ABSTRACT

Title of Thesis:	Development of a New Class of High order Implicit Non-Oscillatory Schemes for Conservation Laws
Degree candidate:	Karthikeyan Duraisamy
Degree and year:	Master of Science, 2004
Thesis directed by:	Dr. Jian-Guo Liu Department of Applied Mathematics and Scientific Computation

Hyperbolic conservation laws allow for discontinuities to develop in the solution. In order to obtain non-oscillatory solutions near discontinuities and high gradient regions, numerical schemes have to satisfy conditions additional to linear stability requirements. These restrictions render high order implicit time integration schemes impractical since the allowable time-step sizes are not much higher than that for explicit schemes. In this work, an investigation is made on the factors that cause these time-step restrictions and two novel schemes are developed in an attempt to overcome the severity of these restrictions. In the first method, the order of accuracy of the *time integration* is lowered in high gradient and discontinuous regions. In the second method, the solution is reconstructed in *space and time* in a non-oscillatory manner. These concepts are evaluated on model scalar and vector hyperbolic conservation equations. The ultimate objective of this work is to develop a scheme that is accurate and unconditionally non-oscillatory.

Development of a New Class of High order Implicit Non-Oscillatory Schemes for Conservation Laws

by

Karthikeyan Duraisamy

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Science
2004

Advisory Committee:

Professor Jian-Guo Liu, Chairman/Advisor
Dr. James D. Baeder
Dr. Ramani Duraiswami

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

Hyperbolic conservation equations can be used to model a wide variety of physical phenomena. A characteristic feature of hyperbolic conservation equations is the fact that these equations allow *discontinuities* to develop in the solution. Though not admitted by the differential form of the equations, these discontinuities can be defined[1]by using the integral (or weak) form. Designing numerical schemes for the solution of hyperbolic equations is complicated. In "smooth" regions of the solution, the scheme should be high order accurate (for instance, in the sense of Taylor series expansion). However, across a discontinuity, the concept of *order of accuracy* breaks down. Hence, an accurate scheme should be able to represent the solution well in both the above mentioned cases. However, it is seen that high order accurate schemes suffer from severe time-step restrictions when discontinuities and high gradients appear in the solution. In this work, an investigation is made on the factors that cause these time-step restrictions and two novel schemes are developed in an attempt to overcome the severity of these restrictions.

1.1 Linear stability restrictions

Numerical schemes for the solution of time-dependant hyperbolic equations can be broadly classified into two categories:

• Explicit schemes, in which the numerics at any time instant depends on the solution at previous time-steps.

[1]The discontinuity jump is given by the Rankine-Hugoniot relations (refer [1])

• Implicit schemes, in which the numerics at any time instant could depend on the solution at previous, present or future time-steps.

Linear stability of a numerical scheme ensures that the error does not grow in an unbounded fashion. In explicit schemes, linear stability places a restriction on the allowable time step size. Hence these methods could become inefficient when the hyperbolic system is highly stiff or when the spatial mesh is largely non-uniform. In such cases, it is desirable to use implicit schemes for which the time-step is usually limited by accuracy and not linear stability. In specific cases, for instance, the Euler equations, implicit schemes will be beneficial for flows in which the dominant time scales are much larger than the acoustic time scales. Also, implicit schemes may be preferred in practical computations when the time-step size required to achieve the desired accuracy may be several times higher than the explicit limit.

1.2 Monotonicity constraints

Accurate numerical solution of hyperbolic equations is further complicated by the fact that linear stability of a numerical scheme does not guarantee well-behaved solutions near discontinuities. Linearly stable schemes that do not satisfy some form of the so-called *monotonicity conditions* give rise to spurious oscillations near discontinuities. Physical solutions to homogeneous hyperbolic conservation equations do not allow the appearance of new extrema and enforcement of the monotonicity conditions is an attempt to mimic that behavior in the numerical solution.

It is well known that solutions of *linear*[2] high-order numerical schemes for conservation laws are necessarily non-monotone near regions of discontinuities and high solution gradients. Research on high resolution finite difference and finite volume schemes has mainly concentrated on *controlling* the spatial interpolant. For example, Total Variation Diminishing (TVD) [5] schemes reduce the order of accuracy of spatial interpolation near discontinuities and extrema, the Essentially Non-Oscillatory (ENO) [13] type schemes use an adaptive stencil for spatial interpolation, etc. These approaches essentially make the spatial scheme *non-linear*, thus transcending the restrictions imposed on linear high-

[2]By linear, it is meant that the formulation of the scheme does not depend on the solution

order schemes by satisfying some form of the monotonicity condition.

However, ensuring the non-oscillatory behavior of high order schemes imposes a severe time step restriction. For explicit schemes, this restriction may not be much more severe than the linear stability limit. Shu et. al. [14] have shown that conventional implicit schemes become conditionally non-oscillatory when the order of accuracy in time is higher than one. It becomes evident that linear high order implicit schemes become impractical when the solution contains discontinuities since the allowable time steps are only slightly higher than that of explicit schemes.

In this work, akin to the spatial discretization, the time discretization is also made *non-linear* in an attempt to circumvent the time-step restrictions that govern the monotonicity of implicit schemes.

1.3 Outline of work

In Chapter 2, the concepts of monotonicity are reviewed. Enforcement of any form of the monotonicity conditions introduces time-step restrictions. These restrictions are analyzed and evaluated for different implicit schemes.

In Chapter 3, the generation of spurious oscillations is examined with the aid of some examples.

Chapter 4 introduces the concept of *time-limiting*, wherein a new class of *time-limited* schemes are introduced. The idea is to take a convex combination of a first order and a higher order scheme and modify the coefficients of the combination, such that the scheme works near the higher order limit in smooth regions of the solution and drops to lower order near regions of high gradients and discontinuities.

Chapter 5 introduces the concept of limited space-time reconstruction. In this case, the reconstructed value at the cell-interface is limited in space *and* time.

Chapter 6 summarizes the performance of the above-mentioned concepts and projects the possibilities of future work in this area.

Chapter 2

Background

2.1 Fully discrete schemes

Consider a scalar conservation law of the form,

$$\frac{\partial u(x,t)}{\partial t} + \frac{\partial f(u(x,t))}{\partial x} = 0 \tag{2.1}$$

with suitable initial and boundary conditions. To discretize this equation, space is divided into cells $[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ and time is divided into intervals $[t^n, t^{n+1}]$ as depicted in fig. 2.1. The integral form of eqn. 2.1 in the control volume $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [t^n, t^{n+1}]$ is given by:

$$\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} [u(x,t^{n+1}) - u(x,t^n)]dx = -\int_{t^n}^{t^{n+1}} [f(u(x_{j+\frac{1}{2}},t)) - f(u(x_{j-\frac{1}{2}},t))]dt$$

This can then be written in the *conservation form*:

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \tau(\hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}}), \tag{2.2}$$

where,

$$\bar{u}_j^n \;=\; \frac{1}{\Delta x}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u(x,t^n)dx, \quad \textit{and} \quad \hat{f}_{j+\frac{1}{2}} = \frac{1}{\Delta t}\int_{t^n}^{t^{n+1}} f(u(x_{j+\frac{1}{2}}),t)dt \tag{2.3}$$

with $\Delta x = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$, $\Delta t = t^{n+1} - t^n$ and $\tau = \frac{\Delta t}{\Delta x}$. By defining a suitable approximation to the flux integral (eqn. 2.3), numerical schemes can be constructed to any order of accuracy. A few well known approximations are shown

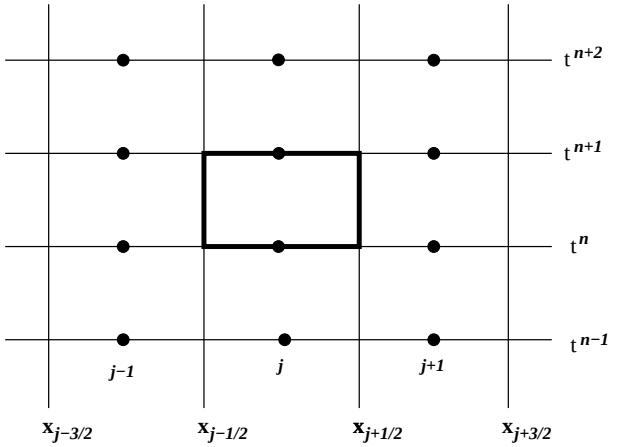


Figure 2.1: Schematic of discretization in time and space

Case	Approximation
Explicit Euler	$\hat{f}_{j+\frac{1}{2}} = f(u(x_{j+\frac{1}{2}}, t^n)) + O(\Delta t)$
Implicit Euler	$\hat{f}_{j+\frac{1}{2}} = f(u(x_{j+\frac{1}{2}}, t^{n+1})) + O(\Delta t)$
Trapezoidal	$\hat{f}_{j+\frac{1}{2}} = \frac{1}{2} f(u(x_{j+\frac{1}{2}}, t^n)) + \frac{1}{2} f(u(x_{j+\frac{1}{2}}, t^{n+1})) + O(\Delta t^2)$

Table 2.1: Commonly used flux definitions

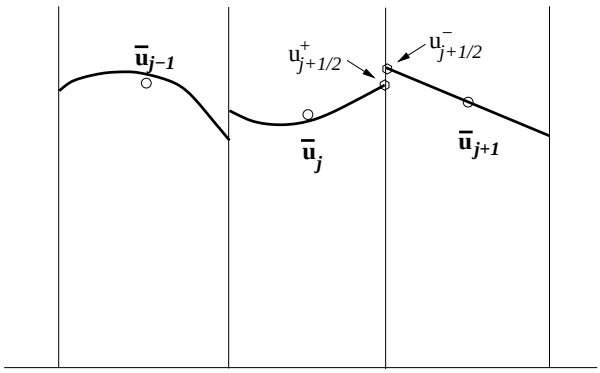


Figure 2.2: Spatial reconstruction

in table 2.1. In this work, Godunov type MUSCL methods [17] are considered. In this method, the flux $f(u(x_{j+\frac{1}{2}}))$ in table 2.1 at a cell interface is evaluated by:

1) Defining *reconstructed* values of the solution to the left (+) and right (−) of the interface (fig. 2.2). These values are usually computed using monotone interpolation of the cell-average values.

2) Using an appropriate upwind solver, treating the left and right states as defining a Riemann problem. For instance, the Roe upwind solver would yield $f(u(x_{j+\frac{1}{2}})) = \frac{1}{2}(f(u^-_{j+\frac{1}{2}}) + f(u^+_{j+\frac{1}{2}})) - \frac{1}{2}|a|(u^-_{j+\frac{1}{2}} - u^+_{j+\frac{1}{2}})$, where a is an approximation to the interface flux jacobian.

Note: There exist reconstruction-evolution schemes in the literature (for instance, the piecewise parabolic method [6] or the ENO methods of Harten et. al. [21]), where the reconstructed profile is convected in an exact manner. Such a procedure will not be addressed in this work.

2.2 Semi-discrete schemes and the method-of-lines

In the method-of-lines approach, temporal and spatial discretizations are decoupled from each other. This is done by replacing the spatial derivative in eqn. 2.1 by:

$$\frac{\partial f(u(x_j,t))}{\partial x} \approx \frac{f_{j+\frac{1}{2}}(t) - f_{j-\frac{1}{2}}(t)}{\Delta x} \tag{2.4}$$

Thus, the semi-discrete form of eqn. 2.1 is:

$$\frac{du_j(t)}{dt} = -\frac{f_{j+\frac{1}{2}}(t) - f_{j-\frac{1}{2}}(t)}{\Delta x} \tag{2.5}$$

The flux $f_{j+\frac{1}{2}}$ can be constructed to any given order of spatial accuracy using the reconstruction approach described in the previous section. The resulting set of ordinary differential equations (ODE) given by

$$u_t = L(u) \tag{2.6}$$

can be integrated in time using single or multi-step time integration methods. This way of decoupling space and time is called the method-of-lines.

2.3 Monotonicity conditions and TVD

A physically acceptable weak solution to eqn. 2.1 has the following monotonicity properties [5]:

(i) No new spatial extrema can be created,

(ii) The value of a local minimum is non-decreasing and the value of a local maximum is non increasing.

Let a numerical solution to 2.1 be given by:

$$u_j^{n+1} = H(u_{j-k}^n, u_{j-k+1}^n, ..., u_{j+k}^n) \tag{2.7}$$

Linear stability of this scheme is not sufficient to guarantee the absence of spurious oscillations near discontinuities. Then, the numerical scheme is said to be *monotone* [5] if

$$\frac{\partial H}{\partial u_i}(u_{j-k}^n, u_{j-k+1}^n, ..., u_{j+k}^n) \geq 0 \quad \forall \quad j-k \leq i \leq j+k \tag{2.8}$$

Note that for a linear scheme of the form:

$$u_j^{n+1} = \sum_k b_k u_{j+k}^n, \tag{2.9}$$

eqn. 2.8 requires all the coefficients $b_k \geq 0$. It has been shown by Harten [4] that converged solutions to monotone schemes always result in physically acceptable states. However, Harten goes on to show that conservative monotone schemes for the solution of eqn. 2.1 are only of first order of accuracy and consequently, a weaker condition, based on the *total variation* was first proposed. The total variation of a numerical solution at a time level t^n is defined by,

$$TV(u(., t^n)) = \sum_{i=-\infty}^{\infty} |u(x_{i+1}, t^n) - u(x_i, t^n)|$$

A numerical scheme for the solution of eqn. 2.1 is said to be Total Variation Diminishing (TVD) if the total variation does not increase in time. That is,

$$TV(u(., t^{n+1})) \leq TV(u(., t^n))$$

TVD schemes are based on the fact that oscillations always add to the total variation and thus oscillations cannot grow indefinitely without violating the TVD

condition[10]. Even though the TVD condition is weaker than the monotonicity condition, it has the advantage that it is simple to apply and is sufficient to guarantee the convergence of a conservative numerical scheme to weak solutions of conservation laws. Contrary to monotone numerical schemes, non-linear TVD schemes can be of order of accuracy higher than one. (But linear TVD schemes are still at most first order accurate). It has to be mentioned that theoretically TVD could still allow spurious oscillations (a hypothetical example can be found in fig. 16.3 of ref. [10]), but practically, with the use of conservative differencing, oscillations will be prevented. As mentioned in ref. [10], *"most attempts at enforcing TVD end up enforcing much stronger non-linear stability conditions and the very least, it establishes an upper bound on oscillations".*

2.4 TVD conditions for fully-discrete schemes

Combining Lemmas 3.1 and 3.2 in [7], if a numerical solution to eqn. 2.1 can be represented by:

$$u_j^{n+1} - \frac{\Delta t}{\Delta x}(B^-_{j+\frac{1}{2}}\Delta u_{j+\frac{1}{2}} + B^+_{j-\frac{1}{2}}\Delta u_{j-\frac{1}{2}})^{n+1} = u_j^n - \frac{\Delta t}{\Delta x}(C^-_{j+\frac{1}{2}}\Delta u_{j+\frac{1}{2}} + C^+_{j-\frac{1}{2}}\Delta u_{j-\frac{1}{2}})^n, \tag{2.10}$$

where $\Delta u_{j+\frac{1}{2}} = u_{j+1} - u_j$, then eqn. 2.10 is TVD if:

$$\begin{aligned} C^+_{j+\frac{1}{2}} &\geq 0 \\ C^-_{j+\frac{1}{2}} &\leq 0 \\ \frac{\Delta t}{\Delta x}(C^+_{j+\frac{1}{2}} - C^-_{j+\frac{1}{2}}) &\leq 1 \\ B^+_{j+\frac{1}{2}} &\leq 0 \\ B^-_{j+\frac{1}{2}} &\geq 0 \end{aligned}$$

2.4.1 Sample TVD schemes

In this section, a few well known spatially second order TVD schemes that use *limiting* are presented. Reference [10], for instance gives a comprehensive survey of such schemes. The linear advection equation corresponds to the flux $f = au$ in eqn. 2.1. In this particular case, $a > 0$ is assumed to be a constant and the numerical solution is given by eqn. 2.2. In the explicit case, the flux in eqn. 2.2

is given by:

$$\hat{f}_{j+\frac{1}{2}} = f(u_{j+\frac{1}{2}}, t^n) \ \ where \tag{2.11}$$

$$u_{j+\frac{1}{2}} = u_j + 0.5\Delta u_{j+\frac{1}{2}}\phi_j(r_j) \tag{2.12}$$

where, $\Delta u_{j+\frac{1}{2}} = u_{j+1} - u_j$, $r_j = \frac{\Delta u_{j-\frac{1}{2}}}{\Delta u_{j+\frac{1}{2}}}$ and $\phi_j(r_j)$ is a *spatial limiter*. Substituting these expressions into eqn. 2.2 and casting the resulting equation in the form of eqn. 2.10,

$$u_j^{n+1} = u_j^n - \nu\left[1 + \frac{1}{2}(\frac{\phi_j}{r_j} - \phi_{j-1})\right]\Delta u_{j-1/2} \tag{2.13}$$

with $\nu = a\frac{\Delta t}{\Delta x}$. Satisfaction of the TVD condition requires:

$$\frac{\phi_j}{r_j} - \phi_{j-1} \geq -2,$$

$$\nu \leq \frac{1}{1 + \frac{1}{2}(\frac{\phi_j}{r_j} - \phi_{j-1})}$$

The former condition can be viewed as one that originates from non-oscillatory spatial interpolation and the latter is essentially a time-step restriction resulting from the higher than first order scheme. These conditions can be met if limiters satisfy $0 \leq \phi(r) \leq min(2, 2r)$ and the time-step satisfies $\nu \leq \frac{1}{2}$. In addition, the limiters may be required to satisfy the second order condtion, requiring $\phi(r) \in I[1, r]$. These two regions are shown in fig. 2.4. The second order TVD region corresponds to the part where the different shaded regions intersect. Several limiters have been designed over the past two decades. Ref. [10] presents a comprehensive review of different limiters. Fig. 2.5 shows the minmod limiter [5] given by, $\phi(r) = minmod(1, r) = 0.5(1+sgn(r))*max(1, abs(r))$. This limiter corresponds to the lower bound of the TVD region and proves to be dissipative in practice. The Van-leer limiter [16] given by $\phi(r) = \frac{r+|r|}{1+r}$ is shown in fig. 2.6. This limiter lies inside the TVD region and in addition, is differentiable - a fact that is known to enhance the steady convergence [10].

It has to be mentioned that the requirement $\phi(r) = 0 \ \ \forall \ \ r < 0$ and $\phi(r) \geq 0$ will result in clipping of true extrema. Various high order extensions like the UNO [8], ENO[21], WENO [13] that satisfy less stringent conditions than the TVD have been developed.

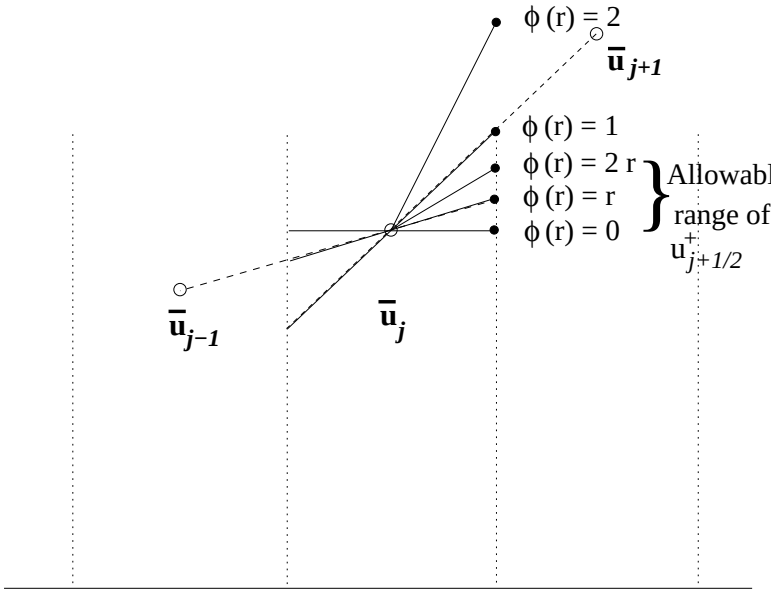


Figure 2.3: Several choices of $\phi(r)$

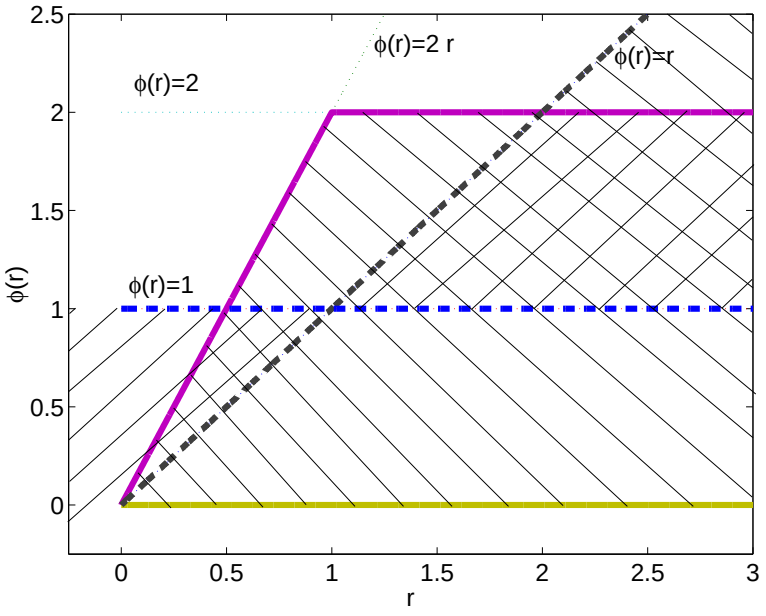


Figure 2.4: Second order accurate and TVD regions. Second order accurate region: lines slanted to the left. TVD region: lines slanted to the right

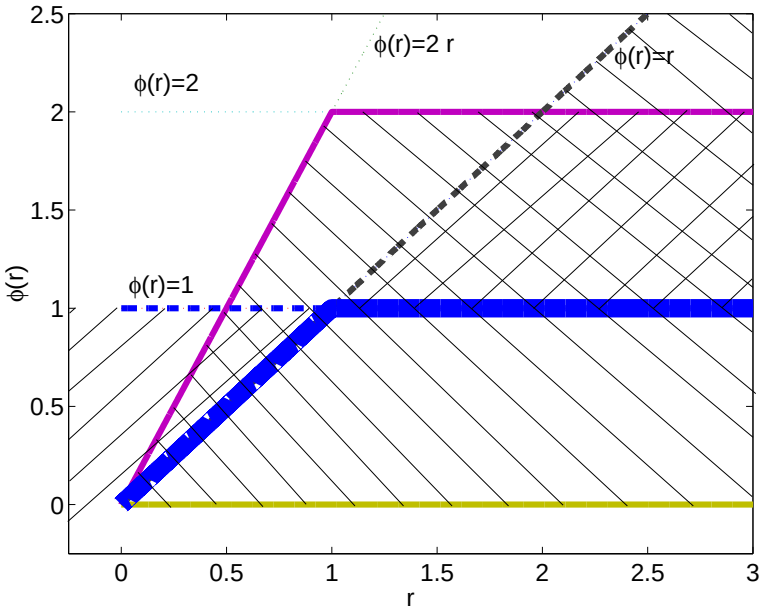


Figure 2.5: Minmod limiter: Thick lines

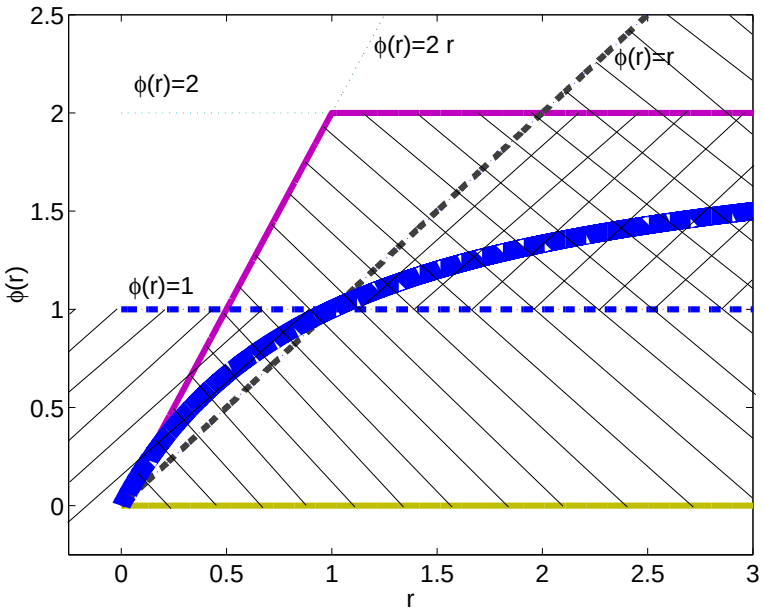


Figure 2.6: Van-leer limiter: Thick lines

2.5 TVD limits in a Method-of-lines framework

In this section, the TVD limits of time integration schemes are reviewed in a method-of-lines framework. This approach follows that of Shu et al. [2], where the analysis is performed for explicit Runge Kutta schemes.

Spatial discretization of eqn. 2.1 would yield a system of ODEs in time, which is represented by eqn. 2.6. As mentioned earlier, a large number of highly accurate non-oscillatory spatial discretizations are available in the literature. These schemes become conditionally TVD when used with an Euler explicit scheme as detailed in the previous section. This restriction is used as a basis for comparison. Let $\Delta t \leq \Delta t_{ee}$ be the allowable time step for which the explicit Euler scheme is TVD. i.e.,

$$\| u^n \| \geq \| u^n + \Delta t L(u^n) \| \quad for \quad \Delta t \leq \Delta t_{ee} \tag{2.14}$$

Note: The norm in question is the Total Variation semi-norm. For example, in the previous section, $a\frac{\Delta t_{ee}}{\Delta x} = 0.5$.

This provides a good basis for evaluating other time-integration schemes for eqn. (2.6). Hence, any time integration scheme for eqn. (2.6) would yield a discrete numerical scheme which will be conditionally TVD under a new time step restriction, which is denoted by $\Delta t \leq k\Delta t_{ee}$. Shu et al. [2], [14] have carried out an analysis for explicit Runge Kutta (RK) schemes. Shu's results show that $k \leq 1$ for high order explicit RK schemes. Further, Shu et al., have also shown in [14] that implicit RK and implicit multi-step methods of order higher than 1 have finite values of k (or in other words, the resulting numerical schemes are conditionally TVD).

In this work, the value of k is derived for some existing implicit schemes (Table 2.2). The details are given in appendix A. It is seen that only the implicit Euler method (which is first order accurate) is unconditionally TVD. The Trapezoidal method becomes non-monotone for a time-step twice that of the explicit Euler method and the implicit BDF2 becomes non-monotone for half the time-step of the explicit Euler method. The SDIRK-2 which is a 2 stage implicit method performs marginally better than the Trapezoidal method. Hence, it becomes obvious that when these schemes (though they are implicit) are used in the numerical solution of conservation equations, they still have severe time-step restrictions.

Method[1]	State update formula	k
Imp. Euler	$u^{n+1} = u^n + \Delta t L(u^{n+1})$	∞
Imp. Trap	$u^{n+1} = u^n + \frac{\Delta t}{2}[L(u^n) + L(u^{n+1})]$	2.0
Imp. BDF2	$u^{n+1} = \frac{2\Delta t L(u^{n+1}) + 4u^n - u^{n-1}}{3}$	0.5
SDIRK-2 ($\gamma = \frac{2-\sqrt{2}}{2}$)	$u^{(1)} = u^n + \Delta t \gamma L(u^{(1)})$ $u^{n+1} = u^n + \Delta t[(1-\gamma)L(u^{(1)}) + \gamma L(u^{n+1})]$	$\frac{1}{1-2\gamma} \approx 2.4142$

Table 2.2: TVD Limits for some existing Implicit schemes

2.6 Large Time-step Godunov methods

There exists some literature on large time step extensions of the Godunov methods for explicit schemes. Refs. [24]-[26] detail methods where the characteristics are traced backward across several computational cells. However, these methods are not general enough since tracking the domain of dependence becomes difficult when strong non-linearities exist in the solution. Leveque [27] presents a method in which interactions of waves from neighboring cells are handled in an approximate linearized manner. Encouraging results are reported for linear equations, but the solutions seem to degrade for non-linear problems as a result of the linear treatment of the interactions. Guinot [28] presents a *time-line interpolation* scheme where the equivalence of space and time is utilized. In this method, information from characteristics originating in the entire domain of dependence is used. This is accomplished by sweeping through the cell interfaces and solving the resulting recurrence relations. All the methods described above are explicit and are essentially based on an extension of the definition of Godunov methods. Forth [20] presents a Godunov-based scheme for scalar equations that switches between the implicit Euler and 2nd order implicit BDF methods in different parts of the domain. This method is similar in concept to the approach to be detailed in Chapter 4. Both of these approaches yielded non-oscillatory schemes but turn out to be diffusive at large CFL numbers.

[1]Imp. BDF2: Implicit second order Backward Difference Method, SDIRK-2: 2-Stage Singly Diagonally Implicit Runge Kutta Method

Chapter 3

Generation of Spurious Oscillations

In this chapter, an attempt is made at understanding the mechanism of generation of spurious oscillations in the solution of hyperbolic conservation laws. The linear convection equation ($u_t + au_x = 0$ with $a = 1$) is chosen to be the representative problem and the numerical schemes are restricted to Godunov-type methods as described in the previous chapter.

3.1 Solution near a discontinuity

Consider the profile shown in fig. 3.1. The filled circles represent the discrete values of the solution, which is to be advanced one time-step. The reconstructed value at the cell-interfaces using different methods[1]is also shown. The un-limited methods treat information without any bias and hence, tend to interpolate across the discontinuity, whereas, the limited method drops to first order (at $u_{j+\frac{1}{2}}$). The updated value u_j^{n+1} is given by:

$$u_j^{n+1} = u_j^n - \tau(u_{j+\frac{1}{2}}^n - u_{j-\frac{1}{2}}^n) \tag{3.1}$$

Fig. 3.2 shows the update $u_j^p = u_j^n - (u_{j+\frac{1}{2}}^n - u_{j-\frac{1}{2}}^n)$: Therefore, for $0 \leq \tau \leq 1$, u_j^{n+1}, any use of information across the discontinuity will lead to undershoots for *any* $\tau \geq 0$.

[1]Any limited method would give the same reconstructed value of $u_{j+\frac{1}{2}} = 0$ for this particular profile

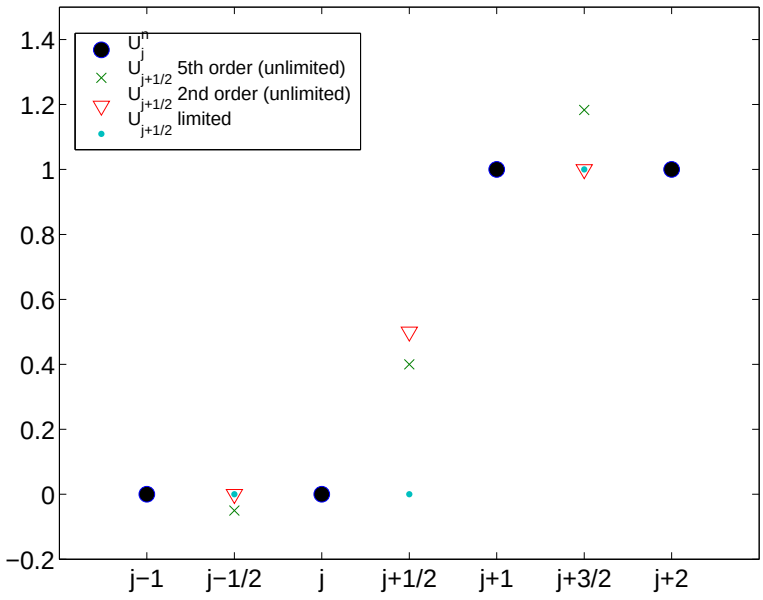


Figure 3.1: Reconstructed values obtained by various spatial discretization methods

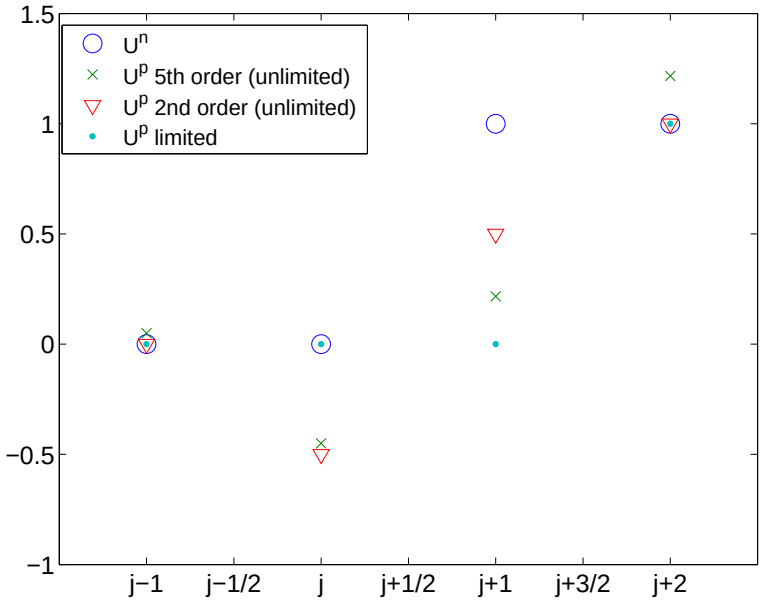


Figure 3.2: Range of updated values for various methods. $u^{n+1} \in I[u^n, u^p] \ \forall \ 0 \leq \tau \leq 1$

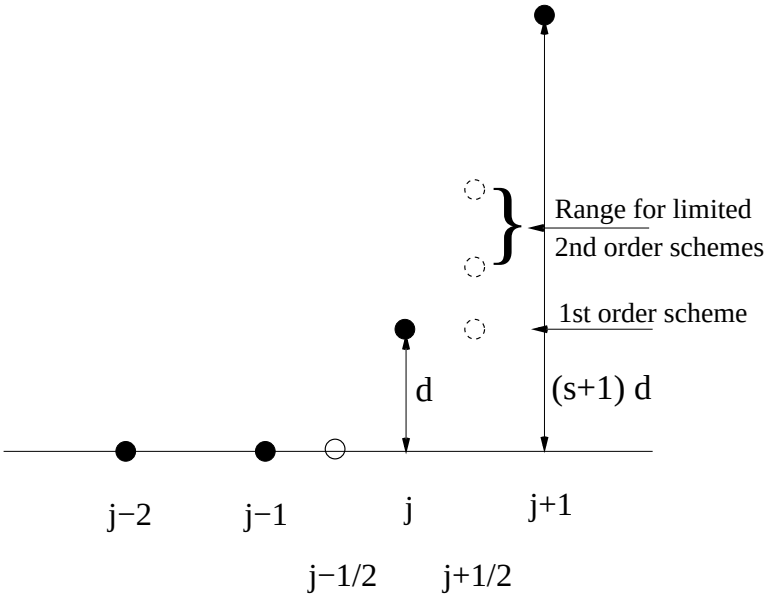


Figure 3.3: Sample data distribution.

3.2 Solution near region of high gradient

In the previous section, limited interpolation yielded a non-oscillatory solution for $\tau \leq 1$. In this section, a different scenario that results in a more severe time-step restriction will be demonstrated. Consider initial data distributed as shown by the filled circles in fig. 3.3. For the data: $u^n_{j-2} = 0, u^n_{j-1} = 0, u^n_j = d, u^n_{j+1} = (s+1)d$. $s \geq 0$ determines the steepness of the profile. Note that these data points could correspond to a smooth profile. (As an example, $u = 0$ if $x \leq x_j$ and $u = (x - x_j)^n$ if $x >= x_j$). Any limited scheme would give $u_{j-\frac{1}{2}} = 0$. Hence, the update is:

$$u^{n+1}_j = d - \tau u^n_{j+\frac{1}{2}} \tag{3.2}$$

An under-shoot will be developed if $u^{n+1}_j < 0$ or equivalently, $\tau > \frac{u^n_{j+\frac{1}{2}}}{d}$. Table 3.1 shows the reconstructed value $u^n_{j+\frac{1}{2}} = u^n_j + \phi(r_j)\Delta u_{j+\frac{1}{2}}$ for various schemes. It is seen that for higher order schemes, $u_{j+\frac{1}{2}} > d$ and hence, the allowable time-step $\tau < 1$. Table 3.1 also gives the maximum time-step τ_{max} for which there is no under-shoot for any s.

[2]It is possible to write the 3 point WENO scheme in the limiter form. The derivation is shown in appendix B

Method	Limiter	$\frac{u_{j+\frac{1}{2}}}{d}$	τ_{max}
First order	$\phi(r) = 0$	1	1.0000
Min-mod	$\phi(r) = minmod(1, r)$	$1 + min(\frac{s}{2}, \frac{1}{2})$	0.6667
Van-leer	$\phi(r) = \frac{r+\|r\|}{1+r}$	$1 + \frac{s}{s+1}$	0.5000
3 point WENO^2	$\phi(r) = \frac{r+2r^4}{1+2r^4}$	$1 + \frac{0.5s^4+s}{s^4+2}$	0.6329

Table 3.1: Different second order reconstructions

.

3.3 Time-step restrictions

In this section, time-step restrictions resulting from general TVD conditions will be explored for 3 point[3]explicit schemes for the linear advection equation.

3.3.1 Explicit schemes

For the linear advection equation with $a > 0$, second order explicit schemes can be cast in the form:

$$u_j^{n+1} = u_j^n - \tau(C^n_{j-\frac{1}{2}} \Delta u^n_{j-\frac{1}{2}}) \tag{3.3}$$

As given in eqn. 2.13, for second order limited schemes,

$$C_{j-\frac{1}{2}} = 1 + \frac{1}{2}(\frac{\phi_j}{r_j} - \phi_{j-1})$$

Equation 3.3 can be written as

$$u_j^{n+1} = u_j^n(1 - \tau C^n_{j-\frac{1}{2}}) + u^n_{j-1}\tau C^n_{j-\frac{1}{2}} \tag{3.4}$$

Therefore, the positivity conditions $C^n_{j-\frac{1}{2}} \geq 0$ and $\tau C^n_{j-\frac{1}{2}} \leq 1$ ensure a convex combination and imply TVD. Hence, to determine the time-step restrictions, it suffices to find the bounds of the function

$$Q(r, s) = \frac{\phi(r)}{r} - \phi(s) \quad \forall \quad r, s \in \mathbf{R} \tag{3.5}$$

[3]This applies to multi-point schemes as well since they can be equivalently cast in the 3 point form

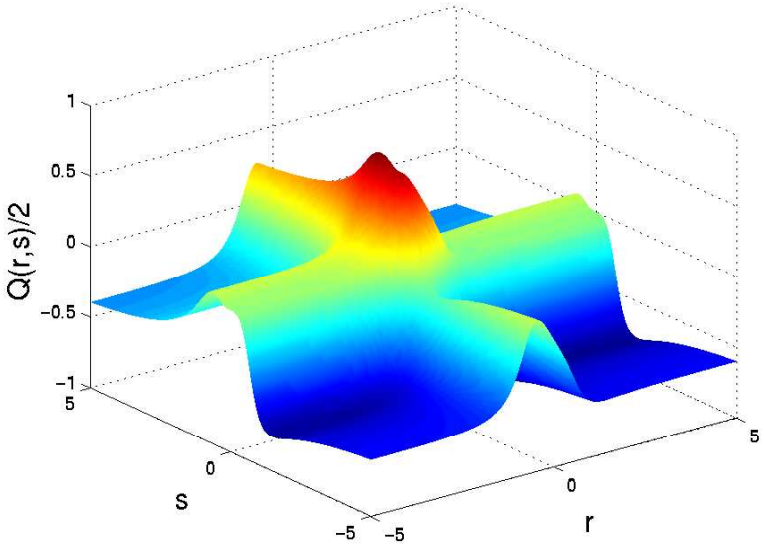


Figure 3.4: 3 point WENO scheme: $Q(r,s) = \frac{1+2r^3}{1+2r^4} - \frac{s+2s^4}{1+2s^4}$

As a sample case, consider the 3 point WENO scheme. For this scheme, Q(r,s) is shown in fig. 3.4. It is possible to find the extrema analytically and these values are given by: $Q_{min} = -1.564$ at $\{r,s\} = \{-1.44, 1.39\}$ and $Q_{max} = 1.4818$ at $\{r,s\} = \{0.657, 0.454\}$. Hence, it is seen that the scheme is TVD for $\tau \leq 0.574$. These results are reflected in the numerical computations also. Fig. 3.5 shows the initial condition for the linear advection equation in a periodic domain of 120 points. Fig. 3.6-3.9 show the solution computed using various time-step sizes to a final time $t = 0.4033$. From the calculations, it becomes evident that small undershoots and overshoots[4]develop for any $\tau > 0.6$. This value of τ is very close to the analytical bound of 0.574. The results of similar analyses for other limiters is shown in Table 3.2.

3.3.2 Implicit schemes

The previous section dealt with the explicit Euler TVD limits for different spatial discretizations. As mentioned in the previous chapter, the limit for implicit time integration schemes can be related to the explicit Euler limit and it was found that these limits are in fact not much different from the explicit limits.

[4]The $\tau = 0.6417$ result generated undershoots at early time-steps, but they were subsequently damped out.

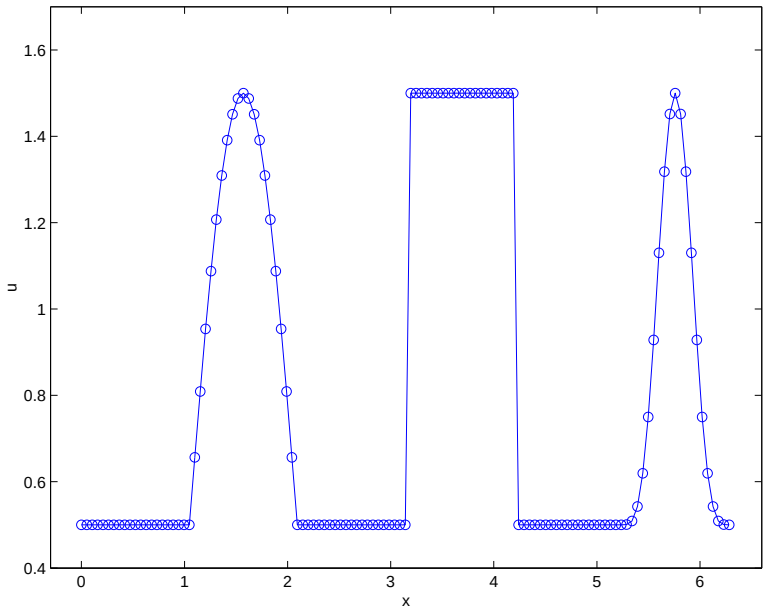


Figure 3.5: Initial condition used for sample calculation

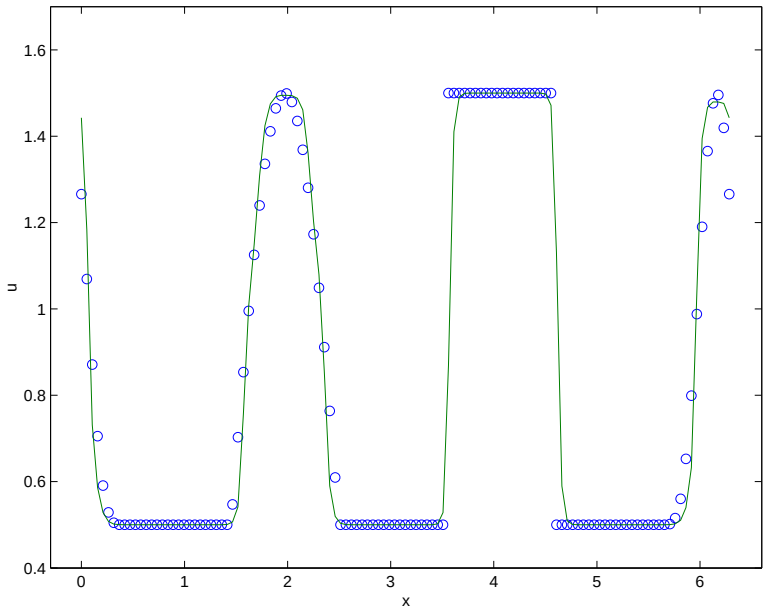


Figure 3.6: Solution after 14 time-steps, $\tau = 0.55$. Circles: exact, Lines: 3 point WENO

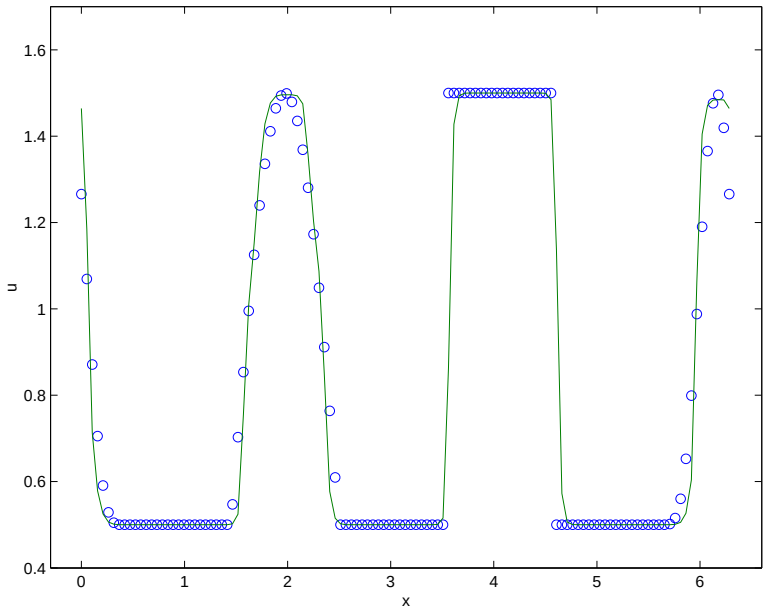


Figure 3.7: Solution after 13 time-steps, $\tau = 0.5923$. Circles: exact, Lines: 3 point WENO

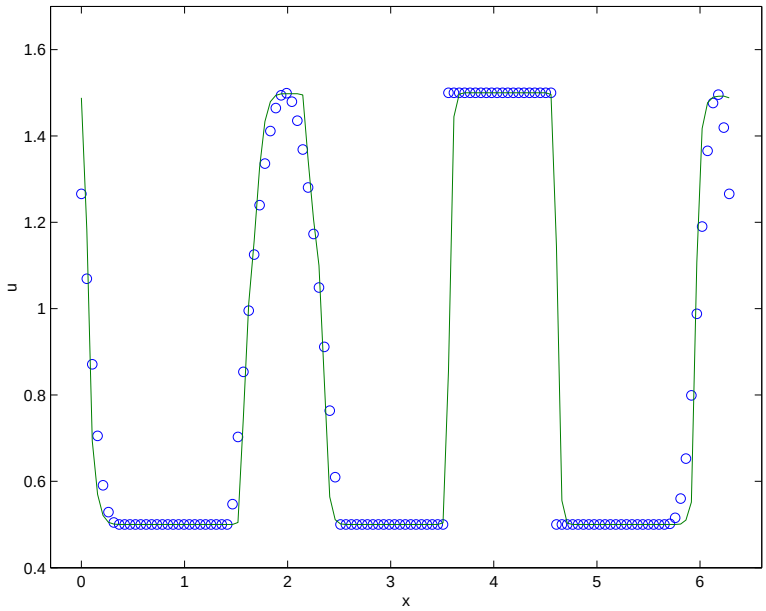


Figure 3.8: Solution after 12 time-steps, $\tau = 0.6417$. Circles: exact, Lines: 3 point WENO

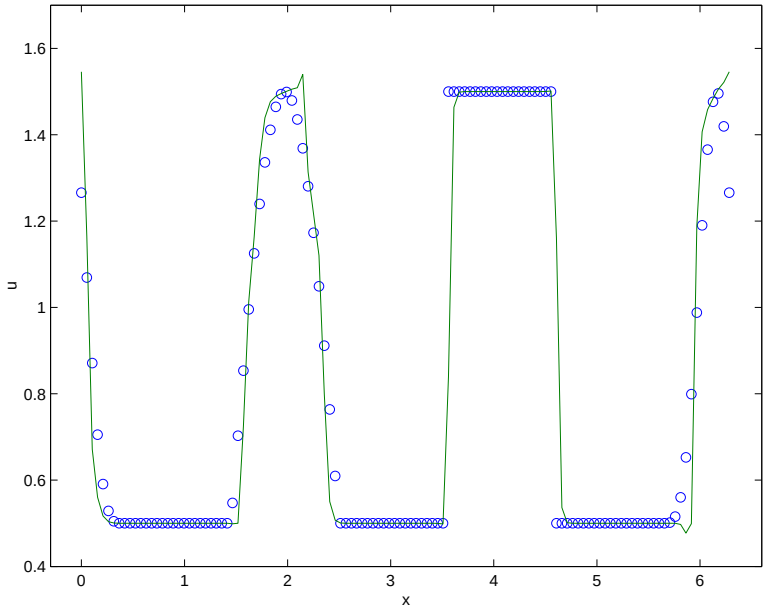


Figure 3.9: Solution after 10 time-steps, $\tau = 0.7$. Circles: exact, Lines: 3 point WENO

Method	τ_{max}
First order	1.0000
Min-mod	0.6667
2nd order ENO	0.5000
Van-leer	0.5000
3 point WENO	0.5742

Table 3.2: TVD limits for different schemes

The underlying physical reason for this is the same as that for the spatial reconstruction. In the spatial reconstruction, it was demonstrated that usage of unbiased information *across the discontinuity in space* would result in spurious oscillations. Similarly, high order implicit time integration schemes can generate spurious oscillations if unbiased information is used across different time-levels in the presence of discontinuities.

In conventional numerical schemes for hyperbolic conservation laws, the spatial reconstruction is made non-linear by the application of a limiter, but the temporal scheme is still *linear* and hence subject to severe time-step restrictions. In the following chapters, two novel methods that *limit* temporal information will be presented.

Chapter 4

Limited Time Integration

4.1 Formulation

4.1.1 The L-TRAP Scheme

Before presenting the details of the new scheme, it is useful to consider the θ method, which is a time-space decoupled method. Consider conservative and consistent numerical flux functions (of any spatial order) corresponding to a non-oscillatory scheme:

$$\begin{aligned} f^n_{j+\frac{1}{2}} &= f(\bar{u}^n_{j-l_1}, \bar{u}^n_{j-l_1+1}, ..., \bar{u}^n_{j-l_2-1}, \bar{u}^n_{j-l_2}), \quad and \\ f^{n+1}_{j+\frac{1}{2}} &= f(\bar{u}^{n+1}_{j-l_1}, \bar{u}^{n+1}_{j-l_1+1}, ..., \bar{u}^{n+1}_{j-l_2-1}, \bar{u}^{n+1}_{j-l_2}). \end{aligned}$$

The well-known θ method for the solution of eqn. (2.1) is given by:

$$\bar{u}^{n+1}_j = \bar{u}^n_j - \tau \left[(1-\theta)(f^n_{j+\frac{1}{2}} - f^n_{j-\frac{1}{2}}) + \theta(f^{n+1}_{j+\frac{1}{2}} - f^{n+1}_{j-\frac{1}{2}}) \right] \tag{4.1}$$

On immediate observation, it is found that if a constant value of θ is used over the whole domain, one obtains schemes as shown in Table 4.1, and the TVD limit is given by $k = \frac{1}{1-\theta}$. This scheme can be written in the conservation form,

$$\bar{u}^{n+1}_j = \bar{u}^n_j - \tau(\hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}}), \tag{4.2}$$

where,

$$\hat{f}_{j\pm\frac{1}{2}} = (1-\theta) f^n_{j\pm\frac{1}{2}} + \theta f^{n+1}_{j\pm\frac{1}{2}}$$

A new numerical method is proposed, for which:

$$\begin{aligned} \hat{f}_{j\pm\frac{1}{2}} &= (1-\theta_{j\pm\frac{1}{2}}) f^n_{j\pm\frac{1}{2}} + \theta_{j\pm\frac{1}{2}} f^{n+1}_{j\pm\frac{1}{2}}, \quad with \\ \theta_{j\pm\frac{1}{2}} &= 0.5(\theta_{j\pm1} + \theta_j), \quad \theta_j, \theta_{j\pm1} \in [0.5, 1] \end{aligned}$$

Condition	Method	Accuracy	TVD limit (k)
$\theta = 1.0$	Imp. Euler	1st order	∞
$\theta = 0.5$	Imp. Trap	2nd order	2.0
$\theta = 0.0$	Exp. Euler	1st order	1.0

Table 4.1: Family of schemes for fixed value of θ

Here, $\theta_j \in I[0.5, 1]$ and is defined for each cell j (the definition of θ_j is given in the following subsections). Now, since $\hat{f}_{j\pm\frac{1}{2}}$ is a convex combination of conservative and consistent numerical fluxes $f^n_{j\pm\frac{1}{2}}$ and $f^{n+1}_{j\pm\frac{1}{2}}$, it is itself conservative and consistent.

The idea then, is to define a θ_j, such that in regions where the solution is smooth, $\theta_j \approx 0.5$ is used, thus attaining second order accuracy in time locally. In regions of high gradients, $\theta_j \approx 1.0$ is used, thus locally dropping the time accuracy to first order. By doing this, *an attempt is made at locally satisfying the TVD condition. The definition of θ_j is based on the time evolution of the solution at j, hence this method is no longer fully space-time decoupled.* This method is termed the Limited-Trapezoidal method or the L-TRAP.

4.1.2 The L-DIRK2 Scheme

A new 2-stage implicit scheme is proposed for the solution of eqn. (2.1). The scheme is given by,

$$\begin{aligned}
\bar{u}^{(1)}_j &= \bar{u}^n_j - \tau\gamma(f^{(1)}_{j+\frac{1}{2}} - f^{(1)}_{j-\frac{1}{2}}) \\
\bar{u}^{n+1}_j &= \bar{u}^n_j - \tau[(a_{21_{j+\frac{1}{2}}} f^{(1)}_{j+\frac{1}{2}} - a_{21_{j-\frac{1}{2}}} f^{(1)}_{j-\frac{1}{2}}) \\
&+ (a_{22_{j+\frac{1}{2}}} f^{n+1}_{j+\frac{1}{2}} - a_{22_{j-\frac{1}{2}}} f^{n+1}_{j-\frac{1}{2}})]
\end{aligned}$$

where,

$$\begin{aligned}
a_{21_{j\pm\frac{1}{2}}} &= \gamma + \theta_{j\pm\frac{1}{2}}(1 - 2\gamma), \\
a_{22_{j\pm\frac{1}{2}}} &= (1 - \gamma) + \theta_{j\pm\frac{1}{2}}(2\gamma - 1), \\
\theta_{j\pm\frac{1}{2}} &= 0.5(\theta_j + \theta_{j\pm1}) \quad , \quad \theta_j \in [0, 1], \\
\gamma &= \frac{2 - \sqrt{2}}{2}
\end{aligned}$$

If $\theta_j = \theta$ is assumed constant at all points in the domain, the TVD limit is given by, $k = \frac{1}{\theta(1-2\gamma)}$. It is seen that this scheme reduces to the second order

(in time) SDIRK [3] for $\theta = 1.0$ giving a $k = 2.4142$ and to a 2-stage first order unconditionally TVD method if $\theta = 0$. (Refer appendix A for details). Note that both these schemes are in a method-of-lines framework.
Again, the basic idea is to locally drop the order of accuracy in regions of high gradients and maintain second order accuracy in smooth regions. Hence, the local θ_j is defined accordingly.

4.1.3 Definition of θ_j

The L-TRAP and L-DIRK2 schemes as defined in the previous sections are general and this section presents just one of the ways by which a limiter could be designed.
The definition for θ_j is inspired by the following lemma by Hyunh [9] for quadratic interpolation: *Given the data* $f(q_1), f(q_2)$ *and the derivative* $f'(q_1)$ *or* $f'(q_2)$ *at points* q_1 *and* q_2*, the resulting quadratic interpolant is monotone in* $I[q_1, q_2]$ *if* $f'(q_1), f'(q_2) \in I[0, 2s]$*, where* $s = \frac{f(q_2)-f(q_1)}{q_2-q_1}$.
Hence, this monotonicity condition is evaluated (with respect to the time derivative) at all points in the domain. For those points at which this condition is satisfied, $\theta_j \approx 0.5$ for L-TRAP and $\theta_j \approx 1.0$ for L-DIRK2, thus allowing the local time accuracy to be second order. For points at which this condition is not satisfied, a value of θ_j ranging from 0.5 to 1.0 for L-TRAP and θ_j ranging from 1.0 to 0.0 for the L-DIRK2 is assigned. The procedure is as follows:
Let

$$\begin{aligned} s_j^{n+\frac{1}{2}} &= \frac{u_j^{n+1} - u_j^n}{\Delta t} \\ L(u_j^n) &= \left(\frac{\partial u}{\partial t}\right)_j^n = -\frac{1}{\Delta x}(f_{j+\frac{1}{2}}^n - f_{j-\frac{1}{2}}^n) \\ L(u_j^{n+1}) &= \left(\frac{\partial u}{\partial t}\right)_j^{n+1} = -\frac{1}{\Delta x}(f_{j+\frac{1}{2}}^{n+1} - f_{j-\frac{1}{2}}^{n+1}) \end{aligned}$$

Define a parameter r_j for each of the domain points with $\theta_j = 1.0 - 0.5r_j$ for the L-TRAP scheme and $\theta_j = r_j$ for the L-DIRK2 scheme. Hence, $r_j = 1$ for second order accuracy and $r_j = 0$ for first order accuracy.
For monotone quadratic interpolation $L(u_j^n)$ and $L(u_j^{n+1})$ should lie in the interval $[0, 2s_j^{n+\frac{1}{2}}]$. This is true if,

$$L(u_j^{n+1}) * (L(u_j^{n+1}) - 2s_j^{n+\frac{1}{2}}) \leq \epsilon \;\; and \;\; L(u_j^n) * (L(u_j^n) - 2s_j^{n+\frac{1}{2}}) \leq \epsilon. \qquad (4.3)$$

$\epsilon = 0$ would strictly be equivalent to the conditions of Hyunh's lemma. An arbitrarily small and positive ϵ (in this case $\epsilon = 1 \times 10^{-10}$) is used to ensure that the limiter does not turn on and off spuriously for small changes in the solution. Hence, *if eqn. (4.3) is satisfied,* $r_j = 1.0$ *is assigned* and local second order accuracy is approached.
If,

$$\frac{L(u_j^{n+1})}{(s_j^{n+\frac{1}{2}} + \epsilon)} \leq -\epsilon_1 \quad or \quad \frac{L(u_j^n)}{(s_j^{n+\frac{1}{2}} + \epsilon)} \leq -\epsilon_1 \tag{4.4}$$

then, either or both of the quantities $L(u_j^n)$ and $L(u_j^{n+1})$ are of the opposite sign as $s_j^{n+\frac{1}{2}}$ and hence the interpolant can become non-monotone. Hence, *if eqn. (4.4) is satisfied,* $r_j = 0$ *is assigned* and hence the local accuracy is dropped to first order. ϵ_1 is again arbitrarily small and positive. A value of 1×10^{-5} is used. If eqns. (4.3) and (4.4) are not satisfied, then one has a case which is analogous to a situation where $f'(q_1)$ and $f'(q_2)$ are of the same sign as s, but either or both are too steep. In such a case, an appropriate $r_j \in [0, 1]$ is assigned as follows:

$$r_j = min \left[\frac{2s_j^{n+\frac{1}{2}}}{L(u_j^n) + \epsilon}, \frac{2s_j^{n+\frac{1}{2}}}{L(u_j^{n+1}) + \epsilon}, 1 \right] \tag{4.5}$$

It becomes clear that $r_j \approx 0$ at discontinuities and at extrema, since these would be indistinguishable by the above procedure. This could introduce undesirable clipping even at smooth extrema. This can be corrected by checking the spatial interpolant. For example, if a MUSCL [17] spatial interpolation is used, the strategy developed by Suresh et. al. is used in [15] to preserve accuracy near extrema. Given a highly accurate interpolated value at $x_{j+\frac{1}{2}}$, this method determines whether spatial-limiting is required at $x_{j+\frac{1}{2}}$ based on a 4 point spatial stencil. Hence if spatial-limiting is not required for the current solution at $x_{j+\frac{1}{2}}$ at time levels n and $n+1$, r_j is reset to 1.0. Note that if one uses the monotonicity preserving scheme of Suresh et. al. [15] for spatial discretization, the extra work required to reset the time-limiter is minimal.

4.2 Linear and Non-linear Stability Analysis

A linear stability and monotonicity analysis is presented for a simple case. The L-TRAP method is compared with the implicit Euler and Trapezoidal schemes for

the linear advection equation $u_t + u_x = 0$ with a first order upwind discretization in periodic space. This discretization is chosen because it is simple, linear and monotone (in a semi-discretized sense). Let the CFL number be represented by $\sigma = \frac{\Delta t}{\Delta x}$ and let the number of spatial points be N. Then, the L-TRAP (and related families of schemes) would be given by:

$$u_j^{n+1} = u_j^n - \sigma \left[(1 - \theta_{j+\frac{1}{2}}) u_j^n + \theta_{j+\frac{1}{2}} u_j^{n+1} - (1 - \theta_{j-\frac{1}{2}}) u_{j-1}^n - \theta_{j-\frac{1}{2}} u_{j-1}^{n+1} \right]$$

This can be represented in matrix form ($\mathbf{D}[\mathbf{a}, \mathbf{b}]$ is a $N \times N$ periodic bi-diagonal matrix with lower-diagonal elements a_j and diagonal elements b_j):

$$\mathbf{A}[-\sigma\theta_{j-\frac{1}{2}}, 1 + \sigma\theta_{j+\frac{1}{2}}] \;\; \mathbf{U}^{n+1} = \mathbf{B}[\sigma(1 - \theta_{j-\frac{1}{2}}), 1 - \sigma(1 - \theta_{j+\frac{1}{2}})] \;\; \mathbf{U}^n$$

This can be represented as:

$$\mathbf{U}^{n+1} = \mathbf{M}\mathbf{U}^n, \quad where, \quad \mathbf{M} = \mathbf{A}^{-1}\mathbf{B} \tag{4.6}$$

Because of the linearity of the problem, one can completely determine the properties of these schemes based on the structure of the matrix $\mathbf{M}$.

4.2.1 Linear Stability

Since the system is periodic, linear stability is equivalent to the spectral radius of $\mathbf{M}$ being less than 1. For an initial condition consisting of a hat function, (N=60, $\sigma = 3.0$, $t = \Delta t$) fig. 4.1 shows the eigen-values of $\mathbf{M}$. Note that $\mathbf{M}$ is a constant matrix for all time-steps for implicit Euler and Trapezoidal methods, whereas, for the L-TRAP case, it changes with time because the limiter is a function of time. It is evident that all 3 schemes are linearly stable. In fact, it is easy to show that all 3 schemes are un-conditionally linearly stable. This figure gives evidence of the dispersion of the Trapezoidal method and the excessive damping of the implicit Euler method. The L-TRAP method stays close to the Trapezoidal method, except for a pair of eigen-values, which show damping. This is because the limiter (refer section 3.3) senses the sharp discontinuities in time. Fig. 4.2 shows the solution after 1 time step. Since the TVD limit of the Trapezoidal method (with 1st order upwind in space) is $\sigma = 2.0$, it exhibits oscillations in the numerical solution. The L-TRAP scheme stays close to the Trapezoidal method in smooth regions and is thus less dissipative than the implicit Euler method.

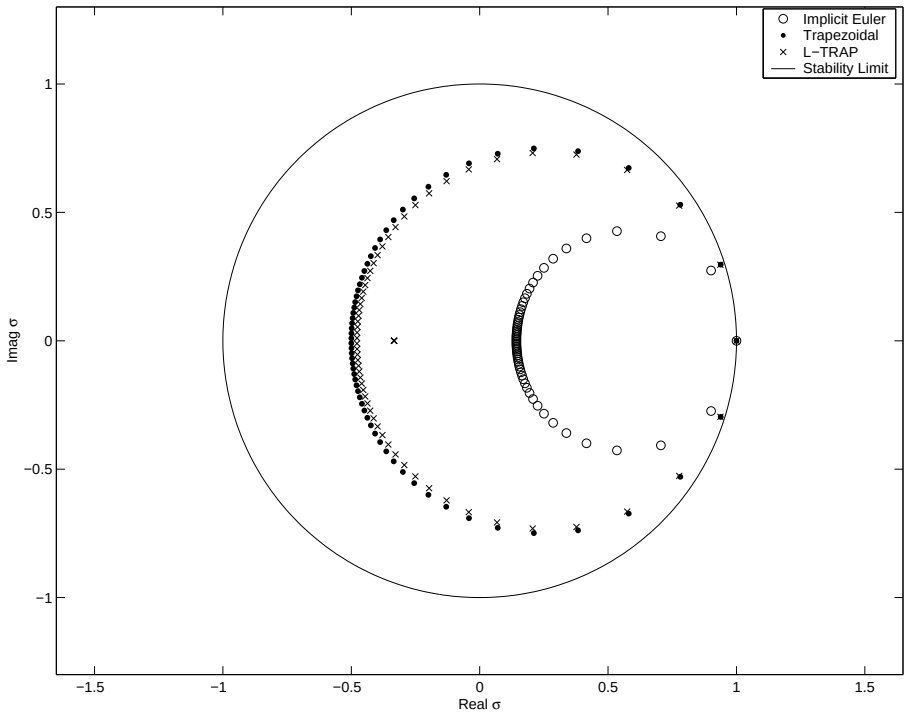


Figure 4.1: Eigen-value plot: $\sigma = 3.0$, N=60, t=Δt

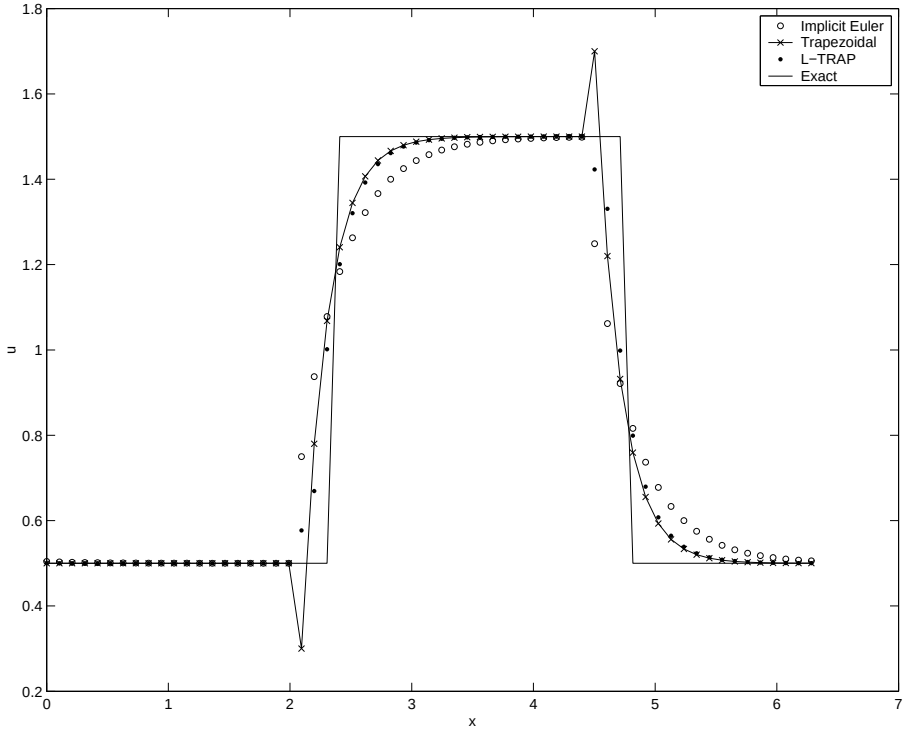


Figure 4.2: Numerical Solution: $\sigma = 3.0$, N=60, t=Δt

4.2.2 Monotonicity Analysis

Consider eqn. (4.6). It is then possible to write:

$$u_i^{n+1} = \sum_{j=1}^{N} m_{i,j} u_j^n$$

Now, monotonicity (refer eqn. (2.7)) is equivalent to ensuring that $m_{i,j} \geq 0, \;\; j = \{1, ..N\}$. It is known that:

• $\sum_{j=1}^{N} m_{i,j} = 1$ from consistency, and

• Therefore, in a global sense, monotonicity is implied by $L^{\infty}(\mathbf{M}) = 1$. If this condition is satisfied, the scheme is rigorously monotone (since along with consistency, this guarantees that $m_{i,j} \geq 0 \;\; \forall \;\; i, j$). Hence the numerical solution will be monotone for any initial condition.

However, *this is a sufficient condition and not a necessary one for the monotone behavior of numerical solutions.* For example, if the initial data is smooth the Trapezoidal method would give a smooth solution even if the time-step exceeds the monotonicity limit. Hence, *global monotonicity is not necessary in this case.* The focal point of the above argument is that one can afford to have a non-monotone scheme (but linear stability is still required) in parts of the domain where the solution is smooth and enforce monotonicity in non-smooth regions. In terms of coefficients, this means that $\sum_{j=1}^{N} |m_{i,j}|$ can be > 1 for some i.

On explicit construction of $\mathbf{M}$, it is found that $m_{i,j} \geq 0 \;\; for \;\; i \neq j$ for all three methods with 1st order upwinding in space. Also, the Determinant of $\mathbf{A}$= $|\mathbf{A}|$ is positive. Hence, the positivity of only the diagonal elements $m_{i,i}$ is required. It is found that these are given by:

$$\begin{aligned}
m_{i,i} &= \frac{1}{|\mathbf{A}|} \left\{ \frac{1 - \sigma(1 - \theta_{i+\frac{1}{2}})}{1 + \sigma\theta_{i+\frac{1}{2}}} \left[\prod_{k=1}^{N} (1 + \sigma\theta_{k+\frac{1}{2}}) \right] + \frac{\sigma(1 - \theta_{i+\frac{1}{2}})}{\sigma\theta_{i+\frac{1}{2}}} \left[\prod_{k=1}^{N} (\sigma\theta_{k+\frac{1}{2}}) \right] \right\} \\
&= \frac{1 - \sigma(1 - \theta_{i+\frac{1}{2}})}{1 + \sigma\theta_{i+\frac{1}{2}}} P_1 + \frac{\sigma(1 - \theta_{i+\frac{1}{2}})}{\sigma\theta_{i+\frac{1}{2}}} P_2, \quad (clearly, \quad 0 < P_2 < P_1) \\
&= \sigma(1 - \theta_{i+\frac{1}{2}}) \left[\frac{P_2}{\sigma\theta_{i+\frac{1}{2}}} - \frac{P_1}{1 + \sigma\theta_{i+\frac{1}{2}}} \right] + \frac{P_1}{1 + \sigma\theta_{i+\frac{1}{2}}}
\end{aligned}$$

From this, the following inferences can be made:

• For the implicit Euler method, $\theta_{i+\frac{1}{2}} = 1 \;\; \forall \;\; i$. Therefore, $m_{i,j} \geq 0 \;\; \forall \;\; i, j$. *Hence,* $L^{\infty}(\mathbf{M}) = 1$ *and this method is unconditionally monotone.*

• For the L-TRAP method, the limiter can ensure that in regions of large gradients, $\theta_{i+\frac{1}{2}} \rightarrow 1$ and hence, $m_{i,i} \geq 0$ can be maintained. Since $m_{i,j} \geq 0, \quad i \neq j$, u_i^{n+1} becomes a convex combination of $u_j^n, \quad j = 1, ..N$ and local monotonicity can be achieved.

• For the Trapezoidal method,

$$m_{i,i} = \frac{(1+\frac{\sigma}{2})^{N-1}(1-\frac{\sigma}{2}) + (\frac{\sigma}{2})^N}{(1+\frac{\sigma}{2})^N - (\frac{\sigma}{2})^N}$$

therefore, when $\sigma > 2 + \epsilon$, ($\epsilon \rightarrow 0$ for large N), $m_{i,i} < 0 \quad \forall i$ and the numerical solution becomes non-monotone.

4.3 Numerical Results

Numerical results are presented for the linear advection equation, the Burger's equation and the one dimensional Euler equations. For all the results presented, LU-SGS [19] with Newton-type sub-iterations [11] is used to solve the implicit set of equations at each time step. Before the start of each sub-iteration, θ_j is determined for all points in the domain and $\theta_{j\pm\frac{1}{2}}$ is updated.

4.3.1 Linear Advection Equation

The first test case is the linear advection equation $u_t + u_x = 0$, with periodic boundary conditions and a smooth initial condition $u_o(x) = sin^4(\frac{x}{2})$ over a domain $[0, 2\pi]$. This initial profile is convected one revolution over the uniform domain. The number of spatial points is represented by N. This test case is chosen to demonstrate the fact that smooth extrema are preserved and uniform second order accuracy in time is achieved. A 5th order monotonicity preserving scheme (MP5) [15] is used for spatial discretization. Figure 4.3 shows the solution after one period of revolution for a domain with N=16 at a CFL number (represented by $\sigma = \frac{\Delta t}{\Delta x}$) of 0.5 using the L-DIRK2 scheme. It is observed that reasonable accuracy is achieved even with this coarse spatial discretization. The extremum is clearly preserved and it is found that the time-limiter sets itself to second order accuracy at all points at all times. Figure 4.4 shows the L_1, L_2 and L_∞ error norms for the L-TRAP and L-DIRK2 schemes for $\sigma = 0.5$ for different levels of spatial discretizations. It is seen that the errors are within range

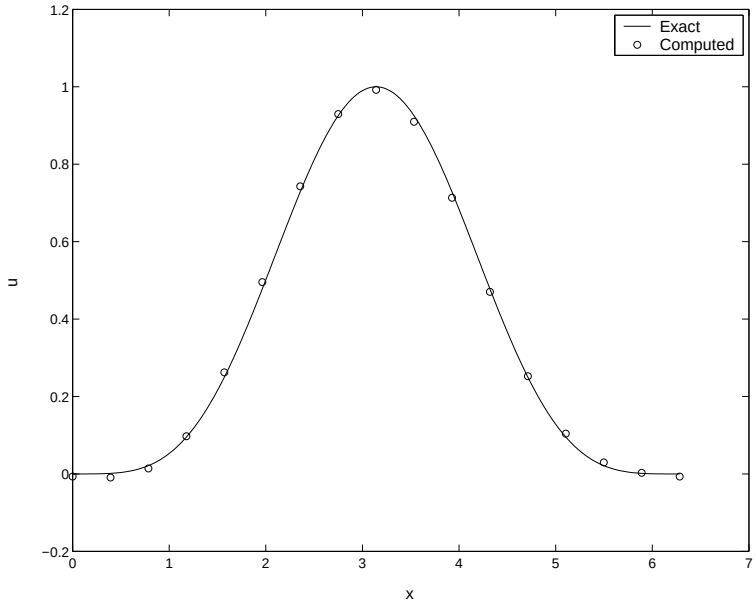


Figure 4.3: Linear advection, $\sigma = 0.5$, MP5 in space, N=16, periodic bc, 1 period of rev

of $(\Delta t)^2$ and $(\Delta x)^5$. The main observation of this exercise is the fact that the schemes are uniformly second order accurate in time for all evaluated cases and no spurious limiting is present at the smooth extremum.
The second case is a domain that comprises of a half-sine wave, a step function and a $sin^4(x)$ distribution over a domain $[0, 2\pi]$, comprising of 360 equally spaced points. This initial condition is convected one revolution over the uniform domain. Spatial discretization is done with the MP5 scheme. The L-TRAP (fig 4.5b) and L-DIRK2 schemes (fig 4.5c) are compared with first and second order implicit schemes. $\sigma = 2.0$ is used because it is high enough to demonstrate the large dissipation exhibited by the first order implicit Euler method and the non-linear instabilities of the second order Trapezoidal and Backward difference methods (fig 4.5a). Note that the CFL number corresponding to the TVD limit $k = 1$ is $\sigma \approx 0.4$ and hence the TVD limit of the Trapezoidal scheme is $\sigma \approx 0.8$ and the limit for the BDF2 is even lower. It is seen that the limited schemes are less dissipative than the first order method and less oscillatory when compared to the linear second order methods.

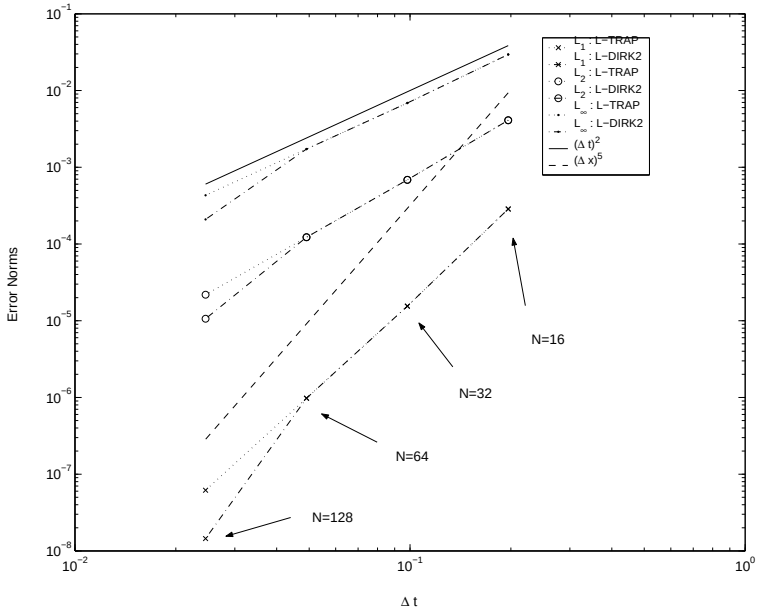


Figure 4.4: Error norms for linear advection $\sigma = 0.5$, MP5 in space, periodic bc, 1 period of rev

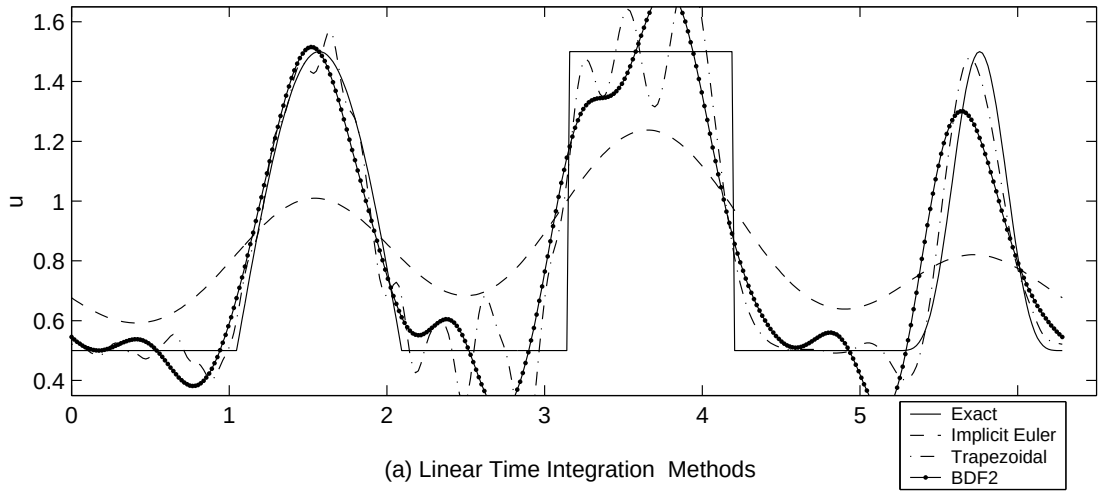
4.3.2 Burger's Equation

The third test case is the inviscid Burger's equation $u_t + (\frac{u^2}{2})_x = 0$, with periodic boundary conditions and a domain $[0, 2\pi]$ of 100 equally spaced points. The initial condition is comprised of an expansion wave and a compression wave. This profile is convected till $t = 2.0$, before which the compression wave becomes a shock. The MP5 spatial discretization scheme is used with $\sigma = \{u\}_{max}\frac{\Delta t}{\Delta x}$. Again, the implicit Euler (fig 4.6a) method shows large dissipation and the linear second order time integration schemes develop oscillations in the vicinity of the shock. It is seen that the limited schemes resolve the expansion wave and the shock well.

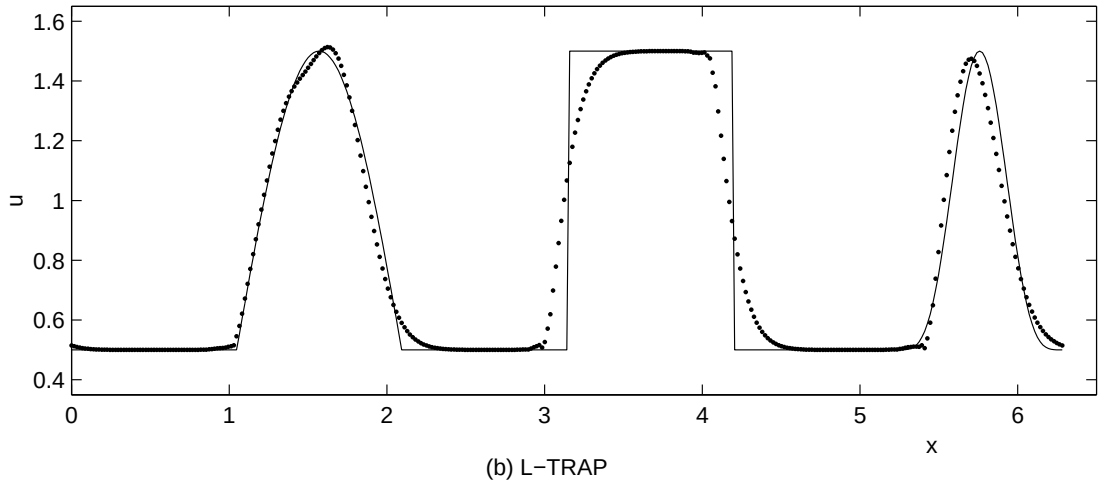
4.3.3 Euler Equations

The one dimensional Euler equations of gas dynamics are given by:

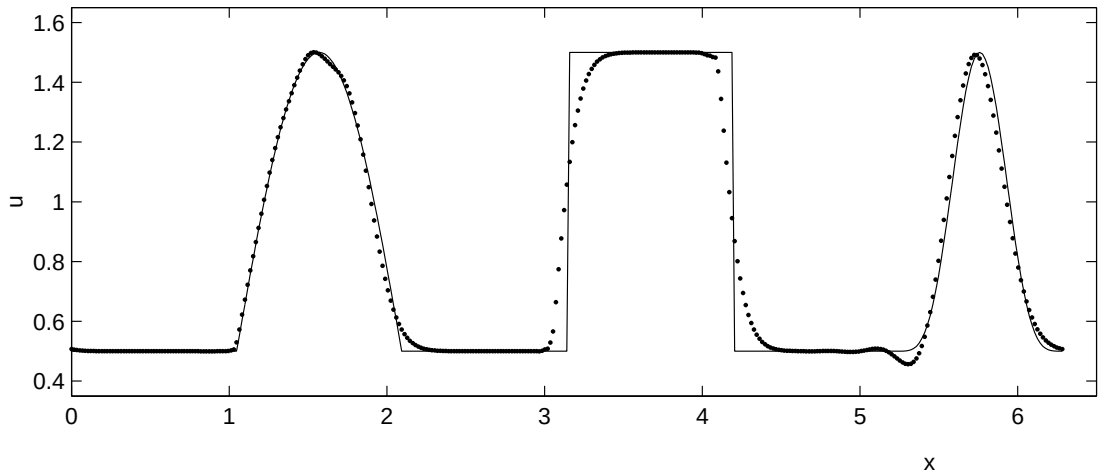
$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0 \tag{4.7}$$



(a) Linear Time Integration Methods



(b) L−TRAP



(c) L−DIRK2

Figure 4.5: Linear advection, $\sigma = 2.0$, MP5 in space, N=360, periodic bc, 1 period of rev

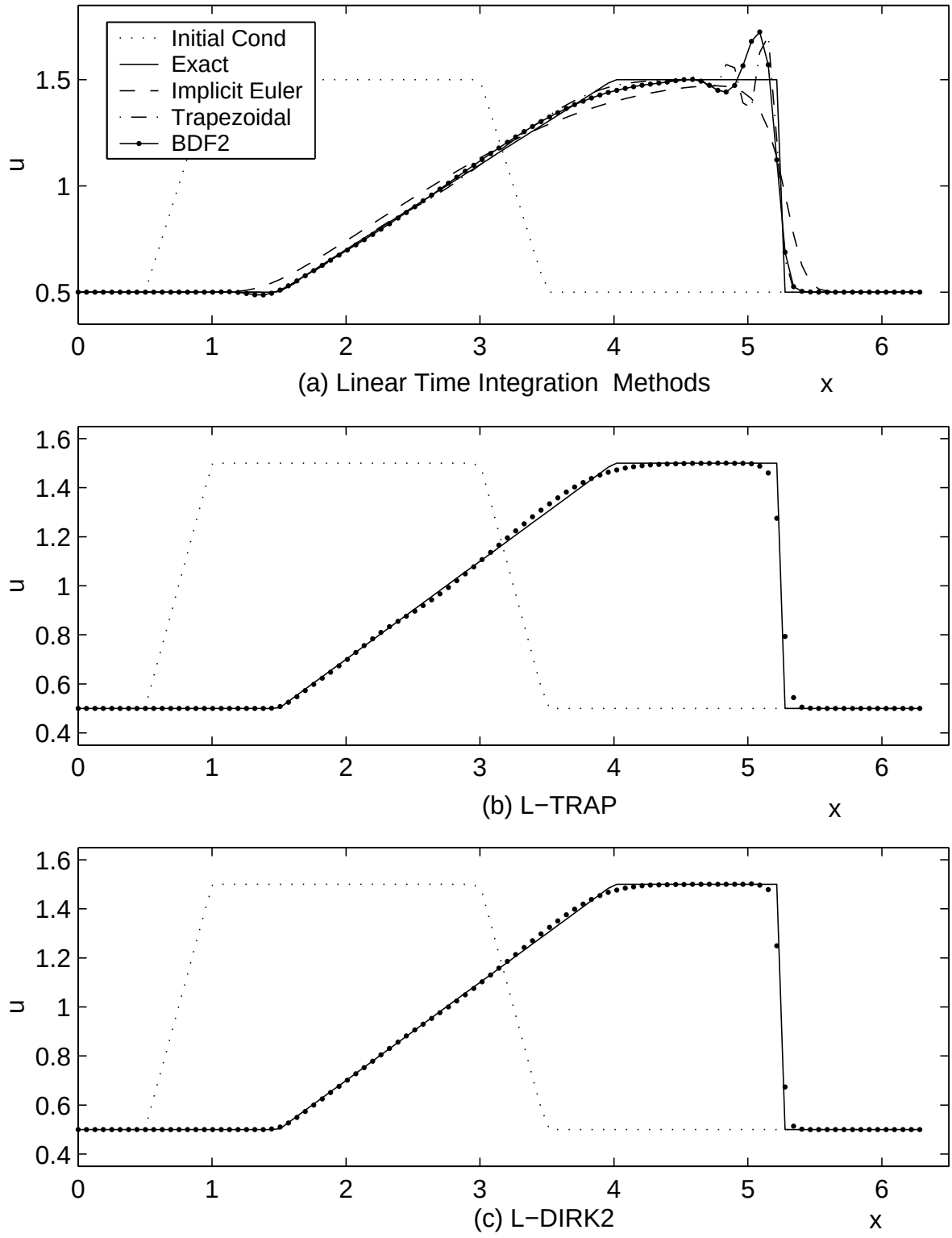


Figure 4.6: Burger's equation, $\sigma = 2.0$, MP5 in space, N=100, periodic bc, t=2.0

where $\mathbf{U}$, the vector of conserved variables and $\mathbf{F}$, the flux vector are defined by,

$$\mathbf{U} = \left\{ \begin{array}{c} \rho \\ \rho u \\ e \end{array} \right\}, \quad \mathbf{F} = \left\{ \begin{array}{c} \rho u \\ p + \rho u^2 \\ (e+p)u \end{array} \right\}$$

ρ, u, p are density, velocity, and pressure respectively. e is the total energy per unit volume given by,

$$e = \frac{p}{\gamma - 1} + \frac{\rho u^2}{2}$$

The concepts of monotonicity as introduced in the earlier sections of the paper cannot be rigorously defined for a non-linear system of equations like the Euler equations. At best, one can extend the concepts of scalar equations and hope to obtain non-oscillatory numerical solutions.

The limiter r_j in eqns. (4.3), (4.4) and (4.5) can be defined in many arbitrary ways for a system of equations. For example, in the present work, a density-based limiter is used, for which,

$$\begin{aligned} s_j^{n+\frac{1}{2}} &= \frac{\rho_j^{n+1} - \rho_j^n}{\Delta t} \\ L(\rho_j^n) &= \left(\frac{\partial \rho}{\partial t} \right)_j^n = -\frac{1}{\Delta x} [(\rho u)_{j+\frac{1}{2}}^n - (\rho u)_{j-\frac{1}{2}}^n] \\ L(\rho_j^{n+1}) &= \left(\frac{\partial \rho}{\partial t} \right)_j^{n+1} = -\frac{1}{\Delta x} [(\rho u)_{j+\frac{1}{2}}^{n+1} - (\rho u)_{j-\frac{1}{2}}^{n+1}] \end{aligned}$$

Since, in a Finite volume framework, these quantities are computed as part of the solution process, these values, when substituted in eqns. (4.3), (4.4) and (4.5) yield the required limiter. Correction at smooth extrema can also be done using the monotonicity preserving strategy as presented in the previous sections. Note that the time derivative of density is computed using the conservation of mass. Another limiter, for instance, can be defined by using a similar procedure on each of the three conserved variables and choosing the one which is closest to first order.

Implementation

In this section, the implementation of the L-TRAP method in the solution of the one dimensional Euler equations will be presented. Consider a discretization of

eqn. (4.7) with an appropriate high order method for computing the interfacial fluxes $\mathbf{F}_{j\pm\frac{1}{2}}$.

$$\mathbf{U}_j^{n+1} - \mathbf{U}_j^n = -\tau \left[((1-\theta_{j+\frac{1}{2}})\mathbf{F}^n_{j+\frac{1}{2}} - (1-\theta_{j-\frac{1}{2}})\mathbf{F}^n_{j-\frac{1}{2}}) + (\theta_{j+\frac{1}{2}}\mathbf{F}^{n+1}_{j+\frac{1}{2}} - \theta_{j-\frac{1}{2}}\mathbf{F}^{n+1}_{j-\frac{1}{2}}) \right] \quad (4.8)$$

The required implicit term $\mathbf{F}(\mathbf{U}^{n+1})$ can replaced by the upwind-split linear approximation:

$$\begin{aligned} \mathbf{F}(\mathbf{U}^{n+1}) &= \mathbf{F}^+(\mathbf{U}^{n+1}) + \mathbf{F}^-(\mathbf{U}^{n+1}) \\ &\approx \left[\mathbf{F}^+(\mathbf{U}^n) + \mathbf{A}^+(\mathbf{U}^n)\Delta\mathbf{U}^n\right] + \left[\mathbf{F}^-(\mathbf{U}^n) + \mathbf{A}^-(\mathbf{U}^n)\Delta\mathbf{U}^n\right] \end{aligned}$$

where, $\mathbf{A}^\pm$ are approximations to the split-flux jacobians $\frac{\partial \mathbf{F}^\pm}{\partial \mathbf{U}}$. In order to remove linearization errors, one can introduce Newton sub-iterations [11] in the variable p and the set of algebraic equations can be reduced to the form given by:

$$\Delta\mathbf{U}^p_{j-1}\left[-\tau\theta_{j-\frac{1}{2}}\mathbf{A}^+(\mathbf{U}^n_{j-1})\right] + \Delta\mathbf{U}^p_j\left[\mathbf{I} + \tau\theta_{j-\frac{1}{2}}\mathbf{A}^+(\mathbf{U}^p_j) - \tau\theta_{j+\frac{1}{2}}\mathbf{A}^-(\mathbf{U}^p_j)\right] + \Delta\mathbf{U}^p_{j+1}\left[\tau\theta_{j+\frac{1}{2}}\mathbf{A}^-(\mathbf{U}^n_{j+1})\right]$$

$$= -(\mathbf{U}^p_j - \mathbf{U}^n_j) - \tau\left[((1-\theta_{j+\frac{1}{2}})\mathbf{F}^n_{j+\frac{1}{2}} - (1-\theta_{j-\frac{1}{2}})\mathbf{F}^n_{j-\frac{1}{2}}) + (\theta_{j+\frac{1}{2}}\mathbf{F}^p_{j+\frac{1}{2}} - \theta_{j-\frac{1}{2}}\mathbf{F}^p_{j-\frac{1}{2}})\right]$$

with $\Delta\mathbf{U}^p = \mathbf{U}^{p+1} - \mathbf{U}^p$. Hence a block tridiagonal inversion is required at each sub-iteration p and $\mathbf{U}^{p+1}$ is updated for every sub-iteration as $\mathbf{U}^{p+1} = \mathbf{U}^p + \Delta\mathbf{U}^p$. The sub-iterations are continued till $||\Delta\mathbf{U}^p|| \rightarrow 0$ in a suitable norm, at which point, the RHS is identical to eqn. (4.8) (with $p = n+1$). Before the start of each sub-iteration, θ_j is determined for all points in the domain. LU-SGS method [19] has been used in the solution of the implicit system of equations. The L-DIRK2 can also be implemented in a similar manner.

Numerical Results

The application of the schemes to the 1-D Euler equations is demonstrated in this section. Both test cases are Riemann problems in a constant area tube. The left and right states are represented by the subscripts $_L$ and $_R$. The domain is [0,1] and the interface is at $x = 0.5$. The number of points in the domain is represented by N. Both these solutions do not involve application of boundary conditions since the final time is chosen such that none of the waves cross the computational domain. A second order upwind MUSCL [17] extrapolation is used

(with Super-bee limiter [12]) for spatial discretization and interfacial fluxes are computed using the AUSMDV [18] flux differencing scheme. The Euler-explicit TVD limit with a Super-bee limiter corresponds to $\sigma = \{|u| + a\}_{max} \frac{\Delta t}{\Delta x} < 0.5$, where u is the local velocity and a is the local sonic velocity.

Sod's problem: Sod's problem is given by $\{p_L, \rho_L, u_L\} = \{1.0,\ 1.0,\ 0.0\}$ and $\{p_R, \rho_R, u_R\} = \{0.1,\ 0.125,\ 0.0\}$. For this case, N=200 and $\sigma = 3.0$. Fig. 4.7 shows the density evolution. It is seen that both time-limited methods resolve the expansion wave well and the shock and contact discontinuity are captured with minimal smearing even at this high CFL number. Fig. 4.8 shows the evolution of pressure and reinforces the observations from the density plot.

Lax's problem: Lax's problem is given by $\{p_L, \rho_L, u_L\} = \{3.528,\ 0.445,\ 0.698\}$ and $\{p_R, \rho_R, u_R\} = \{0.571,\ 0.5,\ 0.0\}$. Figures 4.9 and 4.10 demonstrate the results for N=200, $\sigma = 3.0$.

From these results, it is seen that the L-TRAP and L-DIRK2 methods of time integration are beneficial in the sense that they generate less oscillatory numerical solutions. Some portions of the solution exhibit very small amplitude oscillations. These can be removed by making improvements on the limiter. Analytical expressions for the TVD limits of the L-TRAP scheme assuming a second order MUSCL spatial discretization hint that the local time-step restriction appears to be a function of the ratio of successive θ_j. Hence, a better limiter would correspond to a smoother distribution of θ_j over the domain.

4.4 Summary

A new class of time-limited implicit schemes have been introduced. The main concept behind these schemes is that the order of accuracy in time is dropped locally in regions where the time evolution of the solution is not smooth. Hence, these schemes are essentially *non-linear in time* and can thus circumvent the limits imposed on linear time integration schemes.

There is a lot to learn from monotonicity concepts that are used in spatial discretization in the sense that these concepts can be extended to time integration. A number of improvements can be made to the present schemes. The L-TRAP and L-DIRK2 schemes are general and there are many ways by which one could improve the design of θ_j (a monotone time interpolation approach is

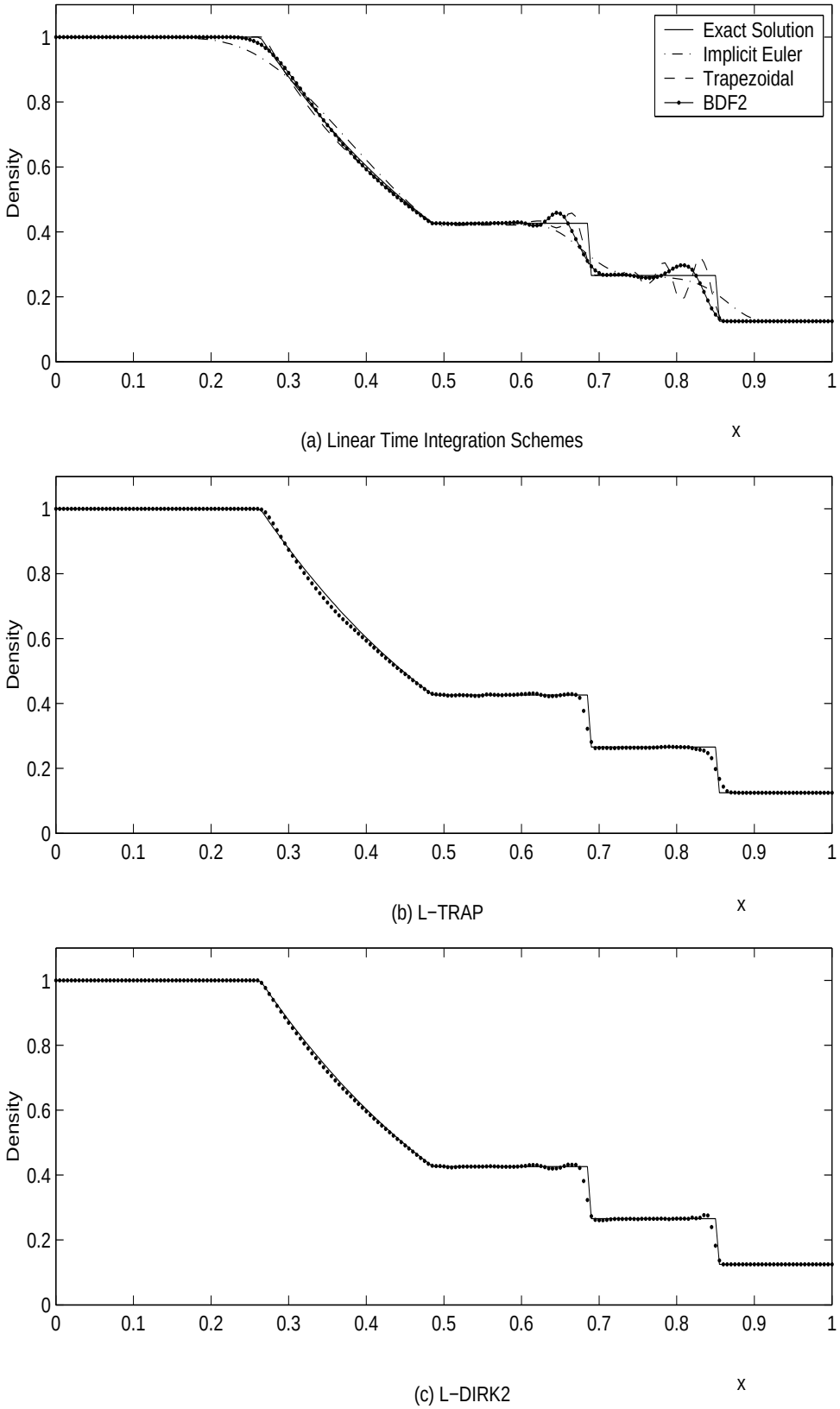


Figure 4.7: Sod's problem, $\sigma = 3.0$, 2nd order MUSCL, N=200, t=0.2

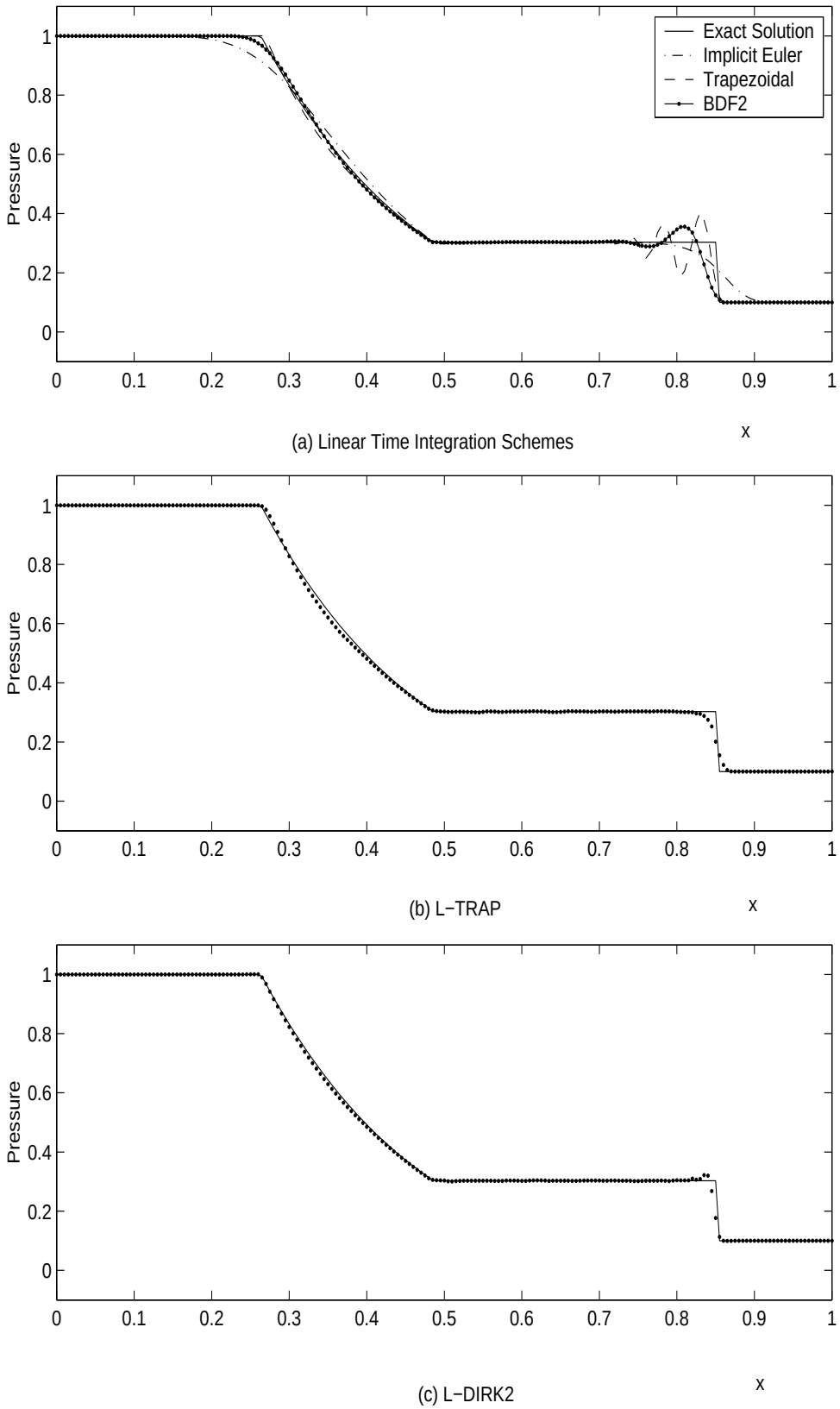


Figure 4.8: Sod's problem, $\sigma = 3.0$, 2nd order MUSCL, N=200, t=0.2

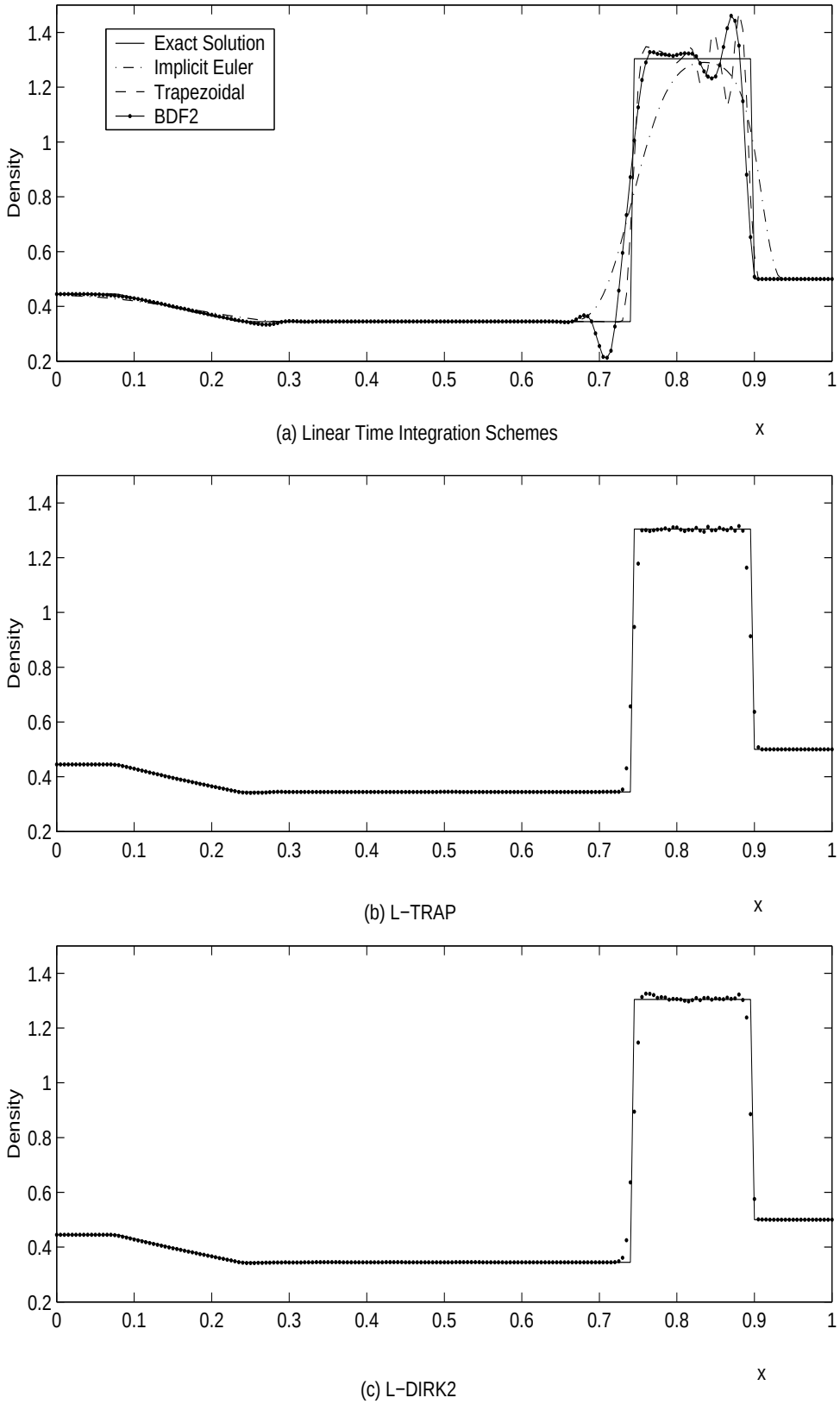
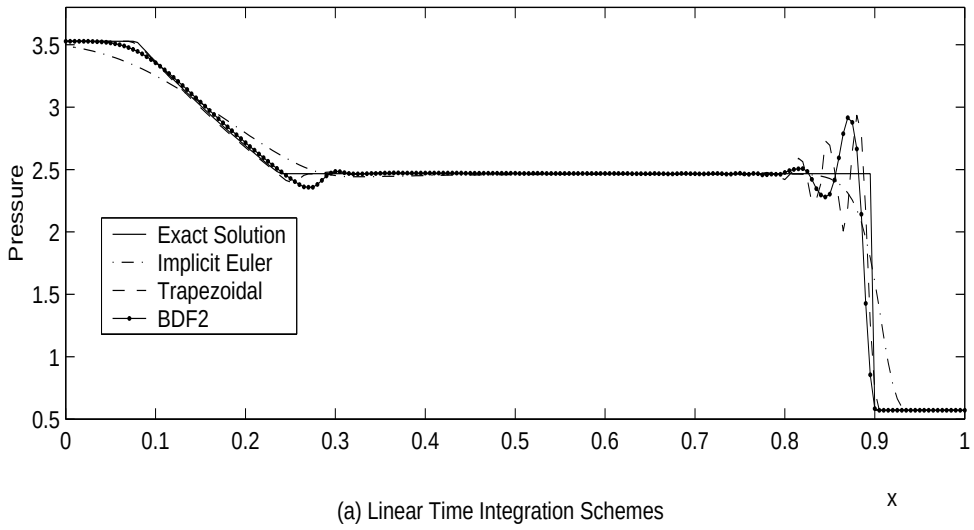
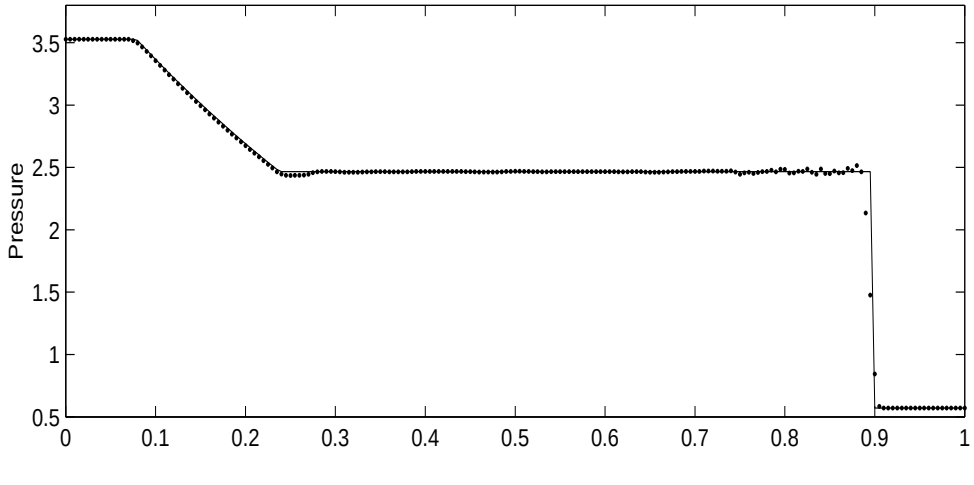


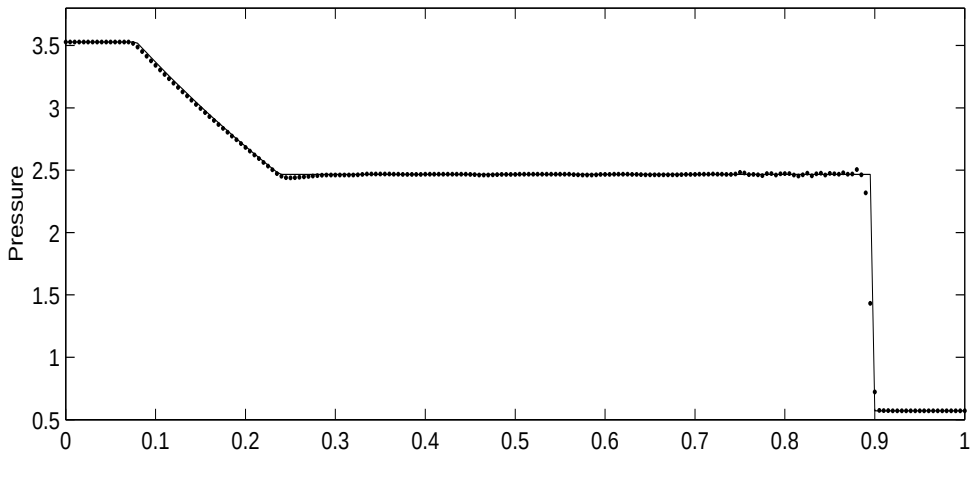
Figure 4.9: Lax's problem, $\sigma = 3.0$, 2nd order MUSCL, N=200, t=0.16



(a) Linear Time Integration Schemes



(b) L−TRAP



(c) L−DIRK2

Figure 4.10: Lax's problem, $\sigma = 3.0$, 2nd order MUSCL, N=200, t=0.16

used in this work) and $\theta_{j+\frac{1}{2}}$ (a simple average has been used).

Numerical results obtained from the solution of scalar and systems of conservation laws have demonstrated that such a concept can work. The solution is more accurate than first order methods and much less oscillatory compared to other second order methods. Note that the TVD property has not been proven for these methods and this explains the slight oscillations found in the solution.

Chapter 5

Limited Space-Time Reconstruction

5.1 Formulation

As mentioned in Chapter 2, to discretize eqn. 2.1, space is divided into cells $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ and time is divided into intervals $[t^n, t^{n+1}]$ as depicted in fig. 2.1. The *conservation form* can be written as:

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \tau(\hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}}), \tag{5.1}$$

where,

$$\bar{u}_j^n = \frac{1}{\Delta x} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u(x, t^n) dx, \quad and \tag{5.2}$$

$$\hat{f}_{j+\frac{1}{2}} = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} f(u(x_{j+\frac{1}{2}}), t) dt \tag{5.3}$$

with $\Delta x = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$, $\Delta t = t^{n+1} - t^n$ and $\tau = \frac{\Delta t}{\Delta x}$. By defining a suitable approximation to the flux integral (eqn. 5.3), numerical schemes can be constructed to any order of accuracy. A few well known approximations are shown in table 2.1. Higher order accurate schemes can be constructed using multiple stages (Runge-Kutta methods) or multiple steps (for instance, the Backward Difference methods). In this work, Godunov type MUSCL methods [17] are considered. (Refer Chapter 2 for further details).

For the new scheme, $\hat{f}_{j+\frac{1}{2}}$ is approximated by:

$$\begin{aligned}\hat{f}_{j+\frac{1}{2}} &= \frac{1}{\Delta t}\int_{t^n}^{t^{n+1}} f(u(x_{j+\frac{1}{2}}),t)dt \\ &= f(u(x_{j+\frac{1}{2}},t^{n+\frac{1}{2}})) + O(\Delta t^2)\end{aligned}$$

Second order accuracy is achieved by considering the mid-point in time. To define the interface values, a Taylor series approximation is used:

$$\begin{aligned}u(x,t) &= \bar{u}(x_j,t^{n+1}) + \frac{\partial u}{\partial x}(x_j,t^{n+1})(x-x_j) \\ &+ \frac{\partial u}{\partial t}(x_j,t^{n+1})(t-t^{n+1}) + O(\Delta x^2,\Delta t^2)\end{aligned}$$

Note that the Taylor series is centered about the new time-level t^{n+1} and hence the resulting numerical scheme would be implicit in time.

Similar approaches with series centered at time-level t^n abound in the literature (for instance, [21]). The usual practice is to approximate the time derivative using the governing equation. For instance,

$$\frac{\partial u}{\partial t}(x_j,t^n) = -\frac{\partial f}{\partial u}(x_j,t^n)\frac{\partial u}{\partial x}(x_j,t^n)$$

Very high order schemes can be constructed by using higher order terms in the Taylor series and replacing the time derivative terms by equivalent spatial derivatives.

The present scheme differs in the sense that it is implicit and the *time derivative is not approximated by spatial derivatives*. This allows a definition of the reconstructed value in terms of spatial and temporal differences. The interface value would then be defined as:

$$u_{j+\frac{1}{2}}^{n+\frac{1}{2}} = \bar{u}_j^{n+1} + \frac{(\Delta_x)_{j+1/2}^{n+1}}{2} - \frac{(\Delta_t)_j^{n+1/2}}{2},$$

where $(\Delta_x)_{j+1/2}^{n+1} = \bar{u}_{j+1}^{n+1} - \bar{u}_j^{n+1}$ and $(\Delta_t)_j^{n+1/2} = \bar{u}_j^{n+1} - \bar{u}_j^n$.

Fig. 5.1 shows results for the new method as applied to the linear convection equation $u_t + u_x = 0$ with $u(x,0) = \sin^4(x/2)$ and periodic boundary conditions. A CFL number $\nu = \Delta t/\Delta x = 3.5$ is used. The shaded profile is the exact evolution of the solution at the half time step. It is evident that the addition

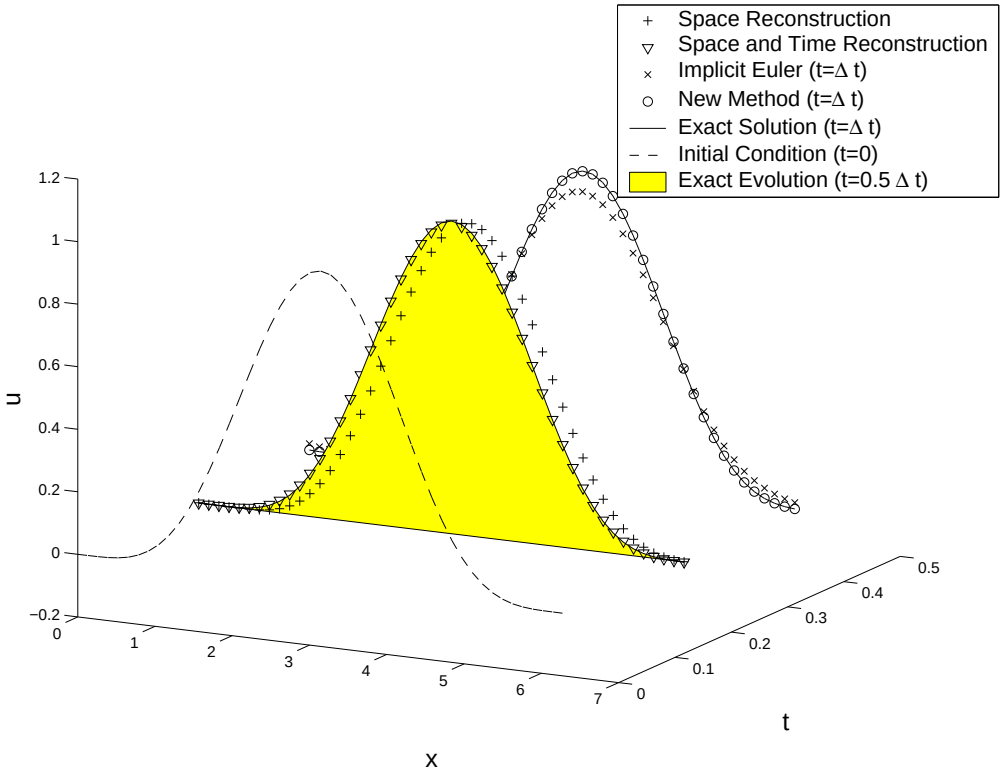


Figure 5.1: Linear advection equation with $\nu = 3.5$

of the time reconstruction greatly improves the final solution when compared to an implicit Euler method which can be viewed as using a purely spatial reconstruction.

However, this reconstruction is not guaranteed to be monotone in the presence of discontinuities and hence limiters are introduced as follows:

$$u_{j+\frac{1}{2}}^{n+\frac{1}{2}} = \bar{u}_j^{n+1} + \phi_j \frac{(\Delta_x)_{j+1/2}^{n+1}}{2} - \psi_j \frac{(\Delta_t)_j^{n+1/2}}{2},$$

The selection of appropriate limiters will be detailed in the following section.

5.2 TVD analysis

Consider the linear advection equation $u_t + au_x = 0$, with $a > 0$. From this point forward, the over-bar for cell-averaged quantities and superscripts for the spatial $(n+1)$ and temporal $(n+1/2)$ differences are dropped for clarity. For the new scheme,

$$\hat{f}_{j+\frac{1}{2}} = a[u_j^{n+1} + \phi_j \frac{(\Delta_x)_{j+1/2}}{2} - \psi_j \frac{(\Delta_t)_j}{2}] \tag{5.4}$$

Therefore,

$$\begin{aligned} u_j^{n+1} &= u_j^n - \tau(\hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}}) \\ &= u_j^n - \nu[(u_j^{n+1} - u_{j-1}^{n+1}) \\ &+ \frac{1}{2}(\frac{\phi_j}{r_j} - \phi_{j-1})(\Delta_x)_{j-1/2} \\ &- \frac{1}{2}(\psi_j - \frac{\psi_{j-1}}{s_{j-1}})(\Delta_t)_j] \end{aligned}$$

where, $r_j = \frac{(\Delta_x)_{j-1/2}}{(\Delta_x)_{j+1/2}}$, $s_{j-1} = \frac{(\Delta_t)_j}{(\Delta_t)_{j-1}}$ and $\nu = a\frac{\Delta t}{\Delta x}$. Finally, this gives:

$$u_j^{n+1} = u_j^n - \nu \left[\frac{1 + \frac{1}{2}(\frac{\phi_j}{r_j} - \phi_{j-1})}{1 + \frac{\nu}{2}(\frac{\psi_{j-1}}{s_{j-1}} - \psi_j)} \right] (\Delta_x)_{j-1/2} \tag{5.5}$$

using the Lemma concerning implicit operators from Harten (refer Chapter 2), the above scheme is TVD if

$$\frac{1 + \frac{1}{2}(\frac{\phi_j}{r_j} - \phi_{j-1})}{1 + \frac{\nu}{2}(\frac{\psi_{j-1}}{s_{j-1}} - \psi_j)} \geq 0 \tag{5.6}$$

For $a < 0$, a similar expression involving $j+1$ is obtained. For the non-linear case $u_t + f(u)_x = 0$, a similar condition involving a local CFL-like number is obtained. Both these expressions and the corresponding TVD conditions are provided in the appendix C.

Notice that the numerator and denominator represent spatial and temporal reconstruction respectively and hence both should be individually positive. For the spatial part, any standard flux limiter can be used, but in this work, the following is used:

$$\phi(r) = \begin{cases} r & if \quad |r| <= 1 \\ 1 & if \quad |r| > 1 \end{cases} \tag{5.7}$$

This is equivalent to second order accurate ENO reconstruction. Note that it is uniformly second order accurate *and* strictly TVD[1]. It is apparent that the temporal terms (denominator) are similar to the spatial terms except for the scaling by the CFL number - hence, in order to satisfy the TVD conditions, the time limiter can be defined as:

$$\psi(s) = \begin{cases} s/\nu & if \quad |s| <= 1 \quad and \quad \nu > 1 \\ 1/\nu & if \quad |s| > 1 \quad and \quad \nu > 1 \\ s & if \quad |s| <= 1 \quad and \quad \nu <= 1 \\ 1 & if \quad |s| > 1 \quad and \quad \nu <= 1 \end{cases} \tag{5.8}$$

Equations 5.7 and 5.8 are sufficient to guarantee an *unconditionally TVD scheme*. It is seen that from eqn. 5.8 that for (local) $\nu > 1$, the TVD region has shrunk because of the scaling and this results in a local drop in temporal accuracy even when the solution is smooth in space and time ($s \approx 1$). This can be overcome by taking advantage of the implicitness of the scheme. From the denominator of eqn. 5.6,

$$\psi_j <= \frac{2}{\nu} + \frac{\psi_{j-1}}{s_{j-1}}$$

In implicit schemes, the system of equations are solved in an approximate linearized form at each time-step, but to restore time accuracy, the so-called *sub-iterations* have to be performed. Refer [11] for a review of the approach. In the current method, the first sub-iteration is done assuming $\psi_j = 0$ throughout the domain. After this, for any sub-iteration $p+1$, ψ can be set using information from p:

$$\psi_j^{p+1} = min\left[1, max\left[\frac{2}{\nu} + \frac{\psi_{j-1}^p}{s_{j-1}^p}, 0\right]\right] \tag{5.9}$$

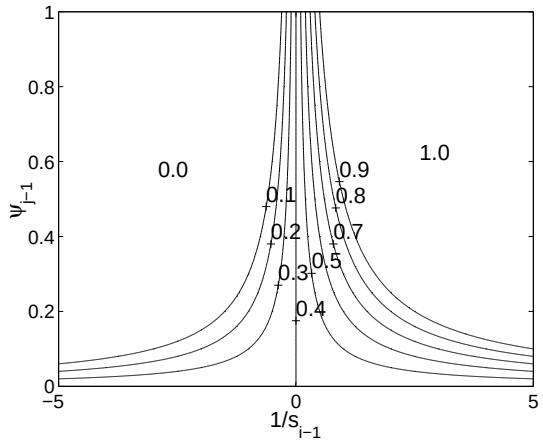


Figure 5.2: ψ_j as a function of ψ_{j-1} and s_{j-1} for $\nu = 5$. For higher ν, there is more limiting in the left half plane

This restores accuracy (refer fig. 5.1) when the solution is smooth. To an extent, this also ensures smooth variation of the time limiter between grid-points. In practice, this marginally degrades the convergence of sub-iterations. Equations 5.7 and 5.9 define the limiters that will be used in this work for space and time reconstruction.

5.3 Numerical Results

Numerical results are presented for the inviscid Burger's equation and the Euler equations in one and two dimensions. For each of the above cases, Newton-type sub-iterations are used to solve the implicit set of equations at each time step. The new scheme is termed the *implicit second order STR* (Space-Time Reconstruction). For comparative purposes, all linear time integration schemes are used with the second order ENO spatial discretization. For this spatial scheme, explicit time integration schemes will be non-oscillatory up to a maximum CFL number $\sigma = max_j\{\nu_j\} = 0.5$.

5.3.1 Burger's Equation

The first test case is the inviscid Burger's equation $u_t + (\frac{u^2}{2})_x = 0$, with periodic boundary conditions and a domain $[0, 2\pi]$ of 100 equally spaced points. The initial condition is comprised of an expansion wave and a compression wave.

This profile is convected to a time $t = 2.0$, before which the compression wave becomes a shock.

For a time-step corresponding to a CFL number $\sigma = max_j\{u_j\frac{\Delta t}{\Delta x}\} = 4.0$, fig. 5.3 shows that the implicit Euler method shows large dissipation and the linear second order time integration schemes develop oscillations in the vicinity of the shock. The trapezoidal method performs particularly poorly since it has little damping for high frequency dispersive errors. It is seen that the new scheme (fig. 5.4) is non-oscillatory. As expected limiter-1 (eqn. 5.8) is very dissipative and is only marginally better than the implicit Euler method. Limiter-2 (eqn. 5.9) is seen to resolve the shock and the expansion wave well and will be used in the rest of the numerical computations. Figures 5.5 and 5.6 assess the performance of the scheme up to $\sigma = 10.0$. It is evident that the new method remains non-oscillatory and reasonably accurate for the entire range of time-step sizes.

5.3.2 1D Euler Equations

For the 1D Euler equations, space and time limiters can be applied to each scalar equation $i = 1..3$ at each point $j = 1..jmax$ in the domain. The reconstruction may be based on primitive, conservative or characteristic variables. However, if primitive or conservative variables are used, the time limiter is somewhat arbitrary and ν_j will have to be based on the maximum eigen value $max\{|\lambda_i|)\}_j$. Based on numerical computations, this approach still seems to work if the time-limiter (ψ_i) for a given cell is restricted to the minimum of the three computed limiters, resulting in increased dissipation. If characteristic variables are used for reconstruction, $(\nu_i)_j$ can be based on $(|\lambda_i|)_j$, and hence different ψ_i can be used. Computations prove that this is indeed much more accurate especially in treating contact discontinuities, which being linear, are more susceptible to dissipation. Reconstruction using characteristic variables involves more work than using primitive or conservative variables. It is also known ([21], [15]) that in many cases (including Lax's problem), even high order explicit schemes operating at very small CFL numbers require characteristic variables for non-oscillatory reconstruction. This approach will be used to obtain the following numerical results. Once the reconstruction is accomplished, Roe's approximate Riemann solver [22] is used to calculate the interfacial fluxes.

Sod's problem: Sod's problem for an infinite length constant area duct is defined by the initial left and right states given by $\{p_L, \rho_L, u_L\} = \{1.0, 1.0, 0.0\}$ and $\{p_R, \rho_R, u_R\} = \{0.1, 0.125, 0.0\}$. Figures 5.7-5.10 confirm that the scheme is less diffusive than the first order method and stable compared to the second order methods.

Lax's problem: Lax's problem is defined by $\{p_L, \rho_L, u_L\} = \{3.528, 0.445, 0.698\}$ and $\{p_R, \rho_R, u_R\} = \{0.571, 0.5, 0.0\}$. This problem is slightly more difficult than the Sod's problem in that the density is not monotone decreasing and hence the contact discontinuity and the shock are more difficult to capture without spurious oscillations. As mentioned earlier, reconstructions based on primitive and conserved variables yield a fair amount of noise at high CFL numbers. Figures 5.11 and 5.12 outline some results for this problem.

5.3.3 2D Euler Equations

The two dimensional Euler equations are given by:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} + \frac{\partial \mathbf{G}}{\partial x} = 0$$

where,

$$\mathbf{U} = \left\{ \begin{array}{c} \rho \\ \rho u \\ \rho v \\ e \end{array} \right\}, \quad \mathbf{F} = \left\{ \begin{array}{c} \rho u \\ p + \rho u^2 \\ \rho uv \\ (e+p)u \end{array} \right\}, \quad \mathbf{G} = \left\{ \begin{array}{c} \rho v \\ \rho uv \\ p + \rho v^2 \\ (e+p)v \end{array} \right\},$$

with $e = p/(\gamma - 1) + \rho(u^2 + v^2)/2$. The test case involves convection of an isentropic vortex in a uniform grid ($0 <= x <= 10, -5 <= y <= 5$) that is periodic in both directions. Perturbations are added to the free-stream in such a way that there is no entropy gradient in the flow-field. Free-stream conditions are ($\rho = 1, u = 0.5, v = 0, p = 1/\gamma$). The perturbations are given by:

$$\begin{aligned} (\delta u, \delta v) &= \frac{\beta}{2\pi} e^{\frac{1-r^2}{2}} (-(y - y_o), (x - x_o)) \\ \rho &= \left[1 - \frac{(\gamma - 1)\beta^2}{8\gamma\pi} e^{1-r^2} \right]^{\frac{1}{\gamma - 1}} \\ p &= \frac{\rho^\gamma}{\gamma} \end{aligned}$$

where, $\beta = 5$ is the vortex strength and r is the distance from the vortex origin (5,0). As a result of the isentropicity, the exact solution corresponds to a pure advection of the vortex at the free-stream speed. Further details can be found in [23]. This test case is used to assess the performance of the new scheme in comparison to existing time integration methods. Primitive variable-based reconstruction is applied in space as well as time. The parameter ν in the definition of the time-limiter (refer appendix C) is taken to be $\nu_x = \frac{\Delta t}{\Delta x}(|u| + a)$ and $\nu_y = \frac{\Delta t}{\Delta y}(|v| + a)$ in the x and y directions respectively.

Figure 5.13 shows the pressure contours after 1 period of revolution (t=20) on a 81×81 grid. The time step ($\Delta t = 0.5$) roughly corresponds to a maximum CFL number of 8.0 in the x-direction. For all methods, 8 sub-iterations, which resulted in 3-4 orders of magnitude convergence were used. Fig. 5.14 shows the swirl velocity on a line passing through the center of the vortex. The implicit Euler scheme is seen to be highly inaccurate and the BDF-2 is less dissipative, but shows a noticeable phase error and a slight overshoot. Fig. 5.15 shows the performance of the new scheme with and without *time* limiting. The unlimited version, though less dissipative, still generates a slight undershoot. Application of the time limiter reduces the peak-to-peak velocity, but removes the overshoot. Figs. 5.16, 5.17 compare the pressure at the same locations.

Figs. 5.18 and 5.19 compare the schemes for a *stiffer* case. Since the time-step Δt is large, it can be expected to dominate the total error. The computation was carried out on a mesh that is twice as fine in each spatial direction. The BDF-2 is seen to generate a more severe over-shoot compared to the coarse grid. The limited version of the new scheme is seen to be more dissipative, but does not generate spurious overshoots or undershoots. It is also evident that the new scheme exhibits a smaller phase error compared to the BDF-2. It has to be mentioned that one of the contributing factors to the poor performance of the BDF-2 could be a result of starting the computations with two-implicit Euler time-steps.

Figure 5.20 shows the evolution for 4 periods at a much smaller time step ($\Delta t = 0.05$ with 5 sub-iterations) on the coarse grid. These experiments were intended to evaluate the concept of space-time reconstruction in 2 dimensions applied to relatively smooth solutions.

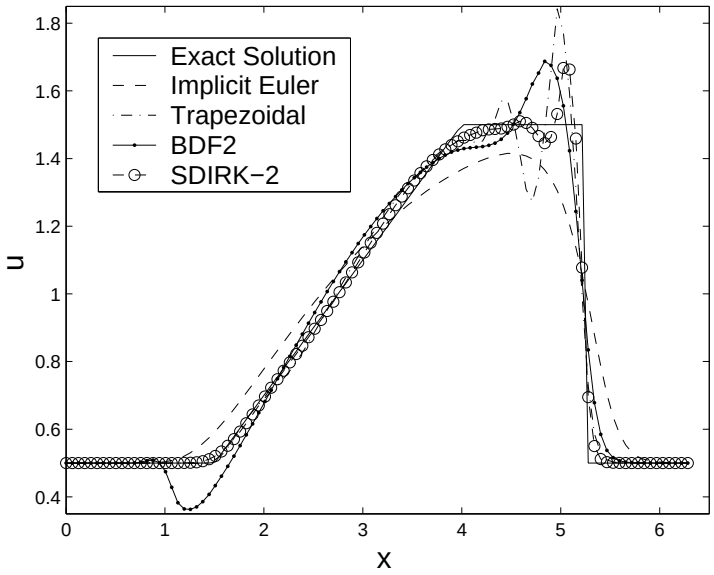


Figure 5.3: Burger's Equation with $\sigma = 4.0$, N=100 and $T_{final} = 2.0$

5.4 Summary

A new single step second order implicit scheme that uses reconstruction in space and time has been presented. The reconstruction is made non-oscillatory by the introduction of spatial and temporal limiters that ensure satisfaction of the TVD conditions. Numerical results on model scalar and vector hyperbolic conservation equations suggest that the scheme remains non-oscillatory and reasonably accurate over a large range of time-step sizes.

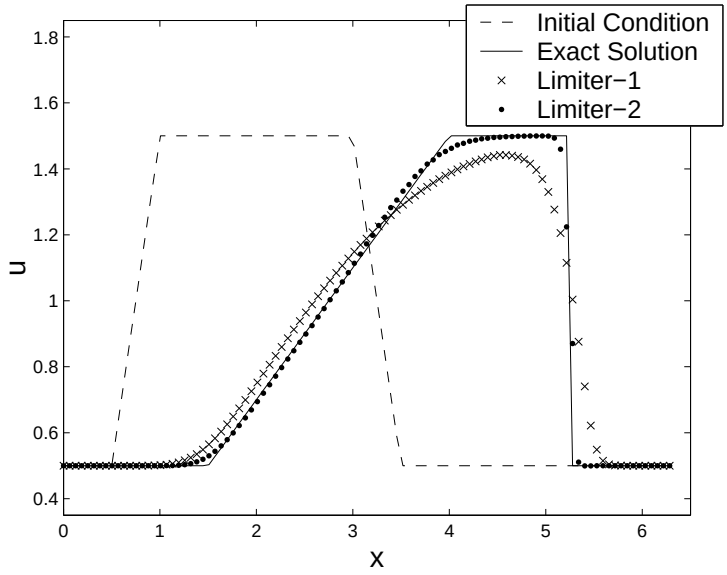


Figure 5.4: Burger's Equation with $\sigma = 4.0$, N=100 and $T_{final} = 2.0$

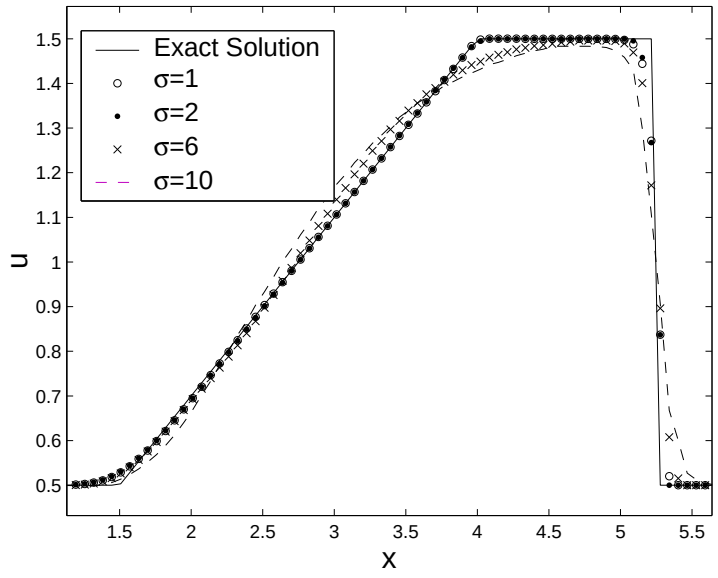


Figure 5.5: Performance of 2nd Order STR for various CFL numbers. N=100 and $T_{final} = 2.0$

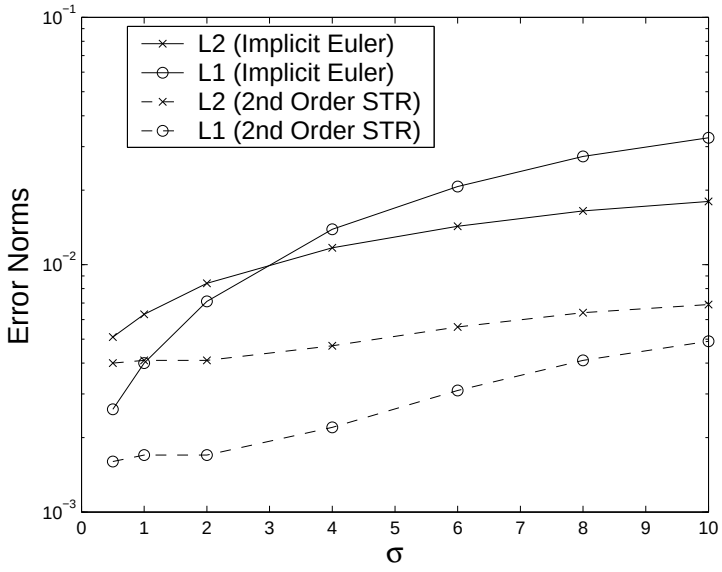


Figure 5.6: L_1 and L_2 norms for Burger's Equation. N=100 and $T_{final} = 2.0$

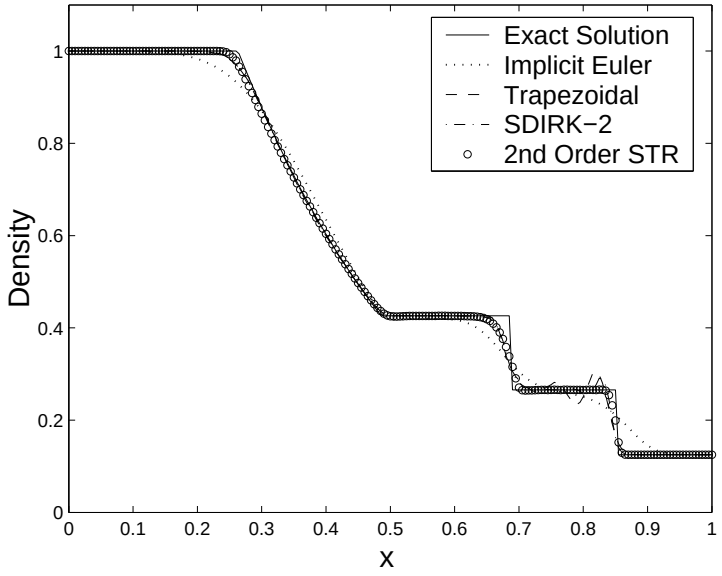


Figure 5.7: Sod's Problem with $\sigma = 4.0$, N=200 and $T_{final} = 0.2$

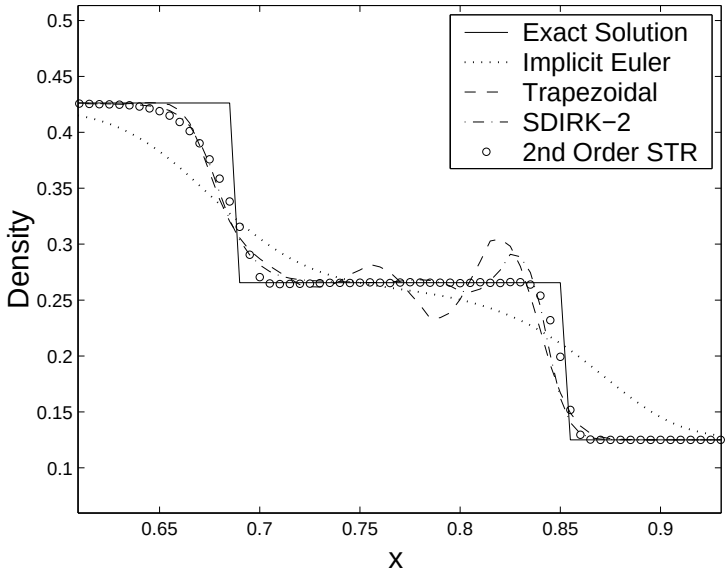


Figure 5.8: Zoomed up version of fig. 5.7

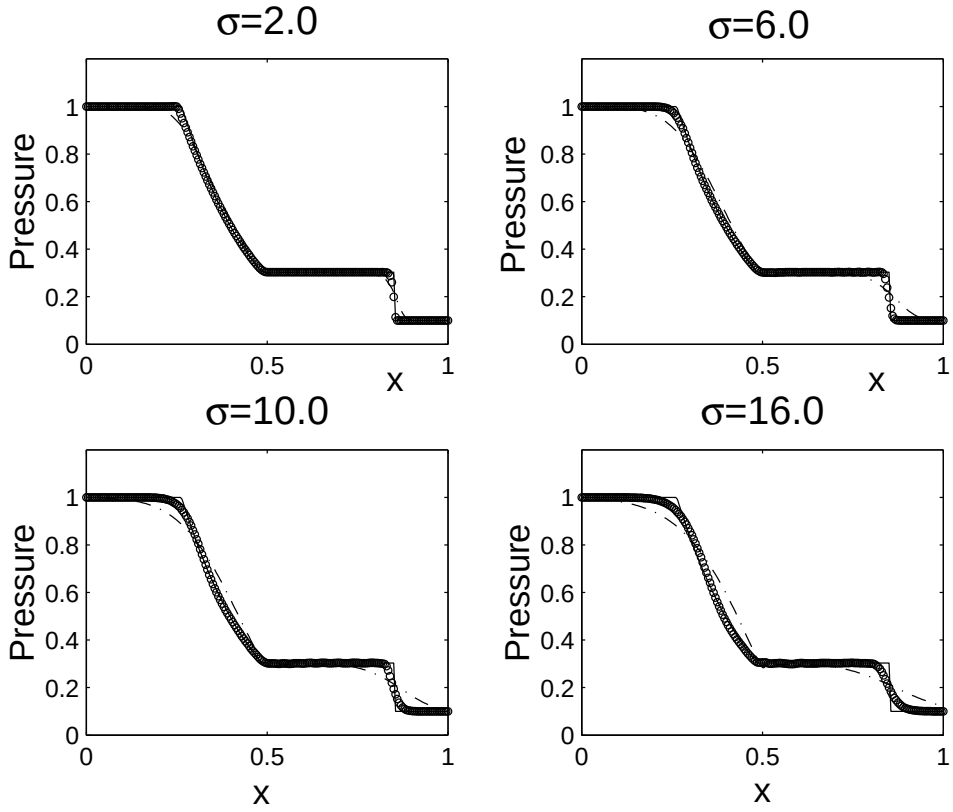


Figure 5.9: Sod's Problem (N=200): Performance of 2nd Order STR (bubbles) and Implicit Euler (broken lines) for various CFL numbers

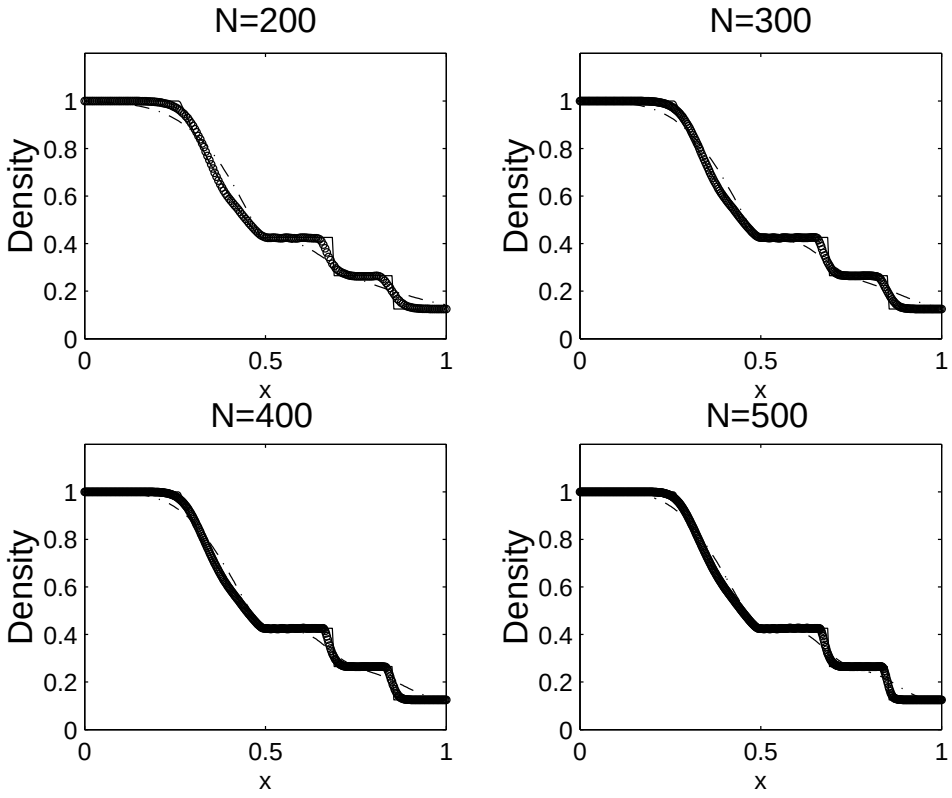


Figure 5.10: Sod's Problem ($\sigma = 16.0$): Performance of 2nd Order STR (bubbles) and Implicit Euler (broken lines) with mesh refinement

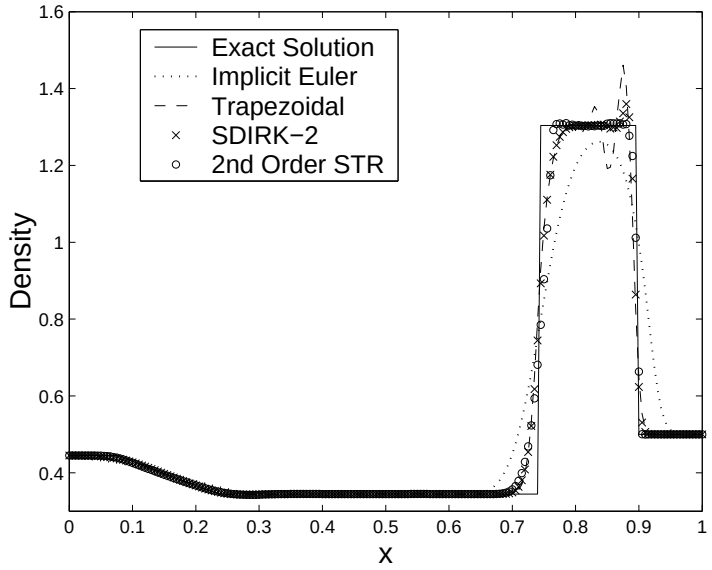


Figure 5.11: Lax's Problem with $\sigma = 4.0$, N=200 and $T_{final} = 0.16$

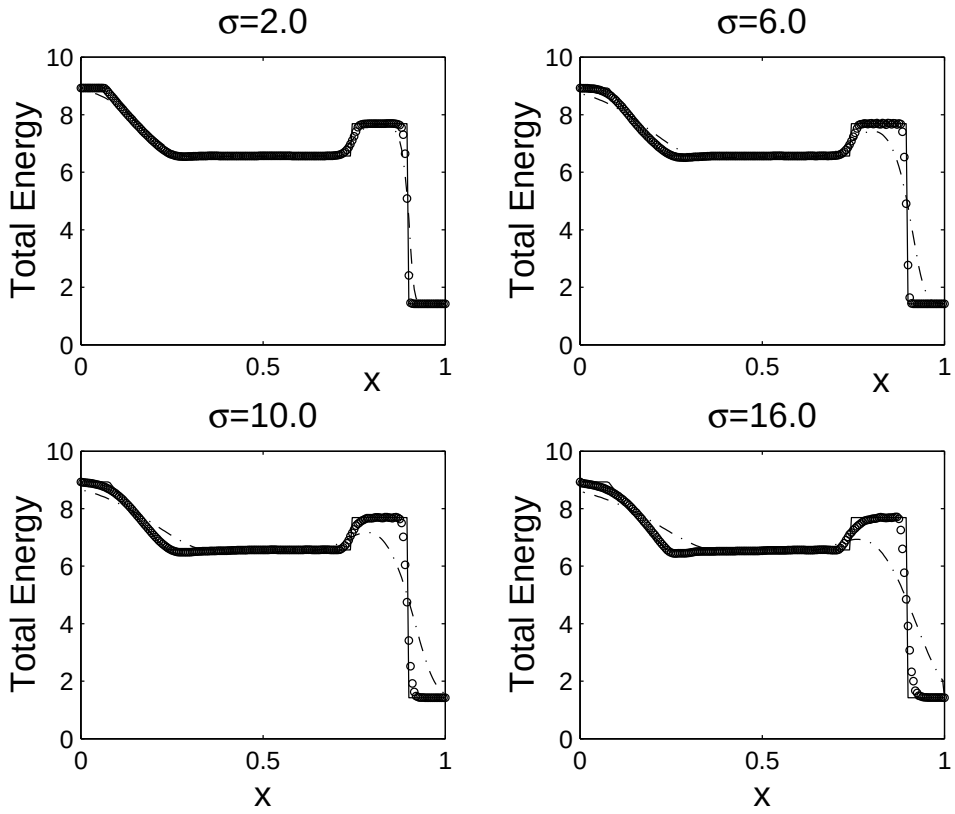


Figure 5.12: Lax's Problem (N=200): Performance of 2nd Order STR (bubbles) and Implicit Euler (broken lines) for various CFL numbers

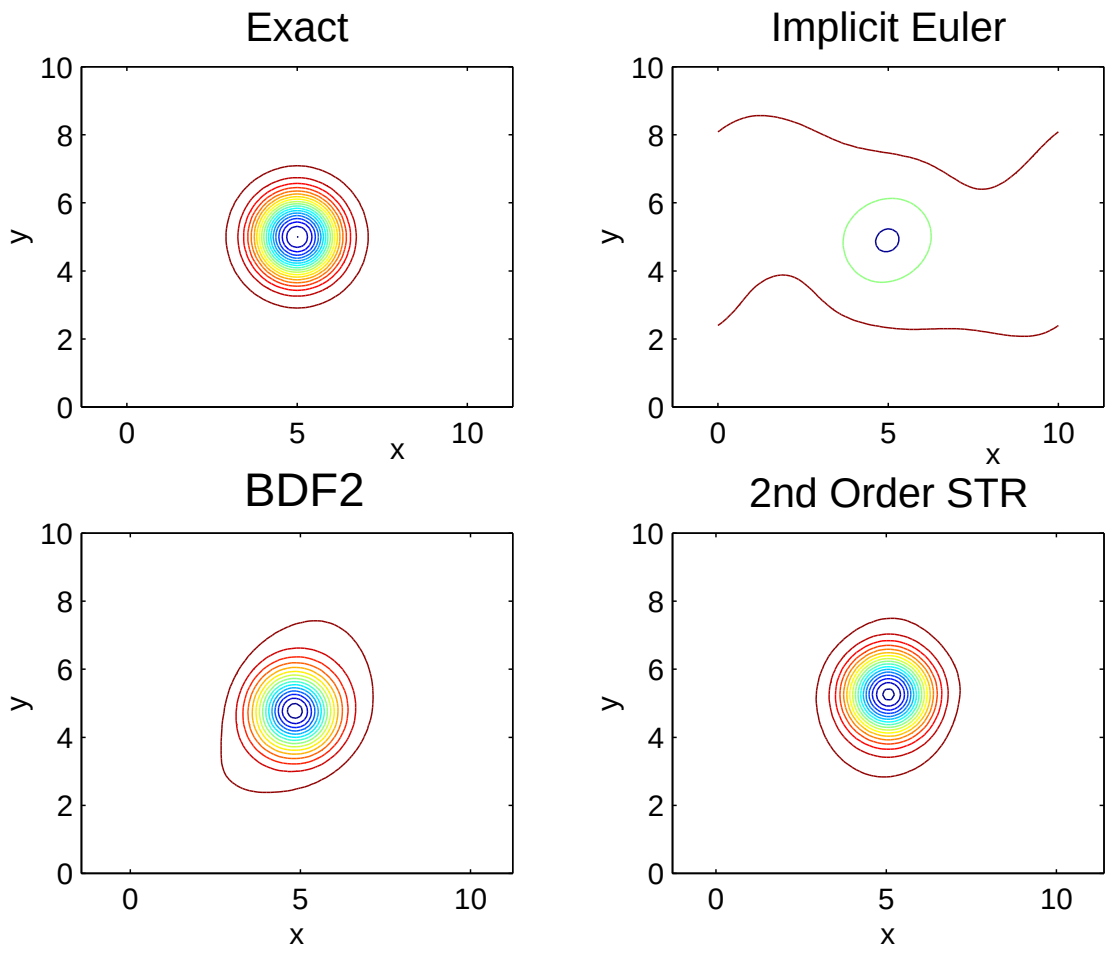


Figure 5.13: 2D Vortex Convection (N=81×81): $\Delta t = 0.5$, 1 period (40 time-steps)

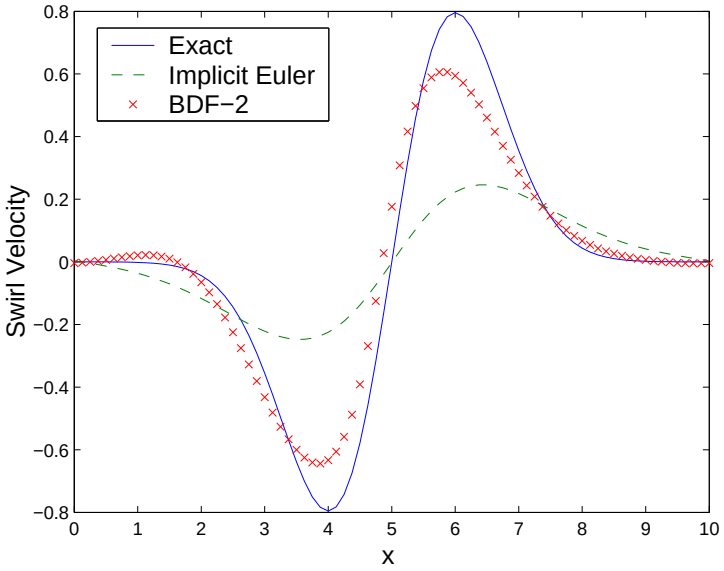


Figure 5.14: 2D Vortex Convection (N=81×81): $\Delta t = 0.5$, 1 period (40 time-steps)

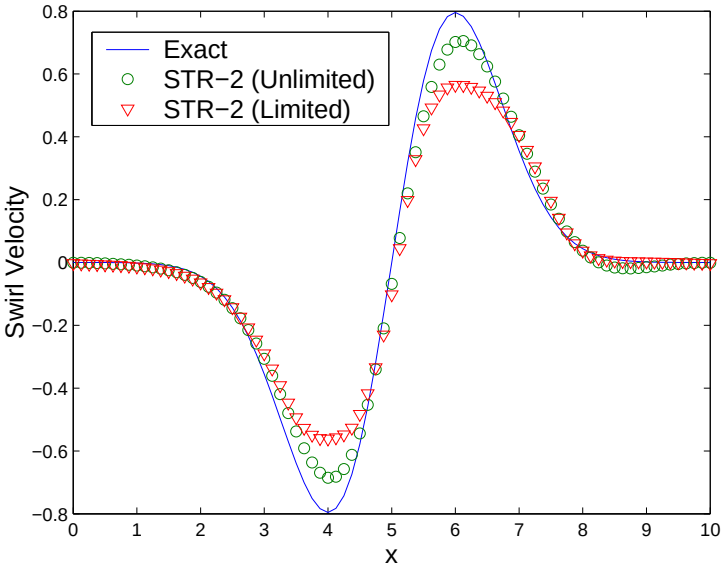


Figure 5.15: 2D Vortex Convection (N=81×81): $\Delta t = 0.5$, 1 period (40 time-steps)

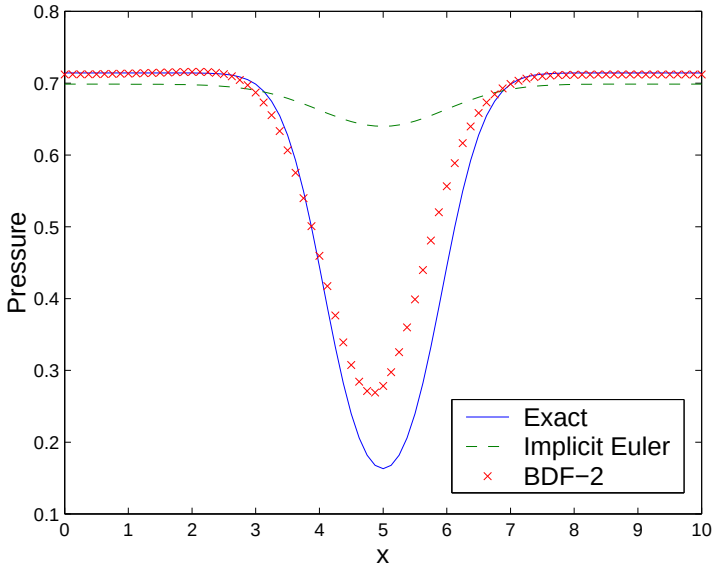


Figure 5.16: 2D Vortex Convection (N=81×81): $\Delta t = 0.5$, 1 period (40 time-steps)

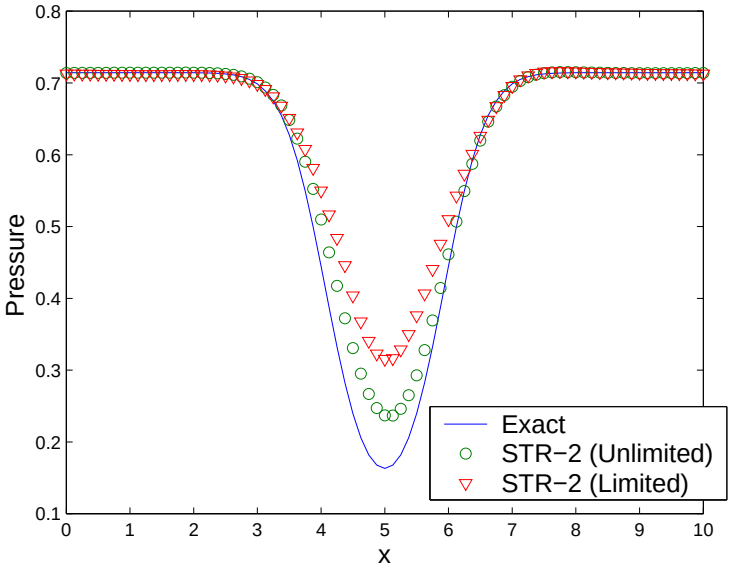


Figure 5.17: 2D Vortex Convection (N=81×81): $\Delta t = 0.5$, 1 period (40 time-steps)

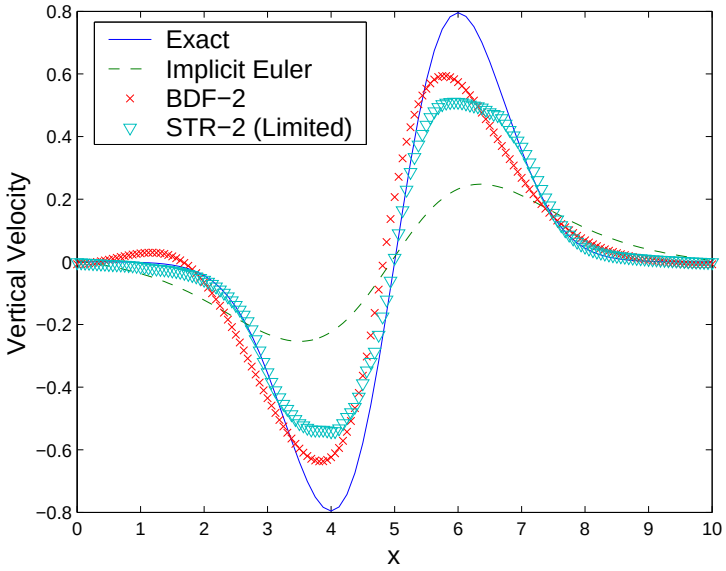


Figure 5.18: 2D Vortex Convection (N=161×161): $\Delta t = 0.5$, 1 period (40 time-steps)

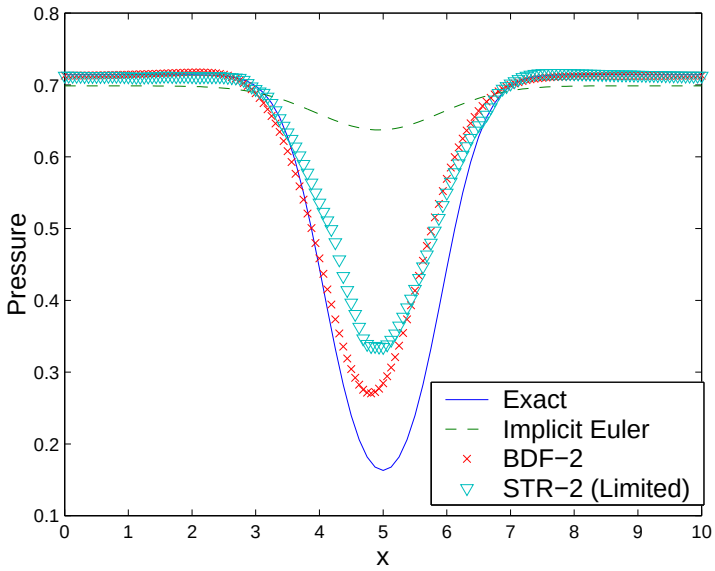


Figure 5.19: 2D Vortex Convection (N=161×161): $\Delta t = 0.5$, 1 period (40 time-steps)

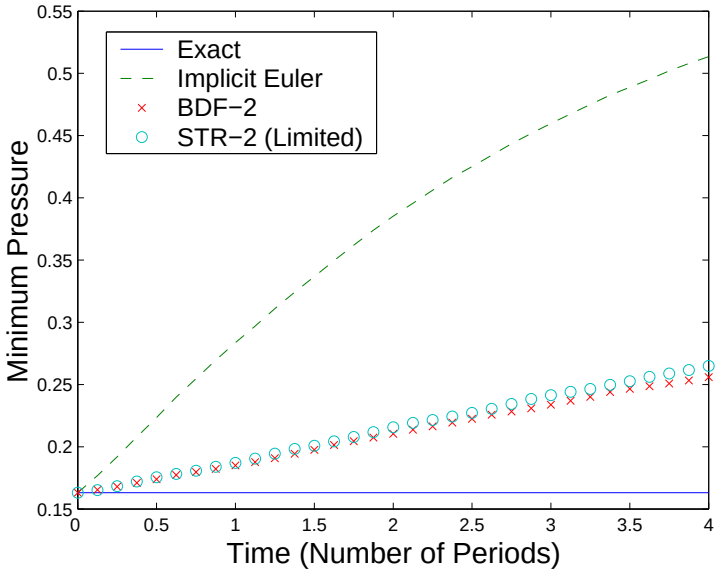


Figure 5.20: 2D Vortex Convection (N=81×81): $\Delta t = 0.05$, 1600 time-steps. Evolution of Pressure at Vortex Center

Chapter 6

Summary and Conclusions

High order numerical schemes for the solution of Hyperbolic conservation equations suffer from time-step restrictions resulting from monotonicity requirements. These restrictions make conventional high order implicit schemes impractical since the allowable time-steps are not significantly greater than that for explicit schemes. In this thesis, two novel high order implicit schemes were developed in an attempt to overcome these restrictions.

In Godunov-based numerical schemes, spurious oscillations can be avoided by properly biasing the information in the stencil used in spatial reconstruction of the solution at the cell-interface. Thus, the high-order barrier facing linear spatial reconstruction is avoided. In this work, this concept is extended to temporal discretizations, in that the temporal *stencil* is also made non-linear.

First, *Time-limited* schemes are developed in a method-of-lines framework. In this approach, any non-oscillatory spatial discretization can be used along with the temporal scheme. The time discretization is performed using a convex combination of a first order and a higher order scheme. The coefficients of the combination are modified such that the scheme works near the higher order (in time) limit in smooth regions of the solution and drops to lower order near the discontinuous regions. Preliminary numerical results on model conservation equations show improvements over conventional implicit schemes in that the solution is more accurate than first order methods and much less oscillatory compared to higher order methods for moderate time step sizes.

Second, a new single step second order implicit scheme that uses *reconstruction in space and time* has been developed. Conventionally, to achieve high order accuracy in single step methods, temporal derivatives are replaced by equivalent

spatial derivatives. In the present approach, the temporal derivative is retained. The resulting space-time reconstruction is made non-oscillatory by the introduction of limiters that satisfy TVD conditions. Numerical experiments on model problems show accurate non-oscillatory solutions over a large range of time-steps. The current scheme is at best second order in space and time. Future work can involve increasing the order of accuracy by possibly having more quadrature points in the time interval.

The ultimate objective is to design an *accurate* scheme that achieves *unconditional non-linear stability*. A complete approach towards satisfying this involves the following components:

• The temporally unlimited method should be as accurate as existing analogous implicit schemes for smooth solutions.

• The temporally unlimited method should be as accurate as existing analogous implicit schemes operating below their non-linear stability limit for discontinuous solutions.

• The limited method should be unconditionally stable in a non-linear sense and it should achieve the same without being too diffusive.

• The limiter should be able to preserve accuracy in smooth regions.

• The cost per time-step should be comparable to efficient linear time integration schemes.

• In relevant cases, it should be more efficient than explicit schemes in achieving a desired level of accuracy.

Preliminary numerical results on model scalar and vector hyperbolic conservation equations suggest that the new schemes hold promise in achieving some of the above mentioned requirements. It has to be mentioned that further research is required before these schemes can be efficiently applied to actual applications in computational physics. This work serves just as an introduction to the concepts of temporal limiting and reconstructions. Further, the TVD condition is known to be very stringent and rigorous satisfaction of the same usually makes it difficult to design highly accurate schemes. Devising efficient schemes could look at alternate monotonicity conditions that are possibly less restrictive.

Bibliography

[1] E. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics, Second Edition*, Springer-Verlag, (1999).

[2] S. Gottlieb and C.-W. Shu, *Total Variation Diminishing Runge-Kutta Schemes*, Mathematics of Computation, 67 (1998).

[3] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations*, vol. 2, Springer-Verlag, (1991).

[4] A. Harten, J. Hyman, and P. Lax, *On Finite Difference Approximations and Entropy Conditions for Shocks*, Communications on Pure and Applied Mathematics, 29 (1976).

[5] A. Harten, *High Resolution Schemes for Hyperbolic Conservation Laws*, Journal of Computational Physics, 49 (1983).

[6] P. Colella and P. Woodward, *The Piecewise Parabolic Method (PPM) for Gas-Dynamical Simulations*, Journal of Computational Physics, 54 (1984).

[7] A. Harten, *On a Class of Total Variation Stable Finite Difference Schemes*, SIAM Journal of Numerical Analysis, 21 (1984).

[8] A. Harten and S. Osher, *Uniformly High-order Accurate Non-oscillatory Schemes 1*, SIAM Journal of Numerical Analysis, 24 (1987).

[9] H. Hyunh, *Accurate Monotone Cubic Interpolation*, SIAM Journal of Numerical Analysis, 30 (1993).

[10] C. B. Laney, *Computational Gasdynamics*, Cambridge University Press, (1998).

[11] T. Pulliam, *Time Accuracy and the use of Implicit Methods*, AIAA Paper 93-3360 (1993).

[12] P. Roe, *Some Contributions to the Modelling of Discontinuous Flows*, in Lectures in Applied Mathematics, 22 (1985).

[13] B. Cockburn, C. Johnson, C.-W. Shu and E. Tadmor, *Essentially Non-oscillatory and Weighted Essentially Non-oscillatory Schemes for Hyperbolic Conservation Laws in Advanced Numerical Approximation of Nonlinear Hyperbolic Equations*, Lecture Notes in Mathematics, 1697 (1998).

[14] S. Gottlieb, C.-W. Shu and E. Tadmor, *Strong Stability Preserving High-order Time Discretization Methods*, SIAM Review, 43 (2001).

[15] A. Suresh and H. Huynh, *Accurate Monotonicity-preserving Schemes with Runge-Kutta Time Stepping*, Journal of Computational Physics, 136 (1997).

[16] B. Van Leer, *Towards the Ultimate Conservative Difference Scheme II. Monotonicity and Conservation combined in a Second Order Scheme*, Journal of Computational Physics, 14 (1974).

[17] B. Van Leer, *Towards the Ultimate Conservative Difference Scheme V. a Second Order Sequel to Godunov's Method*, Journal of Computational Physics, 32 (1979).

[18] W. Wada and M.-S. Liou, *An Accurate and Robust Flux Splitting Scheme for Shock and Contact Discontinuities*, SIAM Journal of Scientific Computing, 18 (1997).

[19] S. Yoon and A. Jameson, *Lower-Upper Symmetric-Gauss-Seidel Method for the Euler and Navier-stokes Equations*, AIAA Journal, 26 (1988).

[20] S. Forth, *A Second Order Accurate, Space-Time Limited BDF Scheme for the Linear Advection Equation*, Godunov Methods: Theory and Applications, Kluwer Academic/Plenum Publishers, (2001).

[21] A. Harten, B. Engquist, S. Osher and S. Chakravarthy, *Uniformly High Order Accurate Essentially Non-oscillatory schemes III*, Journal of Computational Physics, 71 (1987).

[22] P.Roe, *Approximate Riemann Solvers, Parameter Vectors and Difference Schemes*, Journal of Computational Physics, 43 (1981).

[23] H-C. Yee, N. Sandham and M. Djomehri, *Low Dissipative High Order Shock-Capturing Methods using Characteristic-Based Filters*, Journal of Computational Physics, 150 (1999).

[24] D. Goldberg and B. Wylie, *Characteristic Method using Time-line Interpolations*, Journal of Hydraulics Engineering, 109 (1983).

[25] B. Leonard, *Note on the Von Neumann Stability of Explicit One-dimensional Advection Schemes*, Computational Methods in Applied Mechanics and Engineering, 118 (1994).

[26] P. Roache,*A Flux-based Modified Method of Characteristics*, International Journal of Numerical Methods in Fluids, 15 (1992).

[27] R. LeVeque, *Large Time Step Shock-capturing Techniques for Scalar Conservation Laws*, SIAM Journal of Numerical Analysis, 19 (1982).

[28] V. Guinot, *The Time-Line Interpolation Method for Large-Time-Step Godunov-Type Schemes,* Journal of Computational Physics, 177 (2002).

[29] X-D Liu, S. Osher and T. Chan, *Weighted Essentially Non-oscillatory Schemes*, Journal of Computational Physics, 115 (1994).

Appendix A

TVD limits of some implicit schemes

We consider expressing TVD limits for implicit schemes as a ratio of the explicit Euler TVD limit. This closely follows the approach in [2]. Consider an s-stage implicit Runge-Kutta (RK) scheme for eqn 2.6:

$$\begin{aligned} u^{(i)} &= u^n + \Delta t \sum_{j=1}^{i} a_{ij} L(u^{(j)}) \quad i = 1..s \\ u^{n+1} &= u^n + \sum_{i=1}^{s} b_i L(u^{(i)}) \end{aligned}$$

This is represented in the Butcher array form as :

c_1	a_{11}	0	0	.	.	0
c_2	a_{21}	a_{22}	0	.	.	0
.	.	.	.	.	.	.
.	.	.	.	.	.	.
c_s	a_{s1}	a_{s2}	a_{s3}	.	.	a_{ss}
	b_1	b_2	b_3	.	.	b_s

with, $c_i = \sum_j a_{ij}$. For example, the implicit Euler, Trapezoidal and the 2-stage Singly Diagonally Implicit Runge-Kutta (SDIRK-2) methods are given in table A.1.

For convenience of analysis, we write the above equation in the form,

$$u^{(0)} = u^n \tag{A.1}$$

$$u^{(i)} = \sum_{j=0}^{i-1} \alpha_{i,j} u^{(j)} + \Delta t \sum_{j=1}^{i} \beta_{i,j} L(u^{(j)}), \quad i = 1..s+1 \tag{A.2}$$

$$u^{n+1} = u^{(s+1)} \tag{A.3}$$

0	0	0
1	0	1
1	0	1

0	0	0
1	$\frac{1}{2}$	$\frac{1}{2}$
1	$\frac{1}{2}$	$\frac{1}{2}$

γ	γ	
1	$1-\gamma$	γ
	$1-\gamma$	γ

Table A.1: Implicit Euler,Trapezoidal and SDIRK-2 ($\gamma = \frac{2-\sqrt{2}}{2}$)

Theorem *If a spatial discretization is TVD for $\Delta t \leq \Delta t_{ee}$ if used with an explicit Euler time discretization, then it will be TVD if used with the RK method (eqns. A.1 - A.3) under the new time step restriction,*

$$\Delta t \leq k\Delta t_{ee} \quad , \quad k = min_{i,j}\frac{\alpha_{i,j}}{\beta_{i,j}}, i = 1..s+1, j = 0..i-1.$$

Proof Consider any stage in eqn A.3, and for the moment, assume $\alpha_{i,j} \geq 0$ *and* $\beta_{i,j} \geq 0$

$$\begin{aligned}
u^{(i)} &= \sum_{j=0}^{i-1} \alpha_{i,j} u^{(j)} + \Delta t \sum_{j=1}^{i} \beta_{i,j} L(u^{(j)}), \quad i = 1..s+1 \\
u^{(i)} + \beta_{i,i} u^{(i)} &= \sum_{j=0}^{i-1} \alpha_{i,j}[u^{(j)} + \Delta t \frac{\beta_{i,j}}{\alpha_{i,j}} L(u^{(j)})] + \beta_{i,i}[u^{(i)} + \Delta t L(u^{(i)})] \\
\parallel u^{(i)} + \beta_{i,i} u^{(i)} \parallel &= \parallel \sum_{j=0}^{i-1} \alpha_{i,j}[u^{(j)} + \Delta t \frac{\beta_{i,j}}{\alpha_{i,j}} L(u^{(j)})] + \beta_{i,i}[u^{(i)} + \Delta t L(u^{(i)})] \parallel \\
(1 + \beta_{i,i}) \parallel u^{(i)} \parallel &\leq \sum_{j=0}^{i-1} \alpha_{i,j} \parallel [u^{(j)} + \Delta t \frac{\beta_{i,j}}{\alpha_{i,j}} L(u^{(j)})] \parallel + \beta_{i,i} \parallel [u^{(i)} + \Delta t L(u^{(i)})] \parallel \\
\parallel u^{(i)} \parallel + \beta_{i,i} \parallel u^{(i)} \parallel &\leq \sum_{j=0}^{i-1} \alpha_{i,j} \parallel u^{(j)} \parallel + \beta_{i,i} \parallel u^{(i)} \parallel \quad \textbf{if} \quad \mathbf{\Delta t \leq k \Delta t_{ee}} \\
Therefore, \quad \parallel u^{(i)} \parallel &\leq \sum_{j=0}^{i-1} \alpha_{i,j} \parallel u^{(j)} \parallel
\end{aligned}$$

If this is applied recursively from $i = 0$ *to* $s+1$, we get

$$\parallel u^{(i)} \parallel \leq \sum_{j=0}^{i-1} \alpha_{i,j} \parallel u^{(n)} \parallel, \quad i = 0...s+1$$

but, by consistency, $\sum_{j=0}^{i-1} \alpha_{i,j} = 1$. Therefore, we have,

$$\parallel u^{n+1} \parallel \leq \parallel u^{n} \parallel$$

It can also be shown that the assumption that $\beta_{i,j} \geq 0$ can be relaxed under certain conditions, but the positivity of the α 's cannot be relaxed [2]. Hence, Implicit RK schemes should try to maximize $min_{i,j}\frac{\alpha_{i,j}}{\beta_{i,j}}$. Shu et al. [14] have shown that unconditionally TVD implicit schemes of order higher than one do not exist.

A.1 θ schemes

The implicit Euler and Trapezoidal schemes are given by the one-parameter family of schemes given by,

$$u^{n+1} = u^n + \Delta t[(1-\theta)L(u^n) + \theta L(u^{n+1})]$$

Hence, the TVD limit is given by,

$$k = \frac{1}{1-\theta}, \quad or \quad \Delta t \leq \Delta t_{ee}\frac{1}{1-\theta}$$

Hence, the implicit Euler scheme ($\theta = 1$), is unconditionally TVD ($k = \infty$) and the Trapezoidal scheme ($\theta = \frac{1}{2}$) is conditionally TVD ($k = 2$).

A.2 L-DIRK2 with constant θ

Consider the L-DIRK2 scheme with a constant value of θ over the domain. Then, this scheme can be written in the Butcher array format given by,

γ	γ	
1	$\gamma + \theta(1-2\gamma)$	$(1-\gamma) + \theta(2\gamma - 1)$
	$\gamma + \theta(1-2\gamma)$	$(1-\gamma) + \theta(2\gamma - 1)$

In this case,

$$\begin{aligned}
\alpha_{10} &= 1 \\
\beta_{11} &= \gamma \\
\alpha_{20} &= 1 - \alpha_{21} \\
\beta_{21} &= \gamma(1-\alpha_{21}) + \theta(1-2\gamma) \\
\beta_{22} &= (1-\gamma) + \theta(2\gamma - 1)
\end{aligned}$$

Hence, α_{21} is a parameter that can be varied, giving a TVD limit of:

$$\begin{aligned} \Delta t &\leq \frac{\alpha_{21}}{\beta_{21}} \Delta t_{ee} \\ &= \frac{\alpha_{21}}{\gamma(1-\alpha_{21}) + \theta(1-2\gamma)} \Delta t_{ee} \\ &= k \Delta t_{ee} \end{aligned}$$

Where,

$$k = \frac{1}{\theta(1-2\gamma)} \qquad (\alpha_{21} = 1)$$

The 2-stage SDIRK scheme corresponds to $\theta = 1$ and is thus conditionally TVD with $k = \frac{1}{1-2\gamma} \approx 2.4142$.

Appendix B

Slope-limiter form of WENO scheme

The interface value given by the 3 point Weighted Essentially Non-Oscillatory scheme [29] is given by:

$$u_{j+\frac{1}{2}} = w\left[\frac{3}{2}u_j - \frac{1}{2}u_{j-1}\right] + (1-w)\left[\frac{1}{2}u_{j+1} + \frac{1}{2}u_j\right] \tag{B.1}$$

The weighting function w is defined as:

$$\begin{aligned} w &= \frac{\alpha_0}{\alpha_0+\alpha_1}, \quad with \\ \alpha_0 &= \frac{1}{2(u_j - u_{j-1})^4} \quad and \\ \alpha_1 &= \frac{1}{(u_{j+1}-u_j)^4} \end{aligned}$$

In practice, a very small number $\epsilon > 0$ is added to the denominator of the definitions of α_0 and α_1 in order to prevent division by zero.

It is seen that w can be written as:

$$\begin{aligned} w &= \frac{1}{1+\frac{\alpha_1}{\alpha_0}} \\ &= \frac{1}{1+2r_j^4}, \quad where, \\ r_j &= \frac{u_j - u_{j-1}}{u_{j+1}-u_j} \end{aligned}$$

Eqn. B.1 can be cast as:

$$\begin{aligned} u_{j+\frac{1}{2}} &= u_j + \frac{\Delta u_{j+\frac{1}{2}}}{2}\left[(1-w) + \frac{w}{r_j}\right] \\ &= u_j + \frac{\Delta u_{j+\frac{1}{2}}}{2}\left[\frac{r_j + 2r_j^4}{1+2r_j^4}\right] \end{aligned}$$

The slope limiter form is given by

$$u_{j+\frac{1}{2}} = u_j + \frac{\Delta u_{j+\frac{1}{2}}}{2}\phi_j(r_j)$$

Therefore, the limiter $\phi(r)$ for the 3 point WENO scheme can be written as:

$$\phi(r) = \frac{r + 2r^4}{1 + 2r^4}$$

Casting the WENO scheme in this form allows one to prove the TVD properties.

Appendix C

Space-Time Reconstruction

C.1 Limiter in a general frame-work

For the linear convection case $u_t + au_x = 0$ with $a < 0$, the TVD condition is:

$$\frac{1 + \frac{1}{2}(-\frac{\phi_{j+1}}{r_{j+1}} + \phi_j)}{1 + \frac{|\nu|}{2}(\frac{\psi_{j+1}}{s^-_{j+1}} - \psi_j)} \geq 0$$

where, $s^-_{j+1} = \frac{(\Delta_t)_j}{(\Delta_t)_{j+1}}$.

For a non-linear case $u_t + f(u)_x = 0$, with $\frac{\partial f}{\partial u}$ being positive or negative, we get similar conditions except for the fact that ν is replaced by $\nu_j = \frac{\Delta t}{\Delta x}\frac{(f_j^{n+1} - f_j^n)}{(u_j^{n+1} - u_j^n)}$. Hence the time-limiter (eqn. 5.9) for a general case would be:

$$\psi_j^{p+1} = max\left[0, min\left[\frac{2}{|\nu_j|} + \frac{\psi_{j-1}^p}{s_{j-1}^p}, \frac{2}{|\nu_j|} + \frac{\psi_{j+1}^p}{(s^-_{j+1})^p}, 1\right]\right]$$

C.2 Linear stability of the unlimited scheme

A conventional Fourier stability analysis of the unlimited scheme as applied to the linear convection equation ($a > 0$) results in the amplification function G as a function of the spatial wave number β and the CFL number ν given by:

$$G(\beta, \nu) = \frac{1 - \frac{\nu}{2}(1 - e^{-i\beta})}{1 - \frac{\nu}{2}(1 - e^{i\beta})}$$

It is seen that $|G| = 1$ for all possible β and ν. This implies unconditional linear stability and the scheme performs well for smooth solutions. However, for discontinuous solutions, accumulation of dispersion errors will result in spurious oscillations.