Numerical Techniques for Conservation Laws with Source Terms

by Justin Hudson

Project Supervisors

Dr. P.K. Sweby Prof. M.J. Baines

Abstract

In this dissertation we will discuss the finite difference method for approximating conservation laws with a source term present which is considered to be a known function of x, t and u. Finite difference schemes for approximating conservation laws without a source term present are discussed and are adapted to approximate conservation laws with a source term present. First we consider the source term to be a function of x and t only and then we consider the source term to be a function of u also. Some numerical results of the different approaches are discussed throughout the dissertation and an overall comparison of the different approaches is made when the source term is stiff.

This dissertation was funded by the Engineering and Physical Science Research Council.

Contents

Figures

CHAPTER 1:

CHAPTER 2:

CHAPTER 3:

CHAPTER 4:

CHAPTER 5:

CHAPTER 6:

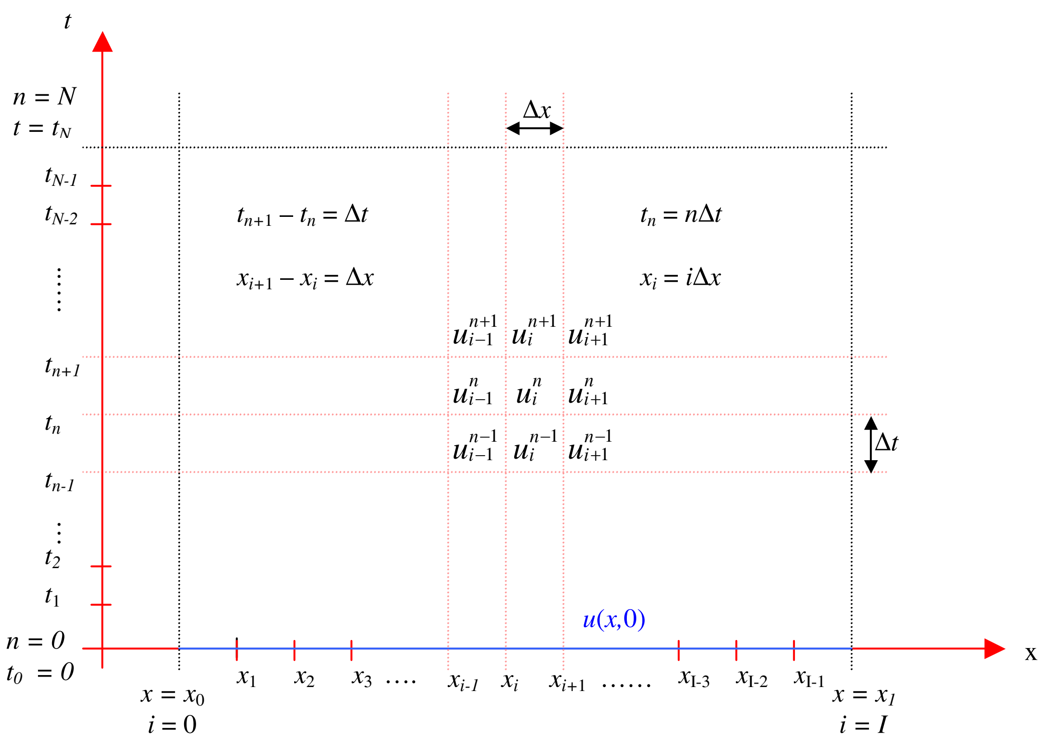
Symbols and Notation

The following is a list of symbols and notation used throughout this project.

Δx	Step-size in x-direction.
Δt	Step-size in t-direction.
i	Integer denoting current step number.
n	Integer denoting current step number.
I	Total number of steps in x-direction.
N	Total number of steps in t-direction.
$x_i = i\Delta x$	Current position in space.
$t_n = n\Delta t$	Current position in time.
$u(x,t)$	The exact solution.
$u(x,0)$	The initial data.
$f(u(x,t))$	The flux.
$R(x,t,u(x,t))$	The source term.
$u_i^n \approx u(i\Delta x, n\Delta t)$	The numerical approximation of the exact solution.
$u_i^{(1)} \approx u(i\Delta x, (n+1)\Delta t)$	The *first order* numerical approximation of the exact solution.
$f_i^n \approx f(u(i\Delta x, n\Delta t))$	The numerical approximation of the flux.
$R_i^n \approx R(i\Delta x, n\Delta t, u(i\Delta x, n\Delta t))$	The numerical approximation of the source term.
c	The wave speed for the advection equation.

$a(u(x,t))$	The wave speed for the conservation law.
$s = \dfrac{\Delta t}{\Delta x}$	
$v = c\dfrac{\Delta t}{\Delta x}$	The Courant number for the advection equation.
$e_i^n = \left\lvert u_i^n - u(i\Delta x, n\Delta t)\right\rvert$	The 'true' error of a scheme at the nodes.
T_i^n	The truncation error of a scheme.
$\phi_i = \phi(\theta_i)$	The flux-limiter of a second order scheme.
$v_{i+1/2} = s\begin{cases} \dfrac{f_{i+1}^n - f_i^n}{u_{i+1}^n - u_i^n} & \text{if } u_{i+1}^n \neq u_i^n \\ a(u_i^n) & \text{if } u_{i+1}^n = u_i^n \end{cases}$	The local Courant number for the conservation law.
$\dfrac{\partial u}{\partial t} \approx \dfrac{u_i^{n+1} - u_i^n}{\Delta t}$	A forward difference approximation.
$\dfrac{\partial u}{\partial x} \approx \dfrac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}$	A central difference approximation.
$\dfrac{\partial u}{\partial x} \approx \dfrac{u_i^n - u_{i-1}^n}{\Delta x}$	A backward difference approximation.

Also, we will be using a fixed mesh, i.e.



1 Introduction

Recently, the numerical solution of conservation laws with a source term, i.e.

$$\frac{\partial u}{\partial t}+\frac{\partial f(u)}{\partial x}=R(x,t,u) \tag{1.1}$$

where $f(u)$ is the flux and $R(x,t,u)$ is the source term, has been in great demand. This is due to the frequency in which conservation laws with source terms arise in mathematical models of physical situations. For example, the 1-D Shallow Water Equations models flow in rivers for a channel of finite depth and requires the numerical solution of a system of equations of the form (1.1). Consider the Shallow Water Equation discussed by Bermudez and Vazques[4]

$$\frac{\partial w}{\partial t}+\frac{\partial F(w)}{\partial x}=R(x,w) \tag{1.2}$$

where

$$w(x,t)=\begin{bmatrix} h \\ q \end{bmatrix}=\begin{bmatrix} h \\ uh \end{bmatrix},\quad F(w)=\begin{bmatrix} q \\ \frac{q^2}{h}+\frac{1}{2}gh^2 \end{bmatrix} \text{ and } R(x,w)=\begin{bmatrix} 0 \\ ghH'(x) \end{bmatrix}.$$

Here, $h(x,t)$ and $u(x,t)$ represent the total height above the bottom of the channel and the fluid velocity, respectively, and $H(x)$ is the depth of the same point but from a fixed reference level (see Figure 1-1). The analytical solution of (1.2) can be extremely difficult to find and sometimes is impossible. Thus, numerical methods are required to approximate the solution of (1.2).

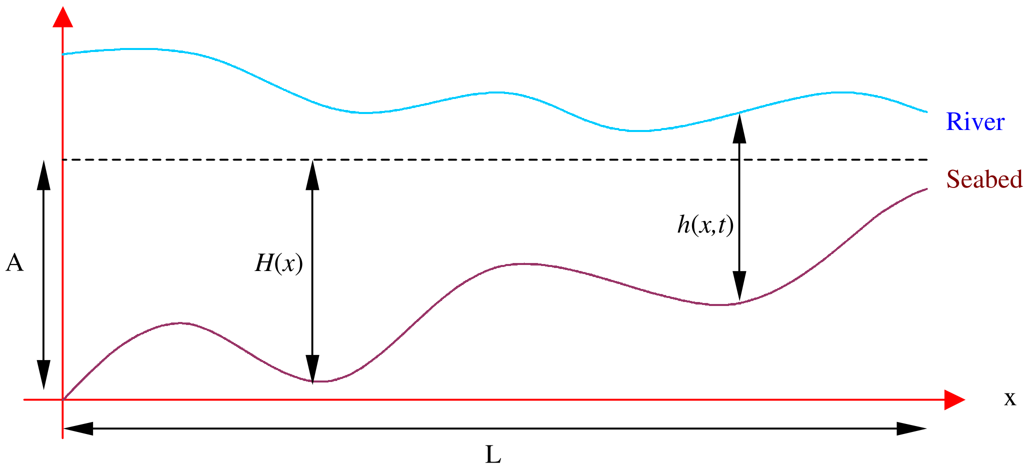


Figure 1-1: Shallow Water Equation.

The solution of (1.1) can be difficult to numerically approximate accurately even when the source term is not present, i.e.

$$\frac{\partial u}{\partial t}+\frac{\partial f(u)}{\partial x}=0. \qquad (1.3)$$

Throughout Chapter 2, we will use finite differences to approximate (1.2) and discuss the accuracy and stability of the schemes derived. We will look at the truncation error and show that first order finite difference schemes are dissipative and second order finite difference schemes are dispersive. Flux-limiter methods will also be discussed so that we can minimise the dispersion present in second order finite difference schemes.

In Chapter 2, we will see that the majority of difficulties encountered when approximating (1.2) can be overcome but we now need to consider how to approximate (1.1), where the source term is now present. A great deal of research has been carried out in conservation laws with source terms but how to handle source terms, especially when they are stiff, is still an open issue. In Chapter 3, we will discuss various approaches for approximating (1.1) but with the source term being only a function of x and t, i.e.

$$\frac{du}{dt}+\frac{df(u)}{dx}=R(x,t).$$

We will consider 'adding' the source term, the Lax-Wendroff approach and the MPDATA approach and we will compare the three approaches for a test problem.

In Chapter 4 we extend the work to consider (1.1) where the source term is also a function of u. This requires an approximation of the source term since we do not know u. In Chapter 4, we will discuss a variety of approaches for numerically approximating (1.1) including the three discussed in Chapter 3. A simple test problem will be used to analyse the different approaches and in Chapter 5, we will compare the different approaches with a test problem whose source term is stiff.

2 1-D Conservation Law

In this chapter, we will look at some numerical schemes for approximating the 1-D scalar conservation law

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

where $u(x,t)$ is the conserved quantity and $f(u)$ is the flux. We can also rearrange (2.1) to obtain the quasi-linear form

$$\frac{\partial u}{\partial t} + a(u)\frac{\partial u}{\partial x} = 0 \tag{2.2}$$

where $a(u) = f'(u)$, which is called the wave-speed. If $a(u) = c$, where c is a constant, then (2.1) becomes the linear advection equation.

2.1 1-D Linear Advection Equation

The most basic form of the conservation law is the linear advection equation

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0 \tag{2.3}$$

where c is a constant and $f(u) = cu$. Here, the constant c is known as the wave speed since $a(u) = c$. There are a variety of numerical techniques for approximating the linear advection equation, such as finite element methods and finite volume methods. Another class of numerical technique used for approximating the linear advection equation are finite difference methods. Finite difference methods involve replacing the derivatives of (2.3) with finite difference approximations. e.g.

$$\frac{\partial u}{\partial t} = \frac{u_i^{n+1} - u_i^n}{\Delta t}$$

which is called the forward difference approximation in time,

$$\frac{\partial u}{\partial t} = \frac{u_i^{n+1} - u_i^{n-1}}{2\Delta t}$$

which is called the central difference approximation in time and

$$\frac{\partial u}{\partial t} = \frac{u_i^n - u_{i-1}^n}{\Delta x}$$

which is called the backward difference approximation in space. The three finite differences can be obtained by using Taylor's theorem, i.e.

$$u_i^{n+1} \approx u_i^n + \Delta t \left[\frac{\partial u}{\partial t} \right]_i^n + \frac{\Delta t^2}{2} \left[\frac{\partial^2 u}{\partial t^2} \right]_i^n + \ldots$$

and by re-arranging we may obtain the forward difference approximation

$$\left[\frac{\partial u}{\partial t} \right]_i^n \approx \frac{u_i^{n+1} - u_i^n}{\Delta t}.$$

Thus, by using finite differences, we can obtain a finite difference scheme that approximates the linear advection equation. For example, if we use a forward difference approximation in space and a central difference approximation in time and assume both of these finite differences to be approximations at (*i,n*), we may obtain

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + c \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta t},$$

and by re-arranging, we obtain

$$u_i^{n+1} = u_i^n - \frac{\nu}{2}\left(u_{i+1}^n - u_{i-1}^n\right),$$

which is a finite difference scheme which approximates the linear advection equation. Unfortunately, this finite difference scheme is unconditionally unstable as we will see later.

2.1.1 First Order Schemes

In order to obtain a first order scheme, we use a forward difference approximation in time and a backward difference approximation in space and assume both of these finite differences to be approximations at (i,n), i.e.

$$\frac{\partial u}{\partial t} = \frac{u_i^{n+1} - u_i^n}{\Delta t} \quad \text{and} \quad \frac{\partial u}{\partial x} = \frac{u_i^n - u_{i-1}^n}{\Delta x}.$$

Substituting these into (2.3) gives

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + c\left[\frac{u_i^n - u_{i-1}^n}{\Delta x}\right] = 0,$$

and hence,

$$u_i^{n+1} = u_i^n - v\left(u_i^n - u_{i-1}^n\right)$$

where $v = c\frac{\Delta t}{\Delta x}$ and is known as the Courant number. This scheme is one of the most basic numerical approximations of the advection equation. However, it can be shown that this scheme is numerically unstable if $c < 0$, in which case we use a forward difference approximation in space and time and assume that both are approximations at (i,n), i.e.

$$\frac{\partial u}{\partial t} = \frac{u_i^{n+1} - u_i^n}{\Delta t} \quad \text{and} \quad \frac{\partial u}{\partial x} = \frac{u_{i+1}^n - u_i^n}{\Delta x},$$

then substituting into (2.1) gives

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + c\left[\frac{u_{i+1}^n - u_i^n}{\Delta x}\right] = 0.$$

Whence,

$$u_i^{n+1} = u_i^n - v\left(u_{i+1}^n - u_i^n\right).$$

This scheme is numerically unstable if $c > 0$. Separately, these schemes can become numerically unstable, but if we combine them

$$u_i^{n+1} = u_i^n - \begin{cases} v\left(u_i^n - u_{i-1}^n\right) & \text{if } v > 0 \\ v\left(u_{i+1}^n - u_i^n\right) & \text{if } v < 0 \end{cases} \tag{2.4}$$

we obtain the Upwind method with switching through $v = 0$. This scheme can still become unstable but only for $|v| > 1$. This will be discussed later.

Alternatively, we could obtain another first order scheme if we use a forward difference approximation in time and a central difference approximation in space and assume that both are approximations at (i,n), i.e.

$$\frac{\partial u}{\partial t} = \frac{u_i^{n+1} - u_i^n}{\Delta t} \quad \text{and} \quad \frac{\partial u}{\partial x} = \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}$$

then substituting into (2.1) gives

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + c\left[\frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}\right] = 0$$

$$\Rightarrow \quad u_i^{n+1} = u_i^n - \frac{v}{2}\left(u_{i+1}^n - u_{i-1}^n\right).$$

Unfortunately this central scheme is unconditionally unstable, but by replacing u_i^n by the average

$$u_i^n = \frac{1}{2}\left(u_{i+1}^n + u_{i-1}^n\right)$$

we obtain the Lax-Friedrichs scheme

$$u_i^n = \frac{1}{2}\left(u_{i+1}^n + u_{i-1}^n\right) - \frac{v}{2}\left(u_{i+1}^n - u_{i-1}^n\right) \tag{2.5}$$

which is stable for $|v| \le 1$. (See later)

2.1.2 Second Order Schemes

One of the most well known second order schemes for approximating the advection equation is the Lax-Wendroff scheme and is derived as follows:

Using Taylor's theorem

$$u_i^{n+1} \approx u_i^n + \Delta t\left[\frac{\partial u}{\partial t}\right]_i^n + \frac{\Delta t^2}{2}\left[\frac{\partial^2 u}{\partial t^2}\right]_i^n + \ldots \tag{2.6}$$

and since

$$\frac{\partial u}{\partial t} = -c\frac{\partial u}{\partial x} \tag{2.7}$$

$$\Rightarrow \quad \frac{\partial^2 u}{\partial t^2} = -c\frac{\partial^2 u}{\partial t \partial x} = -c\frac{\partial^2 u}{\partial x \partial t} = -c\frac{\partial\left(\frac{\partial u}{\partial t}\right)}{\partial x} = -c\frac{\partial\left(-c\frac{\partial u}{\partial x}\right)}{\partial x}$$

so,

$$\frac{\partial^2 u}{\partial t^2} = c^2\frac{\partial^2 u}{\partial x^2}. \tag{2.8}$$

Substituting (2.7) and (2.8) into (2.6) gives

$$u_i^{n+1} \approx u_i^n - c\Delta t\left[\frac{\partial u}{\partial x}\right]_i^n + c^2\frac{\Delta t^2}{2}\left[\frac{\partial^2 u}{\partial x^2}\right]_i^n + \ldots$$

and by using central difference approximations in space and assuming that both are approximations are at (i,n), i.e.

$$\frac{\partial u}{\partial x} = \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x} \quad \text{and} \quad \frac{\partial^2 u}{\partial x^2} = \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{\Delta x^2}$$

we obtain

$$u_i^{n+1} = u_i^n - c\Delta t\frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x} + c^2\frac{\Delta t^2}{2}\frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{\Delta x^2}.$$

Hence, the second order Lax-Wendroff scheme is

$$u_i^{n+1} = u_i^n - \frac{v}{2}\left(u_{i+1}^n - u_{i-1}^n\right) + \frac{v^2}{2}\left[u_{i+1}^n - 2u_i^n + u_{i-1}^n\right]. \tag{2.9}$$

2.1.3 Implicit Schemes

So far, all the schemes we have looked at have been explicit schemes. This is because none of the schemes we have looked at have terms involving time level n+1 on the right hand side of the scheme. For example, the Lax-Wendroff scheme is explicit

$$u_i^{n+1} = u_i^n - \frac{v}{2}\left(u_{i+1}^n - u_{i-1}^n\right) + \frac{v^2}{2}\left[u_{i+1}^n - 2u_i^n + u_{i-1}^n\right]$$

but if we use central difference approximations in space and assume that both are approximations at $(i,n+1)$ instead of approximations at (i,n)

$$u_i^{n+1} = u_i^n - \frac{v}{2}\left(u_{i+1}^{n+1} - u_{i-1}^{n+1}\right) + \frac{v^2}{2}\left[u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}\right] \qquad (2.10)$$

we obtain the implicit Lax-Wendroff scheme. This scheme is implicit since terms involving n+1 appear on the right hand side of the equation. Implicit schemes cause difficulties since we now have to solve a tri-diagonal system at each time step. Re-arranging (2.10)

$$-\frac{v}{2}(1+v)u_{i-1}^{n+1} + \left(1+v^2\right)u_i^{n+1} + \frac{v}{2}(1-v)u_{i+1}^{n+1} = u_i^n$$

hence

$$\begin{bmatrix} b & c & 0 & 0 & 0 & \dots & 0 \\ a & b & c & 0 & 0 & \dots & 0 \\ 0 & a & b & c & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a & b & c & 0 \\ 0 & \dots & 0 & 0 & a & b & c \\ 0 & \dots & 0 & 0 & 0 & a & b \end{bmatrix} \begin{bmatrix} u_0^{n+1} \\ u_1^{n+1} \\ u_2^{n+1} \\ \vdots \\ u_{I-2}^{n+1} \\ u_{I-1}^{n+1} \\ u_I^{n+1} \end{bmatrix} = \begin{bmatrix} u_0^n - a\,u_{-1}^{n+1} \\ u_1^n \\ u_2^n \\ \vdots \\ u_{I-2}^n \\ u_{I-1}^n \\ u_I^n - c\,u_{I+1}^n \end{bmatrix}$$

is the tri-diagonal system, which needs to be solved at each time step, for the implicit Lax-Wendroff scheme where

$$a = -\frac{v}{2}(1+v),\ b = 1+v^2 \ \text{ and } \ c = \frac{v}{2}(1-v).$$

All implicit schemes take the form

$$A\underline{u}^{n+1} = G$$

where A is a $(I+1)$ x $(I+1)$ matrix and G is a $(I+1)$ column vector. In general, implicit schemes can be more accurate than explicit schemes but implicit schemes are harder to implement and require a lot more calculations than explicit methods.

So far we have looked at a few finite difference schemes, of first or second order, which numerically approximate the solution of the advection equation but there are a great deal more and definitely too many to look at in this section. For a more in depth discussion of finite difference schemes for the advection equation, look in Kroner[8], LeVeque[7] and Ames[14].

2.2 1-D Conservation Law

In Section 2.1, we discussed some finite difference schemes for approximating the linear advection equation, which is a form of the scalar conservation law

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

where $f'(u) = a(u)$. However, we can adapt the techniques discussed in Section 2.1 so that we can numerically approximate the solution of the scalar conservation law but we must be careful how we approximate (2.1) since we wish to ensure conservation.

2.2.1 Non-Conservative Schemes

If a scheme is non-conservative, then the scheme will move discontinuities at the incorrect wave speed. For example, if we approximated the quasi-linear form of equation (2.1) by using the finite difference method then we would obtain a non-conservative scheme. Consider inviscid Burger's equation, i.e.

$$\frac{\partial u}{\partial t}+\frac{\partial\left(\frac{1}{2}u^2\right)}{\partial x}=0,$$

re-writing in quasi-linear form gives

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}=0$$

and by using a forward difference approximation in time and a backward difference approximation in space and assuming that both are approximations are at (i,n), i.e.

$$\frac{\partial u}{\partial t} = \frac{u_i^{n+1} - u_i^n}{\Delta t} \quad \text{and} \quad \frac{\partial u}{\partial x} = \frac{u_i^n - u_{i-1}^n}{\Delta x}$$

we obtain

$$u_i^{n+1} = u_i^n - s u_i^n \left[u_i^n - u_{i-1}^n \right],$$

assuming $u_i^n > 0$. This scheme is conservative for smooth data only and if used to numerically approximate discontinuities, the scheme becomes non-conservative moving the discontinuity at the wrong speed.

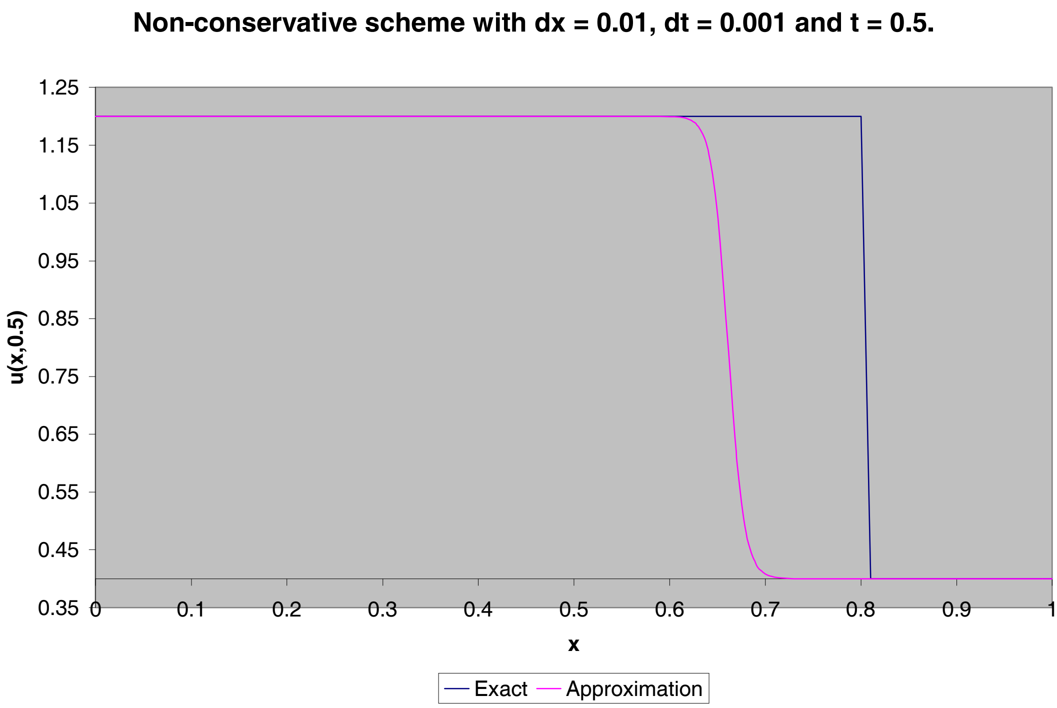


Figure 2-1: Non-conservative scheme.

If we use the non-conservative scheme, which approximates inviscid Burger's equation, with initial data

$$u(x,0) = \begin{cases} 1.2 & \text{if} \quad x < 0.3 \\ 0.4 & \text{if} \quad x \geq 0.3 \end{cases},$$

we may obtain the results in Figure 2-1. Here, we can see that the scheme has moved the discontinuity too slowly which means that the scheme is not conservative.

2.2.2 Conservative Schemes

To ensure conservation, we require that the method be in conservation form, i.e.

$$u_i^{n+1} = u_i^n - s\left[F\left(u_{i-p}^n, u_{i-p+1}^n, \ldots, u_{i+q}^n\right) - F\left(u_{i-p-1}^n, u_{i-p}^n, \ldots, u_{i+q-1}^n\right)\right]$$

where F is called the numerical flux function and is of $p + q + 1$ arguments. We can ensure conservation by numerically approximating (2.1) and using a similar approach as we did in the previous sub-section. For example, when we derived the Upwind scheme, we used a forward difference in time and either a forward or a backward difference in space depending on the value of v. Here, we take a same approach but we will apply finite differences to f instead of u, i.e.

$$\frac{\partial u}{\partial t} = \frac{u_i^{n+1} - u_i^n}{\Delta t}$$

and either $\frac{\partial f}{\partial x} = \frac{f_{i+1}^n - f_i^n}{\Delta x}$ if $v_{i+1/2} > 0$ or $\frac{\partial f}{\partial x} = \frac{f_i^n - f_{1-1}^n}{\Delta x}$ if $v_{i+1/2} < 0$

where

$$v_{i+1/2} = sa(u_{i+1/2}^n).$$

Hence,

$$u_i^{n+1} = u_i^n - s\begin{cases}\left(f_i^n - f_{i-1}^n\right) & \text{if } v_{i+1/2} > 0 \\ \left(f_{i+1}^n - f_i^n\right) & \text{if } v_{i+1/2} < 0\end{cases}$$

is the Upwind scheme for the scalar conservation law where $s = \frac{\Delta t}{\Delta x}$ and $f_i^n = f(u_i^n)$.

However, difficulties arise when approximating $v_{i+1/2}$. This is because in

$$v_{i+1/2} = sa(\ u_{i+1/2}^n\ )$$

$u_{i+1/2}$ is unknown. One approach used to overcome this problem could be to approximate $u_{i+1/2}$ by

$$u_{i+1/2} = \frac{1}{2}\left(u_{i+1}^n + u_i^n\right).$$

Another method, which ensures conservation, is to approximate $v_{i+1/2}$ by replacing $a(u)$ by a "local v" defined at each grid point by

$$v_{i+1/2} = s\begin{cases} \left[\dfrac{f^n_{i+1} - f^n_i}{u^n_{i+1} - u^n_i}\right] & \text{if } u^n_{i+1} \neq u^n_i \\ a(u^n_i) & \text{otherwise} \end{cases}.$$

Problems also occur when adapting the Lax-Wendroff scheme to the non-linear case. This is because (2.8) no longer holds. However, we can overcome this problem by re-writing (2.8) as

$$\frac{\partial^2 u}{\partial t^2} = -\frac{\partial^2 f(u)}{\partial t \partial x} = -\frac{\partial\left(\dfrac{\partial f(u)}{\partial t}\right)}{\partial x} = -\frac{\partial\left(\dfrac{\partial f(u)}{\partial u}\dfrac{\partial u}{\partial t}\right)}{\partial x} = \frac{\partial\left(a(u)\dfrac{\partial f}{\partial x}\right)}{\partial x} \tag{2.11}$$

and by using Taylor's theorem,

$$u^{n+1}_i = u^n_i - \Delta t\frac{\partial f(u)}{\partial x} + \frac{\Delta t^2}{2}\frac{\partial\left(a(u)\dfrac{\partial f}{dx}\right)}{\partial x}$$

whence we may obtain

$$u^{n+1}_i = u^n_i - \frac{s}{2}\left(f^n_{i+1} - f^n_{i-1}\right) + \frac{s}{2}\left[v_{i+1/2}\left(f^n_{i+1} - f^n_i\right) - v_{i-1/2}\left(f^n_i - f^n_{i-1}\right)\right]$$

the Lax-Wendroff scheme for the conservation law. Table 2-1 lists a variety of finite difference schemes for the conservation law.

Name of Scheme	Scheme	Order
Upwind (first order)	$u_i^{n+1} = u_i^n - \begin{cases} s\left(f_i^n - f_{i-1}^n\right) & \text{if } \nu_{i+1/2} > 0 \\ s\left(f_{i+1}^n - f_i^n\right) & \text{if } \nu_{i+1/2} < 0 \end{cases}$	**1**
Lax-Friedrichs	$u_i^{n+1} = \frac{1}{2}\left(u_{i+1}^n + u_{i-1}^n\right) - \frac{s}{2}\left[f_{i+1}^n - f_{i-1}^n\right]$	**1**
Second Order Upwind (Warming and Beam)	$u_i^{n+1} = u_i^n - \frac{s}{2}\begin{cases} \left(3-\nu_{i-1/2}\right)\left(f_i^n - f_{i-1}^n\right) - \left(1-\nu_{i-3/2}\right)\left(f_{i-1}^n - f_{i-2}^n\right) & \text{if } \nu_{i+1/2} > 0 \\ -\left(1+\nu_{i+3/2}\right)\left(f_{i+2}^n - f_{i+1}^n\right) + \left(\nu_{i+1/2} + 3\right)\left(f_{i+1}^n - f_i^n\right) & \text{if } \nu_{i+1/2} < 0 \end{cases}$	**2**
Leapfrog	$u_i^{n+1} = u_i^{n-1} - s\left(f_{i+1}^n - f_{i-1}^n\right)$	**2**
Lax-Wendroff	$u_i^{n+1} = u_i^n - \frac{s}{2}\left(f_{i+1}^n - f_{i-1}^n\right) + \frac{s}{2}\left[\nu_{i+1/2}\left(f_{i+1}^n - f_i^n\right) - \nu_{i-1/2}\left(f_i^n - f_{i-1}^n\right)\right]$	**2**
MacCormack Predictor-Corrector	$u_i^* = u_i^n - s\left(f_{i+1}^n - f_i^n\right)$ $u_i^{n+1} = \frac{1}{2}\left(u_i^n + u_i^*\right) - \frac{s}{2}\left[f_i^* - f_{i-1}^*\right]$	**2**

Table 2-1: Finite difference schemes for the 1-D conservation law.

For all schemes in Table 2-1, $\nu_{i+1/2} = s\begin{cases} \left[\dfrac{f_{i+1}^n - f_i^n}{u_{i+1}^n - u_i^n}\right] & \text{if } u_{i+1}^n \neq u_i^n \\ a(u_i^n) & \text{otherwise} \end{cases}$ and $s = \dfrac{\Delta t}{\Delta x}$.

Here, we can see that adapting the finite difference method to the scalar conservation law can cause minor problems.

2.3 Truncation Error and Stability

2.3.1 Truncation Error

The truncation error of a scheme is very useful, since it tells us whether the scheme is consistent and the order of accuracy of the scheme. To derive the truncation error of a scheme, we assume that the values at the grid points are exact, i.e. $u_i^n = u(i\Delta x, n\Delta t)$, and then use Taylor series expansions. The truncation error is also known as the discretisation error, which is the error caused by using finite difference approximations to approximate the derivatives of (2.3). As an example, consider the Lax-Friedrichs scheme (2.5) for the scalar conservation law

$$u_i^{n+1} = \frac{1}{2}\left(u_{i+1}^n + u_{i-1}^n\right) - \frac{s}{2}\left[f_{i+1}^n - f_{i-1}^n\right].$$

Now, by assuming that the values at the grid points are exact, i.e. $u_i^n = u(i\Delta x, n\Delta t)$, and by using Taylor's theorem

$$u_i^{n+1} \approx u_i^n + \Delta t\left[\frac{\partial u}{\partial t}\right]_i^n + \frac{\Delta t^2}{2}\left[\frac{\partial^2 u}{\partial t^2}\right]_i^n + \ldots$$
$$u_i^{n+1} \approx u_i^n + \Delta x\left[\frac{\partial u}{\partial x}\right]_i^n + \frac{\Delta x^2}{2}\left[\frac{\partial^2 u}{\partial x^2}\right]_i^n + \ldots$$

then by substituting into (2.5) gives

$$\begin{aligned}\Delta t \mathrm{T}_i^n = u + \Delta t\frac{\partial u}{\partial t} - \frac{1}{2}\left(u + \Delta x\frac{\partial u}{\partial x} + \frac{\Delta x^2}{2}\frac{\partial^2 u}{\partial x^2} + u - \Delta x\frac{\partial u}{\partial x} + \frac{\Delta x^2}{2}\frac{\partial^2 u}{\partial x^2}\right) \\ + \frac{\Delta t}{2\Delta x}\left[f + \Delta x\frac{\partial f}{\partial x} + \frac{\Delta x^2}{2}\frac{\partial^2 f}{\partial x^2} + \frac{\Delta x^3}{6}\frac{\partial^3 f}{\partial x^3} - \left(f - \Delta x\frac{\partial f}{\partial x} + \frac{\Delta x^2}{2}\frac{\partial^2 f}{\partial x^2} - \frac{\Delta x^3}{6}\frac{\partial^3 f}{\partial x^3}\right)\right] \\ + O(\Delta x^3) + O(\Delta t^2)\end{aligned}$$

where T_i^n denotes the truncation error. Hence,

$$\mathrm{T}_i^n = \frac{\Delta t}{2}\frac{\partial^2 u}{\partial t^2} - \frac{\Delta x^2}{2\Delta t}\frac{\partial^2 u}{\partial x^2} + \frac{\Delta x^2}{6}\frac{\partial^3 f}{\partial x^3} + O(\Delta x^3) + O(\Delta t^2)$$

is the truncation error of the Lax-Friedrichs scheme. The Lax-Friedrichs scheme is second order in space but only first order in time, which makes the Lax-Friedrichs scheme first order and consistent, since as $\Delta x^2 \to 0$ and $\Delta t \to 0$, the truncation error tends to zero, $\mathrm{T}_i^n \to 0$. Similarly, if we consider the Lax-Wendroff scheme for the advection equation

$$u_i^{n+1} = u_i^n - \frac{v}{2}\left(u_{i+1}^n - u_{i-1}^n\right) + \frac{v^2}{2}\left[u_{i+1}^n - 2u_i^n + u_{i-1}^n\right],$$

we can show that the Lax-Wendroff scheme has a truncation error of

$$\mathrm{T}_i^n = \frac{1}{6}\left(\Delta t^2 \frac{\partial^3 u}{\partial t^3} - c\,\Delta x^2 \frac{\partial^3 u}{\partial x^3} \right) + O(\Delta x^3) + O(\Delta t^3).$$

The Lax-Wendroff scheme is second order and consistent since as $\Delta t^2 \to 0$ and $\Delta \mathrm{x}^2 \to 0$, the truncation error tends to zero, $\mathrm{T}_i^n \to 0$.

In general, if a scheme has a truncation error of order $O(\Delta x^p) + O(\Delta t^q)$, then the scheme is of order *p* in space, *q* in time and of overall order min(*p*,*q*). Also, if *p* and *q* are greater than or equal to 1, then the scheme is consistent.

2.3.2 Stability

We also need to know the interval of absolute stability of a finite difference scheme since, if we choose our step-sizes such that the interval of absolute stability is breached, then the finite difference scheme will become unstable giving very inaccurate results. Now, a numerical scheme is stable provided the error at the nodes

$$e_i^n = \left| u_i^n - u(i\Delta x, n\Delta t) \right|$$

does not blow up. I.e. if the numerical values at the nodes are not exact, then errors begin to creep in the numerical approximation. If those errors blow up, then the scheme becomes numerically unstable. Figure 2-2 shows the Upwind scheme becoming numerically unstable with initial data

$$u(x,0) = \begin{cases} 1 & \text{if } \ x < 0.5 \\ 0 & \text{if } \ x \geq 0.5 \end{cases}.$$

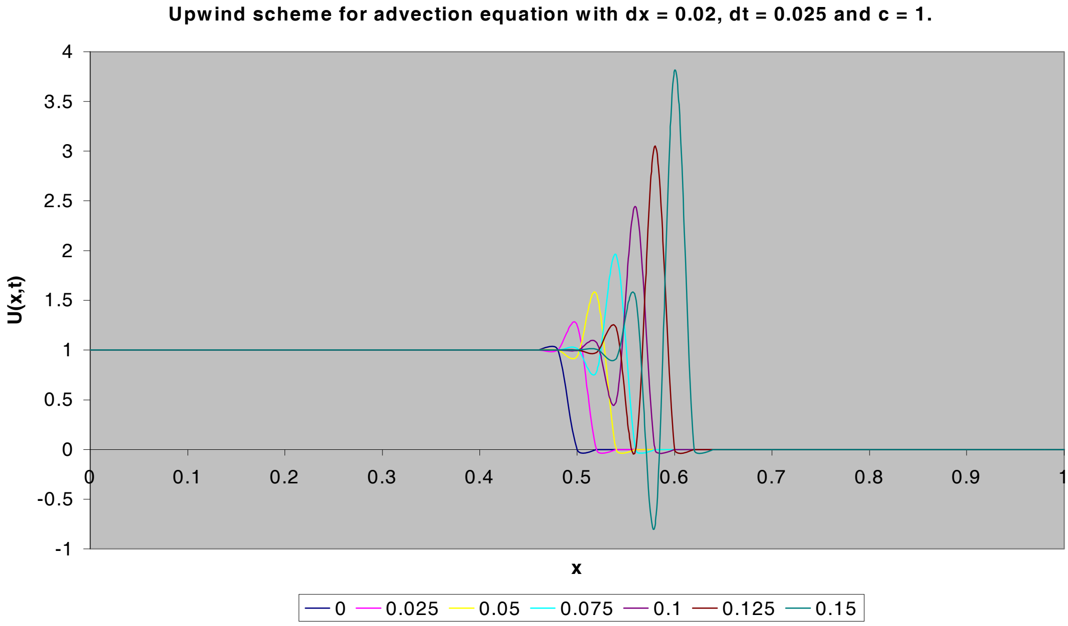


Figure 2-2: The Upwind scheme becoming unstable.

There are several analytical techniques that can be used to see if a scheme is stable, one of which is the Fourier method. The Fourier method consists of substituting a Fourier mode $u_i^n = \xi_n e^{iki\Delta x}$ into the scheme to obtain an expression for the amplification factor ξ. The scheme will then be stable provided

$$|\xi| \leq 1.$$

For example, consider the Lax-Wendroff scheme for the advection equation

$$u_i^{n+1} = u_i^n - \frac{v}{2}\left(u_{i+1}^n - u_{i-1}^n\right) + \frac{v^2}{2}\left[u_{i+1}^n - 2u_i^n + u_{i-1}^n\right].$$

By substituting $u_i^n = \xi_n e^{iki\Delta x}$, we obtain

$$\xi_{n+1} e^{iki\Delta x} = \xi_n e^{iki\Delta x} - \frac{v}{2}\left(\xi_n e^{ik(i+1)\Delta x} - \xi_n e^{ik(i-1)\Delta x}\right)$$
$$+ \frac{v^2}{2}\left[\xi_n^{\ ik(i+1)\Delta x} - 2\xi_n e^{iki\Delta x} + \xi_n e^{ik(i-1)\Delta x}\right].$$

If we now divide by $e^{ijk\Delta x}$ we obtain

$$\xi_{n+1} = \xi_n - \frac{v}{2}\left(\xi_n e^{ik\Delta x} - \xi_n e^{-ik\Delta x}\right) + \frac{v^2}{2}\left[\xi_n^{ik\Delta x} - 2\xi_n + \xi_n e^{-ik\Delta x}\right]$$

and by re-arranging

$$\xi_{n+1} = \left[1 - v^2 + \frac{v^2}{2}\left(e^{ik\Delta x} + e^{-ik\Delta x}\right) - \frac{v}{2}\left(e^{ik\Delta x} - e^{-ik\Delta x}\right)\right]\xi_n .$$

Using the identities

$$e^{jk\Delta x} + e^{-jk\Delta x} = 2\cos k\Delta x$$

and

$$e^{jk\Delta x} - e^{-jk\Delta x} = 2i\sin k\Delta x$$

we may obtain

$$\xi_{n+1} = \left[1 - v^2 + \frac{v^2}{2}(2\cos k\Delta x) - \frac{v}{2}(2i\sin k\Delta x)\right]\xi_n .$$

So, for stability we require

$$\left|1 - v^2 + v^2\cos k\Delta x - vi\sin k\Delta x\right| \le 1 .$$

Here, we can see that the amplification factor lies on an ellipse:

$$\xi = 1 - v^2 + v^2\cos k\Delta x - vi\sin k\Delta x$$

If we let

$$x = 1 - v^2 + v^2\cos k\Delta x$$

and

$$y = -vi\sin k\Delta x$$

and by using the identity

$$\cos^2 k\Delta x + \sin^2 k\Delta x = 1$$

whence

$$\left[\frac{x - (1 - v^2)}{v^2}\right]^2 + \left[\frac{y}{v}\right]^2 = 1 .$$

So, the interval of absolute stability is an ellipse with centre ($1-v^2$) and crosses the x-axis at x = 1 and x = $1-2v^2$. Figure 2-3 shows the unit circle with the ellipse of the amplification factor inside the unit circle. Here, we can see that for the ellipse to stay inside the unit circle, $1 - 2v^2 \ge -1$ and $1 - v^2 \ge 0$. Hence, for the Lax-Wendroff scheme to be stable, $v \le |1|$. This condition on v is called the interval of absolute stability. Notice that if $v = 1$, the ellipse is in fact the unit circle. Table 2-2 lists the stability intervals of a few schemes.

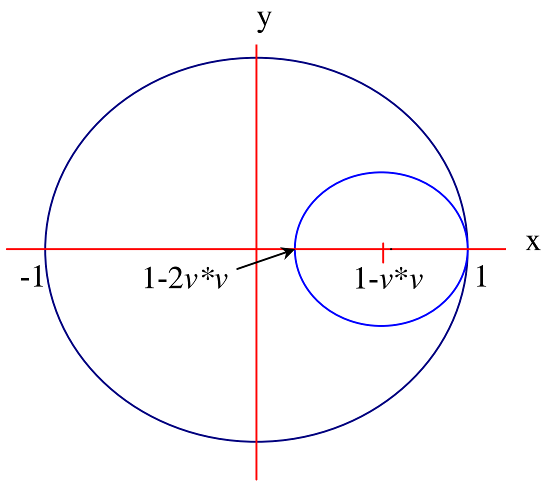


Figure 2-3: Interval of stability for Lax-Wendroff

Name of Scheme	Order (space + time)	Overall Order of Scheme	Interval Of Absolute Stability
Upwind (first order)	1 + 1	**1**	$\lvert \nu \rvert \leq 1$
Lax-Friedrichs	2 + 1	**1**	$\lvert \nu \rvert \leq 1$
Upwind (second order)	2 + 2	**2**	$\lvert \nu \rvert \leq 2$
Leapfrog	2 + 2	**2**	$\lvert \nu \rvert \leq 1$
Lax-Wendroff	2 + 2	**2**	$\lvert \nu \rvert \leq 1$
MacCormack Predictor-Corrector	2 + 2	**2**	$\lvert \nu \rvert \leq 1$

Table 2-2: The interval of absolute stability and the order of some schemes.

Earlier, Figure 2-2 showed the Upwind scheme becoming unstable for $\nu = 1.25$. This is because the Upwind scheme is stable for $0 \leq \nu \leq 1$, when $c > 0$, and since ν lies outside the interval of absolute stability, the scheme will become unstable.

2.4 Dissipation, Dispersion and Oscillations

2.4.1 Dissipation

It can be show that all first order schemes suffer from dissipation which can result in a very inaccurate numerical solution. Dissipation occurs when the travelling wave's amplitude decreases. Figure 2-4 shows some numerical results of the Upwind scheme applied to the advection equation with initial data

$$u(x,0)=\begin{cases}1 & \text{if } x<0.5\\ 0 & \text{if } x>0.5\end{cases}.$$

Figure 2-4 shows us that the Upwind scheme is dissipative since the numerical solution has started to decrease in amplitude.

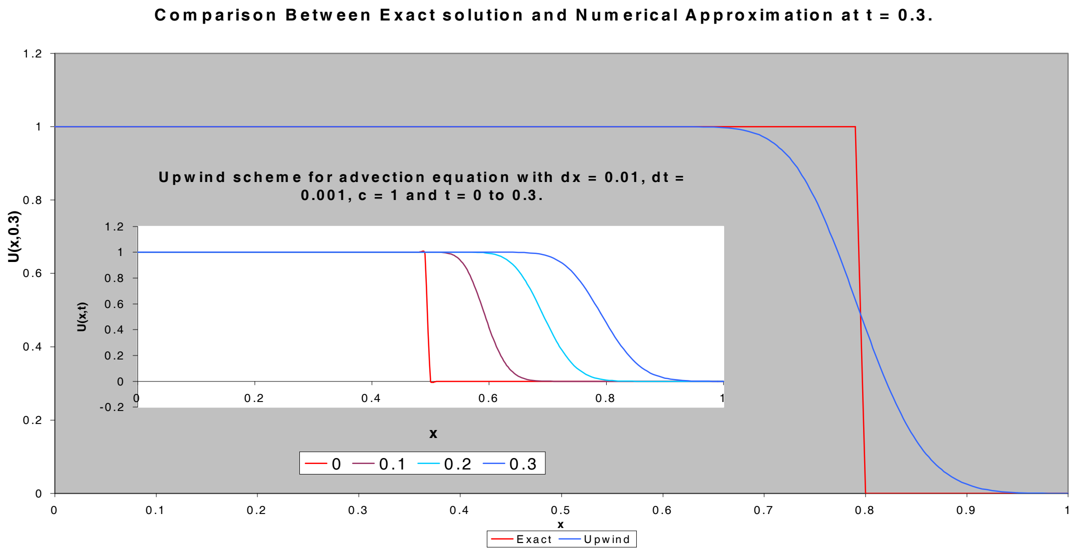


Figure 2-4: Dissipation of the first order Upwind scheme

In order fully understand why dissipation occurs, we will use the analysis of the modified equation, which is discussed by Sweby[13] and LeVeque[7], on the Lax-Friedrichs scheme for the advection equation

$$u_i^{n+1}=\frac{1}{2}\left(u_{i+1}^n+u_{i-1}^n\right)-\frac{\nu}{2}\left[u_{i+1}^n-u_{i-1}^n\right].$$

Earlier, we saw that this scheme had a truncation error of

$$\mathrm{T}_i^n = \frac{\Delta t}{2}\frac{\partial^2 u}{\partial t^2} - \frac{\Delta x^2}{2\Delta t}\frac{\partial^2 u}{\partial x^2} + \frac{\Delta x^2}{6}\frac{\partial^3 u}{\partial x^3} + O(\Delta x^3) + O(\Delta t^2)$$

and by using (2.8)

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$

we may obtain

$$\mathrm{T}_i^n = \left[\frac{\Delta t}{2}c^2 - \frac{\Delta x^2}{2\Delta t}\right]\frac{\partial^2 u}{\partial x^2} + \frac{\Delta x^2}{6}\frac{\partial^3 u}{\partial x^3} + O(\Delta x^3) + O(\Delta t^2).$$

So, the Lax-Friedrichs scheme is a second order approximation to

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = D\frac{\partial^2 u}{\partial x^2} \tag{2.12}$$

where $D = \frac{\Delta x^2}{2\Delta t}\left[1 - v^2\right]$. Equation (2.11) is known as the linear advection-diffusion equation and is ill-posed if $D < 0$. In this case, equation (2.12) is well posed since $\frac{\Delta x^2}{2\Delta t} \geq 0$ so, for (2.12) to be well posed $\left[1 - v^2\right] \geq 0 \Rightarrow |v| \leq 1$. Hence, since for stability, $|v| \leq 1$, equation (2.12) is well posed as long as the scheme is stable. So, the Lax-Friedrichs scheme qualitatively behaves like the solution of (2.12). Now, by using the Fourier Transform of u with respect to x

$$\hat{u}(\xi, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(x,t) e^{ix\xi} dx$$

and substituting into (2.12), we may obtain that (2.12) is an ODE with solution

$$\hat{u}(\xi, t) = \hat{u}(\xi, 0) e^{-D\xi^2 t} e^{ic\xi t}$$

and by using an inverse transform, we may obtain

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{u}(\xi,0) e^{-D\xi^2 t} e^{i(\omega(\xi)t - \xi x)} dx.$$

Here, we can see that the solution is of the form $e^{i(\omega(\xi)t-\xi x)}$, which represents a travelling wave with decreasing amplitude, $\hat{u}(\xi,0)e^{-D\xi^2 t}$. The frequency is ω(ξ) and is dependent on the wave number ξ. In this case the frequency is ω(ξ) = cξ, this is also known as the dispersion relation. Also,

$$\frac{\omega(\xi)}{\xi}$$

is known as the phase velocity and gives us the wave speed of each wave.

For the Lax-Friedrichs scheme, the phase velocity is

$$\frac{\omega(\xi)}{\xi} = c \, .$$

Hence, the waves all travel at the same speed and so, the Lax-Friedrichs scheme is non-dispersive. However, the Lax-Friedrichs scheme suffers from dissipation, due to the wave travelling with decreasing amplitude. Hence, the Lax-Friedrichs scheme suffers from dissipation but not dispersion. We can also show that the Upwind scheme with $v > 0$ suffers from dissipation, since the truncation error of the scheme is

$$\mathrm{T}_i^n = \frac{c}{2}[\Delta t c - \Delta x]\frac{\partial^2 u}{\partial x^2} + O(\Delta x^2) + O(\Delta t^2),$$

the scheme is a second order approximation to (2.12) with

$$D = \frac{c}{2}\Delta x(1-v).$$

Hence, the Upwind scheme is also dissipative and since

$$\frac{c}{2}\Delta x(1-v)\left(1+\frac{1}{v}\right) > \frac{c}{2}\Delta x(1-v),$$

where the left-hand side represents the value of D for the Lax-Friedrichs scheme, we can see that the Lax-Friedrichs scheme is more dissipative than the Upwind scheme for $v > 0$.

2.4.2 Dispersion and Oscillations

Dispersion occurs when waves travel at different wave speeds and is common in all second order schemes. Figure 2-5 shows some numerical results of the Lax-Wendroff scheme applied to the advection equation with initial data

$$u(x,0) = \begin{cases} 1 & \text{if} \quad x < 0.3 \\ 0 & \text{if} \quad x \geq 0.3 \end{cases}.$$

Here, we can see that the Lax-Wendroff scheme suffers from dispersion since oscillations are occurring in the numerical solution behind the discontinuity.

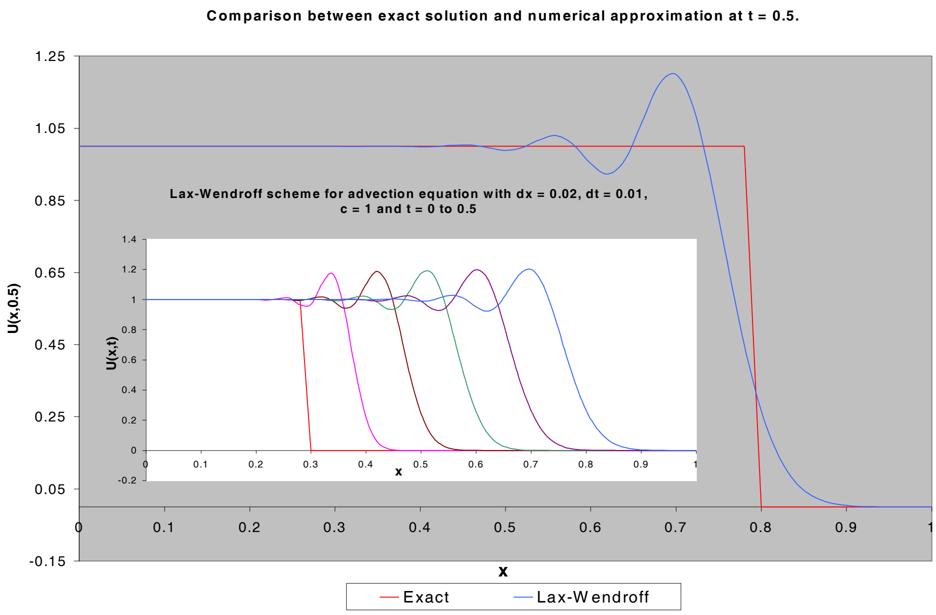


Figure 2-5: Dispersion leading to oscillations of the Lax-Wendroff scheme.

We can see why the Lax-Wendroff scheme suffers from dispersion by taking the same approach as we did for the dissipation case. Consider the Lax-Wendroff scheme for the linear advection equation

$$u_i^{n+1} = u_i^n - \frac{\nu}{2}\left(u_{i+1}^n - u_{i-1}^n\right) + \frac{\nu^2}{2}\left[u_{i+1}^n - 2u_i^n + u_{i-1}^n\right],$$

whose truncation error was

$$T_i^n = \frac{1}{6}\left(\Delta t^2 \frac{d^3 u}{dt^3} - c\,\Delta x^2 \frac{d^3 u}{dx^3} \right) + O(\Delta x^3) + O(\Delta t^3)\,,$$

we can see that the Lax-Wendroff scheme is a better approximation to

$$\frac{du}{dt} + c\frac{du}{dx} = \eta \frac{d^3 u}{dx^3} \quad \text{where} \quad \eta = \frac{c}{6}\Delta x^2 \left(\nu^2 - 1\right)$$

and by using Fourier Transforms, we may obtain

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{u}(\xi,0) e^{i(\omega(\xi)t - \xi x)} dx \quad \text{where} \quad \omega(\xi) = c\xi + \eta \xi^3 .$$

Here, we can see that the solution is of the form $e^{i(\omega(\xi)t - \xi x)}$, which represents a travelling wave with constant amplitude, $\hat{u}(\xi,0)$. This means that the scheme no longer suffers from dissipation, however, consider the phase velocity

$$\frac{\omega(\xi)}{\xi} = \frac{c\xi + \eta\,\xi^3}{\xi} = c + \eta\,\xi^2 .$$

Here, we can see that different wave numbers travel at different speeds and so, the Lax-Wendroff scheme is dispersive. Also, if $\eta < 0$ within the stability region of a scheme, then oscillations will occur behind the discontinuity and if $\eta > 0$ within the stability interval of a scheme, then oscillations will occur in front of the discontinuity. This is because, if $\eta < 0$, then high wave numbers travel with a slower velocity than they should creating oscillations behind the discontinuity, but if $\eta > 0$, then high wave numbers travel with a faster velocity than they should creating oscillations in front of the discontinuity. Figure 2-5 shows that the Lax-Wendroff scheme suffers from oscillations occurring behind the discontinuities, which would imply that $\eta < 0$,

$$\eta = \frac{c}{6}\Delta x^2 \left(\nu^2 - 1\right).$$

For stability, we require $|\nu| \le 1 \Rightarrow \nu^2 \le 1 \Rightarrow \nu^2 - 1 \le 0$, which means that for η < 0,

$\frac{c}{6}\Delta x^2 \ge 0$ and since c = 1 and Δx > 0, verifies that η < 0 creating oscillations behind the discontinuity.

In general, all first order schemes suffer from dissipation but are non-dispersive, and all second order schemes suffer from dispersion but are non-dissipative. For a more in depth discussion on wave theory, see Whitham[9] and Ames[14].

2.5 Flux-limiter Methods

So far we have seen that, in general, all first order schemes suffer from dissipation and all second order schemes suffer from dispersion, which creates oscillations around the discontinuity. However, there is a method which switches between a second order approximation when the region is smooth and a first order approximation when near a discontinuity. This method considerably reduces the size of the oscillations by using a first order approximation near discontinuities and is called the flux-limiter method.

Figure 2-6 shows some numerical results of the Lax-Wendroff scheme with and without the Superbee flux-limiter method applied to the scheme and with the exact solution for initial data

$$u(x,0) = \begin{cases} 1 & \text{if } x < 0.3 \\ 0 & \text{if } x \ge 0.3 \end{cases}.$$

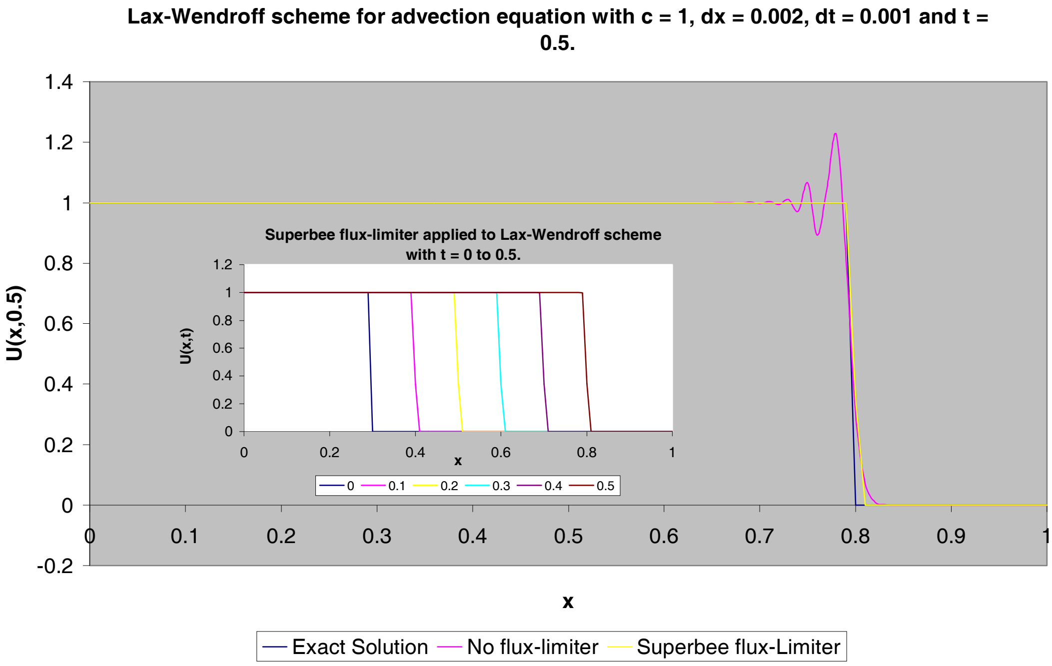


Figure 2-6: Superbee flux-limiter method applied to the Lax-Wendroff scheme.

Here we can see that the Superbee flux-limiter method has eliminated all oscillations from the Lax-Wendroff scheme resulting in an extremely accurate second order scheme. To fully understand flux-limiter methods, we shall closely follow the work of Sweby[13] and LeVeque[7].

Now, we can re-write any second order scheme as

$$u_i^{n+1} = u_i^n - s[F(u;i) - F(u;i-1)] \tag{2.13}$$

where

$$F(u;i) = F_L(u;i) + F_H(u;i). \tag{2.14}$$

Here, $F_L(u;i)$ represents a first order scheme and $F_H(u;i)$ represents a second order correction term. In order to obtain the flux-limiter method for a second order scheme, we re-write (2.14) as

$$F(u;i) = F_L(u;i) + F_H(u;i)\phi_i$$

where ϕ_i represents the flux-limiter, which is yet to be specified. Before we discuss the choice of the flux-limiter, let us re-write the Lax-Wendroff scheme for the scalar conservation law in the form of (2.13).

$$u_i^{n+1} = u_i^n - \frac{s}{2}\left(f_{i+1}^n - f_{i-1}^n\right) + \frac{s}{2}\left[v_{i+1/2}\left(f_{i+1}^n - f_i^n\right) - v_{i-1/2}\left(f_i^n - f_{i-1}^n\right)\right]$$

We can re-write this equation as the first order Upwind scheme plus a second order correction term. Assuming that $v_{i+1/2} > 0$, the Lax-Wendroff scheme can be written as

$$u_i^{n+1} = u_i^n - s\left[f_i^n - f_{i-1}^n\right] - \frac{s}{2}\left(1 - v_{i+1/2}\right)\left(f_{i+1}^n - f_i^n\right) + \frac{s}{2}\left(1 - v_{i-1/2}\right)\left(f_i^n - f_{i-1}^n\right)$$

and we may obtain

$$F_L(u;i) = f_i^n$$

and

$$F_H(u;i) = \frac{1}{2}\left(1 - v_{i+1/2}\right)\left[f_{i+1}^n - f_i^n\right].$$

Here, $F_L(u;i)$ represents the Upwind scheme and $F_H(u;i)$ represents the second order correction term. Similarly, assuming that $v_{i+1/2} < 0$, we may obtain

$$F_L(u;i) = f_{i+1}^n$$

and

$$F_H(u;i) = -\frac{1}{2}\left(1 + v_{i+1/2}\right)\left[f_{i+1}^n - f_i^n\right].$$

Hence, we may obtain

$$u_i^{n+1} = u_i^n - s\left[F(u;i) - F(u;i-1)\right]$$

where

$$F(u;i) = F_L(u;i) + F_H(u;i)\phi_i$$

and

$$F_L(u;i) = \begin{cases} f_i^n & \text{if } v_{i+1/2} > 0 \\ f_{i+1}^n & \text{if } v_{i+1/2} < 0 \end{cases}$$

$$F_H(u;i) = \frac{1}{2}\begin{cases} \left(1 - v_{i+1/2}\right)\left(f_{i+1}^n - f_i^n\right) & \text{if } v_{i+1/2} > 0 \\ -\left(1 + v_{i+1/2}\right)\left(f_{i+1}^n - f_i^n\right) & \text{if } v_{i+1/2} < 0 \end{cases}.$$

We now need to measure the smoothness of the data so that we may choose the flux-limiter to obtain second order accuracy and the TVD property. The TVD property is called the Total Variational Diminishing property and will not be discussed in full in this thesis. However, we will show some regions of TVD for the flux limiter, ϕ_i.

In order to measure the smoothness of the data, we could look at the ratio of consecutive gradients.

$$\theta_i = \frac{u_{j+1} - u_j}{u_{i+1} - u_i}$$

where $j = i - \text{sgn}(v_{i+1/2})$. Here, if θ_i close to 1 then the data is considered to be smooth, but if θ_i is far from 1, then there are kinks in the data at u_i. We can now take ϕ_i to be a function of θ_i, i.e.

$$\phi_i = \phi(\theta_i)$$

where ϕ is a given function. Now, we require the flux-limiter to be of second order and to satisfy the TVD property. If the flux-limiter is to satisfy the TVD property, we must first assume that

$$\phi_i = 0 \text{ if } \theta_i \leq 0$$

and we must choose the flux-limiter to lie in the TVD region shown in Figure 2-7. But to obtain second order accuracy, the flux-limiter must pass through $\phi(1) = 1$ and lie in the region shown in Figure 2-8. Roe's Superbee flux-limiter

$$\phi(\theta_i) = \max(0, \min(2\theta_i, 1), \min(\theta_i, 2))$$

satisfies the second order TVD region as shown in Figure 2-9 and is therefore second order accurate and Figure 2-6 shows that the Lax-Wendroff scheme with Superbee flux-limiter method is considerably more accurate than without the Superbee flux-limiter method. See Sweby[12] for a more in-depth analysis on flux-limiter methods.

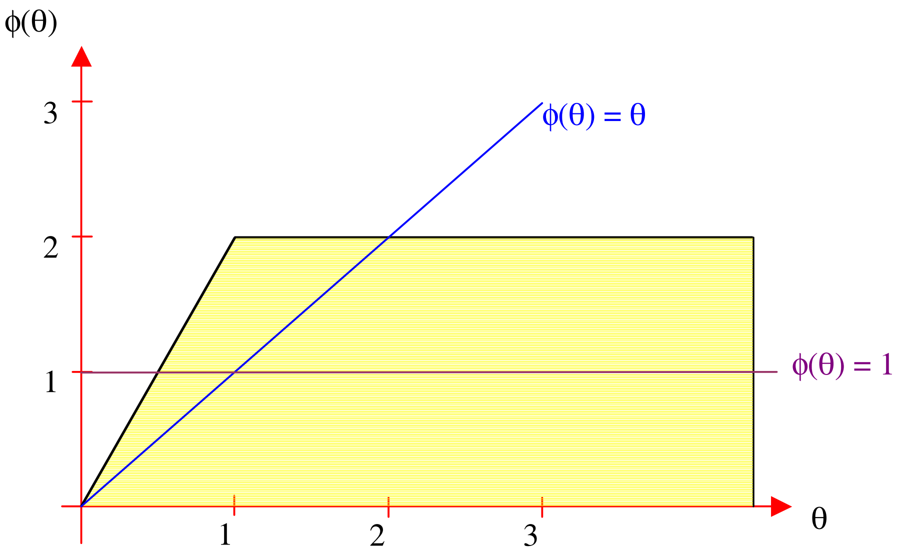


Figure 2-7: TVD region for finite difference schemes.

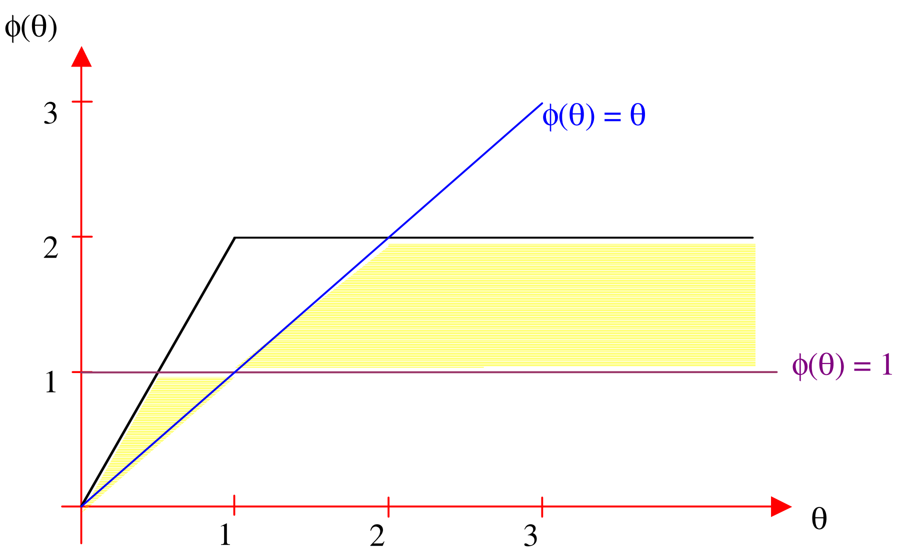


Figure 2-8: Second order TVD region for finite difference schemes.

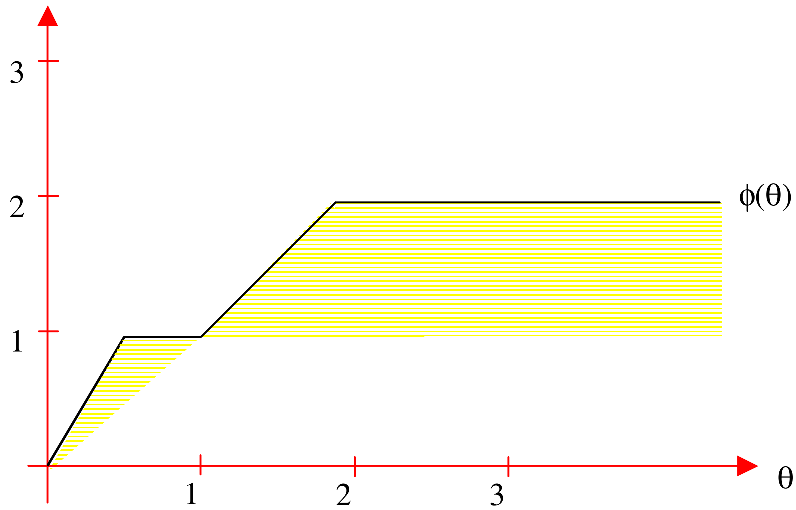


Figure 2-9: Superbee flux-limiter for finite difference schemes.

Table 2-3 lists a few flux-limiters, which satisfy the TVD property and are second order accurate.

Name of Flux-limiter	$\phi(\theta)$
Minmod	$\phi(\theta) = \max(0,\min(1,\theta))$
Roe's Superbee	$\phi(\theta) = \max(0,\min(2\theta,1),\min(\theta,2))$
van Leer	$\phi(\theta) = \frac{\lvert\theta\rvert + \theta}{1+\lvert\theta\rvert}$
van Albada	$\phi(\theta) = \frac{\theta^2 + \theta}{1+\theta^2}$

Table 2-3: Some Flux-limiters for second order schemes.

Throughout Chapter 2, we have discussed the finite difference technique for approximating the scalar conservation law and, in particular, the linear advection equation. However, sometimes the right hand side of (2.1) is not equal to zero but instead, a source term is present which can cause difficulties in approximating the solution accurately. In the next chapter, we will consider such a case, where a source term is now present on the right hand side of (2.1).

3 Conservation Law with Source Term R(x,t)

In Chapter 2, we discussed a variety of finite difference schemes for numerically approximating conservation laws and the linear advection equation. We also discovered that a number of problems occur when numerically approximating conservation laws even when a source term is not present. In this chapter, we will discuss some numerical techniques for solving conservation laws when a source term is present. However, in this chapter we will only consider source terms that are functions of x and t only, i.e.

$$\frac{\partial u}{\partial t}+\frac{\partial f(u)}{\partial x}=R(x,t) \tag{3.1}$$

where $R(x,t)$ is the source term. In this chapter, we will also use the linear advection equation with c = 1 and source term

$$R(x,t)=\begin{cases}-e^{-t} & \text{if } x\le 0.3+t\\ 0 & \text{if } x>0.3+t\end{cases},$$

i.e.

$$\frac{\partial u}{\partial t}+\frac{\partial u}{\partial x}=\begin{cases}-e^{-t} & \text{if } x\le 0.3+t\\ 0 & \text{if } x>0.3+t\end{cases}, \tag{3.2}$$

with initial data

$$u(x,0)=\begin{cases}1 & \text{if } x\le 0.3\\ 0 & \text{if } x>0.3\end{cases} \tag{3.3}$$

whose exact solution is

$$u(x,t)=\begin{cases} e^{-t} & \text{if } x\le 0.3+t \\ 0 & \text{if } x>0.3+t \end{cases}$$

as a test problem to illustrate some numerical results.

3.1 Basic Approach

The most basic finite difference approach used to numerically approximate (3.1) is to 'add' the source term to a scheme that numerically approximates the conservation law without source term (2.1). For example, if we use a forward difference approximation in time, a central difference in space and assume the source term to be an approximation at (*i,n*) then (3.1) becomes

$$\frac{u_i^{n+1}-u_i^n}{\Delta t}+\frac{f_{i+1}^n-f_{i-1}^n}{2\Delta x}=R_i^n$$

and by re-arranging we may obtain

$$u_i^{n+1}=u_i^n-\frac{s}{2}\left[f_{i+1}^n-f_{i-1}^n\right]+\Delta t R_i^n\,.$$

This central scheme is unconditionally unstable, by using the average

$$u_i^n=\frac{1}{2}\left(u_{i+1}^n+u_{i-1}^n\right)$$

we may obtain

$$u_i^{n+1}=\frac{1}{2}\left(u_{i+1}^n+u_{i-1}^n\right)-\frac{s}{2}\left[f_{i+1}^n-f_{i-1}^n\right]+\Delta t R_i^n$$

which is the first order Lax-Friedrichs scheme with the source term 'added' on

$$u_i^{LF}=\frac{1}{2}\left(u_{i+1}^n+u_{i-1}^n\right)-\frac{s}{2}\left[f_{i+1}^n-f_{i-1}^n\right]$$

where

$$u_i^{n+1}=u_i^{LF}+\Delta t R_i^n\,.$$

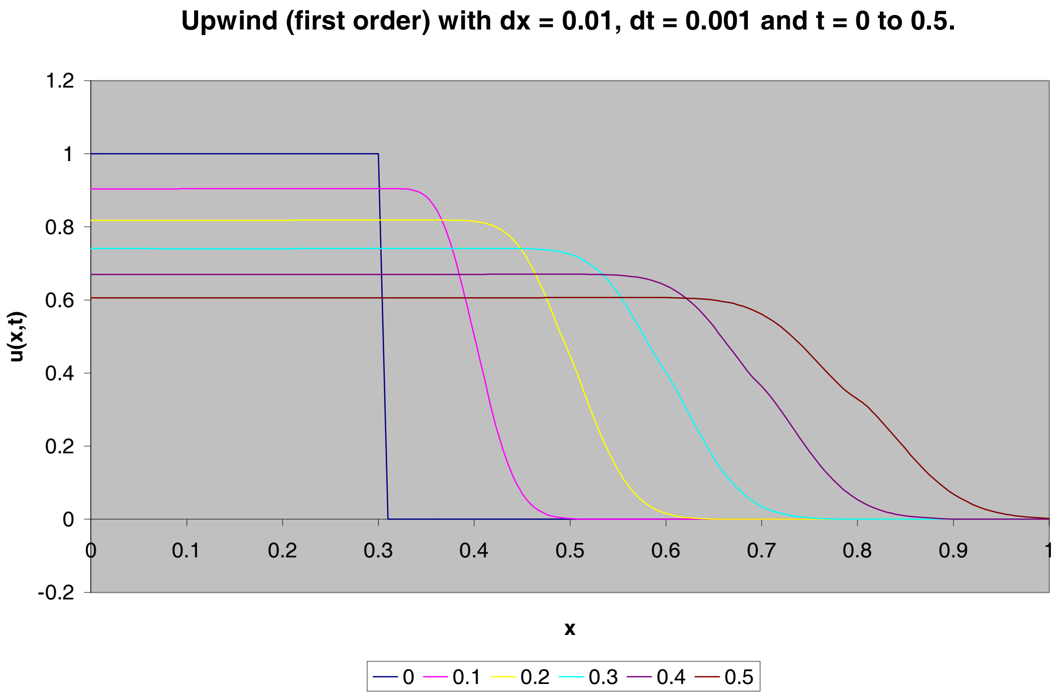


Figure 3-1: The Upwind scheme with source term ‘added’ on.

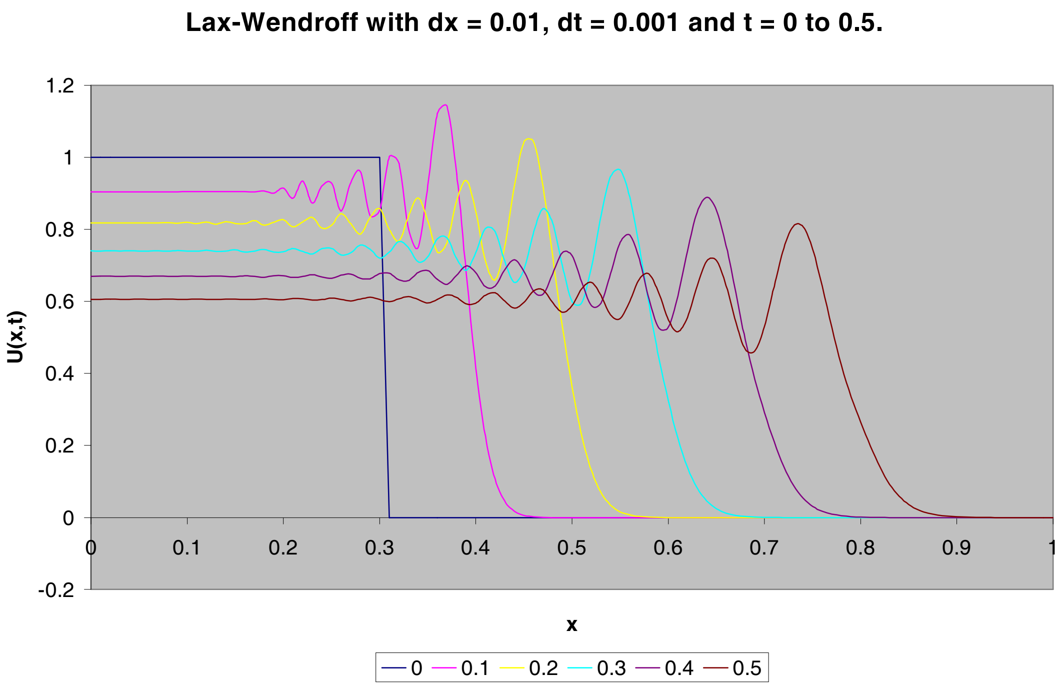


Figure 3-2: The Lax-Wendroff scheme with source term ‘added’ on.

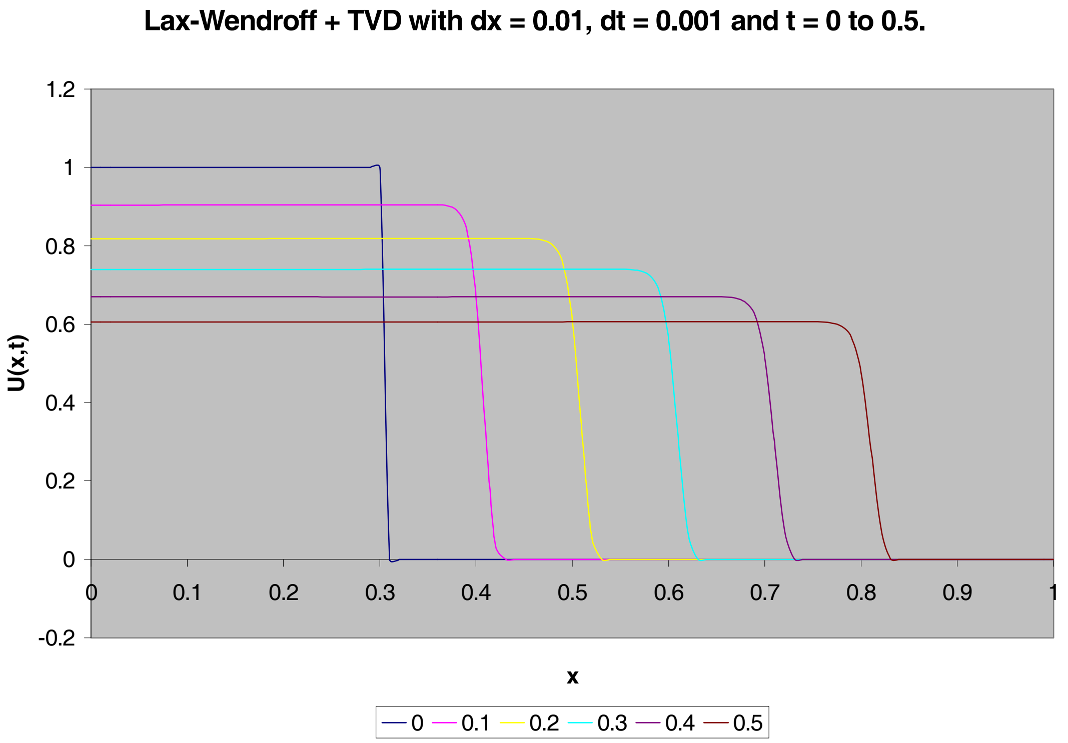


Figure 3-3: The Lax-Wendroff scheme with Superbee flux-limiter and source term 'added'.

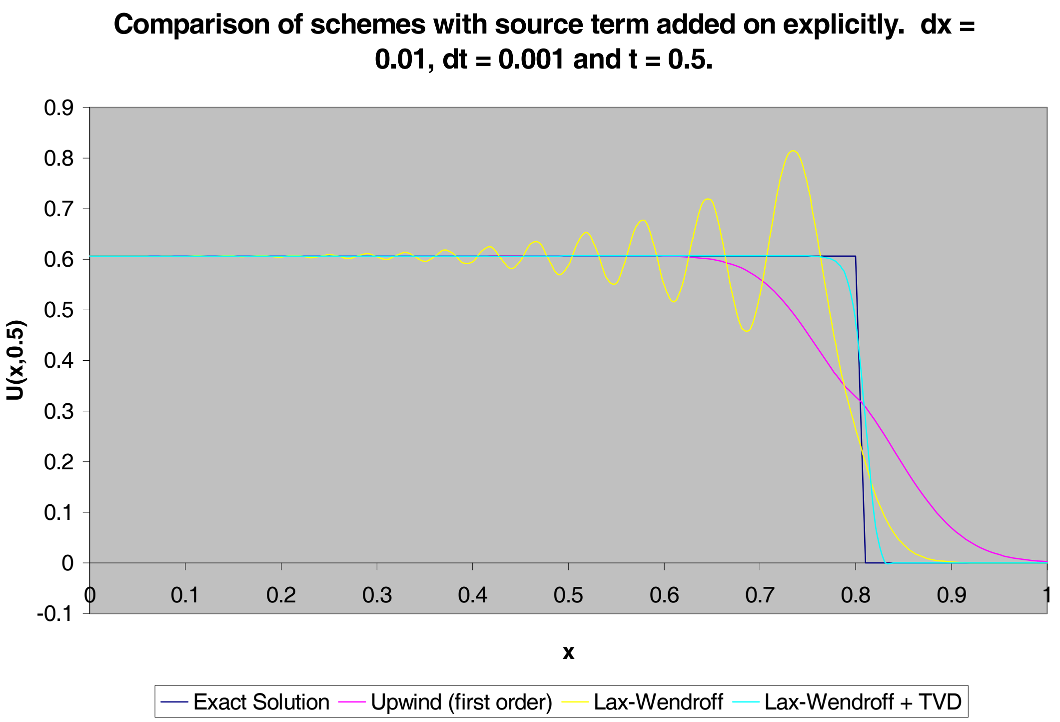


Figure 3-4: Comparison of different schemes with the source term 'added' on.

This approach will work with all schemes discussed in Chapter 2 and, in general

$$u_i^{n+1} = u_i^{SCHEME} + \Delta t R_i^n . \tag{3.4}$$

Here, u_i^{SCHEME} represents a numerical scheme of the conservation law without a source term present. Also, by assuming the source term to be an approximation at ($i,n+1$), we can obtain a semi-implicit scheme

$$u_i^{n+1} = u_i^{SCHEME} + \Delta t R_i^{n+1} . \tag{3.5}$$

Figure 3-1, Figure 3-2, Figure 3-3 and Figure 3-4 are all results of schemes of the form (3.4) applied to (3.2) with initial data (3.3). Figure 3-1 shows the Upwind scheme with the source term ‘added’, Figure 3-2 shows the Lax-Wendroff scheme with source term ’added’ and Figure 3-3 shows the Lax-Wendroff scheme with Superbee flux-limiter and source term ‘added’. Figure 3-4 shows the Upwind scheme, Lax-Wendroff scheme and Lax-Wendroff scheme with Superbee flux-limiter, all with the source term explicitly ‘added’ on. Here, we can see that the Upwind scheme with source term ‘added’ suffers badly from dissipation and that the Lax-Wendroff scheme with source term added suffers badly from dispersion resulting in very large oscillations being present. The most accurate scheme was the Lax-Wendroff scheme with Superbee flux-limiter and source term ‘added’. In addition, we can see all schemes are conservative since the discontinuity was moved at the correct wave speed.

3.2 Lax-Wendroff Approach

We can also use the Lax-Wendroff approach that we used in Chapter 2, Section 2, to approximate the scalar conservation law with source term. However, we must first re-write (2.11) to include the source term. Now, we can re-write (3.1) as

$$\frac{\partial u}{\partial t} = R(x,t) - \frac{\partial f(u)}{\partial x} \tag{3.6}$$

$$\Rightarrow \quad \frac{\partial^2 u}{\partial t^2} = \frac{\partial R}{\partial t} - \frac{\partial^2 f}{\partial t \partial x} = \frac{\partial R}{\partial t} - \frac{\partial^2 f}{\partial x \partial t} = \frac{\partial R}{\partial t} - \frac{\partial\left(\frac{\partial f}{\partial u}\frac{\partial u}{\partial t}\right)}{\partial x} = \frac{\partial R}{\partial t} - \frac{\partial\left(a(u)\left(R - \frac{\partial f}{\partial x}\right)\right)}{\partial x}$$

and we may obtain

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial R}{\partial t} - \frac{\partial(a(u)R)}{\partial x} + \frac{\partial\left(a(u)\frac{\partial f}{\partial x}\right)}{\partial x}. \tag{3.7}$$

Now, by using Taylor's theorem

$$u_i^{n+1} \approx u_i^n + \Delta t\left[\frac{\partial u}{\partial t}\right]_i^n + \frac{\Delta t^2}{2}\left[\frac{\partial^2 u}{\partial t^2}\right]_i^n + \ldots \tag{3.8}$$

and substituting (3.6) and (3.7) into (3.8) gives

$$u_i^{n+1} \approx u_i^n + R\Delta t - \Delta t\frac{\partial f}{\partial x} + \frac{\Delta t^2}{2}\left[\frac{\partial R}{\partial t} - \frac{\partial(a(u)R)}{\partial x} + \frac{\partial\left(a(u)\frac{\partial f}{\partial x}\right)}{\partial x}\right] + \ldots$$

and by using central difference approximations in space and assuming that both are approximations at (i,n) then

$$u_i^{n+1} = u_i^n - \frac{s}{2}\left(f_{i+1}^n - f_{i-1}^n\right) + \frac{s}{2}\left[v_{i+1/2}\left(f_{i+1}^n - f_i^n\right) - v_{i-1/2}\left(f_i^n - f_{i-1}^n\right)\right]$$
$$+ \Delta t\left[R + \frac{\Delta t}{2}\left(\frac{dR}{dt} - \frac{d(a(u)R)}{dx}\right)\right].$$

Hence, by using a forward difference approximation in space and time, we may obtain

$$u_i^{n+1} = u_i^n - \frac{s}{2}\left(f_{i+1}^n - f_{i-1}^n\right) + \frac{s}{2}\left[\nu_{i+1/2}\left(f_{i+1}^n - f_i^n\right) - \nu_{i-1/2}\left(f_i^n - f_{i-1}^n\right)\right]$$
$$+ \Delta t\left[R_i^n + \frac{\Delta t}{2}\left(\frac{R_i^{n+1} - R_i^n}{\Delta t} - \frac{a_{i+1/2}R_{i+1/2}^n - a_{i-1/2}R_{i-1/2}^n}{\Delta x}\right)\right]$$

and by re-arranging

$$u_i^{n+1} = u_i^n - \frac{s}{2}\left(f_{i+1}^n - f_{i-1}^n\right) + \frac{s}{2}\left[\nu_{i+1/2}\left(f_{i+1}^n - f_i^n\right) - \nu_{i-1/2}\left(f_i^n - f_{i-1}^n\right)\right]$$
$$+ \frac{\Delta t}{2}\left[R_i^n + R_i^{n+1}\right] - \frac{\Delta t}{4}\left[\nu_{i+1/2}\left(R_{i+1}^n - R_i^n\right) - \nu_{i-1/2}\left(R_i^n - R_{i-1}^n\right)\right] \tag{3.9}$$

we may obtain a second order approximation to (3.1), which is based on the Lax-Wendroff scheme. We can also apply flux-limiter methods to (3.9) by re-writing (3.9) as

$$u_i^{n+1} = u_i^n - s\left[F(u;i) - F(u;i-1)\right] + \frac{\Delta t}{2}\left[R_i^n + R_i^{n+1}\right] - \frac{\Delta t}{4}\left[\nu_{i+1/2}\left(R_{i+1}^n - R_i^n\right) - \nu_{i-1/2}\left(R_i^n - R_{i-1}^n\right)\right]$$

where

$$F(u;i) = F_L(u;i) + F_H(u;i)\phi_i$$

and

$$F_L(u;i) = \begin{cases} f_i^n & \text{if } \nu_{i+1/2} > 0 \\ f_{i+1}^n & \text{if } \nu_{i+1/2} < 0 \end{cases}$$

$$F_H(u;i) = \frac{1}{2}\begin{cases} \left(1-\nu_{i+1/2}\right)\left(f_{i+1}^n - f_i^n\right) & \text{if } \nu_{i+1/2} > 0 \\ -\left(1+\nu_{i+1/2}\right)\left(f_{i+1}^n - f_i^n\right) & \text{if } \nu_{i+1/2} < 0 \end{cases}.$$

where ϕ_i denotes the flux-limiter method described in Chapter 2, Section 5 and we could use any of the flux-limiters in Table 2-3 to obtain a second order flux-limiter method. If we now apply the Lax-Wendroff approach, without a flux-limiter method, (3.9) to the test problem (3.2) with initial data (3.3), we may obtain the results in Figure 3-5.

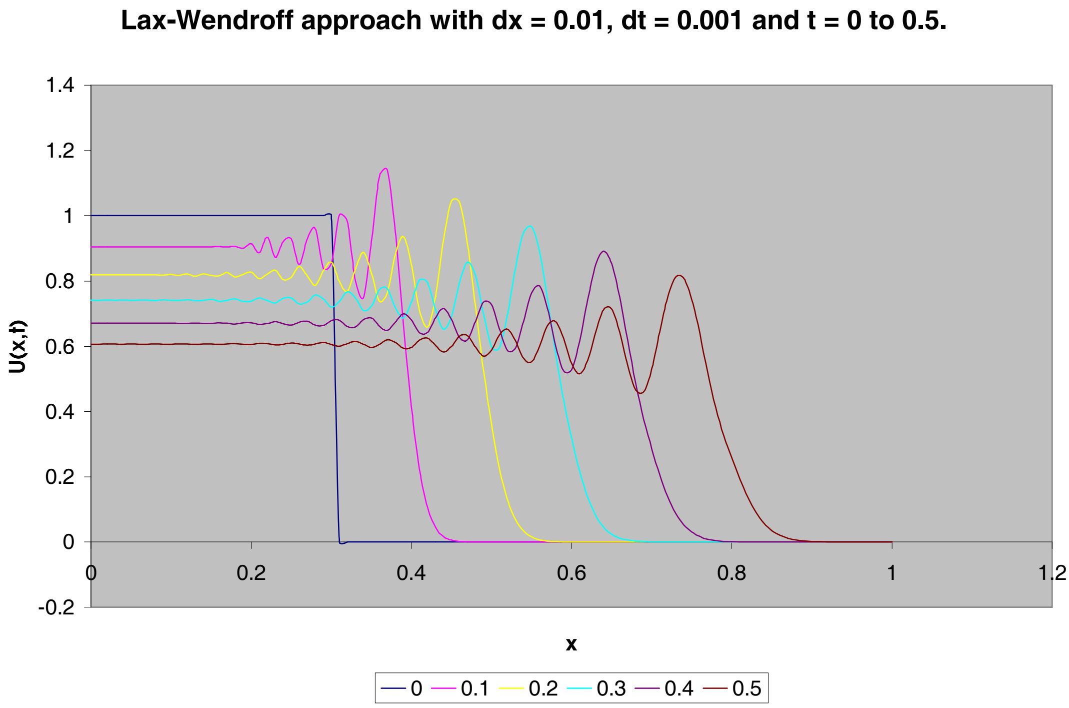


Figure 3-5: The Lax-Wendroff approach for advection equation with source term.

Figure 3-5 shows practically the same results as Figure 3-2 where the source term was 'added' to the Lax-Wendroff scheme. This is because the source term is a known function of x and t so all approximations of the source term will be extremely accurate. However, when the source term is also a function of u, the approximations of the source term are not as accurate making the two schemes accuracy change dramatically, as we will see later.

3.3 MPDATA approach

Smolarkiewicz and Margolin[3] derived an algorithm to approximate the advection transport equation (3.7) called MPDATA. MPDATA is a Multidimensional Positive Definite Advection Transport Algorithm and approximates the advection equation (2.3) with a source term present

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = R(x,t) \tag{3.10}$$

which is also known as the advection transport equation. The author states that this algorithm uses a similar approach to that of the Lax-Wendroff, which can be viewed as the Upwind scheme minus an error estimate, but exploits special properties of the Upwind scheme.

3.3.1 Basic MPDATA

Before we can discuss the MPDATA algorithm for (3.10), we must first look at the most basic MPDATA algorithm, which is based on the advection equation without source term,

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0 \tag{3.11}$$

If we assume that u is nonnegative, then the basic MPDATA algorithm is the Upwind scheme (2.4) re-written in flux form

$$u_i^{n+1} = u_i^n - \left[F\left(u_i^n, u_{i+1}^n, C\right) - F\left(u_{i-1}^n, u_i^n, C\right)\right] \tag{3.12}$$

where

$$F\left(u_L, u_R, C\right) = C^+ u_L + C^- u_R,$$

$$C = c\frac{\Delta t}{\Delta x},\ C^+ = \frac{1}{2}\left(C + |C|\right) \text{ and } C^- = \frac{1}{2}\left(C - |C|\right).$$

This scheme is only first order and we require a second order scheme, but if we look at the truncation error of (3.12)

$$\mathrm{T}_i^n = \frac{\Delta x^2}{2\Delta t}\left(|C| - C^2\right)\frac{\partial^2 u}{\partial x^2} + O(\Delta x^3)$$

and since $c > 0$

$$\mathrm{T}_i^n = \frac{c}{2}\left(\Delta x - c\Delta t\right)\frac{\partial^2 u}{\partial x^2} + O(\Delta x^2) + O(\Delta t^2).$$

Here we can see that (3.12) is a better approximation of the advection-diffusion equation,

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = \frac{\partial\left(K\frac{\partial u}{\partial x}\right)}{\partial x} \quad \text{where} \quad K = \frac{\Delta x^2}{2\Delta t}\left(|C| - C^2\right),$$

and is only first order in space and time. However, we can construct a numerical estimate of the error and subtract it from (3.12) which will make the scheme second order. This approach is similar to that of the Lax-Wendroff scheme for the advection equation, which uses central differences to approximate the right hand side of (3.12) whereas MPDATA uses special properties of the Upwind scheme for approximating and compensating the error. We can re-write the error term as

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = \frac{\partial\left(v^{(1)}u\right)}{\partial x}$$

where

$$v^{(1)} = \frac{\Delta x^2}{2\Delta t}\left(|C| - C^2\right)\frac{1}{u}\frac{\partial u}{\partial x}.$$

is a pseudo velocity. Then by using

$$u_{\mathrm{i}+1/2} = \frac{1}{2}\left(u_{\mathrm{i}+1} + u_i\right) \text{ and } \frac{\partial u}{\partial x} = \frac{u_{i+1}^{(1)} - u_i^{(1)}}{\Delta x},$$

where the superscript $^{(1)}$ denotes the first approximation of the advection equation (3.11), we may obtain the first order accurate approximation

$$V_{i+1/2}^{(1)} = \frac{\Delta x}{\Delta t}\left(|C| - C^2\right)\left[\frac{u_{i+1}^{(1)} - u_i^{(1)}}{u_{i+1}^{(1)} + u_i^{(1)}}\right]$$

of the pseudo velocity. In order to obtain a second order approximation, we subtract the error in the second pass

$$u_i^{n+1} = u_i^{(1)} - \left[F\left(u_i^{(1)}, u_{i+1}^{(1)}, V_{i+1/2}^{(1)}\right) - F\left(u_{i-1}^{(1)}, u_i^{(1)}, V_{i-1/2}^{(1)}\right)\right].$$

Hence, we may now obtain the basic MPDATA algorithm

$$u_i^{n+1} = u_i^{(1)} - \left[F\left(u_i^{(1)}, u_{i+1}^{(1)}, V_{i+1/2}^{(1)}\right) - F\left(u_{i-1}^{(1)}, u_i^{(1)}, V_{i-1/2}^{(1)}\right)\right] \tag{3.13}$$

where the pseudo velocity is

$$V_{i+1/2}^{(1)} = \frac{\Delta x}{\Delta t}\left(|C| - C^2\right)\left[\frac{u_{i+1}^{(1)} - u_i^{(1)}}{u_{i+1}^{(1)} + u_i^{(1)}}\right] \tag{3.14}$$

and the first order approximation is

$$u_i^{(1)} = u_i^n - \left[F\left(u_i^n, u_{i+1}^n, C\right) - F\left(u_{i-1}^n, u_i^n, C\right)\right].$$

So far we have only considered the advection equation with u nonnegative but the basic MPDATA algorithm can also be updated for u to be of variable sign. This is achieved by using the pseudo velocity

$$V_{i+1/2}^{(1)} = \frac{\Delta x}{\Delta t}\left(|C| - C^2\right)\left[\frac{\left|u_{i+1}^{(1)}\right| - \left|u_i^{(1)}\right|}{\left|u_{i+1}^{(1)}\right| + \left|u_i^{(1)}\right|}\right]$$

instead of (3.14). Also, we can apply flux-limiters to the basic MPDATA algorithm by replacing (3.13) with

$$u_i^{n+1} = u_i^{(1)} - \left[F\left(u_i^{(1)}, u_{i+1}^{(1)}, V_{i+1/2}^{(1)}\right)\phi_i - F\left(u_{i-1}^{(1)}, u_i^{(1)}, V_{i-1/2}^{(1)}\right)\phi_{i-1}\right] \tag{3.15}$$

where ϕ_i denotes the flux-limiter method described in Chapter 2, Section 5.

3.3.2 MPDATA Approach for Advection Equation with Source Term R(x,t)

So far we have only discussed the basic MPDATA scheme for the advection equation but MPDATA can also be adapted to approximate the advection transport equation (3.10)

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = R(x,t).$$

If we use a forward difference approximation in time and assume it is an approximation at (i,n). Also, by assuming the source term is an approximation at $(i,n+½)$ we may obtain

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + c\frac{\partial u}{\partial x} = R_i^{n+1/2} \tag{3.16}$$

Now, by using Taylor's theorem

$$u_i^{n+1} \approx u_i^n + \Delta t \left[\frac{\partial u}{\partial t}\right]_i^n + \frac{\Delta t^2}{2}\left[\frac{\partial^2 u}{\partial t^2}\right]_i^n + \ldots$$

$$R_i^{n+1/2} \approx R_i^n + \frac{\Delta t}{2}\left[\frac{\partial R}{\partial t}\right]_i^n + \frac{\Delta t^2}{8}\left[\frac{\partial^2 R}{\partial t^2}\right]_i^n + \ldots$$

and by substituting into (3.16) we may obtain

$$\frac{\partial u}{\partial t} + \frac{\Delta t}{2}\frac{\partial^2 u}{\partial t^2} + c\frac{\partial u}{\partial x} = R_i^n + \frac{\Delta t}{2}\left[\frac{\partial R}{\partial t}\right]_i^n + O(\Delta t^2), \tag{3.17}$$

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = R_i^n + O(\Delta t\ ), \tag{3.18}$$

and

$$\frac{\partial^2 u}{\partial t^2} + c\frac{\partial\left(\frac{\partial u}{\partial t}\right)}{\partial x} = \frac{\partial R}{\partial t} + O(\Delta t\ ). \tag{3.19}$$

Substituting (3.18) into (3.19) and re-arranging gives

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial R}{\partial t} - c\frac{\partial R}{\partial x} + c^2\frac{\partial^2 u}{\partial x^2} + O(\Delta t\ )$$

and substituting this into (3.17) we may obtain

$$\frac{\partial u}{\partial t} + \frac{\Delta t}{2}\left[c^2\frac{\partial^2 u}{\partial x^2} - c\frac{\partial R}{\partial x}\right] + c\frac{\partial u}{\partial x} = R + O(\Delta t^2).$$

Whence we may obtain

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = R + c\frac{\Delta t}{2}\frac{\partial R}{\partial x} - c^2\frac{\Delta t}{2}\frac{\partial^2 u}{\partial x^2} + O(\Delta t^2). \tag{3.20}$$

The first two terms on the right hand side of (3.20) shows the error due to the source term and the third term shows the error due to the method. Here, the second term on the right hand side of (3.20) can blow up creating very inaccurate numerical results, especially if the source term is stiff (see Chapter 5, Section 1), but MPDATA compensates for this term making the scheme considerably more accurate. The

MPDATA approach which numerically approximates (3.10) is derived by assuming the source term approximation to be at (i,n+½) giving

$$u_i^{n+1} = MPDATA\left(u_i^n, C\right) + \Delta t R_i^{n+1/2}$$

where $MPDATA\left(u_i^n, C\right)$ corresponds to the basic MPDATA algorithm discussed in the previous section (3.3.1). Now, by using the average,

$$R_i^{n+1/2} = \frac{1}{2}\left[R_i^{n+1} + R_i^n\right]$$

we may obtain

$$u_i^{n+1} = MPDATA\left(u_i^n, C\right) + \frac{\Delta t}{2}\left(R_i^{n+1} + R_i^n\right)$$

and by advecting $\frac{\Delta t}{2} R_i^n$ we may obtain the MPDATA scheme for approximating the advection transport equation

$$u_i^{n+1} = MPDATA\left(u_i^n + \frac{\Delta t}{2} R_i^n, C\right) + \frac{\Delta t}{2} R_i^{n+1} \qquad (3.21)$$

where $MPDATA\left(u_i^n + \frac{\Delta t}{2} R_i^n, w_{i+1/2}^{n+1/2}\right)$ corresponds to the basic MPDATA algorithm discussed in the previous section (3.3.1). Here, by advecting the auxiliary field, $u_i^n + \frac{\Delta t}{2} R_i^n$, the terms in the truncation error (3.20) due to the source term do not blow up. Now, by applying (3.21) and (3.9) to the test problem (3.2), with initial data (3.3), we may obtain the numerical results in Figure 3-6 and Figure 3-7. Both Figure 3-6 and Figure 3-7 show that the MPDATA scheme suffers from a lot less oscillations behind the discontinuity than the Lax-Wendroff approach, which means that MPDATA is less dispersive than the Lax-Wendroff approach. Also, Figure 3-7 shows that, near the discontinuity, the MPDATA approach is a lot less accurate than the Lax-Wendroff approach.

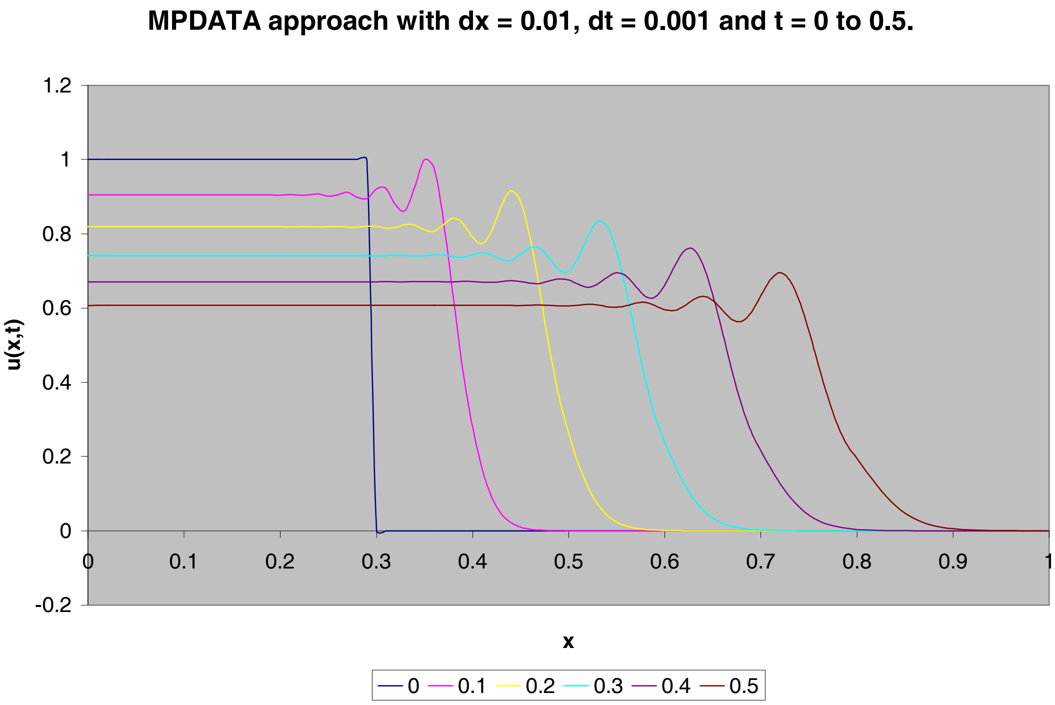


Figure 3-6: MPDATA approach for advection transport equation.

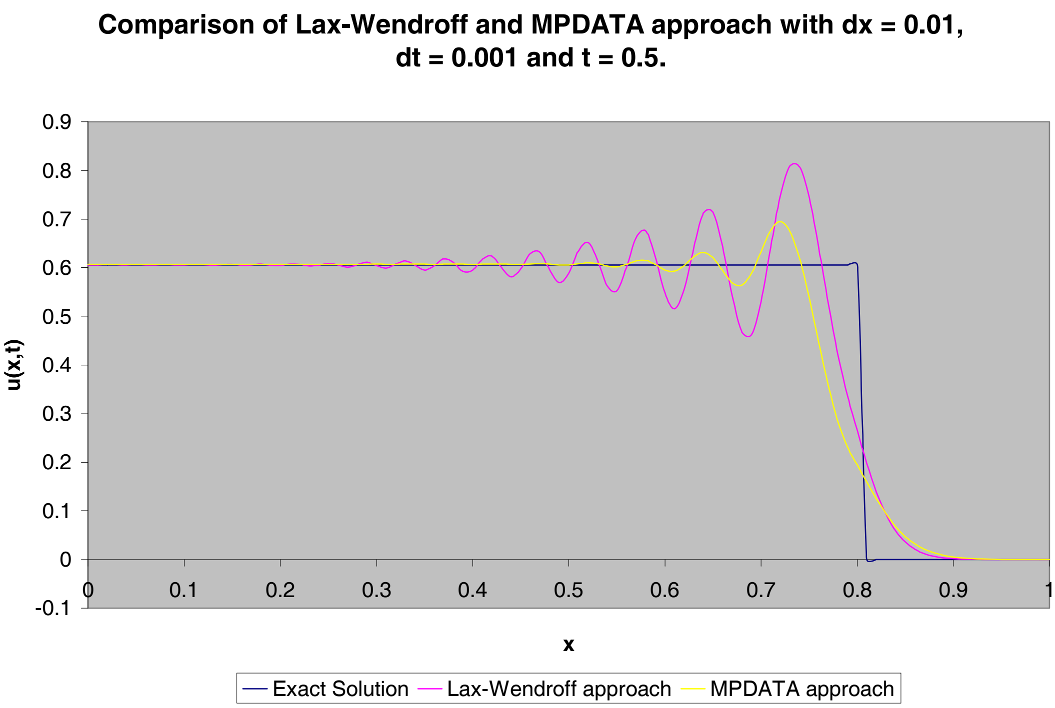


Figure 3-7: Comparison between Lax-Wendroff approach and MPDATA.

So, the MPDATA approach is considerably less dispersive than the Lax-Wendroff approach but, near the discontinuity, the Lax-Wendroff is considerably more accurate. But overall, the MPDATA approach is a lot more accurate than the Lax-Wendroff approach for approximating (3.10). In general, the MPDATA approach (3.21) is very accurate when numerically approximating the advection-transport equation (3.10).

3.3.3 MPDATA Approach for Conservation Law with Source Term R(x,t)

So far we have only looked at MPDATA algorithms for the advection-transport equation. Let us now consider MPDATA algorithms for the scalar conservation law with source term present (3.1), i.e.

$$\frac{\partial u}{\partial t}+\frac{\partial f(u)}{\partial x}=R(x,t)$$

MPDATA can be adapted to approximate (3.1) by considering the velocity c of the advection-transport equation to no longer be a constant but to be a function of u instead, i.e.

$$\frac{\partial u}{\partial t}+\frac{\partial \left(w(u)u\right)}{\partial x}=R(x,t)$$

where $w(u)=\frac{u}{2}$ for inviscid burgers equation, etc. The basic MPDATA algorithm for the conservation law without source term now takes the form

$$u_i^{n+1}=u_i^{(1)}-\left[F\left(u_i^{(1)},u_{i+1}^{(1)},V_{i+1/2}^{(1)}\right)-F\left(u_{i-1}^{(1)},u_i^{(1)},V_{i-1/2}^{(1)}\right)\right] \tag{3.22}$$

where the pseudo velocity is

$$V_{i+1/2}^{(1)}=\frac{\Delta x}{\Delta t}\left(\left|w_{i+1/2}^{n+1/2}\right|-\left[w_{i+1/2}^{n+1/2}\right]^2\right)\left[\frac{u_{i+1}^{(1)}-u_i^{(1)}}{u_{i+1}^{(1)}+u_i^{(1)}}\right]-w_{i+1/2}^{n+1/2}\left[w_{i+3/2}^{n+1/2}-w_{i-1/2}^{n+1/2}\right] \tag{3.23}$$

and the first order approximation is

$$u_i^{(1)}=u_i^n-\left[F\left(u_i^n,u_{i+1}^n,w_{i+1/2}^{n+1/2}\right)-F\left(u_{i-1}^n,u_i^n,w_{i-1/2}^{n+1/2}\right)\right].$$

However, $w_{i+1/2}^{n+1/2}$ is unknown since w is a function of u and u is only known at the grid points (i,n). We could approximate $w_{i+1/2}^{n+1/2}$ by using the average

$$w_{i+1/2}^{n+1/2} = \frac{1}{2}\left(w_{i+1/2}^{n+1} + w_{i+1/2}^{n}\right)$$

or by using linear interpolation

$$w_{i+1/2}^{n+1/2} = \frac{1}{2}\left(3\,w_{i+1/2}^{n} - w_{i+1/2}^{n-1}\right).$$

If we approximate by using linear interpolation, the method would require another scheme to initially start the algorithm off, since we require a value of u at $(i,n\text{-}1)$, but if we use the average, the algorithm becomes impractical since we require the value of u at $(i,n+1)$. So far we have only considered the most basic MPDATA algorithm for the conservation law without source term and have encountered a lot of difficulties. If we now consider a source term then the corresponding MPDATA scheme is

$$u_i^{n+1} = MPDATA\left(u_i^n + \frac{\Delta t}{2} R_i^n, w_{i+1/2}^{n+1/2}\right) + \frac{\Delta t}{2} R_i^{n+1} \tag{3.24}$$

where $MPDATA\left(u_i^n + \frac{\Delta t}{2} R_i^n, w_{i+1/2}^{n+1/2}\right)$ corresponds to the basic MPDATA algorithm, for the conservation law without source term, discussed in Section 3.1. However, care must be taken when using this scheme since if the source term is a function of u then even more difficulties arise when using this algorithm as we will see later.

3.4 Comparison of Schemes Using Test Problem

Now, by using the test problem (3.2) with initial data (3.3), we can obtain the numerical results in Figure 3-8 and Figure 3-9 and compare the numerical solution of the three approaches discussed throughout this chapter with the exact solution. The three approaches are:

1. Upwind (first order) with the source term 'added' on.
2. Lax-Wendroff approach with or without Superbee flux-limiter.
3. MPDATA approach with or without Superbee flux-limiter.

Figure 3-8 compares the numerical solution of the three different approaches with the exact solution at t = 0.5. Figure 3-9 compares the 'true' error of the numerical solution of the three different approaches at t = 0.5. Figure 3-8 shows us that the Upwind scheme suffers from dissipation as expected and that the MPDATA approach without TVD gives less oscillations behind the discontinuity than the Lax-Wendroff scheme without TVD. Figure 3-9 shows that the most accurate approach overall is the Lax-Wendroff approach with Superbee flux-limiter followed by the MPDATA approach with Superbee flux-limiter. We can also see that near the discontinuity, the MPDATA approach is less accurate than the Lax-Wendroff approach, but away from the discontinuity, the MPDATA approach is considerably more accurate than the Lax-Wendroff approach. However, the MPDATA approach will be not so accurate when applied to the scalar conservation law since the MPDATA approach would require special starting procedures and approximates approximations. This is due to the MPDATA approach being specifically derived to approximate the linear advection equation with source term present and not the scalar conservation law.

Throughout this chapter, we have discussed three main approaches which approximate the scalar conservation law with source term, which is a function of x and t and we have obtained some very accurate results. However, we have only considered known source terms and when the source term has to be approximated, the different approaches discussed in this chapter are not so accurate. In the next chapter, we will consider some methods for approximating the scalar conservation law with the source term being a function of u also. Since the source term is now a function of u as well, we will now need to approximate the source term as well.

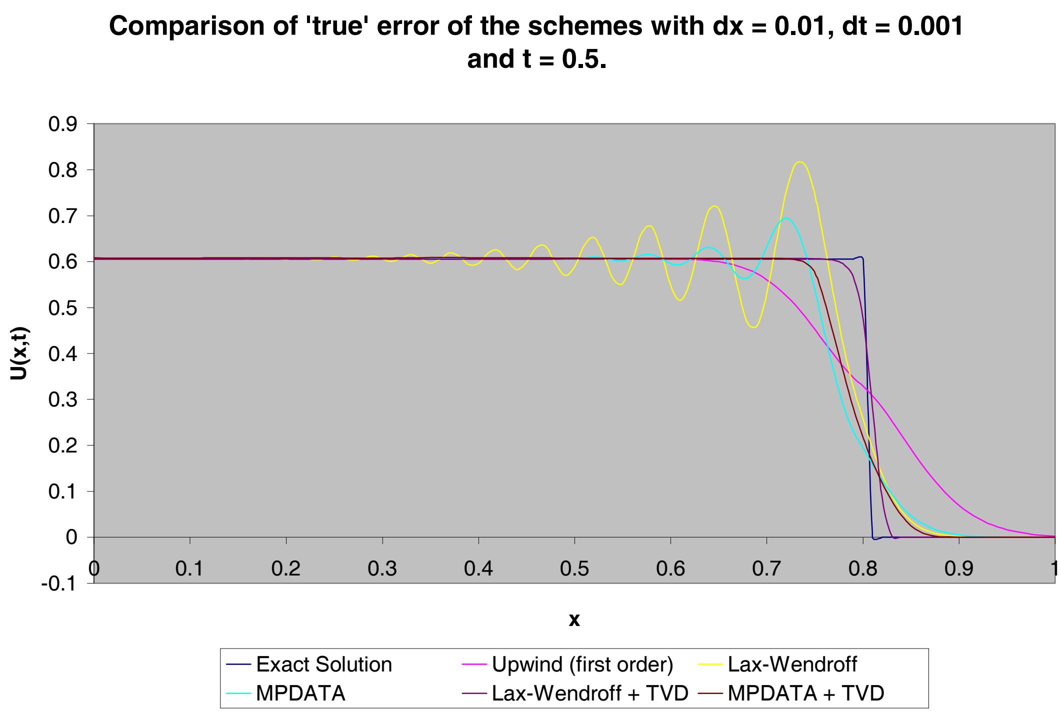


Figure 3-8: Comparison of the three approaches discussed in this chapter.

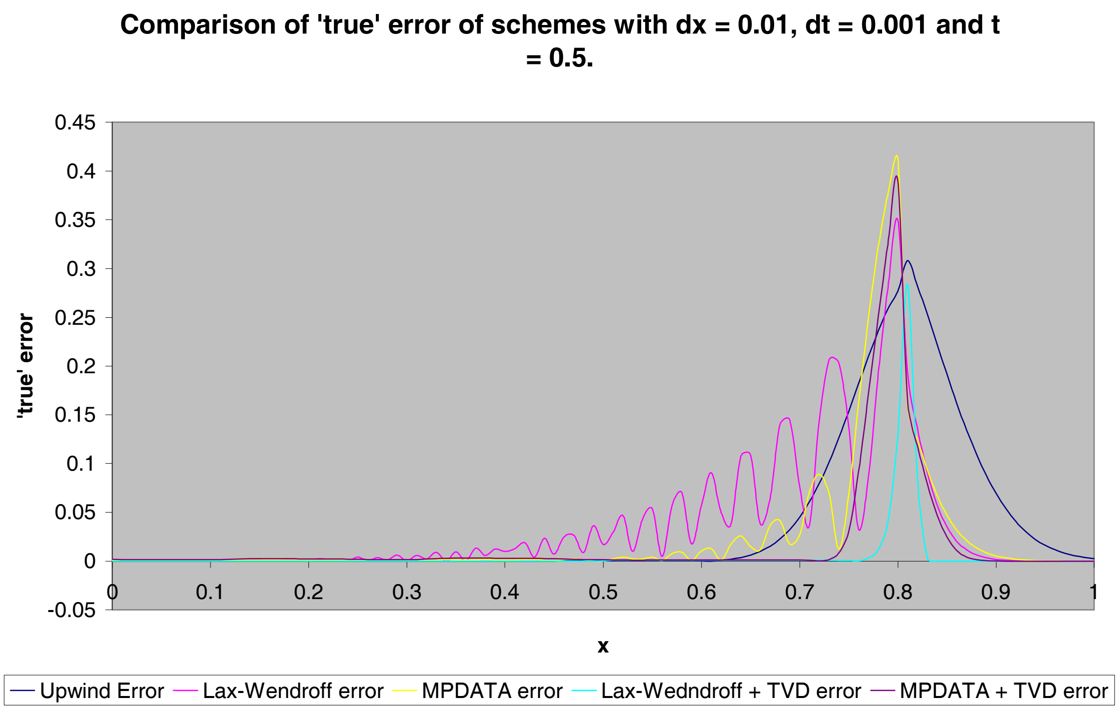


Figure 3-9: Comparison of 'true' error of the three approaches discussed in this chapter.

4 Conservation Law with Source Term R(x,t,u)

In Chapter 3, we discussed some finite difference schemes that numerically approximate conservation laws with a source term which is a function of x and t. In this chapter, we will discuss some finite difference schemes that numerically approximate conservation laws with a source term which is now a function of x, t and u, i.e.

$$\frac{\partial u}{\partial t}+\frac{\partial f(u)}{\partial x}=R(x,t,u) \tag{4.1}$$

where $R(x,t,u)$ is the source term. We shall see that difficulties will arise since the source term is now a known function of u as well as x and t, resulting in the numerical approximation of the source term no longer being exact. Throughout this chapter, we will be using the following test problem considered by LeVeque and Yee[1].

$$\frac{\partial u}{\partial t}+\frac{\partial u}{\partial x}=R(u), \tag{4.2}$$

where

$$R(u)=-u(u-1)\left(u-\frac{1}{2}\right),$$

with initial data

$$u(x,0)=\begin{cases}1 & \text{if } x\le 0.3\\ 0 & \text{if } x>0.3\end{cases}$$

and whose exact solution is

$$u(x,t)=\begin{cases}1 & \text{if } x\le 0.3+t\\ 0 & \text{if } x>0.3+t\end{cases} \tag{4.3}$$

to illustrate some numerical results.

4.1 Adaptation of the Schemes for the Conservation Law with Source Term R(x,t)

In this section, we will discuss how the different approaches for conservation laws with a source term of the form $R(x,t)$, which were discussed in Chapter 3, can be adapted to numerically approximate (4.1).

4.1.1 Basic Approach

The 'adding' of the source term can be easily adapted to numerically approximate equation (4.1). We do not need to adapt scheme (3.4) since we can re-write the source term approximation to include u, i.e. $R_i^n = R\left(i\Delta x, n\Delta t, u_i^n\right)$ which is known since the values of u_i^n are known so, the scheme remains as

$$u_i^{n+1} = u_i^{SCHEME} + \Delta t R_i^n . \tag{4.4}$$

However, scheme (3.5) is semi-implicit since $R_i^{n+1} = R\left(i\Delta x, n\Delta t, u_i^{n+1}\right)$ but the values of u_i^{n+1} are not yet known so we need to re-write

$$u_i^{n+1} = u_i^{SCHEME} + \Delta t R_i^{n+1} . \tag{4.5}$$

One approach is to use Taylor's theorem to obtain

$$R_i^{n+1} \approx R_i^n + \Delta t \left[\frac{\partial R}{\partial t}\right]_i^n + O(\Delta t^2) \tag{4.6}$$

and by substituting (4.6) into (4.5), we may obtain

$$u_i^{n+1} = u_i^{SCHEME} + \Delta t \left[R_i^n + \Delta t \left[\frac{\partial R}{\partial t}\right]_i^n \right] .$$

Here, we could use finite differences to approximate $\frac{\partial R}{\partial t}$, i.e.

$$\frac{\partial R}{\partial t} = \frac{R_i^n - R_i^{n-1}}{\Delta t}, \quad \frac{\partial R}{\partial t} = \frac{R_i^{n+1} - R_i^{n-1}}{2\Delta t} \quad \text{or} \quad \frac{\partial R}{\partial t} = \frac{R_i^{n+1} - R_i^n}{\Delta t},$$

but we would then encounter other difficulties since we only know the values of u_i^n and u_i^{n-1}, except initially when we do not know the values of u_i^{-1}. We could also calculate the derivative analytically and then approximate the derivative, i.e. $\left[\frac{\partial R}{\partial t}\right]_i^n$ but the derivative of the source term may be extremely difficult to find since the source term is a function of *u* and *u* is a function of *x* and *t*. Another approach we could take is to re-arrange $\left[\frac{\partial R}{\partial t}\right]_i^n$ in (4.6) by using the chain rule, i.e.

$$\left[\frac{\partial f}{\partial t}\right]_i^n = \left[\frac{\partial f}{\partial u}\right]_i^n \left[\frac{\partial u}{\partial t}\right]_i^n = \left[\frac{\partial f}{\partial u}\right]_i^n \frac{\left(u_i^{n+1} - u_i^n\right)}{\Delta t}.$$

Substituting into (4.6) gives

$$R_i^{n+1} \approx R_i^n + \left(u_i^{n+1} - u_i^n\right)\left[\frac{\partial R}{\partial u}\right]_i^n + \ldots$$

and by substituting into (4.5), we may obtain

$$\left(1 - \Delta t\left[\frac{\partial R}{\partial u}\right]_i^n\right) u_i^{n+1} = u_i^{SCHEME} + \Delta t R_i^n - \Delta t\, u_i^n \left[\frac{\partial R}{\partial u}\right]_i^n, \tag{4.7}$$

a semi-implicit scheme which 'adds' the source term implicitly. Here, since the source term is a known function of *x*, *t* and *u*, we can calculate the derivatives analytically and then approximate the derivatives, i.e. $\left[\frac{\partial R}{\partial u}\right]_i^n$.

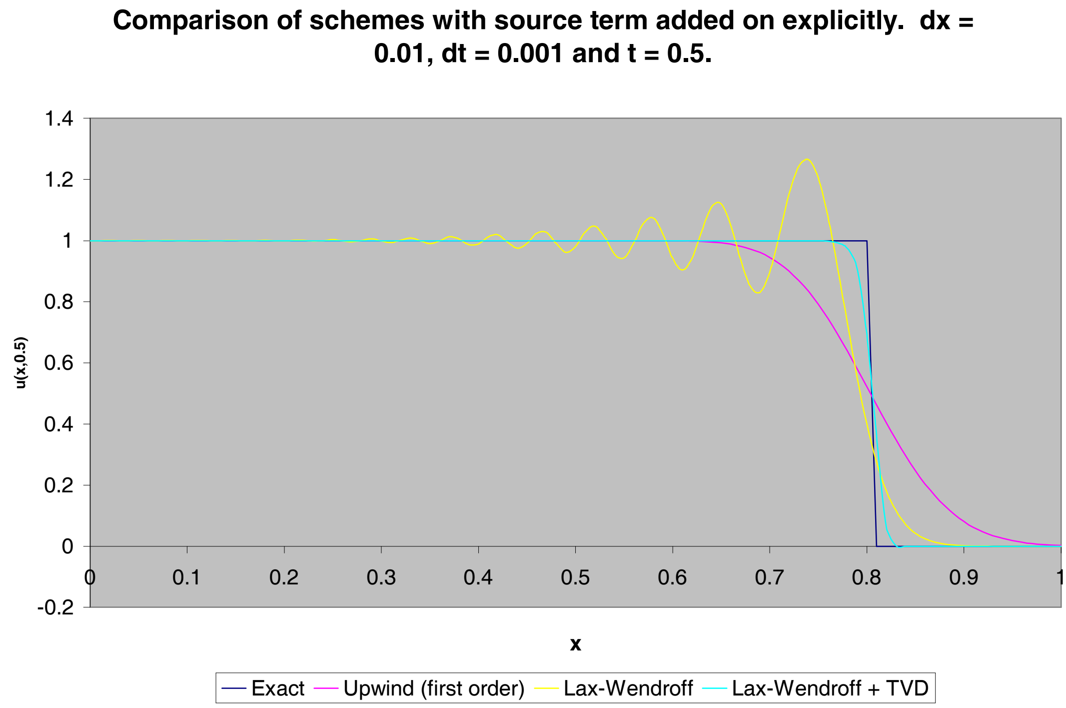


Figure 4-1: Comparison of different schemes with the source term 'added' on.

If we use (4.4) to numerically approximate the test problem (4.2), we may obtain the results shown in Figure 4-1. Figure 4-1 shows that the Upwind scheme with the source term 'added' is still giving the least accurate results and that the Lax-Wendroff with Superbee flux-limiter is giving the most accurate results.

4.1.2 Lax-Wendroff Approach

The Lax-Wendroff approach can be adapted to numerically approximate (4.1) but with difficulty. During the derivation of the Lax-Wendroff approach for numerically approximating (3.1), we obtained

$$u_i^{n+1} = u_i^n - \frac{s}{2}\left(f_{i+1}^n - f_{i-1}^n\right) + \frac{s}{2}\left[v_{i+1/2}\left(f_{i+1}^n - f_i^n\right) - v_{i-1/2}\left(f_i^n - f_{i-1}^n\right)\right]$$
$$+ \Delta t\left[R + \frac{\Delta t}{2}\left(\frac{\partial R}{\partial t} - \frac{\partial (a(u)R)}{\partial x}\right)\right].$$

We then used a forward difference approximation in space and time, to obtain

$$u_i^{n+1} = u_i^n - \frac{s}{2}\left(f_{i+1}^n - f_{i-1}^n\right) + \frac{s}{2}\left[v_{i+1/2}\left(f_{i+1}^n - f_i^n\right) - v_{i-1/2}\left(f_i^n - f_{i-1}^n\right)\right]$$
$$+ \Delta t\left[R_i^n + \frac{\Delta t}{2}\left(\frac{R_i^{n+1} - R_i^n}{\Delta t} - \frac{a_{i+1/2}R_{i+1/2}^n - a_{i-1/2}R_{i-1/2}^n}{\Delta x}\right)\right]$$

and hence,

$$u_i^{n+1} = u_i^n - \frac{s}{2}\left(f_{i+1}^n - f_{i-1}^n\right) + \frac{s}{2}\left[v_{i+1/2}\left(f_{i+1}^n - f_i^n\right) - v_{i-1/2}\left(f_i^n - f_{i-1}^n\right)\right]$$
$$+ \frac{\Delta t}{2}\left[R_i^n + R_i^{n+1}\right] - \frac{\Delta t}{4}\left[v_{i+1/2}\left(R_{i+1}^n - R_i^n\right) - v_{i-1/2}\left(R_i^n - R_{i-1}^n\right)\right].$$

However, if the source term is now also a function of u, then (4.8) becomes semi-implicit since we no longer know the value of R_i^{n+1}. We could replace R_i^{n+1} with (4.6) as we did in the previous sub-section, but this would only create more problems. However, we could replace R_i^{n+1} with

$$R_i^{n+1} \approx R_i^n + \left(u_i^{n+1} - u_i^n\right)\left[\frac{\partial R}{\partial u}\right]_i^n + \ldots$$

and obtain

$$\left(1 - \frac{\Delta t}{2}\left[\frac{\partial R}{\partial u}\right]_i^n\right)u_i^{n+1} = u_i^n - \frac{s}{2}\left(f_{i+1}^n - f_{i-1}^n\right) + \frac{s}{2}\left[v_{i+1/2}\left(f_{i+1}^n - f_i^n\right) - v_{i-1/2}\left(f_i^n - f_{i-1}^n\right)\right]$$
$$+ \Delta t\left[R_i^n - \frac{u_i^n}{2}\left[\frac{\partial R}{\partial u}\right]_i^n\right] - \frac{\Delta t}{4}\left[v_{i+1/2}\left(R_{i+1}^n - R_i^n\right) - v_{i-1/2}\left(R_i^n - R_{i-1}^n\right)\right]$$

the Lax-Wendroff approach for approximating (4.1). We can also apply flux-limiter methods to the Lax-Wendroff approach and obtain

$$\left(1 - \frac{\Delta t}{2}\left[\frac{\partial R}{\partial u}\right]_i^n\right)u_i^{n+1} = u_i^n - s\left[F(u;i) - F(u;i-1)\right] + \Delta t\left[R_i^n - \frac{u_i^n}{2}\left[\frac{\partial R}{\partial u}\right]_i^n\right]$$
$$- \frac{\Delta t}{4}\left[v_{i+1/2}\left(R_{i+1}^n - R_i^n\right) - v_{i-1/2}\left(R_i^n - R_{i-1}^n\right)\right] \qquad (4.8)$$

where

$$F(u;i) = F_L(u;i) + F_H(u;i)\phi_i$$

and

$$F_L(u;i) = \begin{cases} f_i^n & \text{if} \quad v_{i+1/2} > 0 \\ f_{i+1}^n & \text{if} \quad v_{i+1/2} < 0 \end{cases}$$

$$F_H(u;i) = \frac{1}{2} \begin{cases} (1 - v_{i+1/2})(f_{i+1}^n - f_i^n) & \text{if} \quad v_{i+1/2} > 0 \\ -(1 + v_{i+1/2})(f_{i+1}^n - f_i^n) & \text{if} \quad v_{i+1/2} < 0 \end{cases}.$$

Here ϕ_i denotes the flux-limiter method, which can be any of the flux-limiters in Table 2-3. If we use (4.4) to numerically approximate the test problem (4.2), we may obtain the results shown in Figure 4-2. Here, we can see that the Lax-Wendroff approach has numerically approximated (4.2) very accurately. Also, the numerical results in Figure 4-2 are very similar to the numerical results in Figure 4-1, where we 'added' the source term. In general, the Lax-Wendroff approach is more accurate than explicitly 'adding' the source term as we shall see later.

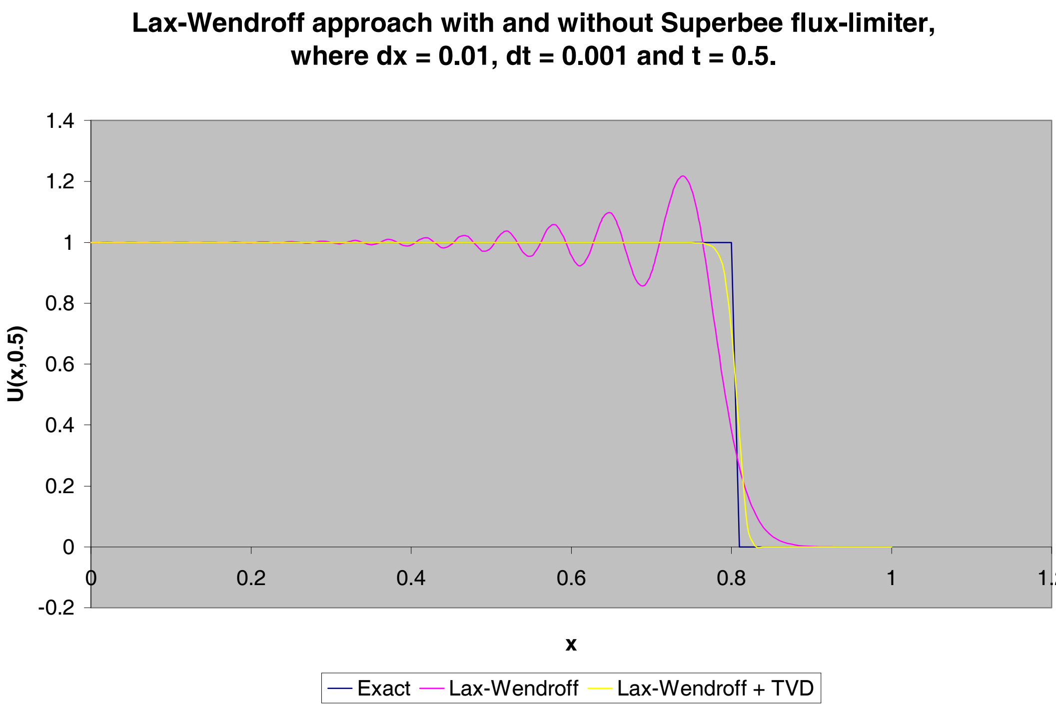


Figure 4-2: Lax-Wendroff approach with and without Superbee flux-limiter.

4.1.3 MPDATA approach

The MPDATA approach also creates difficulties when the source term is also a function of u. This is because (3.21) requires the value of R_i^{n+1} and we must re-write (3.21) so that it numerically approximates the conservation law with source term instead of the advection-transport equation, which was discussed in Chapter 3, Section 3.3. Also, the MPDATA approach requires another scheme to start it off initially, since we need to re-write (4.1) as

$$\frac{\partial u}{\partial t} + \frac{\partial \left(w(u)u \right)}{\partial x} = R(x,t,u)$$

but by using linear interpolation

$$w_{i+1/2}^{n+1/2} = \frac{1}{2}\left(3\, w_{i+1/2}^{n} - w_{i+1/2}^{n-1} \right).$$

Here, we would require the initial values of $w_{i+1/2}^{0}$ and $w_{i+1/2}^{-1}$ but we only know the values of $w_{i+1/2}^{0}$. But, if we used the average,

$$w_{i+1/2}^{n+1/2} = \frac{1}{2}\left(w_{i+1/2}^{n+1} + w_{i+1/2}^{n} \right)$$

the algorithm would become impractical since by using

$$w_i^{n+1} \approx w_i^n + \left(u_i^{n+1} - u_i^n \right) \left[\frac{\partial w}{\partial u} \right]_i^n + \ldots$$

we obtain

$$w_{i+1/2}^{n+1/2} = \frac{1}{2}\left(2 w_{i+1/2}^{n} + \frac{\left(u_{i+1/2}^{n+1} - u_{i+1/2}^{n} \right)}{2} \left[\frac{\partial w}{\partial u} \right]_{i+1/2}^{n} \right)$$

but the values of $u_{i+1/2}^{n+1}$ are unknown. We can overcome the difficulty of the value of R_i^{n+1} being unknown by using

$$R_i^{n+1} \approx R_i^n + \left(u_i^{n+1} - u_i^n \right) \left[\frac{\partial R}{\partial u} \right]_i^n + \ldots$$

and by substituting into (3.24), to obtain

$$\left(1-\frac{\Delta t}{2}\left[\frac{\partial R}{\partial u}\right]_i^n\right)u_i^{n+1} = MPDATA\left(u_i^n + \frac{\Delta t}{2}R_i^n, w_{i+1/2}^{n+1/2}\right) + \frac{\Delta t}{2}\left[R_i^n - u_i^n\left[\frac{\partial R}{\partial u}\right]_i^n\right] \qquad (4.9)$$

where $MPDATA\left(u_i^n, C\right)$ corresponds to the basic MPDATA algorithm with flux-limiter

$$u_i^{n+1} = u_i^{(1)} - \left[F\left(u_i^{(1)}, u_{i+1}^{(1)}, V_{i+1/2}^{(1)}\right)\phi_i - F\left(u_{i-1}^{(1)}, u_i^{(1)}, V_{i-1/2}^{(1)}\right)\phi_{i-1}\right]$$

whose pseudo velocity is

$$V_{i+1/2}^{(1)} = \frac{\Delta x}{\Delta t}\left(\left|w_{i+1/2}^{n+1/2}\right| - \left[w_{i+1/2}^{n+1/2}\right]^2\right)\left[\frac{u_{i+1}^{(1)} - u_i^{(1)}}{u_{i+1}^{(1)} + u_i^{(1)}}\right] - w_{i+1/2}^{n+1/2}\left[w_{i+3/2}^{n+1/2} - w_{i-1/2}^{n+1/2}\right],$$

the first order approximation is

$$u_i^{(1)} = u_i^n - \left[F\left(u_i^n, u_{i+1}^n, w_{i+1/2}^{n+1/2}\right) - F\left(u_{i-1}^n, u_i^n, w_{i-1/2}^{n+1/2}\right)\right],$$

$$w_{i+1/2}^{n+1/2} = \frac{1}{2}\left(3w_{i+1/2}^n - w_{i+1/2}^{n-1}\right)$$

and ϕ_i can be any of the flux-limiters listed in Table 2-3. Here, we can see that the MPDATA approach for numerically approximating (4.1) is becoming very impractical. This is because we are approximating approximations resulting in the accuracy of the algorithm reducing rapidly and we also require another scheme to start the algorithm off. However, MPDATA can be used to accurately numerically approximate the advection-transport equation with source term, $R(x,t,u)$. If we use (4.4) to numerically approximate the test problem (4.2), we may obtain the results shown in Figure 4-3. Here, we can see that the MPDATA approach is quite accurate but not as accurate as the results obtained in Figure 4-1 and Figure 4-2. Also, MPDATA will not be so accurate for the inviscid burgers case with source term.

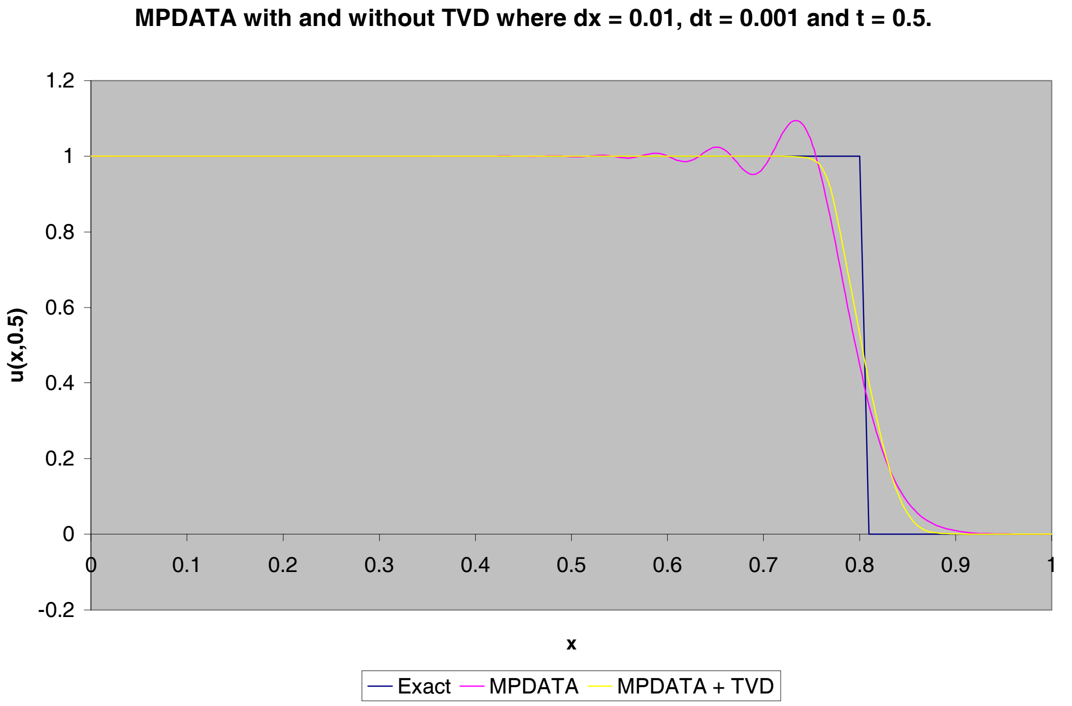


Figure 4-3: MPDATA approach with and without Superbee flux-limiter.

4.2 Roe's Upwind Approach

4.2.1 Advection Equation with Source Term R(x)

Roe[6] derived a finite difference scheme which numerically approximates

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = R(x), \tag{4.10}$$

where $c > 0$ and $R(x)$ is the source term, with second order accuracy. If we consider the initial-value problem of (4.10) with initial data $u(x,0) = u_o(x)$, we may obtain the general solution

$$u(x,t) = u_o(x-ct) + \frac{1}{c}\int_{x-ct}^{x} R(x)\,dx$$

which can be re-written as

$$u_i^{n+1} = u((i-v)\Delta x, n\Delta t) + \frac{1}{c}\int_{(i-v)\Delta x}^{i\Delta x} R(x)\,dx\,. \tag{4.11}$$

Here we can see that the first term on the right hand side of (4.11) can cause difficulties if the Courant number is not an integer. This is because we are using a mesh where we only know the values at the grid points ($i\Delta x, n\Delta t$) and if v is not an integer, then the value of u required no longer lies on the mesh and is thus unknown. However, Roe[6] deduced that the only reasonable way to approximate this term is to use

$$u\big((i-v)\Delta x, n\Delta t\big) = u_i^n - v\big(u_i^n - u_{i-1}^n\big), \tag{4.12}$$

which is the Upwind approach, since no other formula is consistent with (4.10). Also, (4.12) gives the smallest truncation error of all possible choices where the truncation error has positive coefficients and depends only on v.

The integral term present in (4.11) can be approximated in numerous ways. Since we know $R(x)$, we could integrate the source term and use the exact values but a more general approach is to use a two-point approximation

$$\frac{1}{c}\int_{(i-v)\Delta x}^{i\Delta x} R(x)\,dx = \Delta t\big[(1-\alpha)R_i^n + \alpha R_{i-1}^n\big]. \tag{4.13}$$

Here, the value of α is arbitrary and must be chosen such that $0 \le \alpha \le 1$. Hence, by substituting (4.13) and (4.12) into (4.11), we may obtain

$$u_i^{n+1} = u_i^n - v\big(u_i^n - u_{i-1}^n\big) + \Delta t\big[(1-\alpha)R_i^n + \alpha R_{i-1}^n\big] \tag{4.14}$$

which is Roe's Upwind approach for numerically approximating (4.10). However, this scheme is only stable for $c > 0$, but if $c < 0$ then we may obtain

$$u_i^{n+1} = u_i^n - v\big(u_{i+1}^n - u_i^n\big) + \Delta t\big[(1-\alpha)R_i^n + \alpha R_{i+1}^n\big] \tag{4.15}$$

and by combining (4.14) and (4.15) gives

$$u_i^{n+1} = u_i^n - \begin{cases} v\big(u_i^n - u_{i-1}^n\big) - \Delta t\big[(1-\alpha)R_i^n + \alpha R_{i-1}^n\big] & \text{if } v > 0 \\ v\big(u_{i+1}^n - u_i^n\big) - \Delta t\big[(1-\alpha)R_i^n + \alpha R_{i+1}^n\big] & \text{if } v < 0 \end{cases}$$

which is Roe's Upwind approach. Moreover, Roe[6] found that if we take

$$\alpha = \frac{1}{2}$$

then we may obtain a scheme that is second order accurate in the steady state, i.e.

$$u_i^{n+1} = u_i^n - \begin{cases} v\left(u_i^n - u_{i-1}^n\right) - \dfrac{\Delta t}{2}\left[R_i^n + R_{i-1}^n\right] & \text{if } \ v > 0 \\ v\left(u_{i+1}^n - u_i^n\right) - \dfrac{\Delta t}{2}\left[R_i^n + R_{i+1}^n\right] & \text{if } \ v < 0 \end{cases}.$$

Unfortunately, this scheme is only a first order approximation to (4.10) but we can also obtain a second order accurate scheme by using van Leer's MUSCL approach [10] to obtain

$$u_i^{n+1} = u_i^n - v\left(u_i^n - u_{i-1}^n\right) - \frac{v}{2}(1-v)\left[u_{i+1}^n - 2u_i^n + u_{i-1}^n\right] + \Delta t\left[(1-\alpha)R_i^n + \alpha R_{i-1}^n\right]$$

where $c > 0$. And hence we may obtain

$$u_i^{n+1} = u_i^n - \begin{cases} v\left(u_i^n - u_{i-1}^n\right) + \dfrac{v}{2}(1-v)\left[u_{i+1}^n - 2u_i^n + u_{i-1}^n\right] - \Delta t\left[(1-\alpha)R_i^n + \alpha R_{i-1}^n\right] & \text{if } \ v > 0 \\ v\left(u_{i+1}^n - u_i^n\right) + \dfrac{v}{2}(1+v)\left[u_{i+1}^n - 2u_i^n + u_{i-1}^n\right] - \Delta t\left[(1-\alpha)R_i^n + \alpha R_{i+1}^n\right] & \text{if } \ v < 0 \end{cases} \qquad (4.16)$$

which is a second order approximation to (4.10). Notice that (4.16) is the Lax-Wendroff scheme for numerically approximating the advection equation without a source term with a source term approximation added.

4.2.2 Conservation Law with Source Term R(x,t,u)

Bermudez and Vazquez[4] adapted Roe's Upwind approach for numerically approximating (4.10) to numerically approximate the advection equation with source term $R(x,u)$, i.e.

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = R(x,u) \qquad (4.17)$$

where $c > 0$. They used a similar approach as in the previous sub-section to obtain

$$u_i^{n+1} = u\big((i-\nu)\Delta x, n\Delta t\big) + \frac{1}{c}\int_{t_n}^{t_{n+1}} R\big(x_i - c(t_{n+1}-s), u\big(x_i - c(t_{n+1}-s), s\big)\big)\,ds \qquad (4.18)$$

and hence,

$$u_i^{n+1} = u_i^n - \nu\big(u_i^n - u_{i-1}^n\big) + \Delta t\big[(1-\alpha)R_i^n + \alpha R_{i-1}^n\big],$$

which numerically approximates (4.17) and is identical to (4.14). They also discussed another approach, which was to approximate the integral term of (4.18) with

$$\frac{1}{c}\int_{t_n}^{t_{n+1}} R\big(x_i - c(t_{n+1}-s), u\big(x_i - c(t_{n+1}-s), s\big)\big)\,ds = \Delta t R\big((1-\alpha)x_i + \alpha x_{i-1}, (1-\alpha)u_i^n + \alpha u_{i-1}^n\big)$$

instead of

$$\frac{1}{c}\int_{t_n}^{t_{n+1}} R\big(x_i - c(t_{n+1}-s), u\big(x_i - c(t_{n+1}-s), s\big)\big)\,ds = \Delta t\big[(1-\alpha)R_i^n + \alpha R_{i-1}^n\big].$$

Hence, Bermudez and Vazquez[4] obtained two approaches to numerically approximate (4.17)

$$u_i^{n+1} = u_i^n - \begin{cases} \nu\big(u_i^n - u_{i-1}^n\big) - \Delta t\big[(1-\alpha)R_i^n + \alpha R_{i-1}^n\big] & \text{if } \nu > 0 \\ \nu\big(u_{i+1}^n - u_i^n\big) - \Delta t\big[(1-\alpha)R_i^n + \alpha R_{i+1}^n\big] & \text{if } \nu < 0 \end{cases}, \qquad (4.19)$$

which is Roe's Upwind approach and

$$u_i^{n+1} = u_i^n - \begin{cases} \nu\big(u_i^n - u_{i-1}^n\big) - \Delta t R\big((1-\alpha)x_i + \alpha x_{i-1}, (1-\alpha)u_i^n + \alpha u_{i-1}^n\big) & \text{if } \nu > 0 \\ \nu\big(u_{i+1}^n - u_i^n\big) - \Delta t R\big((1-\alpha)x_i + \alpha x_{i+1}, (1-\alpha)u_i^n + \alpha u_{i+1}^n\big) & \text{if } \nu < 0 \end{cases} \qquad (4.20)$$

Here, α = ½ also gives second order accuracy in the steady state for both schemes. Vazquez and Bermudez[4] also discuss various choices of α and give some intervals of absolute stability and positivity, where $c > 0$, for the different values of α, which are listed in Table 4-1.

Scheme	Interval Of Absolute Stability ($c>0$)	Interval Of Absolute Positivity ($c>0$)
$\alpha=0$	$v+\frac{\lambda\Delta t}{2}\le 1$	$v+\lambda\Delta t\le 1$
$\alpha=v$	$v\le 1$ and $\frac{\lambda\Delta t}{2}\le 1$	$v\le 1$ and $\lambda\Delta t\le 1$
$\alpha=\frac{1}{2}v$ (for $R(x)$ only)	$v\le 1$ and $\frac{\lambda\Delta t}{2}\le 1$	$\frac{\lambda\Delta t}{2}\le 1$ and $v\le\frac{1-\lambda\Delta t}{1-\frac{\lambda\Delta t}{2}}$
$\alpha=\frac{1}{2}$	$v\le 1$ and $\frac{\lambda\Delta t}{2}\le 1$	$v+\frac{\lambda\Delta t}{2}\le 1$ and $\frac{\lambda\Delta t}{2}\le v$

Table 4-1: Intervals of absolute stability and positivity for $R(x,u) = -\lambda u$.

Notice that when $\alpha = 0$, both (4.19) and (4.20) become the Upwind scheme with source term 'added' as discussed in Chapter 3, Section 1.

This approach can be easily adapted to numerically approximate (4.1) by re-writing (4.19) and (4.20) as

$$u_i^{n+1}=u_i^n-\begin{cases} s\left(f_i^n-f_{i-1}^n\right)-\Delta t\left[(1-\alpha)R_i^n+\alpha R_{i-1}^n\right] & \text{if } v_{i+1/2}>0 \\ s\left(f_{i+1}^n-f_i^n\right)-\Delta t\left[(1-\alpha)R_i^n+\alpha R_{i+1}^n\right] & \text{if } v_{i-1/2}<0 \end{cases}, \quad (4.21)$$

which is Roe's Upwind approach, and

$$u_i^{n+1}=u_i^n-\begin{cases} s\left(f_i^n-f_{i-1}^n\right)-\Delta t R\left((1-\alpha)x_i+\alpha x_{i-1},(1-\alpha)u_i^n+\alpha u_{i-1}^n\right) & \text{if } v_{i+1/2}>0 \\ s\left(f_{i+1}^n-f_i^n\right)-\Delta t R\left((1-\alpha)x_i+\alpha x_{i+1},(1-\alpha)u_i^n+\alpha u_{i+1}^n\right) & \text{if } v_{i-1/2}<0 \end{cases}. \quad (4.22)$$

Here, both (4.21) and (4.22) are first order accurate schemes but we can obtain second order accurate schemes by using

$$u_i^{n+1}=u_i^n-\begin{cases} s\left(f_i^n-f_{i-1}^n\right)+\frac{s}{2}\left[(1-v_{i+1/2})\left(f_{i+1}^n-f_i^n\right)-(1-v_{i-1/2})\left(f_i^n-f_{i-1}^n\right)\right] \\ \qquad -\Delta t\left[(1-\alpha)R_i^n+\alpha R_{i-1}^n\right] \quad \text{if } v_{i+1/2}>0 \\ s\left(f_{i+1}^n-f_i^n\right)-\frac{s}{2}\left[(1+v_{i+1/2})\left(f_{i+1}^n-f_i^n\right)-(1+v_{i-1/2})\left(f_i^n-f_{i-1}^n\right)\right] \\ \qquad -\Delta t\left[(1-\alpha)R_i^n+\alpha R_{i+1}^n\right] \quad \text{if } v_{i-1/2}<0 \end{cases}$$

and

$$u_i^{n+1} = u_i^n - \begin{cases} s\left(f_i^n - f_{i-1}^n\right) + \frac{s}{2}\left[\left(1 - \nu_{i+1/2}\right)\left(f_{i+1}^n - f_i^n\right) - \left(1 - \nu_{i-1/2}\right)\left(f_i^n - f_{i-1}^n\right)\right] \\ \qquad - \Delta t R\left(\left(1-\alpha\right)x_i + \alpha x_{i-1}, \left(1-\alpha\right)u_i^n + \alpha u_{i-1}^n\right) \quad \text{if} \quad \nu_{i+1/2} > 0 \\ s\left(f_{i+1}^n - f_i^n\right) - \frac{s}{2}\left[\left(1 + \nu_{i+1/2}\right)\left(f_{i+1}^n - f_i^n\right) - \left(1 + \nu_{i-1/2}\right)\left(f_i^n - f_{i-1}^n\right)\right] \\ \qquad - \Delta t R\left(\left(1-\alpha\right)x_i + \alpha x_{i+1}, \left(1-\alpha\right)u_i^n + \alpha u_{i+1}^n\right) \quad \text{if} \quad \nu_{i-1/2} < 0 \end{cases}$$

Also, we can apply the flux-limiter method to obtain

$$u_i^{n+1} = u_i^n - s\left[F\left(u;i\right) - F\left(u;i-1\right)\right] + \Delta t \begin{cases} \left(1-\alpha\right)R_i^n + \alpha R_{i-1}^n & \text{if } \nu_{i+1/2} > 0 \\ \left(1-\alpha\right)R_i^n + \alpha R_{i+1}^n & \text{if } \nu_{i-1/2} < 0 \end{cases} \tag{4.23}$$

and

$$u_i^{n+1} = u_i^n - s\left[F\left(u;i\right) - F\left(u;i-1\right)\right] \\ + \Delta t \begin{cases} R\left(\left(1-\alpha\right)x_i + \alpha x_{i-1}, \left(1-\alpha\right)u_i^n + \alpha u_{i-1}^n\right) & \text{if } \nu_{i+1/2} > 0 \\ R\left(\left(1-\alpha\right)x_i + \alpha x_{i+1}, \left(1-\alpha\right)u_i^n + \alpha u_{i+1}^n\right) & \text{if } \nu_{i-1/2} < 0 \end{cases} \tag{4.24}$$

where

$$F\left(u;i\right) = F_L\left(u;i\right) + F_H\left(u;i\right)\phi_i$$

and

$$F_L\left(u;i\right) = \begin{cases} f_i^n & \text{if } \nu_{i+1/2} > 0 \\ f_{i+1}^n & \text{if } \nu_{i+1/2} < 0 \end{cases}$$

$$F_H\left(u;i\right) = \frac{1}{2}\begin{cases} \left(1 - \nu_{i+1/2}\right)\left(f_{i+1}^n - f_i^n\right) & \text{if } \nu_{i+1/2} > 0 \\ -\left(1 + \nu_{i+1/2}\right)\left(f_{i+1}^n - f_i^n\right) & \text{if } \nu_{i+1/2} < 0 \end{cases}.$$

Here, ϕ_i represents the flux-limiter, which can be any of the second order flux-limiters in Table 2-3.

4.2.3 Some Numerical Results for the Explicit Upwind Approach

Now, by using (4.24) to numerically approximate the test problem (4.2), we may obtain the numerical results in Figure 4-4.

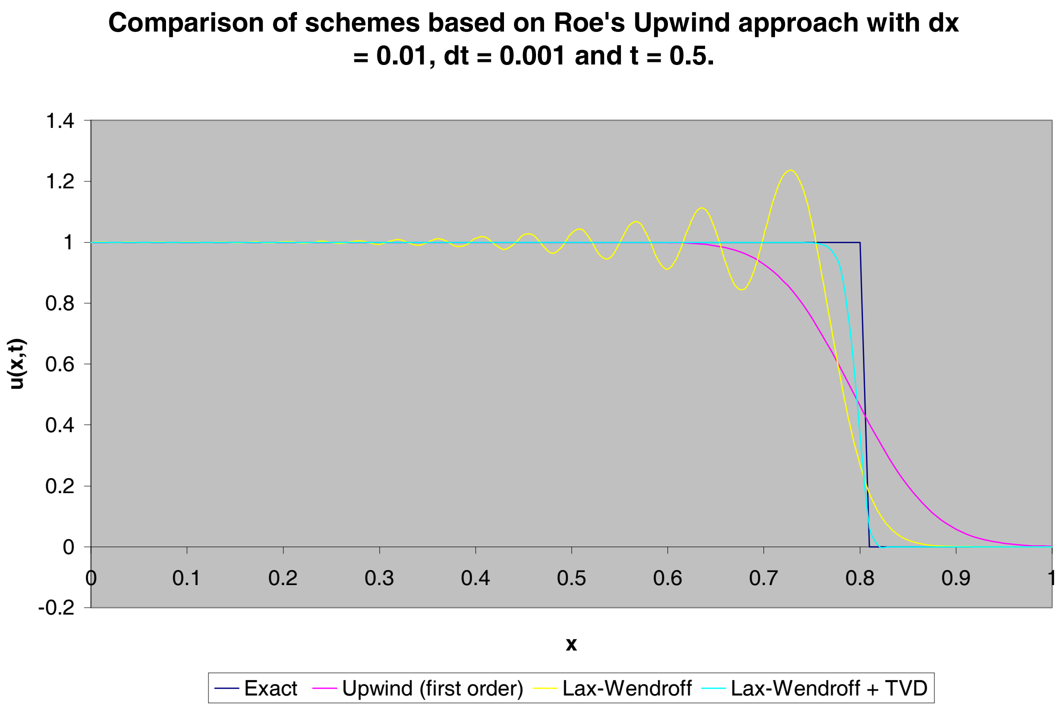


Figure 4-4: Comparison of schemes based on Roe's Upwind approach.

Here, we can see that Roe's upwind approach is giving some very accurate results, especially for the second order Lax-Wendroff plus Superbee flux-limiter, but the results are not as accurate as in Figure 4-1, where we 'added' the source term, and Figure 4-2, where we used the Lax-Wendroff approach. However, we will see later that, in general, Roe's Upwind approach is a lot more accurate at numerically approximating (4.1) than 'adding' the source term and the Lax-Wendroff approach, especially when the source term is stiff.

4.3 Implicit Upwind Approach

Embid, Goodman and Majda[2] discussed some different approaches for numerically approximating

$$\frac{\partial u}{\partial t}+\frac{\partial f(u)}{\partial x}=R(x,u) \tag{4.25}$$

where the source term must be of the form

$$R(x,u)=e(x)g(u).$$

They discussed the first order Engquist-Osher scheme, with switching through zero, and a second order Upwind approach based on the Engquist-Osher approach. Here, we will use the analysis of Embid, Goodman and Majda[2] to derive a first and second order implicit Upwind scheme with the source term 'added' implicitly.

4.3.1 First Order Implicit Upwind Approach

The first scheme that we will discuss is the implicit first order Upwind approach with the source term 'added' implicitly, i.e.

$$u_i^{n+1}=u_i^n+\Delta t R_i^{n+1}-s\begin{cases}\left(f_i^{n+1}-f_{i-1}^{n+1}\right) & \text{if} \quad v_{i+1/2}>0\\ \left(f_{i+1}^{n+1}-f_i^{n+1}\right) & \text{if} \quad v_{i+1/2}<0\end{cases}. \tag{4.26}$$

Here, we will need to re-arrange (4.26) into the system

$$A\underline{u}^{n+1}=G,$$

where A is a $(I+1)$ x $(I+1)$ matrix and G is a $(I+1)$ column vector, and solve this system at every time step. However, difficulties can arise when re-arranging (4.25) into system form. Consider (4.26) when $v_{i+1/2}>0$

$$u_i^{n+1}=u_i^n+\Delta t R_i^{n+1}-s\left(f_i^{n+1}-f_{i-1}^{n+1}\right)$$

and, by re-arranging we may obtain

$$u_i^{n+1} + s\left(f_i^{n+1} - f_{i-1}^{n+1}\right) - \Delta t R_i^{n+1} = u_i^n$$

and since $R(x,u) = e(x)g(u)$

$$u_i^{n+1} + s\left(f_i^{n+1} - f_{i-1}^{n+1}\right) - \Delta t e_i g_i^{n+1} = u_i^n \, . \tag{4.27}$$

However, the second and third terms on the left-hand side of (4.27) create difficulties since they are functions of u and so we cannot re-arrange (4.27) into system form. We can overcome this problem by using Taylor's theorem, i.e.

$$f_i^{n+1} \approx f_i^n + \left(u_i^{n+1} - u_i^n\right)\left[\frac{\partial f}{\partial u}\right]_i^n + \ldots$$

and

$$g_i^{n+1} \approx g_i^n + \left(u_i^{n+1} - u_i^n\right)\left[\frac{\partial g}{\partial u}\right]_i^n + \ldots$$

Now by substituting into (4.27)

$$u_i^{n+1} + s\left(f_i^n + \left(u_i^{n+1} - u_i^n\right)\left[\frac{\partial f}{\partial u}\right]_i^n - \left[f_{i-1}^n + \left(u_{i-1}^{n+1} - u_{i-1}^n\right)\left[\frac{\partial f}{\partial u}\right]_{i-1}^n\right]\right)$$
$$- \Delta t e_i\left(g_i^n + \left(u_i^{n+1} - u_i^n\right)\left[\frac{\partial g}{\partial u}\right]_i^n\right) = u_i^n$$

and by re-arranging we may obtain

$$\left(u_i^{n+1} - u_i^n\right)\left[1 + s\left[\frac{\partial f}{\partial u}\right]_i^n - \Delta t e_i\left[\frac{\partial g}{\partial u}\right]_i^n\right] - s\left(u_{i-1}^{n+1} - u_{i-1}^n\right)\left[\frac{\partial f}{\partial u}\right]_{i-1}^n = \Delta t e_i g_i^n - s\left(f_i^n - f_{i-1}^n\right) \tag{4.28}$$

where $v_{i+1/2} > 0$. Similarly, we may obtain

$$\left(u_i^{n+1} - u_i^n\right)\left[1 - s\left[\frac{\partial f}{\partial u}\right]_i^n - \Delta t e_i\left[\frac{\partial g}{\partial u}\right]_i^n\right] + s\left(u_{i-1}^{n+1} - u_{i-1}^n\right)\left[\frac{\partial f}{\partial u}\right]_{i+1}^n = \Delta t e_i g_i^n - s\left(f_{i+1}^n - f_i^n\right) \tag{4.29}$$

where $v_{i+1/2} < 0$. Hence, by combining (4.28) and (4.29), we may now obtain the system form

$$\begin{bmatrix} b_0 & d_0 & 0 & 0 & 0 & \dots & 0 \\ a_1 & b_1 & d_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & b_2 & d_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{I-2} & b_{I-2} & d_{I-2} & 0 \\ 0 & \dots & 0 & 0 & a_{I-1} & b_{I-1} & d_{I-1} \\ 0 & \dots & 0 & 0 & 0 & a_I & b_I \end{bmatrix} \begin{bmatrix} u_0^{n+1} - u_0^n \\ u_1^{n+1} - u_1^n \\ u_2^{n+1} - u_2^n \\ \vdots \\ u_{I-2}^{n+1} - u_{I-2}^n \\ u_{I-1}^{n+1} - u_{I-1}^n \\ u_I^{n+1} - u_I^n \end{bmatrix} = \begin{bmatrix} G_0 - a_0\left(u_{-1}^{n+1} - u_{-1}^n\right) \\ G_1 \\ G_2 \\ \vdots \\ G_{I-2} \\ G_{I-1} \\ G_I - d_I\left(u_{I+1}^n - u_{I+1}^n\right) \end{bmatrix} \tag{4.30}$$

where

$$G_i = \Delta te_i\, g_i^n - s \begin{cases} \left(f_i^n - f_{i-1}^n\right) & \text{if } \; v_{i+1/2} > 0 \\ \left(f_{i+1}^n - f_i^n\right) & \text{if } \; v_{i+1/2} < 0 \end{cases},$$

$$a_i = \begin{cases} -s\left[\dfrac{\partial f}{\partial u}\right]_{i-1}^n & \text{if } \; v_{i+1/2} > 0 \\ 0 & \text{if } \; v_{i+1/2} < 0 \end{cases}, \quad d_i = \begin{cases} 0 & \text{if } \; v_{i+1/2} > 0 \\ s\left[\dfrac{\partial f}{\partial u}\right]_{i+1}^n & \text{if } \; v_{i+1/2} < 0 \end{cases}$$

and

$$b_i = 1 + \operatorname{sgn}(v_{i+1/2})\, s\left[\frac{\partial f}{\partial u}\right]_i^n - \Delta te_i\left[\frac{\partial g}{\partial u}\right]_i^n .$$

Here, A is a tri-diagonal matrix and so, this system does not require too many calculations and since f and g are known functions of u, we can approximate $\left[\frac{\partial f}{\partial u}\right]_i^n$ and $\left[\frac{\partial g}{\partial u}\right]_i^n$.

4.3.2 Second Order Implicit Upwind Approach

We can also obtain a second order approximation by using the implicit second order Upwind approach with the source term 'added' implicitly, i.e.

$$u_i^{n+1} = u_i^n + \Delta t R_i^{n+1} - \frac{s}{2} \begin{cases} (3 - v_{i-1/2})\left(f_i^{n+1} - f_{i-1}^{n+1}\right) - (1 - v_{i-3/2})\left(f_{i-1}^{n+1} - f_{i-2}^{n+1}\right) & \text{if } \; v_{i+1/2} > 0 \\ -(1 + v_{i+3/2})\left(f_{i+2}^{n+1} - f_{i+1}^{n+1}\right) + (v_{i+1/2} + 3)\left(f_{i+1}^{n+1} - f_i^{n+1}\right) & \text{if } \; v_{i+1/2} < 0 \end{cases} \tag{4.31}$$

We will need to re-arrange (4.31) into system form and solve for each time step in the same way we did in the previous sub-section to obtain

$$\left(u_i^{n+1}-u_i^n\right)\left[1+\frac{s}{2}\left(3-v_{i-1/2}\right)\left[\frac{\partial f}{\partial u}\right]_i^n-\Delta te_i\left[\frac{\partial g}{\partial u}\right]_i^n\right]$$

$$-\frac{s}{2}\left(u_{i-1}^{n+1}-u_{i-1}^n\right)\left[4-v_{i-1/2}-v_{i-3/2}\right]\left[\frac{\partial f}{\partial u}\right]_{i-1}^n+\frac{s}{2}\left(u_{i-2}^{n+1}-u_{i-2}^n\right)\left[1-v_{i-3/2}\right]\left[\frac{\partial f}{\partial u}\right]_{i-2}^n \quad (4.32)$$

$$=\Delta te_i g_i^n-\frac{s}{2}\left[\left(3-v_{i-1/2}\right)\left(f_i^{n+1}-f_{i-1}^{n+1}\right)-\left(1-v_{i-3/2}\right)\left(f_{i-1}^{n+1}-f_{i-2}^{n+1}\right)\right]$$

where $v_{i+1/2}>0$ and

$$\left(u_i^{n+1}-u_i^n\right)\left[1-\frac{s}{2}\left(v_{i+1/2}+3\right)\left[\frac{\partial f}{\partial u}\right]_i^n-\Delta te_i\left[\frac{\partial g}{\partial u}\right]_i^n\right]$$

$$+\frac{s}{2}\left(u_{i+1}^{n+1}-u_{i+1}^n\right)\left[4+v_{i+1/2}+v_{i+3/2}\right]\left[\frac{\partial f}{\partial u}\right]_{i+1}^n-\frac{s}{2}\left(u_{i+2}^{n+1}-u_{i+2}^n\right)\left(1+v_{i+3/2}\right)\left[\frac{\partial f}{\partial u}\right]_{i+2}^n \quad (4.33)$$

$$=\Delta te_i g_i^n-\frac{s}{2}\left[-\left(1+v_{i+3/2}\right)\left(f_{i+2}^n-f_{i+1}^n\right)+\left(v_{i+1/2}+3\right)\left(f_{i+1}^n-f_i^n\right)\right]$$

where $v_{i+1/2}<0$. Hence, by combining (4.32) and (4.33), we may now obtain the system form

$$\begin{bmatrix} b_0 & d_0 & k_0 & 0 & 0 & 0 & \dots & 0 \\ a_1 & b_1 & d_1 & k_1 & 0 & 0 & \dots & 0 \\ l_2 & a_2 & b_2 & d_2 & k_2 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & l_{I-2} & a_{I-2} & b_{I-2} & d_{I-2} & k_{I-2} \\ 0 & \dots & 0 & 0 & l_{I-1} & a_{I-1} & b_{I-1} & d_{I-1} \\ 0 & \dots & 0 & 0 & 0 & l_I & a_I & b_I \end{bmatrix} \begin{bmatrix} u_0^{n+1}-u_0^n \\ u_1^{n+1}-u_1^n \\ u_2^{n+1}-u_2^n \\ \vdots \\ \vdots \\ u_{I-2}^{n+1}-u_{I-2}^n \\ u_{I-1}^{n+1}-u_{I-1}^n \\ u_I^{n+1}-u_I^n \end{bmatrix} = \begin{bmatrix} G_0-a_0\left(u_{-1}^{n+1}-u_{-1}^n\right)-l_0\left(u_{-2}^{n+1}-u_{-2}^n\right) \\ G_1-l_1\left(u_{-1}^{n+1}-u_{-1}^n\right) \\ G_2 \\ \vdots \\ \vdots \\ G_{I-2} \\ G_{I-1}-k_{I-1}\left(u_{I+1}^{n+1}-u_{I+1}^n\right) \\ G_I-d_I\left(u_{I+1}^{n+1}-u_{I+1}^n\right)-k_I\left(u_{I+2}^{n+1}-u_{I+2}^n\right) \end{bmatrix}$$

where

$$G_i=\Delta te_i\, g_i^n-\frac{s}{2}\begin{cases}\left(3-v_{i-1/2}\right)\left(f_i^{n+1}-f_{i-1}^{n+1}\right)-\left(1-v_{i-3/2}\right)\left(f_{i-1}^{n+1}-f_{i-2}^{n+1}\right) & \text{if } v_{i+1/2}>0 \\ -\left(1+v_{i+3/2}\right)\left(f_{i+2}^{n+1}-f_{i+1}^{n+1}\right)+\left(v_{i+1/2}+3\right)\left(f_{i+1}^{n+1}-f_i^{n+1}\right) & \text{if } v_{i+1/2}<0\end{cases},$$

$$l_i=\begin{cases}\frac{s}{2}\left[1-v_{i-3/2}\right]\left[\frac{\partial f}{\partial u}\right]_{i-2}^n & \text{if } v_{i+1/2}>0 \\ 0 & \text{if } v_{i+1/2}<0\end{cases},$$

$$a_i = \begin{cases} -\frac{s}{2}\left[4 - v_{i-1/2} - v_{i-3/2}\right]\left[\frac{\partial f}{\partial u}\right]_{i-1}^{n} & \text{if } v_{i+1/2} > 0 \\ 0 & \text{if } v_{i+1/2} < 0 \end{cases},$$

$$b_i = 1 - \Delta t e_i \left[\frac{\partial g}{\partial u}\right]_i^n - \frac{s}{2}\left[\frac{\partial f}{\partial u}\right]_i^n \begin{cases} v_{i-1/2} - 3 & \text{if } v_{i+1/2} > 0 \\ v_{i+1/2} + 3 & \text{if } v_{i+1/2} < 0 \end{cases},$$

$$d_i = \begin{cases} 0 & \text{if } v_{i+1/2} > 0 \\ \frac{s}{2}\left[4 + v_{i+1/2} + v_{i+3/2}\right]\left[\frac{\partial f}{\partial u}\right]_{i+1}^{n} & \text{if } v_{i+1/2} < 0 \end{cases}$$

and

$$k_i = \begin{cases} 0 & \text{if } v_{i+1/2} > 0 \\ -\frac{s}{2}\left[1 + v_{i+3/2}\right]\left[\frac{\partial f}{\partial u}\right]_{i+2}^{n} & \text{if } v_{i+1/2} < 0 \end{cases}.$$

Here, A is a penta-diagonal matrix and unfortunately requires a lot more calculations than before resulting in the interval of absolute stability and the accuracy of the scheme being reduced. However, Embid, Goodman and Majda[2] discussed using the first order tri-diagonal matrix for the second order Upwind approach based on the Engquist-Oscher scheme to increase the interval of absolute stability. Using the same approach, we can obtain the second order implicit Upwind approach, i.e.

$$\begin{bmatrix} b_0 & d_0 & 0 & 0 & 0 & \dots & 0 \\ a_1 & b_1 & d_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & b_2 & d_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{I-2} & b_{I-2} & d_{I-2} & 0 \\ 0 & \dots & 0 & 0 & a_{I-1} & b_{I-1} & d_{I-1} \\ 0 & \dots & 0 & 0 & 0 & a_I & b_I \end{bmatrix} \begin{bmatrix} u_0^{n+1} - u_0^n \\ u_1^{n+1} - u_1^n \\ u_2^{n+1} - u_2^n \\ \vdots \\ u_{I-2}^{n+1} - u_{I-2}^n \\ u_{I-1}^{n+1} - u_{I-1}^n \\ u_I^{n+1} - u_I^n \end{bmatrix} = \begin{bmatrix} G_0 - a_0\left(u_{-1}^{n+1} - u_{-1}^n\right) \\ G_1 \\ G_2 \\ \vdots \\ G_{I-2} \\ G_{I-1} \\ G_I - d_I\left(u_{I+1}^n - u_{I+1}^n\right) \end{bmatrix} \qquad (4.34)$$

where

$$G_i = \Delta t e_i\, g_i^n - \frac{s}{2}\begin{cases} \left(3 - v_{i-1/2}\right)\left(f_i^{n+1} - f_{i-1}^{n+1}\right) - \left(1 - v_{i-3/2}\right)\left(f_{i-1}^{n+1} - f_{i-2}^{n+1}\right) & \text{if } v_{i+1/2} > 0 \\ -\left(1 + v_{i+3/2}\right)\left(f_{i+2}^{n+1} - f_{i+1}^{n+1}\right) + \left(v_{i+1/2} + 3\right)\left(f_{i+1}^{n+1} - f_i^{n+1}\right) & \text{if } v_{i+1/2} < 0 \end{cases},$$

$$a_i = \begin{cases} -s\left[\dfrac{\partial f}{\partial u}\right]_{i-1}^{n} & \text{if } v_{i+1/2} > 0 \\ 0 & \text{if } v_{i+1/2} < 0 \end{cases}, \quad d_i = \begin{cases} 0 & \text{if } v_{i+1/2} > 0 \\ s\left[\dfrac{\partial f}{\partial u}\right]_{i+1}^{n} & \text{if } v_{i+1/2} < 0 \end{cases}$$

and

$$b_i = 1 + \text{sgn}(v_{i+1/2})s\left[\frac{\partial f}{\partial u}\right]_i^n - \Delta te_i\left[\frac{\partial g}{\partial u}\right]_i^n.$$

They also state that by using the first order matrix, the interval of absolute stability increases resulting in a more robust scheme. We can also apply flux-limiter methods to (4.34) to minimise any oscillations present in the numerical solution. This is obtained by replacing G_i in (4.34) with

$$G_i = \Delta te_i g_i^n - [F(u;i) - F(u;i-1)]$$

where

$$F(u;i) = F_L(u;i) + F_H(u;i)\phi_i$$

and

$$F_L(u;i) = \begin{cases} f_i^n & \text{if } v_{i+1/2} > 0 \\ f_{i+1}^n & \text{if } v_{i+1/2} < 0 \end{cases}$$

$$F_H(u;i) = \frac{1}{2}\begin{cases} (1 - v_{i-1/2})(f_i^n - f_{i-1}^n) & \text{if } v_{i+1/2} > 0 \\ -(1 + v_{i+3/2})(f_{i+2}^n - f_{i+1}^n) & \text{if } v_{i+1/2} < 0 \end{cases}.$$

where ϕ_i represents the flux-limiter, which is described in more detail in Chapter 2, Section 5.

4.3.3 Some Numerical Results for the Implicit Upwind Approach

If we apply (4.30) and (4.34) with and without flux-limiter to the test problem (4.2), we may obtain the numerical results in Figure 4-5. Here, we can see that the results of the first order implicit Upwind approach are quite accurate but the method suffers from dissipation. Also, notice that the second order implicit Upwind approach produced the most accurate results.

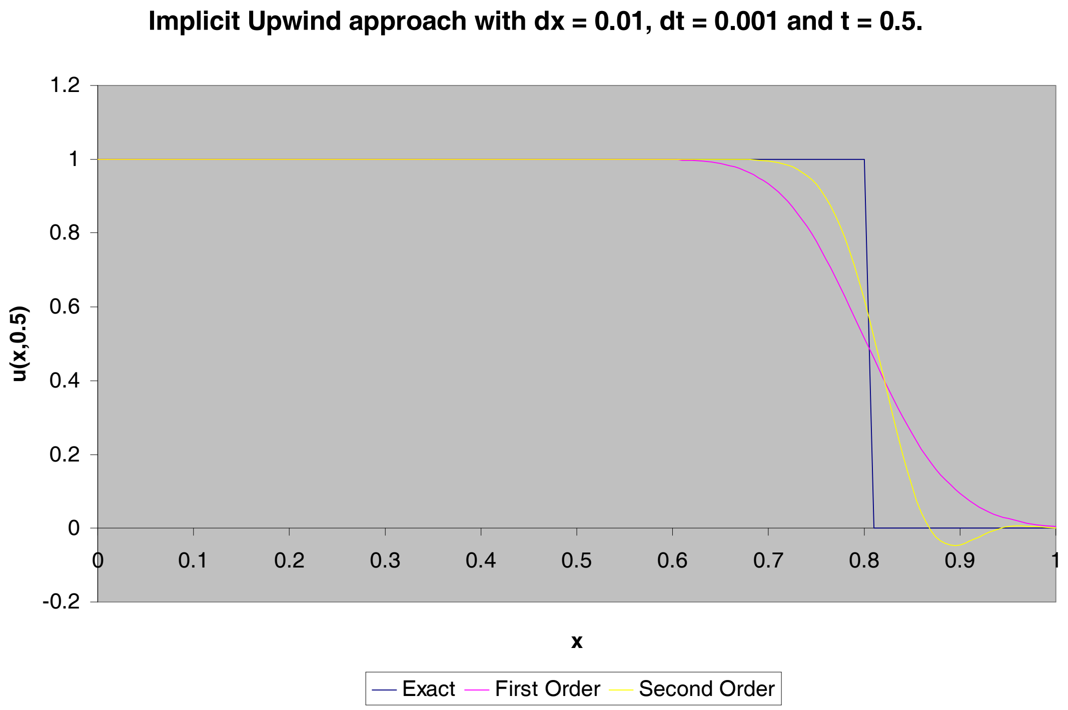


Figure 4-5: Comparison of schemes based on the implicit Upwind approach.

4.4 LeVeque and Yee's MacCormack Approach

In this sub-section we will look at how the MacCormack scheme, which is listed in Table 2-1, can be adapted to numerically approximate (4.1). This approach is frequently used and was discussed by Yee[5], LeVeque and Yee[1] and Embid, Goodman and Majda[2].

4.4.1 Explicit MacCormack Approach

We can approximate (4.1) by expanding on the explicit MacCormack scheme. The MacCormack method is the Lax-Wendroff scheme re-written in predictor-corrector form, i.e.

$$u_i^{n+1} = \frac{1}{2}\left(u_i^n + u_i^{(1)}\right) - \frac{s}{2}\left[f_i^{(1)} - f_{i-1}^{(1)}\right] \tag{4.35}$$

where

$$u_i^{(1)} = u_i^n - s\left(f_{i+1}^n - f_i^n\right)$$

for the conservation law without source term. We can adapt (4.35) to include the source terms explicitly and still maintain second order accuracy, i.e.

$$u_i^{n+1} = \frac{1}{2}\left(u_i^n + u_i^{(1)}\right) - \frac{s}{2}\left[f_i^{(1)} - f_{i-1}^{(1)}\right] + \Delta t \frac{R_i^{(1)}}{2} \tag{4.36}$$

where

$$u_i^{(1)} = u_i^n - s\left(f_{i+1}^n - f_i^n\right) + \Delta t R_i^n .$$

Here, if $f(u) = 0$, then (4.35) reduces to the standard two-stage Runge-Kutta method. Yee[5] also discussed a modified-flux approach for (4.36) which is

$$u_i^{n+1} = u_i^{(2)} + \left[\phi_{i+1/2}^{(2)} - \phi_{i-1/2}^{(2)}\right] \tag{4.37}$$

where

$$u_i^{(2)} = \frac{1}{2}\left(u_i^n + u_i^{(1)}\right) - \frac{s}{2}\left[f_i^{(1)} - f_{i-1}^{(1)}\right] + \Delta t \frac{R_i^{(1)}}{2} ,$$

$$u_i^{(1)} = u_i^n - s\left(f_{i+1}^n - f_i^n\right) + \Delta t R_i^n ,$$

$$\phi_{i+1/2}^{(2)} = \frac{1}{2}\left[\left|v_{i+1/2}\right| - v_{i+1/2}^2\right]\left(u_{i+1}^{(2)} - u_i^{(2)} - Q_{i+1/2}\right)$$

and $Q_{i+1/2}$ is chosen from Table 4-2.

Some choices of $Q_{i+1/2}$ where $\Delta_{i+1/2} = u_{i+1}^n - u_i^n$.
$Q_{i+1/2} = \min\text{mod}\left(\Delta_{i+1/2}, \Delta_{i-1/2}\right) + \min\text{mod}\left(\Delta_{i+1/2}, \Delta_{i+3/2}\right) - \Delta_{i+1/2}$
$Q_{i+1/2} = \min\text{mod}\left(\Delta_{i-1/2}, \Delta_{i+1/2}, \Delta_{i+3/2}\right)$
$Q_{i+1/2} = \min\text{mod}\left(2\Delta_{i-1/2}, 2\Delta_{i+1/2}, 2\Delta_{i+3/2}, \frac{1}{2}\left(\Delta_{i-1/2} + \Delta_{i+3/2}\right)\right)$

Table 4-2: Some choices of $Q_{i+1/2}$.

Here, (4.37) is only TVD if $R(x,t,u) = 0$, otherwise Yee[5] states that (4.37) satisfies the TVD properties as far as the numerical results are concerned, but is extremely difficult to prove that it is TVD.

4.4.2 Semi-Implicit MacCormack Approach

Yee[5] and LeVeque and Yee[1] also discuss an approach which considers the source term approximation to be at $(i,n+1)$ but still uses the explicit MacCormack scheme resulting in a semi-implicit scheme. This approach is obtained by re-writing (4.36) as

$$u_i^{n+1} = u_i^n + \frac{1}{2}\left[\left(u_i^{(2)} - u_i^{(1)}\right) + \left(u_i^{(1)} - u_i^n\right)\right]$$

where

$$\left(u_i^{(2)} - u_i^{(1)}\right) = -\frac{s}{2}\left[f_i^{(1)} - f_{i-1}^{(1)}\right] + \Delta t\left[R_i^{n+1}\right]^{(1)}$$

and

$$\left(u_i^{(1)} - u_i^n\right) = -s\left(f_{i+1}^n - f_i^n\right) + \Delta t R_i^{n+1}.$$

Now by using Taylor's theorem

$$R_i^{n+1} \approx R_i^n + \left(u_i^{n+1} - u_i^n\right)\left[\frac{\partial R}{\partial u}\right]_i^n + \ldots$$

we may obtain

$$u_i^{n+1} = u_i^n + \frac{1}{2}\left[\left(u_i^{(2)} - u_i^{(1)}\right) + \left(u_i^{(1)} - u_i^n\right)\right]$$

where

$$\left(u_i^{(2)} - u_i^{(1)}\right) = -\frac{s}{2}\left[f_i^{(1)} - f_{i-1}^{(1)}\right] + \Delta t R_i^{(1)} + \Delta t\bar{\theta}\left(u_i^{(2)} - u_i^{(1)}\right)\left[\frac{\partial R}{\partial u}\right]_i^{(1)}$$

and

$$u_i^{(1)} = u_i^n - s\left(f_{i+1}^n - f_i^n\right) + \Delta t R_i^n + \Delta t\bar{\theta}\left(u_i^{(1)} - u_i^n\right)\left[\frac{\partial R}{\partial u}\right]_i^n$$

where $0 \le \bar{\theta} \le 1$. Hence, by re-arranging we may obtain

$$u_i^{n+1} = u_i^n + \frac{1}{2}\left[\left(u_i^{(2)} - u_i^{(1)}\right) + \left(u_i^{(1)} - u_i^n\right)\right] \tag{4.38}$$

the semi-implicit MacCormack approach where

$$\left[1 - \Delta t\bar{\theta}\left[\frac{\partial R}{\partial u}\right]_i^n\right]\left(u_i^{(2)} - u_i^{(1)}\right) = -s\left(f_{i+1}^{(1)} - f_i^{(1)}\right) + \Delta t R_i^{(1)}$$

and

$$\left[1 - \Delta t\bar{\theta}\left[\frac{\partial R}{\partial u}\right]_i^n\right]\left(u_i^{(1)} - u_i^n\right) = -s\left(f_{i+1}^n - f_i^n\right) + \Delta t R_i^n.$$

Yee[5] discusses various choices of $\bar{\theta}$ and deduces that we can obtain second order by setting $\bar{\theta} = \frac{1}{2}$. We can also apply the modified flux described in the previous sub-section by re-writing (4.38) as

$$u_i^{n+1} = u_i^{(2)} + \left[\phi_{i+1/2}^{(2)} - \phi_{i-1/2}^{(2)}\right] \tag{4.39}$$

where

$$\left[1 - \Delta t\bar{\theta}\left[\frac{\partial R}{\partial u}\right]_i^n\right]\left(u_i^{(2)} - u_i^{(1)}\right) = -s\left(f_{i+1}^{(1)} - f_i^{(1)}\right) + \Delta t R_i^{(1)},$$

$$\left[1 - \Delta t\bar{\theta}\left[\frac{\partial R}{\partial u}\right]_i^n\right]\left(u_i^{(1)} - u_i^{n}\right) = -s\left(f_{i+1}^{n} - f_i^{n}\right) + \Delta t R_i^{n}$$

and

$$\phi_{i+1/2}^{(2)} = \frac{1}{2}\left[\left|v_{i+1/2}\right| - v_{i+1/2}^2\right]\left(u_{i+1}^{(2)} - u_i^{(2)} - Q_{i+1/2}\right)$$

and $Q_{i+1/2}$ is chosen from Table 4-2.

4.4.3 LeVeque and Yee's Splitting Method for the MacCormack Approach

LeVeque and Yee[1] also discuss a splitting method for the semi-implicit MacCormack approach discussed in this sub-section. The splitting method alternates between solving the conservation law with no source term

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0 \tag{4.40}$$

and then solving the ordinary differential equation

$$\frac{\partial u}{\partial t} = R(x,t,u), \tag{4.41}$$

i.e.

$$u_i^{n+1} = S_f(\Delta t) S_\psi(\Delta t) u_i^n$$

where $S_f(\Delta t)$ denotes the numerical solution of (4.40) and $S_\psi(\Delta t)$ denotes the numerical solution of (4.42). LeVeque and Yee[1] also state that in order to obtain second order accuracy, we can use the Strang splitting [11] to obtain

$$u_i^{n+1} = S_\psi\left(\frac{\Delta t}{2}\right) S_f(\Delta t) S_\psi\left(\frac{\Delta t}{2}\right) u_i^n \tag{4.42}$$

where $S_f(\Delta t)$ denotes the numerical solution of (4.40) and $S_\psi\left(\frac{\Delta t}{2}\right)$ denotes the numerical solution of (4.42). They also give a splitting method of the form (4.42) for the semi-implicit MacCormack approach with TVD discussed in the previous sub-section:

$S_\psi\left(\frac{\Delta t}{2}\right)$:

$$\left[1 - \frac{\Delta t}{4}\left[\frac{\partial R}{\partial u}\right]_i^n\right]\left(u_i^* - u_i^n\right) = \frac{\Delta t}{2} R_i^n$$

$$u_i^* = u_i^n + \left(u_i^* - u_i^n\right).$$

$S_k(\Delta t)$:

$$\left(u_i^{(1)} - u_i^*\right) = -s\left(f_i^* - f_{i-1}^*\right)$$

$$u_i^{(1)} = u_i^* + \left(u_i^{(1)} - u_i^*\right)$$

$$\left(u_i^{(2)} - u_i^{(1)}\right) = -s\left(f_{i+1}^{(1)} - f_i^{(1)}\right)$$

$$u_i^{(2)} = u_i^* + \frac{1}{2}\left[\left(u_i^{(2)} - u_i^{(1)}\right) + \left(u_i^{(1)} - u_i^*\right)\right]$$

$$u_i^{**} = u_i^{(2)} + \left[\phi_{i+1/2}^* - \phi_{i-1/2}^*\right].$$

$S_\psi\left(\frac{\Delta t}{2}\right)$:

$$\left[1 - \frac{\Delta t}{4}\left[\frac{\partial R}{\partial u}\right]_i^{**}\right]\left(u_i^{**} - u_i^{(2)}\right) = \frac{\Delta t}{2} R_i^{**}$$

$$u_i^{n+1} = u_i^{**} + \left(u_i^{**} - u_i^{(2)}\right). \tag{4.43}$$

Here

$$\phi_{i+1/2}^* = \frac{1}{2}\left[\left|v_{i+1/2}\right| - v_{i+1/2}^2\right]\left(u_{i+1}^* - u_i^* - Q_{i+1/2}\right)$$

where $Q_{i+1/2}$ is chosen from Table 4-2. We can also replace $\phi_{i+1/2}^*$ with $\phi_{i+1/2}^{(2)}$.

4.4.4 Some Numerical Results for the MacCormack Approach

If we apply (4.37), (4.39) and (4.43) with and without TVD to the test problem (4.2), we may obtain the numerical results in Figure 4-6 and Figure 4-7. Here, we can see that all three approaches give practically the same results but, as with the Lax-Wendroff approach, this will not always be the case.

Throughout this chapter, we have seen that there are a variety of methods used for approximating conservation laws with a source term present, which is a function of x, t and u. We have also obtained some very accurate results but in the next chapter, we will see that the different approaches discussed throughout this project are not so accurate when the source term becomes stiff.

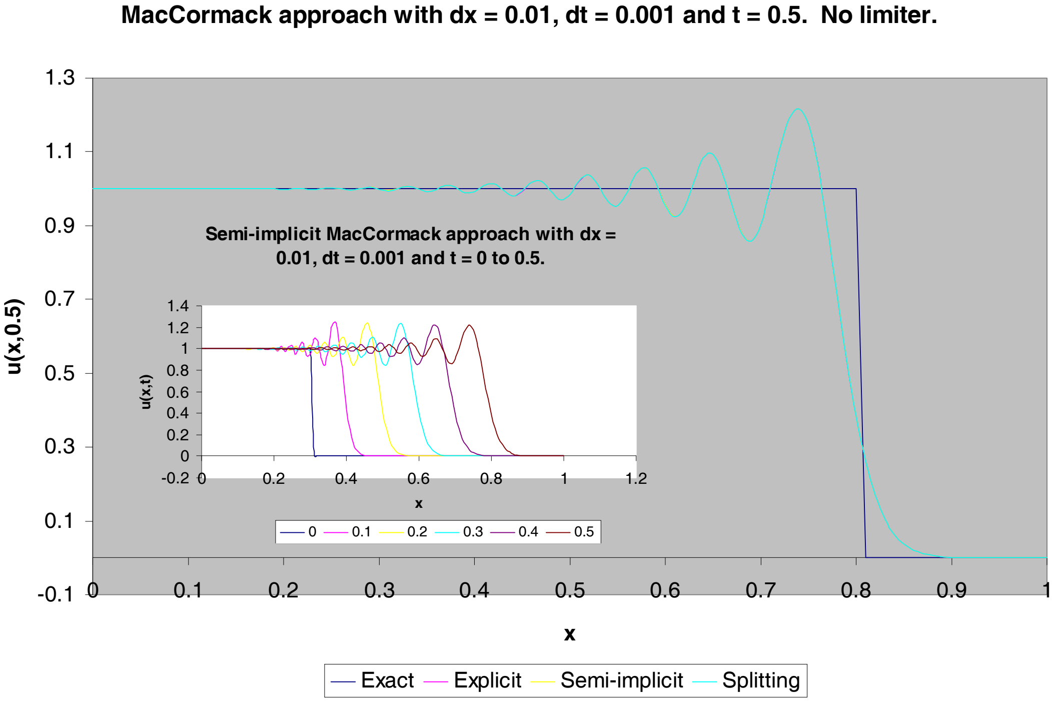


Figure 4-6: Comparison of explicit, semi-implicit and splitting method for MacCormack approach.

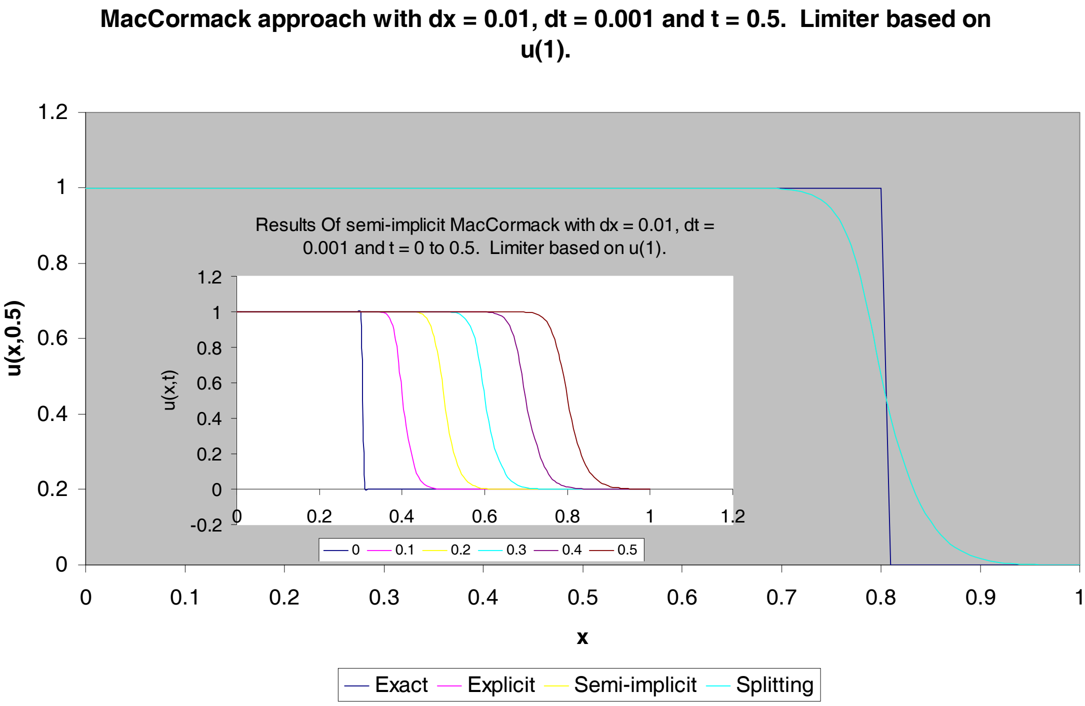


Figure 4-7: Comparison of explicit, semi-implicit and splitting method for MacCormack approach with TVD.

5 Some Numerical Results

In this chapter, we will apply the different approaches discussed throughout this dissertation to a specific test problem (5.1) which was considered by LeVeque and Yee[1], i.e.

$$\frac{\partial u}{\partial t}+\frac{\partial u}{\partial x}=R(u), \tag{5.1}$$

where

$$R(u)=-\mu u(u-1)\left(u-\frac{1}{2}\right),$$

with initial data

$$u(x,0)=\begin{cases}1 & \text{if } x\le 0.3\\ 0 & \text{if } x>0.3\end{cases}$$

and whose exact solution, which is shown in Figure 5-1, is

$$u(x,t)=\begin{cases}1 & \text{if } x\le 0.3+t\\ 0 & \text{if } x>0.3+t\end{cases}. \tag{5.2}$$

Here, $\Delta t\mu$ determines the stiffness of (5.1) and as sμ becomes greater than 1 the propagation speed of some approaches can be greatly affected. When $\Delta t\mu > 1$, the source term is said to be stiff since, for most approaches, we can no longer choose an adequate step-size in time to produce accurate results. A stiff source term moves the discontinuity to a cell boundary for each time step resulting in the discontinuity being moved at entirely the wrong speed. For example, if we apply the Lax-Wendroff approach (4.8) to the test problem (5.1), with μ = 1, 10, 100 and 1000, then we may obtain the numerical results in Figure 5-2, Figure 5-3, Figure 5-4 and Figure 5-5 respectively. Here, we can see that as $\Delta t\mu$ increases, the source term becomes stiff and the numerical approximation becomes less and less accurate. This is because as $\Delta t\mu$ increases, the discontinuity moves slower and slower which means that when the source term is stiff, the scheme is no longer conservative. However, not all of the schemes discussed in Chapter 4 will exhibit this behaviour, as we will see later.

We will use test problem (5.1) to compare the results of some of the methods discussed throughout this dissertation to ascertain which approach produces the most accurate results by seeing which approaches are conservative as the source term becomes stiff.

Name Of Approach	Reference No.	Order	Paper
Explicit 'adding'	(4.5)	1 / 2	-
Semi-implicit 'adding'	(4.7)	1 / 2	-
Lax-Wendroff	(4.8)	2	-
MPDATA	(4.9)	2	Smolarkiewicz + Margolin[3]
Roe's Explicit Upwind I	(4.21)	1	Roe[6], Vazquez + Bermudez[4]
Roe's Explicit Upwind II	(4.23)	2	Roe[6], Vazquez + Bermudez[4]
Implicit Upwind I	(4.30)	1	Embid, Goodman + Majda[2]
Implicit Upwind II	(4.34)	2	Embid, Goodman + Majda[2]
Explicit MacCormack	(4.37)	2	Yee[5], LeVeque + Yee[1], Embid, Goodman + Majda[2]
Semi-Implicit MacCormack	(4.39)	2	Yee[5], LeVeque + Yee[1]
Splitting Method (MacCormack)	(4.43)	2	LeVeque + Yee[1]

Table 5-1: Some different approaches for numerically approximating (5.1).

We will be discussing the results of the schemes listed in Table 5-1 which can also be found in Appendix A where they are written in full.

5.1 Explicit and Implicit 'Adding' Approach.

Now, by applying (4.5) and (4.7) to the test problem (5.1), we may obtain the numerical results in Figure 5-6, Figure 5-7 and Figure 5-8. Here, Figure 5-6 and Figure 5-7 show similar results, where the Lax-Wendroff with source term 'added' is the least accurate since the method moved the discontinuity the slowest resulting in the discontinuity being at approximately $x = 0.45$ at $t = 0.5$ instead of at $x = 0.8$ at $t = 0.5$. The Upwind with source term 'added' was the second most accurate since the discontinuity was at approximately $x = 0.63$ at $t = 0.5$ instead of at $x = 0.8$ at $t = 0.5$. The Lax-Wendroff with Superbee flux-limiter and source term 'added' gave the most accurate results since the discontinuity was at approximately $x = 0.73$ instead of at $x = 0.8$ at $t = 0.5$. Also notice how the Upwind, Lax-Wendroff and Lax-Wendroff with Superbee flux-limiter all gave very 'steep' discontinuities with no dissipation present but each method varied considerably as to where the discontinuity was at $t = 0.5$. Figure 5-8 shows us that even though the numerical results in Figure 5-6 and Figure 5-7 look similar, they are not. Here, we can see that the semi-implicit approach is more accurate than the explicit approach since the discontinuity of the semi-implicit approach is nearer to $x = 0.8$ at $t = 0.5$ than the discontinuity of the explicit approach. Hence, in general the semi-implicit approach is more accurate than the explicit approach.

Notice how the 'adding' approach is no longer conservative when the source term is stiff. This is because the interval of absolute stability of the source term approximation with $\mu = 1000$ is very small and has been breached resulting in the discontinuity being moved at the incorrect wave speed. Thus a very small step-size would be required to ensure stability of the source term approximation resulting in the scheme being impractical for approximating conservation laws with a stiff source term.

5.2 Lax-Wendroff approach

Now, by applying (4.8) to the test problem (5.1), we may obtain the numerical results in Figure 5-9. Also, by using the numerical results in the previous subsection, we may obtain the numerical results in Figure 5-10. Figure 5-9 shows similar results to that of Figure 5-6 and Figure 5-7 where the source term was 'added' explicitly or semi-implicitly. But if we look at Figure 5-10, we can see that the Lax-Wendroff approach is more accurate than the explicit 'adding' approach but less accurate than the semi-implicit 'adding' approach. However, there is very little difference in these approaches since they all placed the discontinuity at approximately $x = 0.73$ at $t = 0.5$ instead of at $x = 0.8$ at $t = 0.5$. So these schemes are no longer conservative when the source term is stiff.

5.3 MPDATA Approach

Now, by applying (4.9) to the test problem (5.1), we may obtain the numerical results in Figure 5-11. Here, we can see that the MPDATA approach has numerically approximated (5.2) considerably more accurately, placing the discontinuity near $x = 0.77$ at $t = 0.5$, than the previous two approaches, which placed the discontinuity at approximately $x = 0.73$ at $t = 0.5$. The MPDATA approach with Superbee flux-limiter is not much more accurate than without Superbee flux-limiter whereas in the previous two cases, the results with TVD were considerably more accurate than without TVD.

The MPDATA approach is the first method that has ensured conservation when the source term is stiff. This is because the MPDATA approach compensates for the terms in the truncation error due to the source term approximation resulting in a conservative method even when the source term is stiff.

5.4 Roe's Upwind Approach

Now, by applying (4.21) and (4.23) to the test problem (5.1), we may obtain the numerical results in Figure 5-12. Here, we can see that by using (4.21), the method has moved the discontinuity too fast. The discontinuity should be at $x = 0.8$ at $t = 0.5$ but the first order explicit Upwind approach has placed the discontinuity at approximately $x = 0.85$ at $t = 0.5$. However, the first order explicit Upwind is the most accurate numerical approximation out of the three displayed in Figure 5-12. The second order Upwind method failed to move the discontinuity at all and produced oscillations on both sides of the discontinuity. The second order Upwind method with Superbee flux-limiter moved the discontinuity too fast resulting in the discontinuity being at approximately $x = 0.9$ at $t = 0.5$ instead of at $x = 0.8$ at $t = 0.5$. This shows us that Roe's Upwind approach is no longer conservative when the source term is stiff.

5.5 Implicit Upwind Approach

Now, by applying (4.30) and (4.34) to the test problem (5.1), we may obtain the numerical results in Figure 5-13. Here, we can see that the second order implicit Upwind has produced the most accurate numerical results seen so far. Also notice how the first order implicit Upwind approach has given the least accurate results due to the discontinuity moving too slow resulting in the discontinuity being at $x = 0.6$ when $t = 0.5$ instead of at $x = 0.8$ when $t = 0.5$. Hence, the first order implicit Upwind approach is no longer conservative when the source term is stiff but the second order implicit Upwind approach is conservative when the source term is stiff and produces very accurate results.

5.6 MacCormack Approach

Now, by applying (4.37), (4.39) and (4.43) to the test problem (5.1), we may obtain the numerical results in Figure 5-14, Figure 5-15, Figure 5-16 and Figure 5-17. Here, we can see that Figure 5-14 and Figure 5-15 are showing similar results where the numerical results obtained without a limiter are the least accurate, the numerical results obtained with a $u^{(1)}$ limiter give the second most accurate results and the $u^{(2)}$ limiter gives the most accurate results. However, in Figure 5-16 even though the results are similar to Figure 5-14 and Figure 5-15, the most accurate numerical results are with the limiter $u^{(*)}$ and the second most accurate numerical results are with the limiter $u^{(2)}$. Figure 5-17 also shows us that the Splitting method produces the most accurate numerical results followed by the semi-implicit MacCormack approach and then the least accurate was the explicit MacCormack approach. So, overall the Splitting method is the most accurate but all methods are no longer conservative when the source term is stiff.

LeVeque and Yee[1] also observed that the explicit MacCormack approach, the semi-implicit MacCormack approach and the splitting method were no longer conservative when the source term is stiff. Their results, with $\Delta x = 0.02$ and $\Delta t = 0.0015$, showed that the splitting method moved the discontinuity too fast and the semi-implicit method moved the discontinuity too slow, if at all, when the source term was stiff.

5.7 Overall Comparison

So far, we have looked at each approach individually but we will now compare all of the different approaches listed in Table 5-1 to see which approach produced the most accurate numerical results when applied to the test problem (5.1).

5.7.1 First Order Comparison

If we apply all the first order approaches listed in Table 5-1 to the test problem (5.1), then we may obtain the numerical results in Figure 5-18. Here, we can see that Roe's Upwind approach has obtained the most accurate numerical approximation. However, Roe's Upwind approach is not very accurate since the numerical approximation moved the discontinuity too fast resulting in the discontinuity being at $x = 0.85$ when $t = 0.5$ instead of at $x = 0.8$ when $t = 0.5$. The explicit 'adding' approach and the implicit Upwind approach both gave very similar results and were the least accurate due to both schemes moving the discontinuity too slow resulting in the discontinuity being at approximately $x = 0.6$ when $t = 0.5$. The semi-implicit 'adding' approach was the second most accurate but also moved the discontinuity too slow resulting in the discontinuity being at approximately $x = 0.65$ when $t = 0.5$. Hence, overall all first order schemes either moved the discontinuity too fast or too slow when the source term is stiff resulting in an inaccurate numerical approximation of the test problem (5.1).

5.7.2 Second Order Comparison

If we apply all the first order approaches listed in Table 5-1 to the test problem (5.1), then we may obtain the numerical results in Figure 5-19. Here we can see that the most accurate second order approach was the implicit Upwind followed by the MPDATA approach. The semi-implicit 'adding' approach, semi-implicit

MacCormack approach, explicit 'adding' approach, Lax-Wendroff approach, explicit MacCormack approach and Splitting method based on the MacCormack approach all gave similar inaccurate results. They all moved the discontinuity too slow resulting in the discontinuity being at approximately x = 0.45 when t = 0.5 instead of at x = 0.8 when t = 0.5. Also, notice how Roe's Upwind approach failed to move the discontinuity at all. Hence, the most accurate second order scheme was the implicit Upwind approach followed by the MPDATA approach with the implicit Upwind giving very accurate results and the MPDATA approach giving accurate results. Here, most of the schemes were not conservative except for the second order implicit Upwind approach and the MPDATA approach.

5.7.3 Second Order with TVD Comparison

If we apply all the first order approaches listed in Table 5-1 to the test problem (5.1), then we may obtain the numerical results in Figure 5-20. Here, we can see that the explicit MacCormack approach, semi-implicit MacCormack approach and the splitting method, based on the semi-implicit MacCormack approach, all produced the least accurate results. This is because the MacCormack approach moved the discontinuity too slowly resulting in the discontinuity being at approximately x = 0.6 when t = 0.5 instead of at x = 0.8 when t = 0.5. The explicit 'adding' approach, semi-implicit 'adding' approach and the Lax-Wendroff approach all produced the second least accurate results. This is because the 'adding' approach and the Lax-Wendroff approach moved the discontinuity too slowly resulting in the discontinuity being at approximately x = 0.73 when t = 0.5. The implicit Upwind approach, which produced the most accurate results in the second order comparison, produced the second least accurate results. Here, the method has moved the discontinuity too fast resulting in the discontinuity being at x = 0.95 when t = 0.5. Roe's Upwind approach, which failed to move the discontinuity at all in the second order comparison, also produced the second most accurate set of results. However, the results of Roe's upwind approach were not very accurate since the method moved the discontinuity too fast resulting in the discontinuity being at approximately $x = 0.9$ when $t = 0.5$. The most accurate method for the second order approach with TVD was the MPDATA approach. The MPDATA approach moved the discontinuity too slow resulting in the discontinuity being at approximately $x = 0.78$ when $t = 0.5$. All of the schemes with TVD are no longer conservative when the source term is stiff.

5.7.4 Conclusion

Hence, overall the second order approach with TVD did not necessarily produce more accurate results than without TVD. In fact the most accurate results were obtained by not using TVD where two of the approaches were conservative when the source term was stiff. However, some of the approaches improved when TVD was applied and others became less accurate. This is because in most cases, when TVD was applied the discontinuity would move faster. In addition, the majority of first order approaches produced extremely inaccurate results except for Roe's Upwind approach which slightly overshot the discontinuity.

5.8 Changing the Step-Size when the Source Term is Stiff

Throughout this section, we have only considered the numerical results using $\Delta x = 0.01$ and $\Delta t = 0.001$, which implies that the Courant number is $s = 0.1$. However, when the source term is stiff, the accuracy of some of the schemes can vary if the step-size is changed. For example, if we use the first order explicit Upwind approach

(4.21) on the test problem (5.1) with $\Delta x = 0.02$ and $\Delta t = 0.0025$, which implies that the Courant number is $s = 0.125$, then we may obtain the results in Figure 5-21. Here we would expect the results to be less accurate than the results shown in Figure 5-12 but Figure 5-21 shows that the results of the first order explicit Upwind approach are more accurate since the approach moved the discontinuity slower than in Figure 5-12. I.e. when we used $\Delta x = 0.01$ and $\Delta t = 0.001$ the explicit first order Upwind approach moved the discontinuity too fast resulting in the discontinuity being at approximately $x = 0.85$ at $t = 0.5$ instead of at $x = 0.8$ at $t = 0.5$. However when we used $\Delta x = 0.02$ and $\Delta t = 0.0025$ the explicit first order Upwind approach moved the discontinuity a little slower than with $\Delta x = 0.01$ and $\Delta t = 0.001$ and approximated the discontinuity at $x = 0.8$ at $t = 0.5$. This shows us that the speed of the discontinuity depends greatly on the Courant number and as the Courant number decreases, the speed of the discontinuity increases. Also notice how the results in Figure 5-21 of the explicit second order Upwind approach with or without Superbee flux-limiter are similar to the results in Figure 5-12. Hence, for the explicit first order Upwind approach a small step-size does not always give the most accurate numerical approximation but Figure 5-22 shows us that if the Courant number becomes too large then the numerical solution becomes unstable.

Throughout this chapter, we have seen that if the source term is stiff then the majority of approaches discussed in this dissertation are no longer conservative, i.e. most of the approaches moved the discontinuity too slow or too fast. However, we have obtained some very accurate numerical results when the source term is stiff. In Chapter 6, we will compare the most accurate approaches of the first order, second order and second order with TVD to see which approach is the most accurate overall.

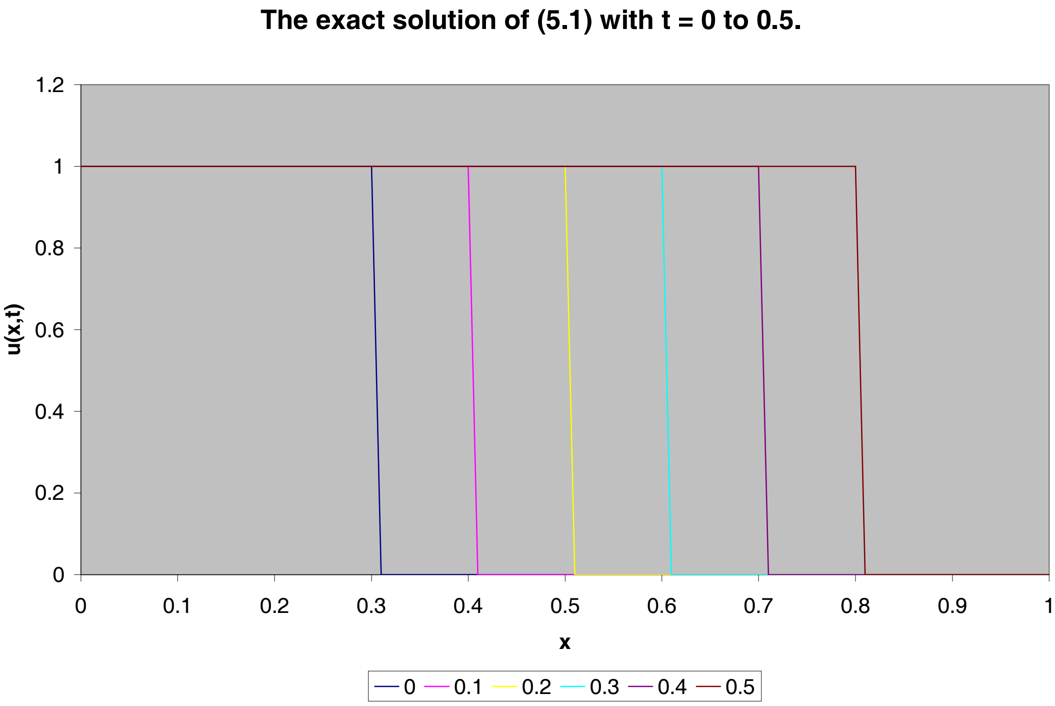


Figure 5-1: The exact solution (5.2).

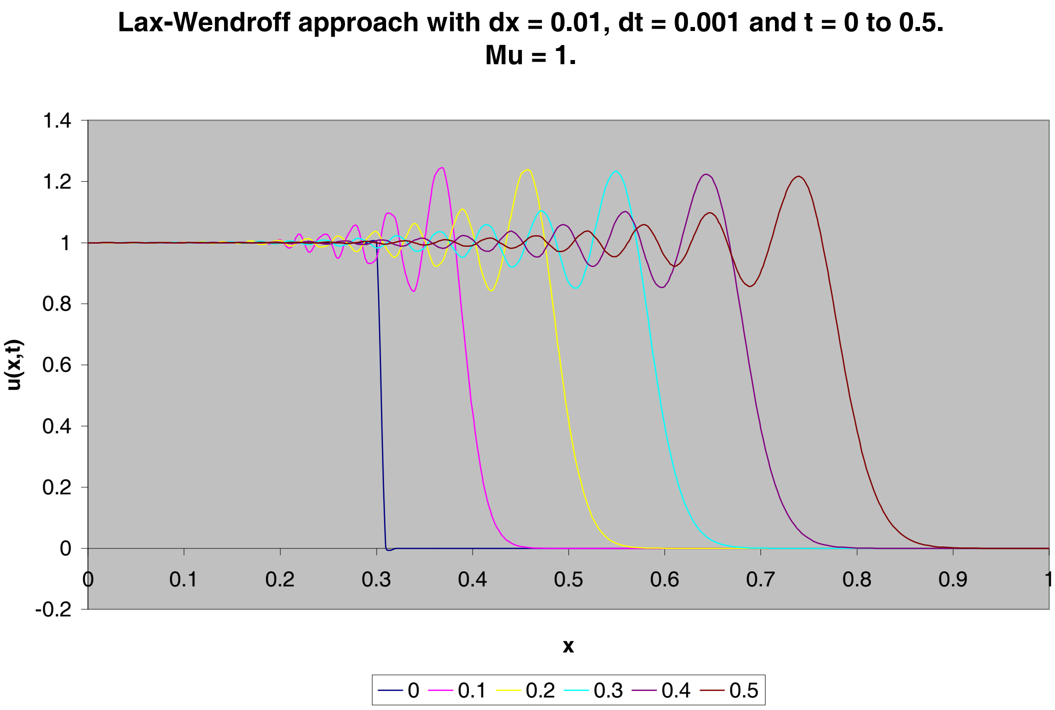


Figure 5-2: Lax-Wendroff approach applied to (5.1) with $\mu = 1$.

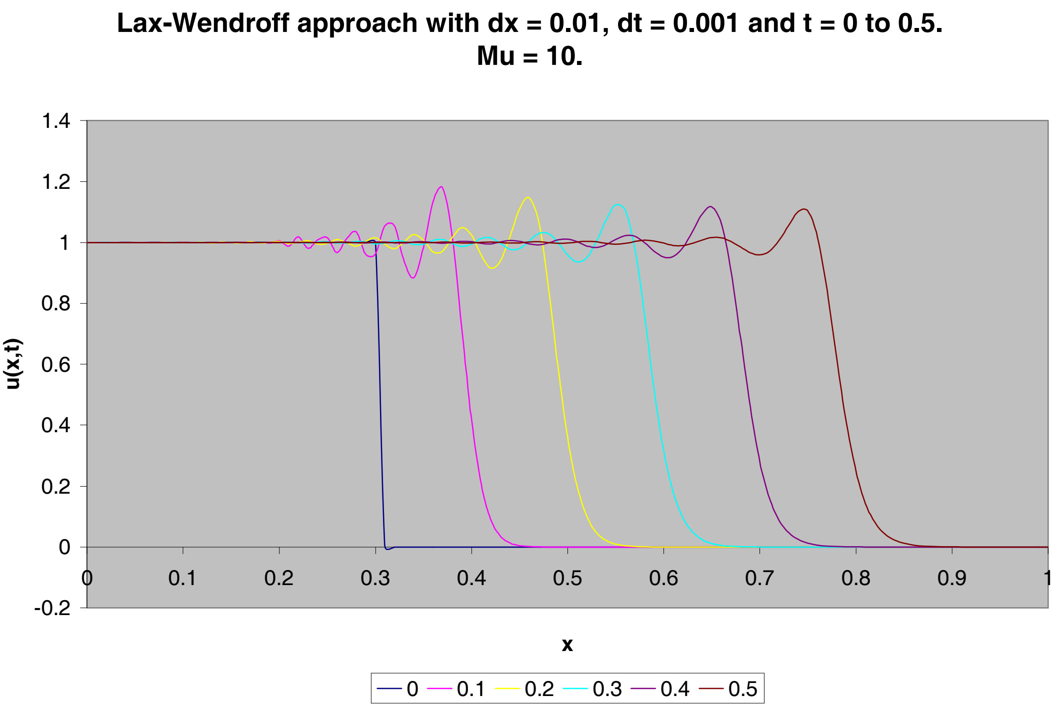


Figure 5-3: Lax-Wendroff approach applied to (5.1) with μ = 10.

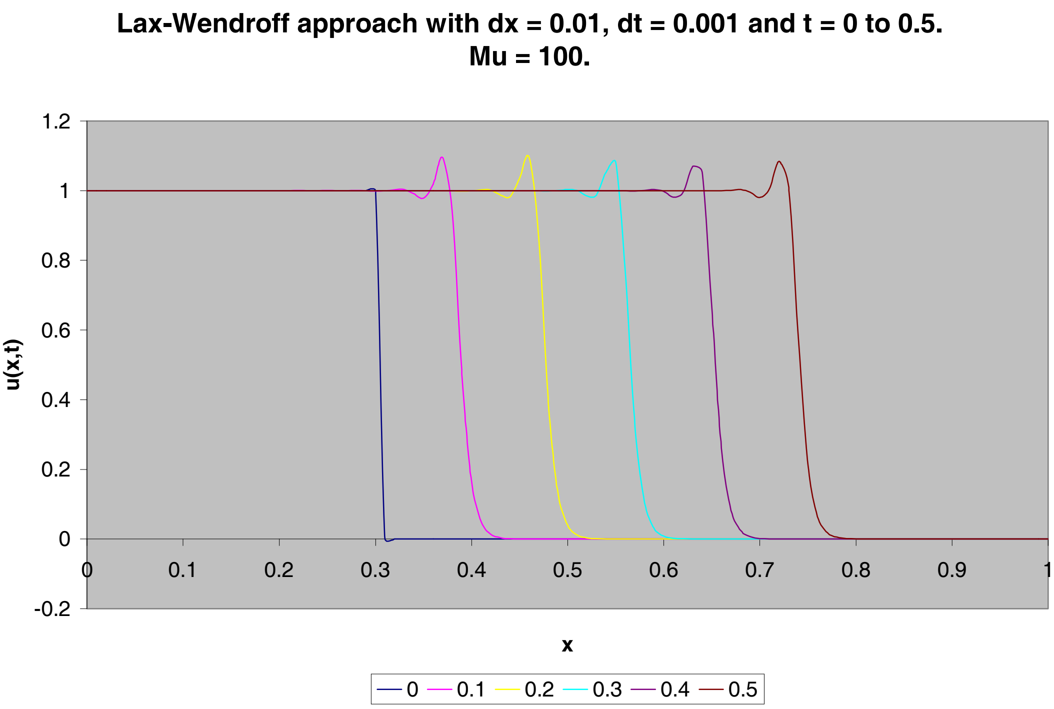


Figure 5-4: Lax-Wendroff approach applied to (5.1) with μ = 100.

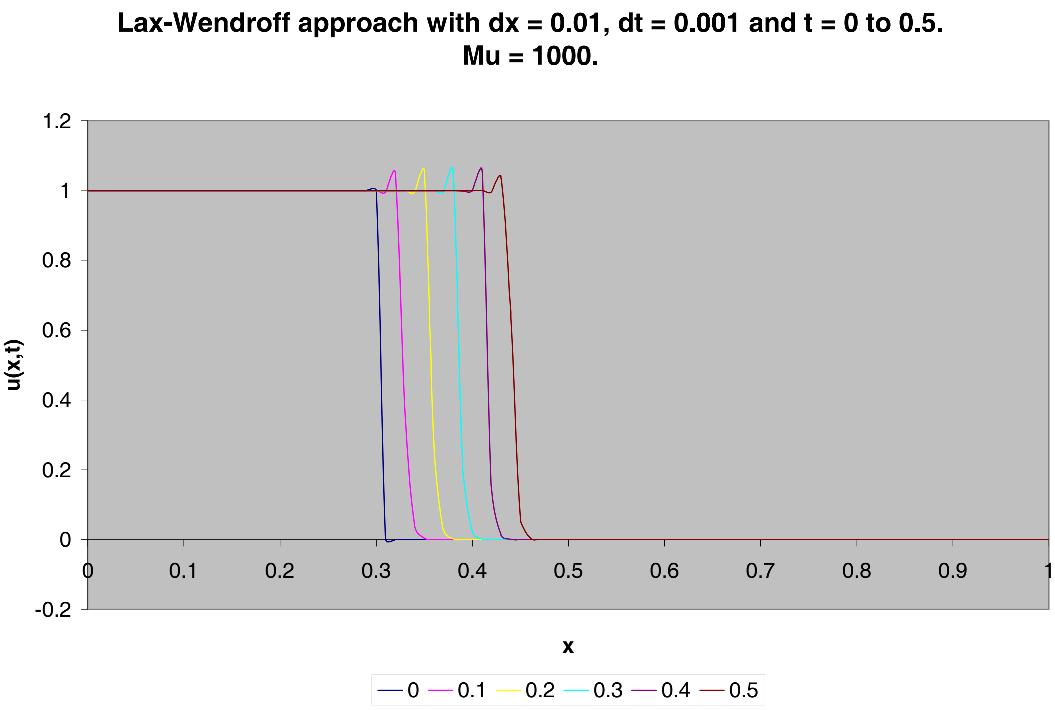


Figure 5-5: Lax-Wendroff approach applied to (5.1) with μ = 1000.

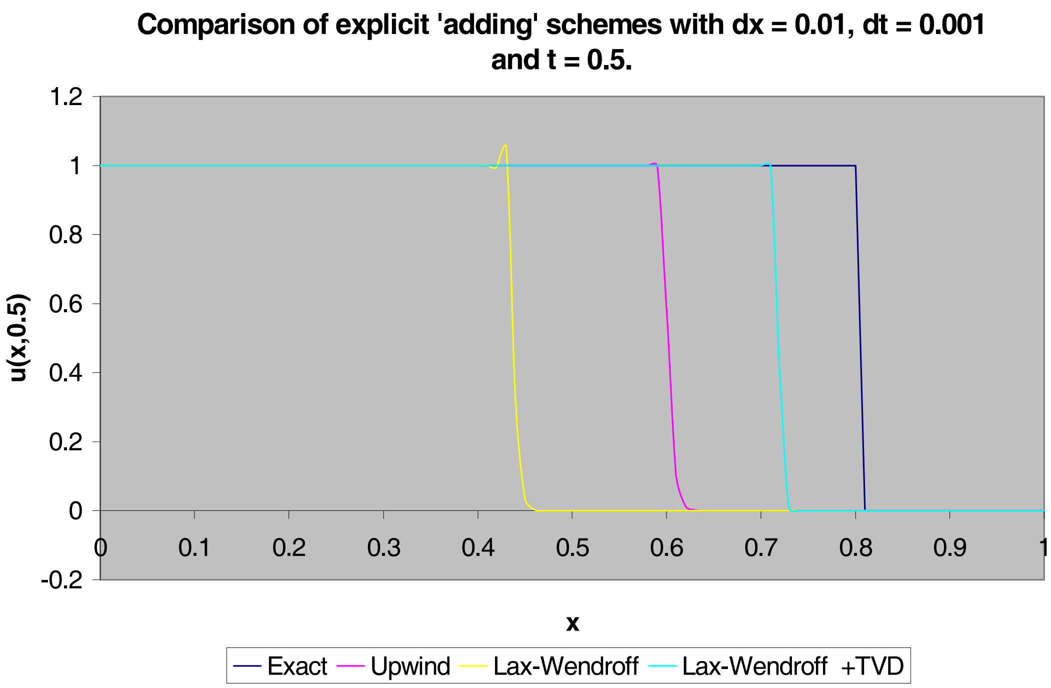


Figure 5-6: Explicit ‘adding’ approach with stiff source term.

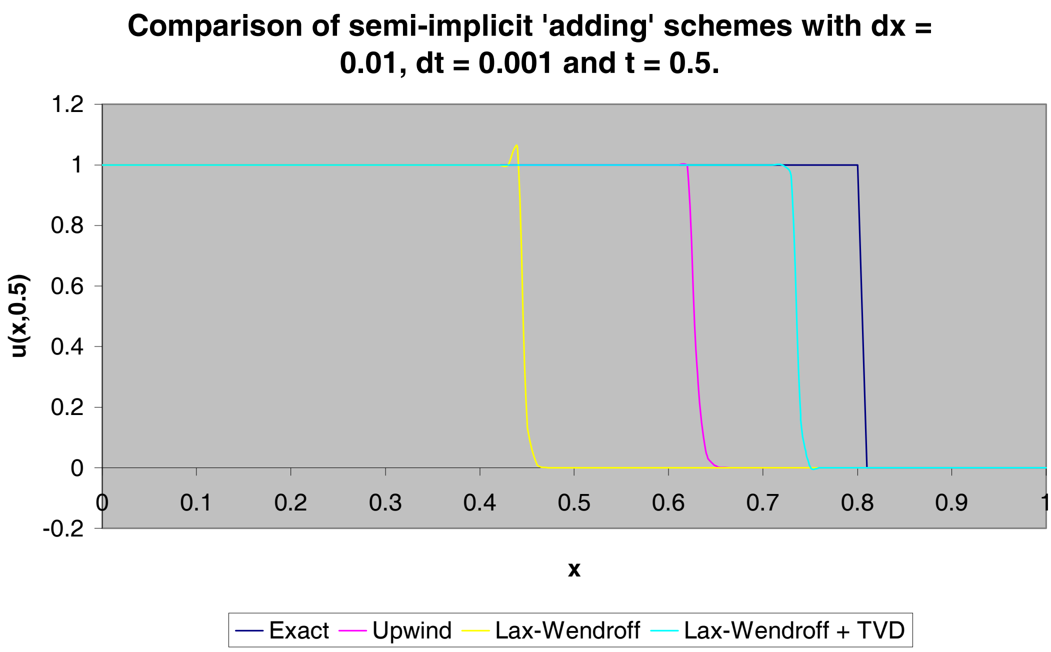


Figure 5-7: Semi-implicit ‘adding’ approach with stiff source term.

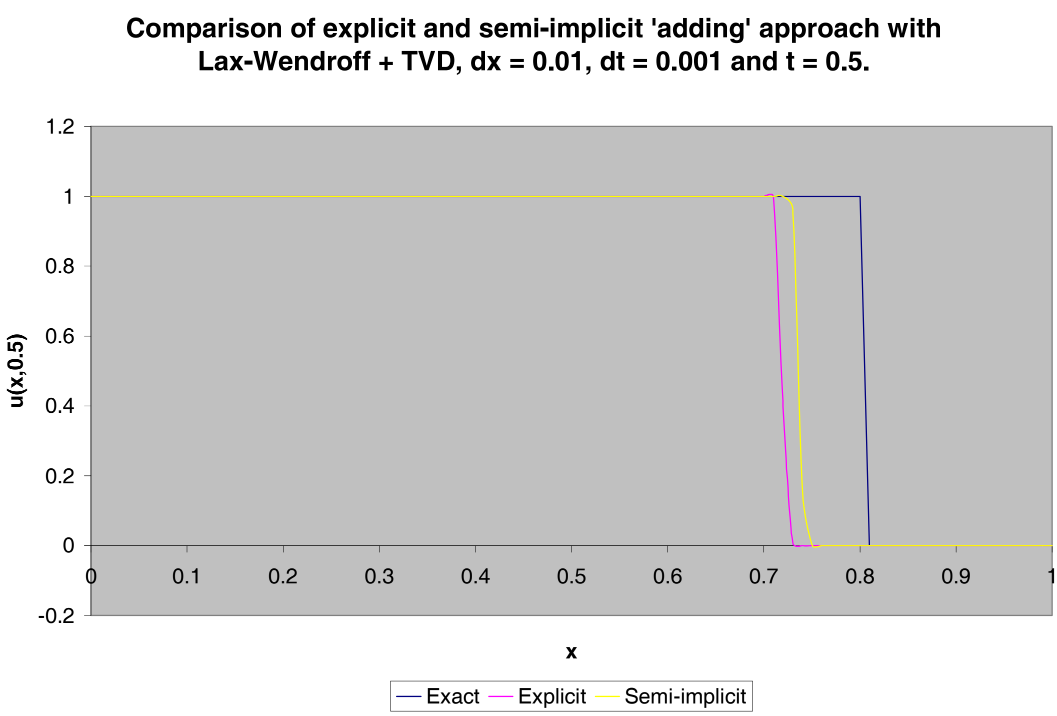


Figure 5-8: Comparison of explicit and semi-implicit ‘adding’ approach.

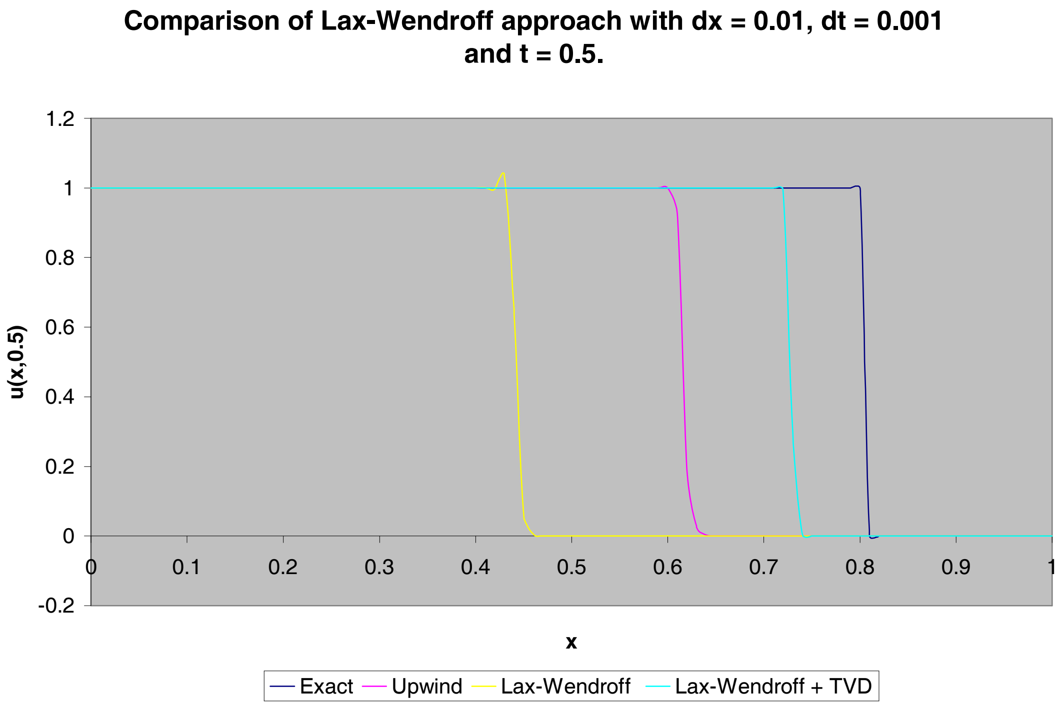


Figure 5-9: Comparison of Lax-Wendroff approach.

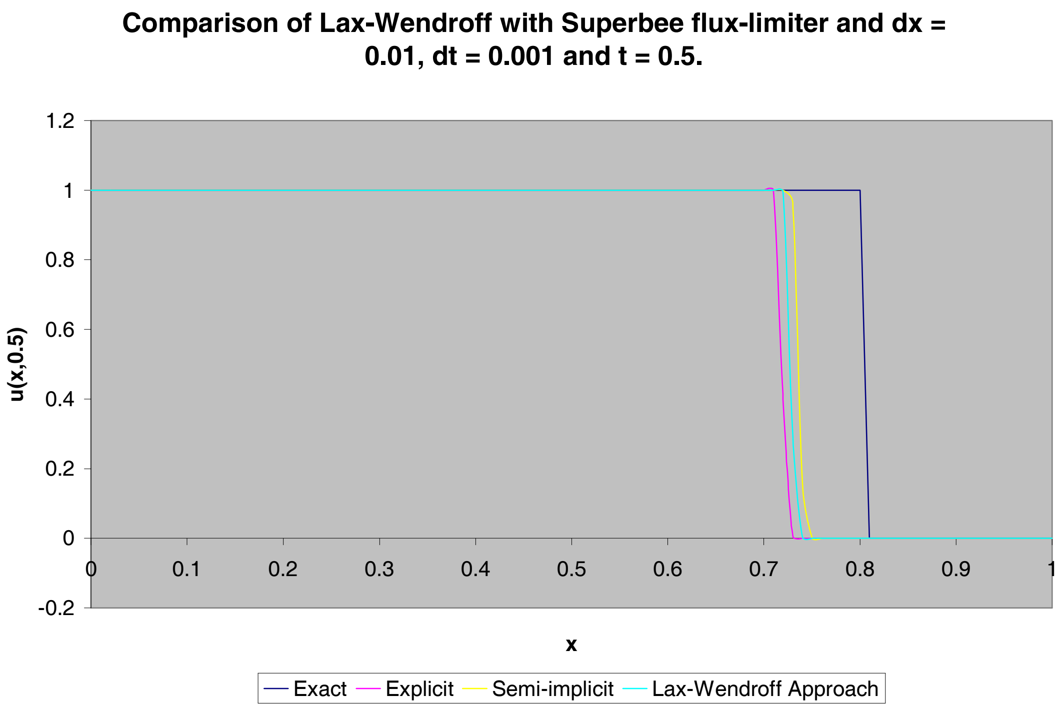


Figure 5-10: Comparison of Lax-Wendroff with Superbee flux-limiter.

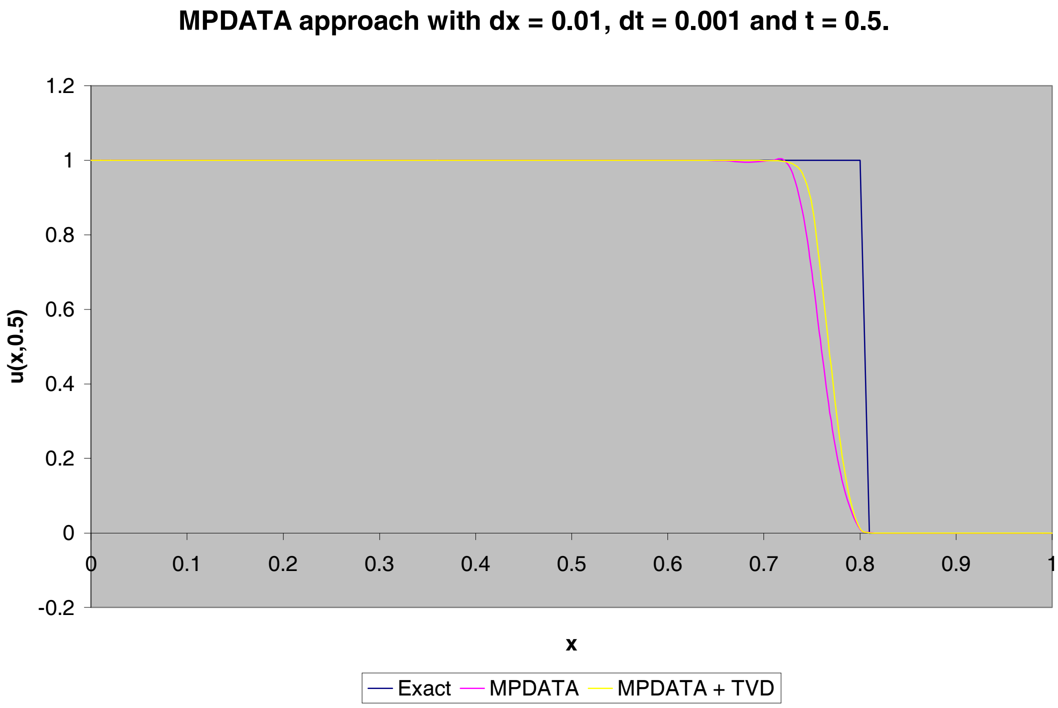


Figure 5-11: MPDATA approach for stiff source term.

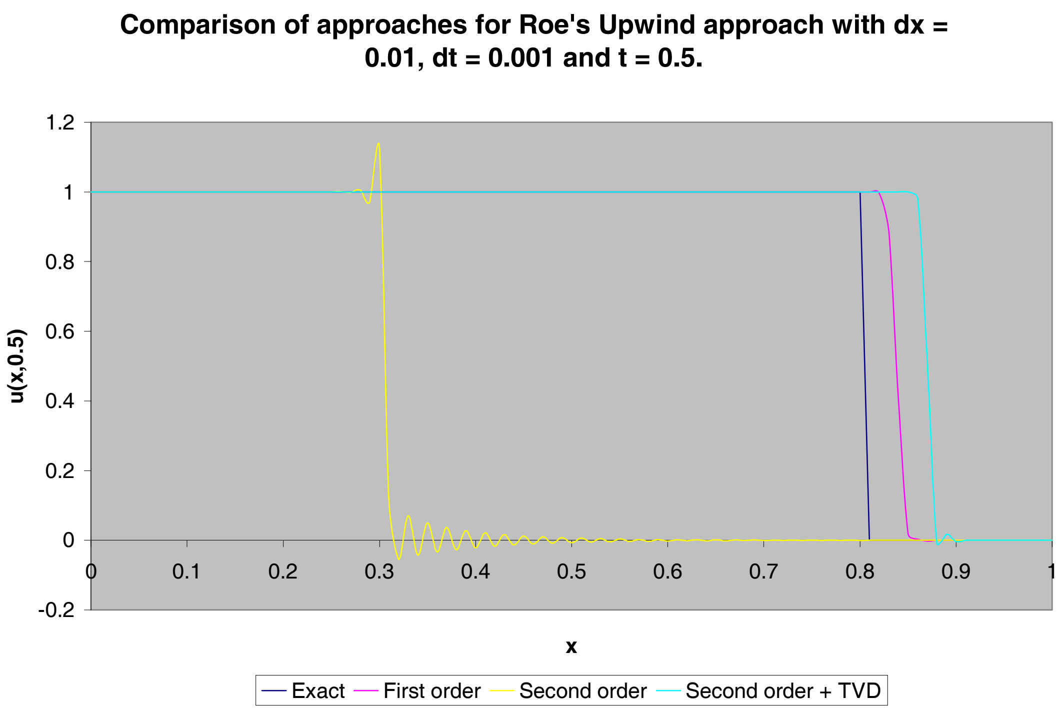


Figure 5-12: Roe's Upwind approach with stiff source term.

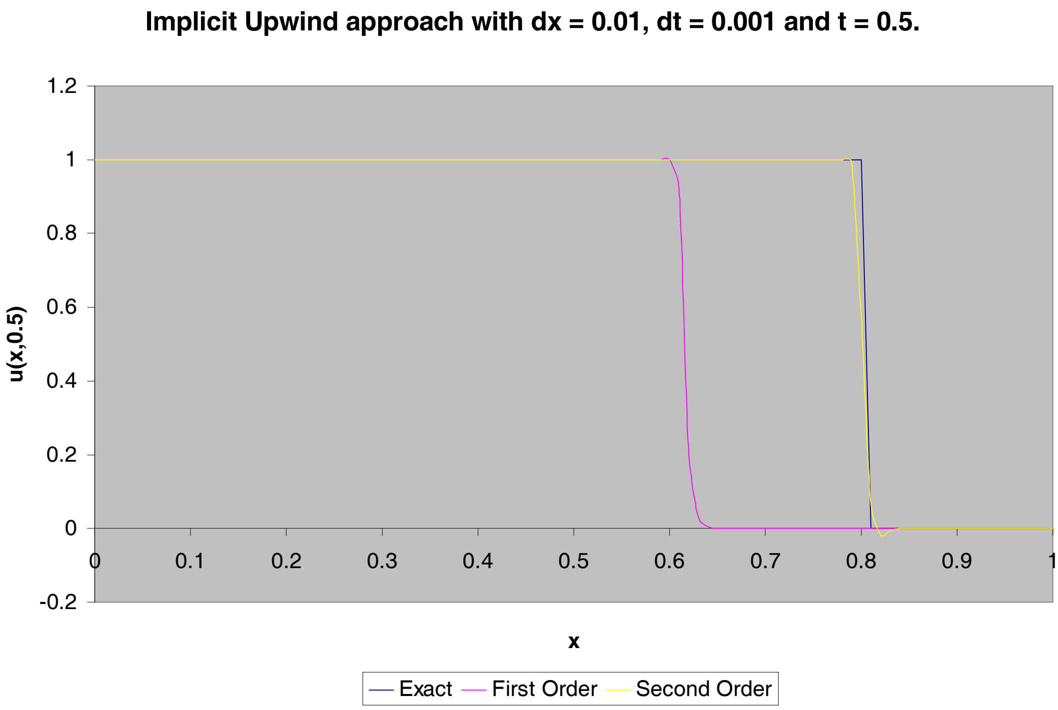


Figure 5-13: Implicit Upwind approach with stiff source term.

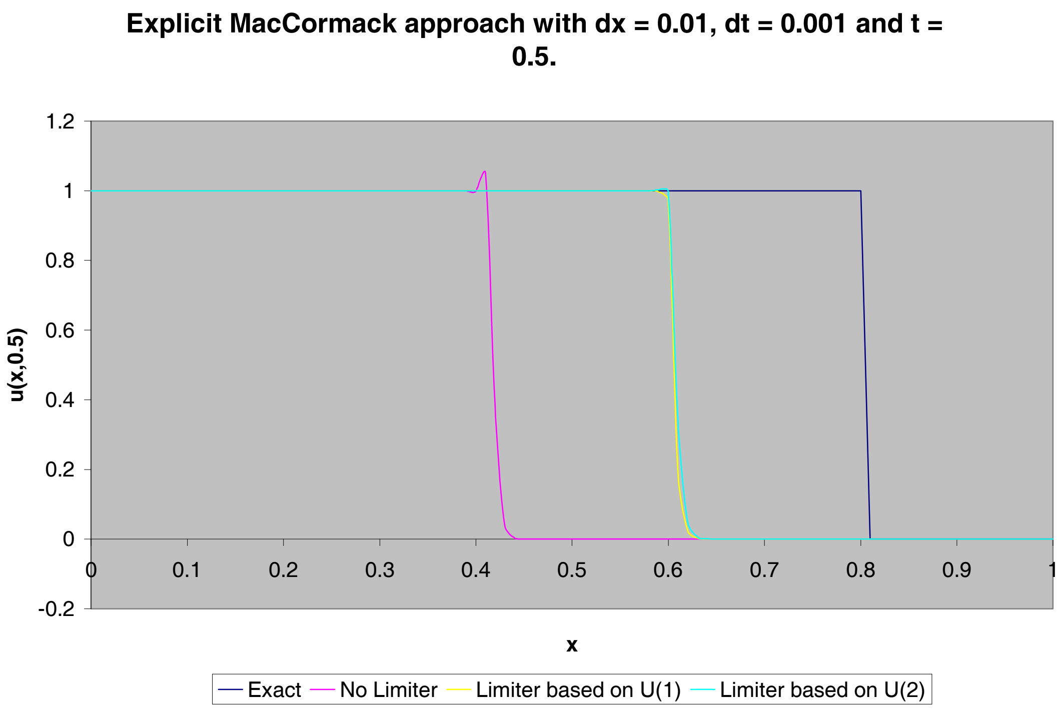


Figure 5-14: Explicit MacCormack approach with stiff source term.

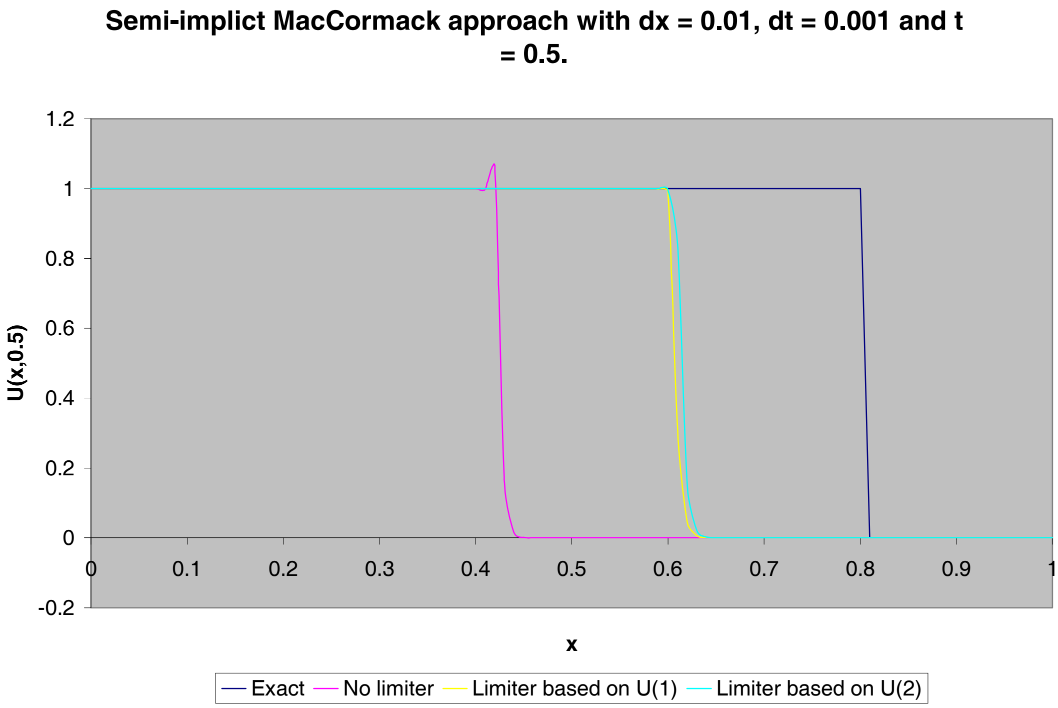


Figure 5-15: Semi-implicit MacCormack approach with stiff source term.

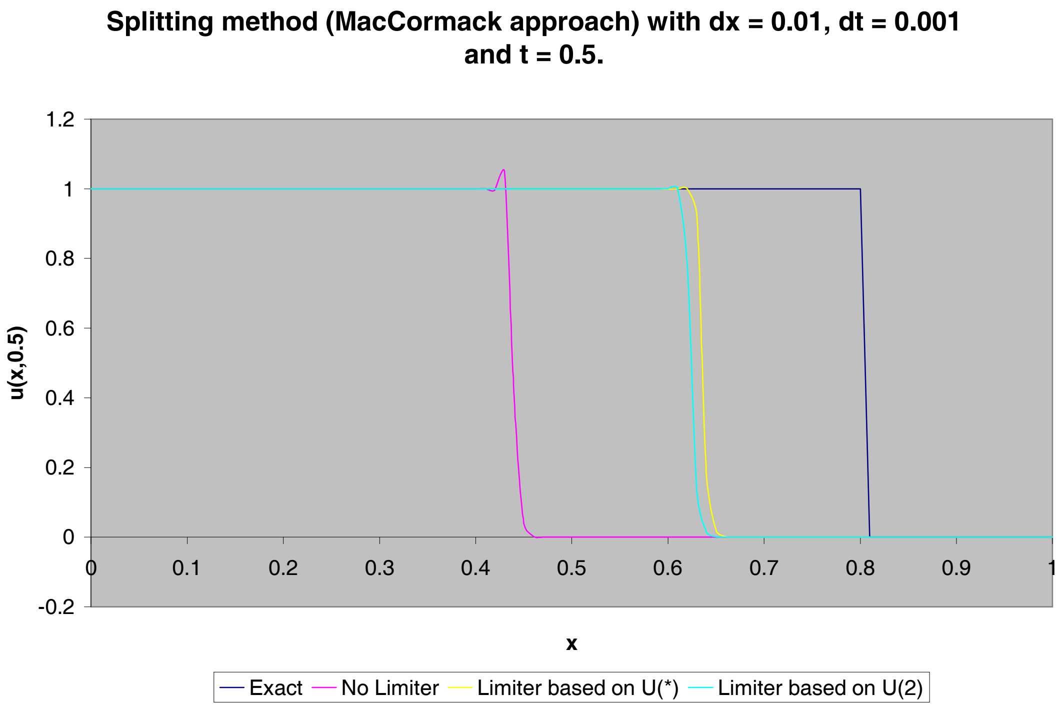


Figure 5-16: Splitting method (MacCormack approach) with stiff source term.

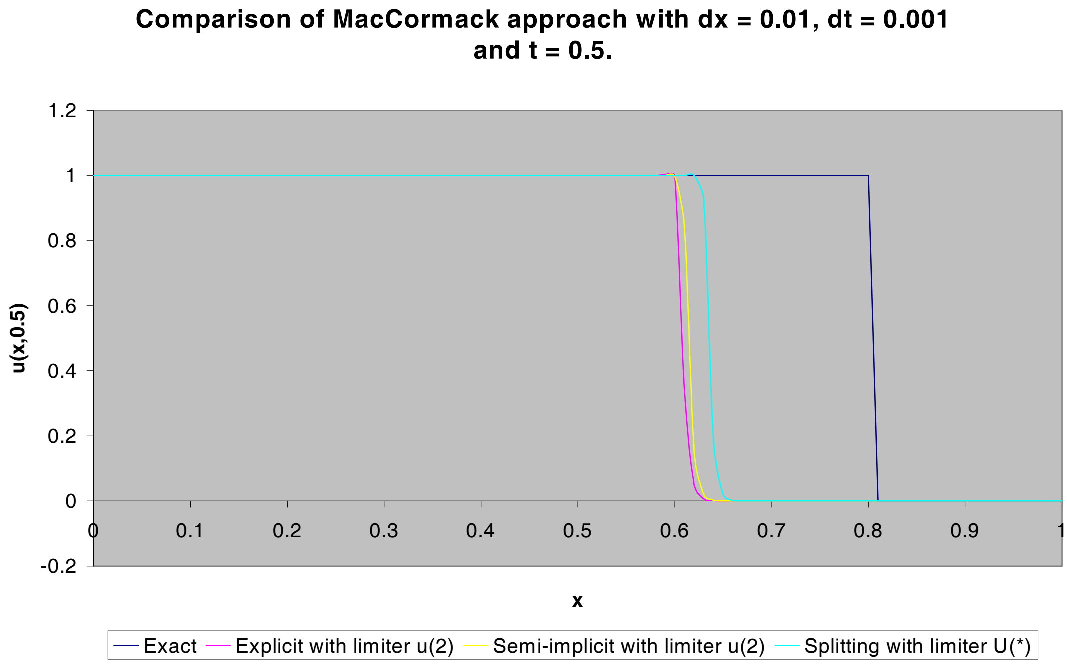


Figure 5-17: Comparison of MacCormack approach with stiff source term.

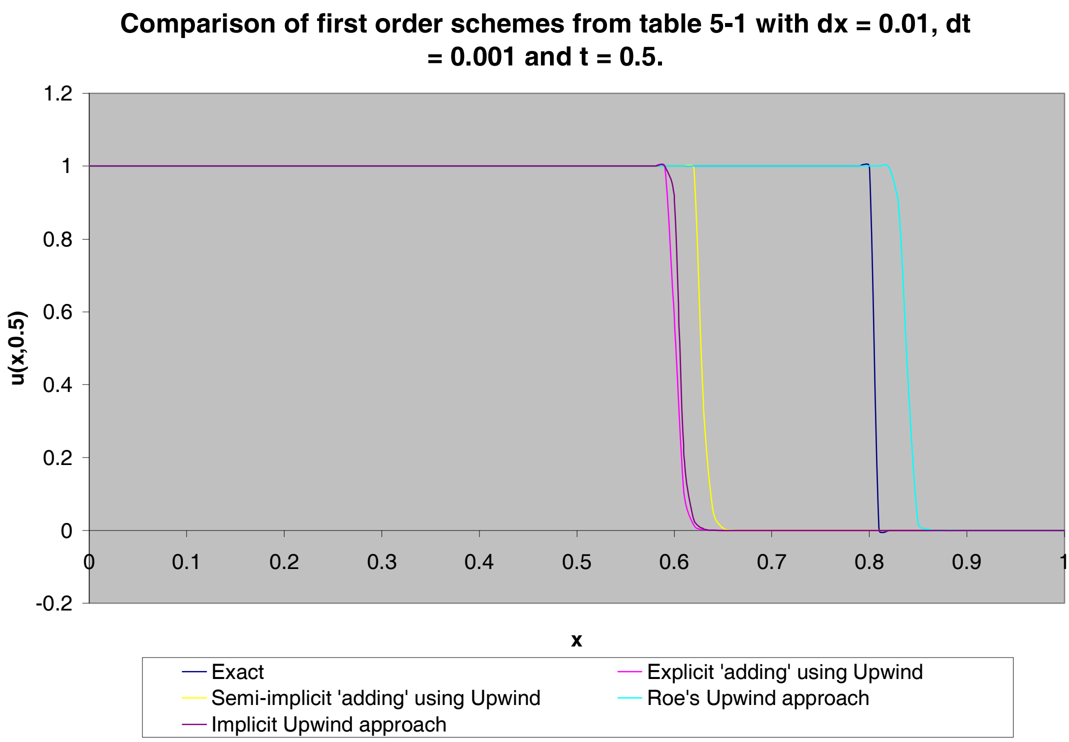


Figure 5-18: Comparison of first order schemes listed in Table 5-1 with stiff source term.

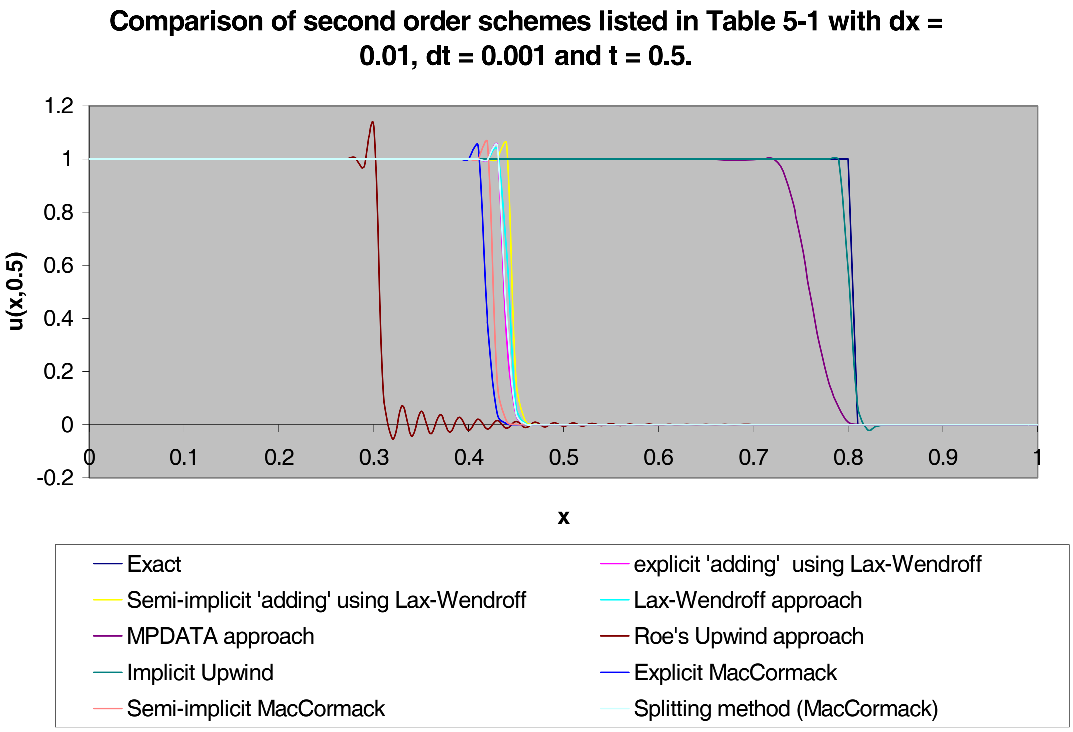


Figure 5-19: Comparison of second order schemes listed in Table 5-1 with stiff source term.

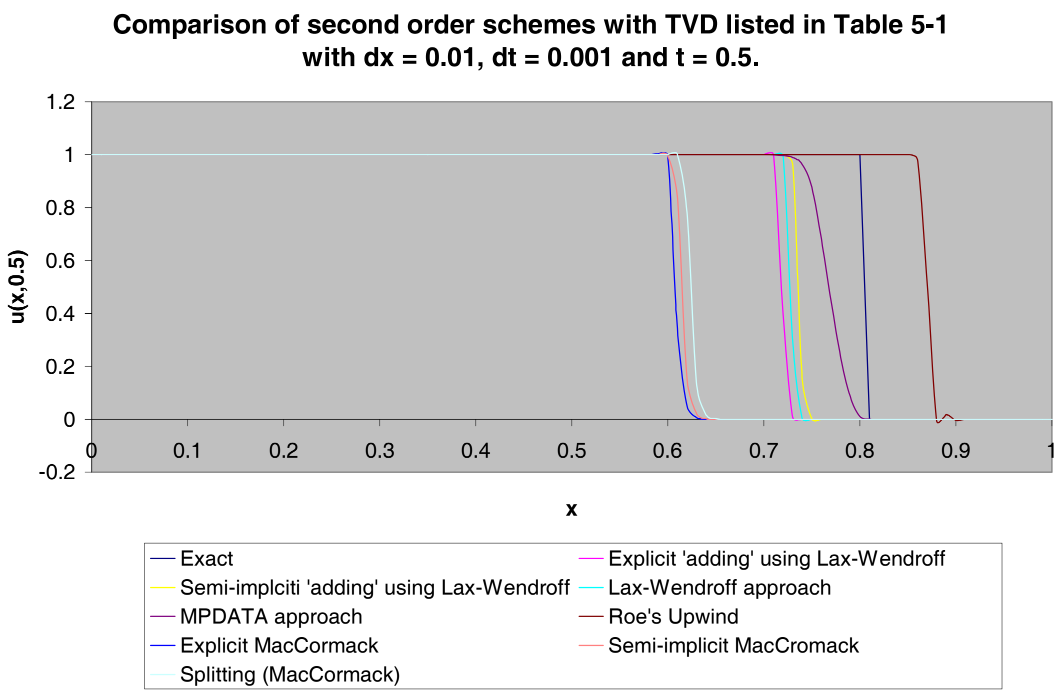


Figure 5-20: Comparison of second order schemes with TVD listed in Table 5-1 with stiff source term.

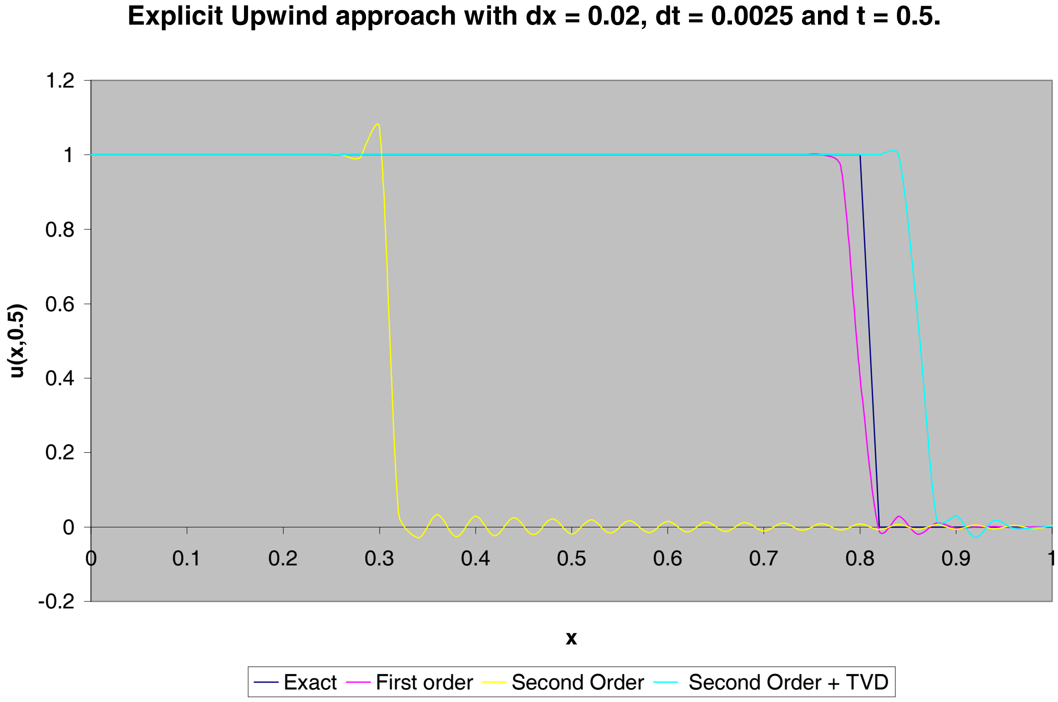


Figure 5-21: Explicit Upwind approach with stiff source term.

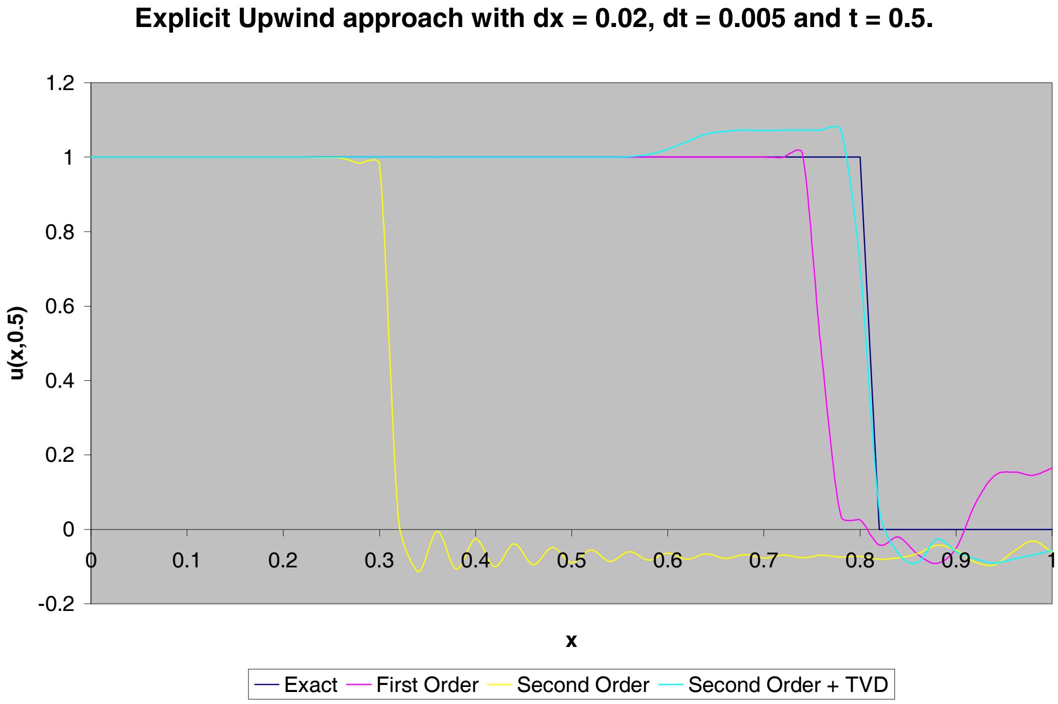


Figure 5-22: Explicit Upwind approach with stiff source term.

6 Conclusion

6.1 Final Comparison

Throughout this dissertation, we have discussed many techniques for numerically approximating the conservation law with and without source term, i.e.

$$\frac{\partial u}{\partial t}+\frac{\partial f(u)}{\partial x}=R(x,t,u) \tag{6.1}$$

and encountered many difficulties, especially when the source term is a function of u. We have also seen that numerically approximating source terms accurately can be extremely difficult to do. However, we have managed to overcome the majority of the difficulties encountered and we have obtained some very accurate finite difference schemes, even when the source term is stiff.

For example, in Chapter 5, we applied the different approaches to the advection-transport equation with a stiff source term, test problem (5.2), and compared the numerical results to obtain the most accurate first order approach, second order approach and second order approach with TVD. These three most accurate approaches are compared in Figure 6-1. Figure 6-1 shows us that the most accurate approach discussed in this project was the second order implicit Upwind approach. Roe's first order upwind approach moved the discontinuity too fast but this was due to a small Courant number. If we increased the step-size, Roe's first order Upwind approach would give us more accurate results but not as accurate as the second order implicit Upwind approach. Notice how the second order MPDATA approach with TVD gave more accurate results than Roe's first order Upwind but less accurate than the second order implicit Upwind approach.

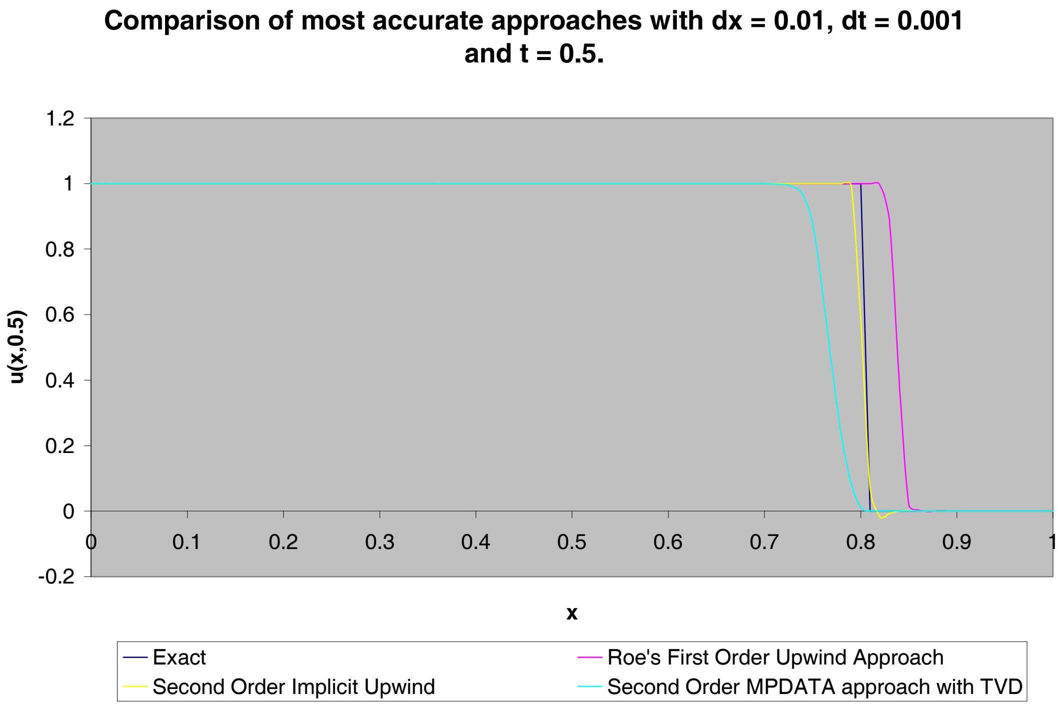


Figure 6-1: Comparison of most accurate approaches with stiff source term.

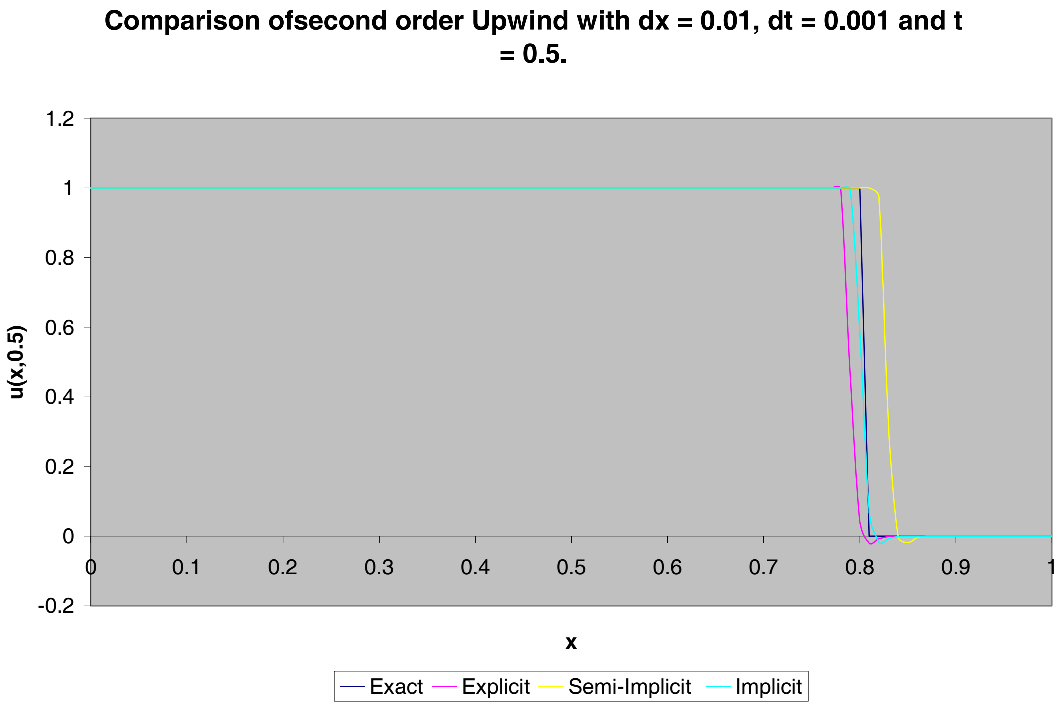


Figure 6-2: Comparison of explicit, semi-implicit and implicit second order Upwind with stiff source term.

This may be due to TVD causing the discontinuity to move faster, when the source term is stiff, or may be due to an implementation problem.

So far, we have seen that the second order implicit Upwind approach has produced the most accurate results. We have looked at a variety of techniques for numerically approximating the source term but we wish to know which technique produces the most accurate results. Figure 6-2 shows some numerical results using the second order Upwind approach applied to the test problem (5.2) but with:

1. The source term and the conservation law approximated explicitly (Explicit).
2. The source term approximated implicitly and the conservation law approximated explicitly (Semi-implicit).
3. The source term and the conservation law approximated implicitly (Implicit).

Here, we can see that the semi-implicit approach produced the least accurate results due to the method moving the discontinuity too fast and the explicit approach produced the second most accurate numerical results. This is unusual since we would expect the semi-implicit approach to be more accurate than the explicit approach. However, when we used the Lax-Wendroff approach, we saw that the semi-implicit approach was more accurate than the explicit due to the discontinuity being moved slightly faster for the semi-implicit approach, see Figure 5-8. Thus, the semi-implicit approach moves the discontinuity slightly faster which makes all approaches which move the discontinuity too slow, i.e. Lax-Wendroff with source term ‘added’, more accurate but all approaches which move the discontinuity at the correct speed or too fast, i.e. the second order Upwind approach, less accurate. The implicit approach produced the most accurate numerical results and also moved the discontinuity faster than the explicit approach, but only a little.

So we can see that numerically approximating the conservation law with source term can be very difficult to approximate accurately since the size of the Courant number greatly influences the accuracy of the numerical approximation especially when the source term is stiff. I.e. applying TVD can cause the discontinuity to move faster when the Courant number is too small. Also, we must be careful when choosing whether to use an implicit, semi-implicit or explicit approach since this also affects the speed of the discontinuity.

6.2 Further Work

In this dissertation, we have only considered a small amount of numerical techniques for numerically approximating the conservation law with source term. We could apply finite volume methods, finite element methods or a whole range of other techniques to numerically approximate the conservation law with source term. We have only looked at numerical results for the advection-transport equation and advection equation and not even considered the inviscid burger equation, etc. Also, we have only briefly looked at splitting methods and high resolution methods, i.e. flux-limiter methods. We have also only considered the courant number for $v = 0.1$ where $\Delta x = 0.01$, $\Delta t = 0.001$ and $c = 1$. Also, we have not considered a system form of the conservation law with source term, i.e. the Shallow Water Equation (1.2) and we have only considered the one-dimensional case. As we can see there is a considerable amount of further work to discuss.

References

1. R.J. LeVeque and H.C. Yee, A Study of Numerical Methods for Hyperbolic Conservation Laws with Stiff Source Terms, *J. Comput. Phys.* **86**, 187 – 210 (1990).

2. P. Embid, J. Goodman and A. Majda, Multiple Steady States for 1-D Transonic Flow, *SIAM J. Sci. Stat. Comput.* **March 1984**, 21 – 41.

3. P.K. Smolarkiewicz and L.G. Margolin, MPDATA: A Finite-Difference Solver for Geophysical Flows, *J. Comput. Phys.* **140**, 459 – 480 (1998).

4. A. Bermudez and M^Elena Vazquez, Upwind Methods for Hyperbolic Conservation Laws with Source Terms, *Computers and Fluids* **Vol. 23, No. 8**, 1049 – 1071(1994).

5. H.C. Yee, Upwind and Symmetric Shock-Capturing Schemes, *NASA Ames Research Center Technical Memoranda* **89464** (1987).

6. P.L. Roe, Upwind Differencing Schemes for Hyperbolic Conservation Laws with Source Terms, in Nonlinear Hyperbolic Problems, C. Carraso, P.-A. Raviart, and D.Serre, eds., *Springer-Verlag, Lecture Notes in Mathematics* **1270**, 41 – 51 (1986).

7. R.J. LeVeque, Numerical Methods for Conservation Laws, *Birkhauser-Verlag* (1990).

8. D. Kroner, Numerical Schemes for Conservation Laws, *Wiley-Teubner series* (1997).

9. G. Whitham, Linear and Nonlinear Waves, *Wiley-interscience* (1974).

10. B. van Leer, Towards the Ultimate Conservative Difference Scheme, V. A second-order sequel to Godunov's method, *J. Comput. Phys*. **14**, 361 – 376 (1979)

11. G. Strang, *SIAM J. Num. Anal.* **5**, 506 (1968).

12. P.K. Sweby, High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws, *SIAM J. Num. Anal.* **21**, 995 (1968).

13. P.K. Sweby, Finite Difference Schemes for Conservation Laws, Lecture Notes for Computer Techniques and Modelling, *University of Reading*. (Unpublished)

14. W.F. Ames, Numerical Methods for Partial Differential Equations, Third Edition, *Academic Press* (1992).

15. K.W. Morton and D.F. Mayers, Numerical Solution of Partial Differential Equations, *Cambridge University Press* (1994).

Appendix A

A Listing of all Numerical Schemes Discussed in Chapters 4 and 5.

All approaches numerically approximate conservation laws with a source term present, i.e.

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = R(x, t, u)$$

1. Explicit 'Adding' of Source Term: (First / Second Order)

$$u_i^{n+1} = u_i^{SCHEME} + \Delta t R_i^n$$

where u_i^{SCHEME} represents a numerical scheme which approximates the conservation law without a source term present and is of first / second order.

2. Semi-Implicit Adding of Source Term: (First/Second Order)

$$\left(1 - \Delta t \left[\frac{\partial R}{\partial u}\right]_i^n\right) u_i^{n+1} = u_i^{SCHEME} + \Delta t R_i^n - \Delta t\, u_i^n \left[\frac{\partial R}{\partial u}\right]_i^n$$

where u_i^{SCHEME} represents a numerical scheme which approximates the conservation law without a source term present and is of first / second order.

3. Lax-Wendroff Approach: (Second Order)

$$\left(1-\frac{\Delta t}{2}\left[\frac{dR}{du}\right]_i^n\right)u_i^{n+1}=u_i^n-s\left[F(u;i)-F(u;i-1)\right]+\Delta t\left[R_i^n-\frac{u_i^n}{2}\left[\frac{dR}{du}\right]_i^n\right]$$
$$-\frac{\Delta t}{4}\left[\nu_{i+1/2}\left(R_{i+1}^n-R_i^n\right)-\nu_{i-1/2}\left(R_i^n-R_{i-1}^n\right)\right]$$

where

$$F(u;i)=F_L(u;i)+F_H(u;i)\phi_i\,,$$

$$F_L(u;i)=\begin{cases} f_i^n & \text{if } \nu_{i+1/2}>0 \\ f_{i+1}^n & \text{if } \nu_{i+1/2}<0 \end{cases},$$

$$F_H(u;i)=\frac{1}{2}\begin{cases} (1-\nu_{i+1/2})\left(f_{i+1}^n-f_i^n\right) & \text{if } \nu_{i+1/2}>0 \\ -(1+\nu_{i+1/2})\left(f_{i+1}^n-f_i^n\right) & \text{if } \nu_{i+1/2}<0 \end{cases}$$

and ϕ_i denotes the flux-limiter which can be any of the flux-limiters in Table A-1.

4. MPDATA Approach: (Second Order)

$$\left(1-\frac{\Delta t}{2}\left[\frac{\partial R}{\partial u}\right]_i^n\right)u_i^{n+1}=MPDATA\left(u_i^n+\frac{\Delta t}{2}R_i^n,\,w_{i+1/2}^{n+1/2}\right)+\frac{\Delta t}{2}\left[R_i^n-u_i^n\left[\frac{\partial R}{\partial u}\right]_i^n\right]$$

where $MPDATA\left(u_i^n,C\right)$ corresponds to the basic MPDATA algorithm with flux-limiter:

$$u_i^{n+1}=u_i^{(1)}-\left[F\left(u_i^{(1)},u_{i+1}^{(1)},V_{i+1/2}^{(1)}\right)\phi_i-F\left(u_{i-1}^{(1)},u_i^{(1)},V_{i-1/2}^{(1)}\right)\phi_{i-1}\right]$$

where

$$V_{i+1/2}^{(1)}=\frac{\Delta x}{\Delta t}\left(\left|w_{i+1/2}^{n+1/2}\right|-\left[w_{i+1/2}^{n+1/2}\right]^2\right)\left[\frac{u_{i+1}^{(1)}-u_i^{(1)}}{u_{i+1}^{(1)}+u_i^{(1)}}\right]-w_{i+1/2}^{n+1/2}\left[w_{i+3/2}^{n+1/2}-w_{i-1/2}^{n+1/2}\right],$$

$$u_i^{(1)}=u_i^n-\left[F\left(u_i^n,u_{i+1}^n,w_{i+1/2}^{n+1/2}\right)-F\left(u_{i-1}^n,u_i^n,w_{i-1/2}^{n+1/2}\right)\right],$$

$$w_{i+1/2}^{n+1/2}=\frac{1}{2}\left(3w_{i+1/2}^n-w_{i+1/2}^{n-1}\right)$$

and ϕ_i denotes the flux-limiter which can be any of the flux-limiters in Table A-1.

5. Roe's Explicit Upwind I: (First Order)

$$u_i^{n+1} = u_i^n - \begin{cases} s\left(f_i^n - f_{i-1}^n\right) - \Delta t\left[(1-\alpha)R_i^n + \alpha R_{i-1}^n\right] & \text{if } \nu_{i+1/2} > 0 \\ s\left(f_{i+1}^n - f_i^n\right) - \Delta t\left[(1-\alpha)R_i^n + \alpha R_{i+1}^n\right] & \text{if } \nu_{i-1/2} < 0 \end{cases}.$$

where $0 \le \alpha \le 1$ and if $\alpha = ½$ then the scheme is second order accurate in space.

6. Explicit Upwind II: (First Order)

$$u_i^{n+1} = u_i^n - \begin{cases} s\left(f_i^n - f_{i-1}^n\right) - \Delta t R\left((1-\alpha)x_i + \alpha x_{i-1}, (1-\alpha)u_i^n + \alpha u_{i-1}^n\right) & \text{if } \nu_{i+1/2} > 0 \\ s\left(f_{i+1}^n - f_i^n\right) - \Delta t R\left((1-\alpha)x_i + \alpha x_{i+1}, (1-\alpha)u_i^n + \alpha u_{i+1}^n\right) & \text{if } \nu_{i-1/2} < 0 \end{cases}.$$

where $0 \le \alpha \le 1$ and if $\alpha = ½$ then the scheme is second order accurate in space.

7. Roe's Explicit Upwind III: (Second Order)

$$u_i^{n+1} = u_i^n - s\left[F(u;i) - F(u;i-1)\right] + \Delta t \begin{cases} (1-\alpha)R_i^n + \alpha R_{i-1}^n & \text{if } \nu_{i+1/2} > 0 \\ (1-\alpha)R_i^n + \alpha R_{i+1}^n & \text{if } \nu_{i-1/2} < 0 \end{cases}$$

where

$$F(u;i) = F_L(u;i) + F_H(u;i)\phi_i,$$

$$F_L(u;i) = \begin{cases} f_i^n & \text{if } \nu_{i+1/2} > 0 \\ f_{i+1}^n & \text{if } \nu_{i+1/2} < 0 \end{cases},$$

$$F_H(u;i) = \frac{1}{2}\begin{cases} (1-\nu_{i+1/2})\left(f_{i+1}^n - f_i^n\right) & \text{if } \nu_{i+1/2} > 0 \\ -(1+\nu_{i+1/2})\left(f_{i+1}^n - f_i^n\right) & \text{if } \nu_{i+1/2} < 0 \end{cases}$$

and $0 \le \alpha \le 1$. If $\alpha = ½$ then the scheme is second order accurate in space. Also, ϕ_i denotes the flux-limiter which can be any of the flux-limiters in Table A-1.

8. Explicit Upwind IV: (Second Order)

$$u_i^{n+1} = u_i^n - s\left[F(u;i) - F(u;i-1)\right] \\ + \Delta t \begin{cases} R\left((1-\alpha)x_i + \alpha x_{i-1}, (1-\alpha)u_i^n + \alpha u_{i-1}^n\right) & \text{if } v_{i+1/2} > 0 \\ R\left((1-\alpha)x_i + \alpha x_{i+1}, (1-\alpha)u_i^n + \alpha u_{i+1}^n\right) & \text{if } v_{i-1/2} < 0 \end{cases}$$

where

$$F(u;i) = F_L(u;i) + F_H(u;i)\phi_i,$$

$$F_L(u;i) = \begin{cases} f_i^n & \text{if } v_{i+1/2} > 0 \\ f_{i+1}^n & \text{if } v_{i+1/2} < 0 \end{cases},$$

$$F_H(u;i) = \frac{1}{2}\begin{cases} (1 - v_{i+1/2})\left(f_{i+1}^n - f_i^n\right) & \text{if } v_{i+1/2} > 0 \\ -(1 + v_{i+1/2})\left(f_{i+1}^n - f_i^n\right) & \text{if } v_{i+1/2} < 0 \end{cases},$$

and $0 \le \alpha \le 1$. If $\alpha = ½$ then the scheme is second order accurate in space. Also, ϕ_i denotes the flux-limiter which can be any of the flux-limiters in Table A-1.

9. Implicit Upwind I: (First Order)

$$\begin{bmatrix} b_0 & d_0 & 0 & 0 & 0 & \cdots & 0 \\ a_1 & b_1 & d_1 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & b_2 & d_2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{I-2} & b_{I-2} & d_{I-2} & 0 \\ 0 & \cdots & 0 & 0 & a_{I-1} & b_{I-1} & d_{I-1} \\ 0 & \cdots & 0 & 0 & 0 & a_I & b_I \end{bmatrix} \begin{bmatrix} u_0^{n+1} - u_0^n \\ u_1^{n+1} - u_1^n \\ u_2^{n+1} - u_2^n \\ \vdots \\ u_{I-2}^{n+1} - u_{I-2}^n \\ u_{I-1}^{n+1} - u_{I-1}^n \\ u_I^{n+1} - u_I^n \end{bmatrix} = \begin{bmatrix} G_0 - a_0\left(u_{-1}^{n+1} - u_{-1}^n\right) \\ G_1 \\ G_2 \\ \vdots \\ G_{I-2} \\ G_{I-1} \\ G_I - d_I\left(u_{I+1}^n - u_{I+1}^n\right) \end{bmatrix}$$

where

$$G_i = \Delta t e_i g_i^n - s\begin{cases} \left(f_i^n - f_{i-1}^n\right) & \text{if } v_{i+1/2} > 0 \\ \left(f_{i+1}^n - f_i^n\right) & \text{if } v_{i+1/2} < 0 \end{cases},$$

$$a_i = \begin{cases} -s\left[\dfrac{\partial f}{\partial u}\right]_{i-1}^n & \text{if } v_{i+1/2} > 0 \\ 0 & \text{if } v_{i+1/2} < 0 \end{cases}, \quad d_i = \begin{cases} 0 & \text{if } v_{i+1/2} > 0 \\ s\left[\dfrac{\partial f}{\partial u}\right]_{i+1}^n & \text{if } v_{i+1/2} < 0 \end{cases}$$

$$b_i = 1 + \mathrm{sgn}(v_{i+1/2}) s \left[\frac{\partial f}{\partial u}\right]_i^n - \Delta t e_i \left[\frac{\partial g}{\partial u}\right]_i^n$$

and $R(x,t,u) = e(x) g(u)$.

10. Implicit Upwind II: (Second Order)

$$\begin{bmatrix} b_0 & d_0 & 0 & 0 & 0 & \dots & 0 \\ a_1 & b_1 & d_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & b_2 & d_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{I-2} & b_{I-2} & d_{I-2} & 0 \\ 0 & \dots & 0 & 0 & a_{I-1} & b_{I-1} & d_{I-1} \\ 0 & \dots & 0 & 0 & 0 & a_I & b_I \end{bmatrix} \begin{bmatrix} u_0^{n+1} - u_0^n \\ u_1^{n+1} - u_1^n \\ u_2^{n+1} - u_2^n \\ \vdots \\ u_{I-2}^{n+1} - u_{I-2}^n \\ u_{I-1}^{n+1} - u_{I-1}^n \\ u_I^{n+1} - u_I^n \end{bmatrix} = \begin{bmatrix} G_0 - a_0\left(u_{-1}^{n+1} - u_{-1}^n\right) \\ G_1 \\ G_2 \\ \vdots \\ G_{I-2} \\ G_{I-1} \\ G_I - d_I\left(u_{I+1}^n - u_{I+1}^n\right) \end{bmatrix}$$

where

$$G_i = \Delta t e_i g_i^n - \left[F(u;i) - F(u;i-1)\right],$$

$$F(u;i) = F_L(u;i) + F_H(u;i)\phi_i,$$

$$F_L(u;i) = \begin{cases} f_i^n & \text{if } v_{i+1/2} > 0 \\ f_{i+1}^n & \text{if } v_{i+1/2} < 0 \end{cases},$$

$$F_H(u;i) = \frac{1}{2}\begin{cases} (1 - v_{i-1/2})\left(f_i^n - f_{i-1}^n\right) & \text{if } v_{i+1/2} > 0 \\ -(1 + v_{i+3/2})\left(f_{i+2}^n - f_{i+1}^n\right) & \text{if } v_{i+1/2} < 0 \end{cases},$$

$$a_i = \begin{cases} -s\left[\frac{\partial f}{\partial u}\right]_{i-1}^n & \text{if } v_{i+1/2} > 0 \\ 0 & \text{if } v_{i+1/2} < 0 \end{cases}, \quad d_i = \begin{cases} 0 & \text{if } v_{i+1/2} > 0 \\ s\left[\frac{\partial f}{\partial u}\right]_{i+1}^n & \text{if } v_{i+1/2} < 0 \end{cases},$$

$$b_i = 1 + \mathrm{sgn}(v_{i+1/2}) s \left[\frac{\partial f}{\partial u}\right]_i^n - \Delta t e_i \left[\frac{\partial g}{\partial u}\right]_i^n,$$

and $R(x,t,u) = e(x) g(u)$. Also, ϕ_i denotes the flux-limiter which can be any of the flux-limiters in Table A-1.

11. Explicit MacCormack: (Second Order)

$$u_i^{n+1} = u_i^{(2)} + \left[\phi_{i+1/2}^{(2)} - \phi_{i-1/2}^{(2)}\right]$$

where

$$u_i^{(2)} = \frac{1}{2}\left(u_i^n + u_i^{(1)}\right) - \frac{s}{2}\left[f_i^{(1)} - f_{i-1}^{(1)}\right] + \Delta t \frac{R_i^{(1)}}{2} \ ,$$

$$u_i^{(1)} = u_i^n - s\left(f_{i+1}^n - f_i^n\right) + \Delta t R_i^n \ ,$$

$$\phi_{i+1/2}^{(2)} = \frac{1}{2}\left[\left|v_{i+1/2}\right| - v_{i+1/2}^2\right]\left(u_{i+1}^{(2)} - u_i^{(2)} - Q_{i+1/2}\right)$$

and, $Q_{i+1/2}$ can be any of the values in Table A-2.

12. Semi-Implicit MacCormack: (Second Order)

$$u_i^{n+1} = u_i^{(2)} + \left[\phi_{i+1/2}^{(2)} - \phi_{i-1/2}^{(2)}\right]$$

where

$$\left[1 - \Delta t \bar{\theta}\left[\frac{\partial R}{\partial u}\right]_i^n\right]\left(u_i^{(2)} - u_i^{(1)}\right) = -s\left(f_{i+1}^{(1)} - f_i^{(1)}\right) + \Delta t R_i^{(1)},$$

$$\left[1 - \Delta t \bar{\theta}\left[\frac{\partial R}{\partial u}\right]_i^n\right]\left(u_i^{(1)} - u_i^n\right) = -s\left(f_{i+1}^n - f_i^n\right) + \Delta t R_i^n$$

$$\phi_{i+1/2}^{(2)} = \frac{1}{2}\left[\left|v_{i+1/2}\right| - v_{i+1/2}^2\right]\left(u_{i+1}^{(2)} - u_i^{(2)} - Q_{i+1/2}\right).$$

and $0 \le \bar{\theta} \le 1$. Also $Q_{i+1/2}$ can be any of the values in Table A-2.

13. Splitting Method (MacCormack): (Second Order)

$S_\psi\left(\frac{\Delta t}{2}\right)$:

$$\left[1 - \frac{\Delta t}{4}\left[\frac{\partial R}{\partial u}\right]_i^n\right]\left(u_i^* - u_i^n\right) = \frac{\Delta t}{2} R_i^n$$

$$u_i^* = u_i^n + \left(u_i^* - u_i^n\right).$$

$S_k(\Delta t)$:

$$\left(u_i^{(1)} - u_i^*\right) = -s\left(f_i^* - f_{i-1}^*\right)$$

$$u_i^{(1)} = u_i^* + \left(u_i^{(1)} - u_i^*\right)$$

$$\left(u_i^{(2)} - u_i^{(1)}\right) = -s\left(f_{i+1}^{(1)} - f_i^{(1)}\right)$$

$$u_i^{(2)} = u_i^{*} + \frac{1}{2}\left[\left(u_i^{(2)} - u_i^{(1)}\right) + \left(u_i^{(1)} - u_i^{*}\right)\right]$$

$$u_i^{**} = u_i^{(2)} + \left[\phi_{i+1/2}^{*} - \phi_{i-1/2}^{*}\right].$$

$S_\psi\left(\frac{\Delta t}{2}\right)$:
$$\left[1 - \frac{\Delta t}{4}\left[\frac{\partial R}{\partial u}\right]_i^{**}\right]\left(u_i^{**} - u_i^{(2)}\right) = \frac{\Delta t}{2} R_i^{**}$$

$$u_i^{n+1} = u_i^{**} + \left(u_i^{**} - u_i^{(2)}\right).$$

Here,

$$\phi_{i+1/2}^{(*)} = \frac{1}{2}\left[\left|\nu_{i+1/2}\right| - \nu_{i+1/2}^{2}\right]\left(u_{i+1}^{(*)} - u_i^{(*)} - Q_{i+1/2}\right)$$

and $Q_{i+1/2}$ can be any of the values in Table A-2.

Name of Flux-limiter	$\phi(\theta)$
Minmod	$\phi(\theta) = \max(0, \min(1, \theta))$
Roe's Superbee	$\phi(\theta) = \max(0, \min(2\theta, 1), \min(\theta, 2))$
van Leer	$\phi(\theta) = \frac{\lvert\theta\rvert + \theta}{1 + \lvert\theta\rvert}$
van Albada	$\phi(\theta) = \frac{\theta^2 + \theta}{1 + \theta^2}$

Table A-1: Some second order flux-limiters.

Some choices of $Q_{i+1/2}$ where $\Delta_{i+1/2} = u_{i+1}^n - u_i^n$.
$Q_{i+1/2} = \min\text{mod}\left(\Delta_{i+1/2}, \Delta_{i-1/2}\right) + \min\text{mod}\left(\Delta_{i+1/2}, \Delta_{i+3/2}\right) - \Delta_{i+1/2}$
$Q_{i+1/2} = \min\text{mod}\left(\Delta_{i-1/2}, \Delta_{i+1/2}, \Delta_{i+3/2}\right)$
$Q_{i+1/2} = \min\text{mod}\left(2\Delta_{i-1/2}, 2\Delta_{i+1/2}, 2\Delta_{i+3/2}, \frac{1}{2}\left(\Delta_{i-1/2} + \Delta_{i+3/2}\right)\right)$

Table A-2: Some choices of $Q_{i+1/2}$ for the MacCormack approach.

Acknowledgements

I would like to thank my supervisors Dr. P.K. Sweby and Professor M.J. Baines for their guidance and support throughout this dissertation. In addition, I would like to thank the Engineering and Physical Science Research Council for their funding and Dr M.E. Hubbard for his input.