%
MC2FB	1.35	1.84	0.15	3.35	1.23	1,791	1.79	15	2%
	1.48	2.27	0.17	3.92	0.23	4,987	4.99	50	1%
MC2TA	1.36	1.86	0.15	3.37	1.24	1,791	1.79	15	2%
	1.49	2.30	0.17	3.96	0.23	4,989	4.99	50	1%

Table 4: Comparison of MC2BF, MC2FB and MC2A's RMSEs in the smooth region of the solution within 2 and 1 percent of the equilibrium values with a Courant number of 0.5, using reference solution as the initial-value distribution. .

References

[1] Book, D.L., Boris, J.P., Hain, K.H.: 1975, *J. Computational Phys.* **18**, 248

[2] Boris, J.P., Book, D.L.: 1973, *J. Computational Phys.* **11**, 38

[3] Chung, T.J.: 2002, *Computational Fluid Dynamics*

[4] Godunov, S.K.: 1959, *Mat. Sb.* **47**, 271 (also see Cornell Aeronautical Lab. ransl.)

[5] Harten, A., Lax, R.D., Van Leer, B.: 1981, ICASE Report (in preparation) - unavailable

[6] Harten, A., Lax, R.D., Van Leer, B.: 1983, SIAM Review, 25:35-61

[7] Laney, B. Culbert.: 1998, *Computational Gas Dynamics*

[8] LeVeque, Randall, J.: 1992, *Numerical methods for conservation laws*

[9] Lax, P.D., Wendroff, B.: 1960, *Comm. Pure Appl. Math.* **13**, 217

[10] Liebovitch, L.S.: 1978, Ph. D. Thesis, Harvard University

[11] Leo, C.J., Pretty, R.L.: 1992, *Transpn. Res.-B* Vol. 26B, No. 3, pp. 207-220

[12] MacCormack, R.W.: 1969, AIAA Paper No. 69-354

[13] Mulder, W.A., van Leer, B.: 1982, 'Implicit upwind methods for for the Euler equations'

[14] Roberts, W.W.: 1969, *Astrophys. J.* **158**, 123

[15] Roe, P.L.: 1981, *J. Comutational Phys.* 43:357-372

[16] Sanders, R.H., Prendergast, K.H.: 1974, *Astrophys. J.* **188**, 489

[17] Sweby, P.K.: 1984, SIAM *J. on Numerical Analysis*, 21:995-1011

[18] Sweby, P.K.: *Lecture Notes in Conservation Laws*, Reading University, Reading, Berkshire

[19] Shu, F.H., Milione, V., Roberts, W.W.: 1973, *Astrophys. J.* **183**, 819

[20] Toro E.F.: 1999, (Editor) *Godunov Methods: Theory and Applications*

[21] Toro E.F.: 2006, (Editor) *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*

[22] Van Leer, B.: 1977, *J. Computational Phys.* **23**, 276

[23] Van Albada, G. D., van Leer, B., Roberts, Jr., W. W.: 1982, *Astron. Astrophys.*, **108**, pp. 76-84

[24] Van Leer, B.: 1981a, ICASE Report No. 81-11

[25] Van Leer, B.: 1981b, ICASE Report

[26] Woodward, P.R.: 1975, *Astrophys. J.* **195**, 61

[27] Zalesak, S.T.: 1981 (private communication to van Albada *et al*.)

[28] Zalesak, S.T.: 1979, *J. Computational Phys*. 31:335-362.