Version 1.2

Numerical Techniques for the Shallow Water Equations

Justin Hudson

The University of Reading, Department of Mathematics, P.O.Box 220,

Whiteknights, Reading, Berkshire, RG6 6AX, UK

E-mail: yha88@dial.pipex.com

Numerical Analysis Report 2/99

Abstract

In this report we will discuss some numerical techniques for approximating the Shallow Water equations. In particular we will discuss finite difference schemes, adaptations of Roe's approximate Riemann solver and the Q-Schemes of Bermudez & Vazquez with the objective of accurately approximating the solution of the Shallow Water equations. We consider four different test problems for the Shallow Water equations with each test problem making the source term more significant, i.e. the variation of the Riverbed becomes more pronounced, so that the different approaches discussed in this report can be rigorously tested. A comparison of the different approaches discussed in this report will also be made so that we may determine which approach produced the most accurate numerical results overall.

The work contained in this report has been carried out as part of the Oxford / Reading Institute for Computational Fluid Dynamics and was funded by the Engineering and Physical Science Research Council and HR Wallingford under a CASE award.

1 Introduction

Throughout this report, we will be discussing some numerical techniques for approximating the Saint-Venant equations, i.e.

$$\frac{\partial \mathbf{w}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{w})}{\partial x} = \mathbf{R}(x, \mathbf{w}) \tag{1.1}$$

where

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

Here, $h(x,t)$ and $u(x,t)$ represent the water depth and the fluid velocity respectively and $H(x)$ is the bed depth from a fixed reference level, see Figure 1-1.

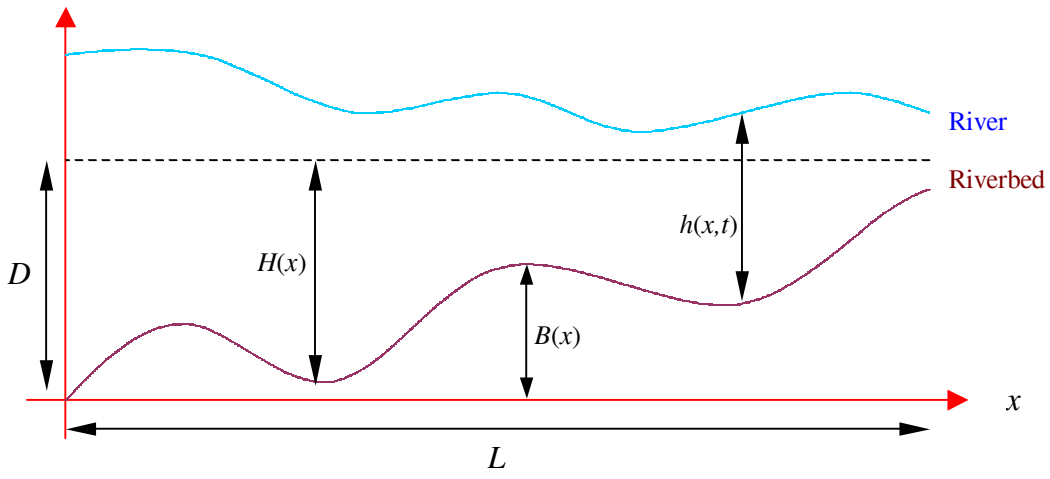

Figure 1-1: Shallow Water Variables.

Notice that (1.1) has a source term present, $\mathbf{R}(x,\mathbf{w})$, which can cause difficulties in accurately approximating (1.1), especially if the source term is stiff (see Hudson[7] and LeVeque & Yee[9]). Sometimes, it is convenient to re-write the source term in terms of the bed height

$$H(x) = D - B(x) \quad \Rightarrow \quad H'(x) = -B'(x),$$

hence,

$$\mathbf{R}(x,\mathbf{w}) = \begin{bmatrix} 0 \\ ghH'(x) \end{bmatrix} = \begin{bmatrix} 0 \\ -ghB'(x) \end{bmatrix}.$$

In Chapter 2, we will discuss four test problems for the Shallow Water Equations (1.1) where the source term becomes more significant for each test problem, i.e. the variation of the riverbed becomes more pronounced. The first test problem will have no source term present and the second and third test problem will have a source term present, with the third test problem having a source term more significant than the second test problem. The final test problem will also have a source term which is the most significant out of all of the four test problems and thus is the most difficult to approximate accurately.

The four test problems will allow us to rigorously test the different numerical approaches for approximating the Shallow Water Equations (1.1), which are discussed in Chapter 3, and determine if the approaches are accurate when the source term becomes difficult to approximate accurately. The numerical approaches we will discuss in Chapter 3 are the finite difference approach, adaptations of Roe's approximate Riemann solver and the Q-Schemes of Bermudez & Vazquez[1]. Some of these approaches require the Jacobian matrix of $\mathbf{F}(\mathbf{w})$

$$\mathbf{A}(\mathbf{w}) = \frac{\partial \mathbf{F}}{\partial \mathbf{w}} = \begin{bmatrix} 0 & 1 \\ gh - u^2 & 2u \end{bmatrix},$$

which has eigenvalues

$$\lambda_1 = u + \sqrt{gh} \quad \text{and} \quad \lambda_2 = u - \sqrt{gh}$$

and eigenvectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ u + \sqrt{gh} \end{bmatrix} \quad \text{and} \quad \mathbf{e}_2 = \begin{bmatrix} 1 \\ u - \sqrt{gh} \end{bmatrix}.$$

All of the numerical approaches discussed will derive numerical schemes that are either first order, second order or flux-limited second order schemes (see LeVeque[10], Kroner[8] and Sweby[14]). In Chapter 4 the different numerical approaches discussed in Chapter 3 will be compared by using the four test problems so that we may determine which approach produced the most accurate numerical results overall.

2 Test Problems

In this chapter, we will discuss four test problems, for the Shallow Water Equations (1.1). The first test problem is the dam-break problem and was discussed by Glaister[3] and Stoker[13]. This problem is one of the most basic problems for (1.1) since no source term is present. The second test problem contains a source term and represents a dam breaking on a variable depth riverbed. The third problem was discussed by LeVeque[11] and represents a square pulse, which breaks up into two waves travelling in opposite directions, on a variable depth riverbed. The final test problem was discussed by Bermudez & Vazquez[1] and represents a tidal wave propagating on a variable depth riverbed. For each test problem, notice how the source term becomes more significant, i.e. the variation of the bed depth becomes more pronounced.

2.1 Problem A - The Dam-Break Problem

In this test problem, (1.1) has no source term present, i.e.

$$\frac{\partial \mathbf{w}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{w})}{\partial x} = 0 . \tag{2.1}$$

This is due to the riverbed being of constant depth resulting in $H'(x) = 0 \ \forall x$. We also have initial conditions

$$u(x,0) = 0 \quad \text{and} \quad h(x,0) = \begin{cases} 1 & \text{if} \quad 0 \le x \le \frac{1}{2} \\ \phi_0 & \text{if} \quad \frac{1}{2} < x \le 1 \end{cases},$$

which are illustrated in Figure 2-1. In Figure 2-1, the discontinuity at x = 0.5 represents a barrier, which separates the two initial river heights and is removed at t = 0. Walls are present at $x = 0$ and at $x = 1$ resulting in reflection at both boundaries.

Note that, for this test problem, if $\frac{1}{\phi_0} > 7.2$ then both eigenvalues of $\mathbf{A}(\mathbf{w})$ are of the same sign and the downstream flow is supercritical. If $\frac{1}{\phi_0} < 7.2$ then the eigenvalues of $\mathbf{A}(\mathbf{w})$ are of opposite sign and the downstream flow is subcritical. If the downstream flow is supercritical then difficulties can arise in accurately numerically approximating the wave speed of the discontinuity at x = 0. In Figure 2-1, $\frac{1}{\phi_0} = \frac{1}{0.5} = 2 < 7.2$, hence the downstream flow is subcritical.

We may obtain an exact solution of the dam-break problem by using the analysis of Stoker[13]. The following exact solution is illustrated in Figure 2-2 and Figure 2-3 and is only valid for $\phi_0 = 0.5$ and $\phi_1 = 1$,

$$u(x,t) = \begin{cases} 0 & \text{if} \quad x < \frac{1}{2} - t\sqrt{g} \\ \frac{1}{3t}\left(2x - 1 + 2t\sqrt{g}\right) & \text{if} \quad \frac{1}{2} - t\sqrt{g} \le x \le (u_2 - c_2)t + \frac{1}{2} \\ u_2 & \text{if} \quad (u_2 - c_2)t + \frac{1}{2} < x \le \mathrm{S}t + \frac{1}{2} \\ 0 & \text{if} \quad x > \mathrm{S}t + \frac{1}{2} \end{cases}$$

and

$$h(x,t) = \begin{cases} 1 & \text{if} \quad x < \frac{1}{2} - t\sqrt{g} \\ \frac{1}{9g}\left(2\sqrt{g} - \frac{2x-1}{2t}\right)^2 & \text{if} \quad \frac{1}{2} - t\sqrt{g} \le x \le (u_2 - c_2)t + \frac{1}{2} \\ \frac{1}{4}\left(\sqrt{1 + \frac{16S^2}{g}} - 1\right) & \text{if} \quad (u_2 - c_2)t + \frac{1}{2} < x \le \mathrm{S}t + \frac{1}{2} \\ \frac{1}{2} & \text{if} \quad x > \mathrm{S}t + \frac{1}{2} \end{cases}$$

where

$$u_2 = S - \frac{g}{8S}\left(1 + \sqrt{1 + \frac{16S^2}{g}}\right),$$

$$c_2 = \sqrt{\frac{g}{4}\left(\sqrt{1 + \frac{16S^2}{g}} - 1\right)}$$

and the wave speed of the discontinuity created at $x = 0$ is

$$S = 2.957918120187525.$$

For a more in depth analysis on how the value of S was obtained see Glaister[3] and Stoker[13].

2.2 Problem B - The Dam-Break Problem on a Variable Depth Riverbed

This test problem is similar to Problem A but the riverbed is no longer of constant depth, resulting in $H'(x) \neq 0$ for some values of x, which means that a source term is present, i.e.

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

For this test problem, the riverbed is defined as

$$B(x) = \begin{cases} \frac{1}{8}\left(\cos\left(10\pi\left(x - \frac{1}{2}\right)\right) + 1\right) & \text{if} \quad \frac{2}{5} \le x \le \frac{3}{5} \\ 0 & \text{otherwise} \end{cases}$$

and we have initial conditions

$$u(x,0) = 0 \quad \text{and} \quad h(x,0) = \begin{cases} 1 - B(x) & \text{if} \quad 0 \le x \le \frac{1}{2} \\ \phi_0 - B(x) & \text{if} \quad \frac{1}{2} < x \le 1 \end{cases},$$

which is illustrated in Figure 2-4. In Figure 2-4, the discontinuity at $x = 0.5$ represents a barrier, which separates the two initial river heights and is removed at $t = 0$. Walls are again present at $x = 0$ and at $x = 1$ giving reflection at these boundaries. Since a source term is now present, difficulties can arise when numerically approximating this problem, as we will see later.

2.3 Problem C - A Problem Discussed by LeVeque[11]

In this test problem, the riverbed is again of variable depth, resulting in $H'(x) \neq 0$ for some values of x, which means that a source term is present, i.e.

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

For this test problem, the riverbed is defined as

$$B(x) = \begin{cases} \frac{1}{4}\left(\cos\left(10\pi\left(x - \frac{1}{2}\right)\right) + 1\right) & \text{if} \quad \frac{2}{5} \le x \le \frac{3}{5} \\ 0 & \text{otherwise} \end{cases}$$

and we have initial conditions

$$u(x,0) = 0 \text{ and } h(x,0) = \begin{cases} 1 - B(x) & \text{if} \quad x < 0.1 \\ 1.2 - B(x) & \text{if} \quad 0.1 \le x \le 0.2 \\ 1 - B(x) & \text{if} \quad x > 0.2 \end{cases},$$

which is illustrated in Figure 2-5. This problem represents an initial pulse, which breaks up into two waves moving in opposite directions. The right-going square-wave pulse passes the hump in the riverbed and becomes partially reflected, causing a disturbance behind the hump. In this problem, there are no walls present at $x = 0$ and $x = 1$ and reflection does not occur.

2.4 Problem D - Tidal Wave Propagation on a Variable Depth Bed

This test problem was discussed by Bermudez and Vazquez[1] and also has a source term present, i.e.

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

We have a variable depth given by

$$H(x) = 50.5 - \frac{40x}{L} + 10\sin\left(\pi\left(\frac{4x}{L} + \frac{1}{2}\right)\right),$$

with initial conditions

$$u(x,0) = 0 \quad \text{and} \quad h(x,0) = H(x),$$

which are illustrated in Figure 2-6, and boundary conditions

$$h(0,t) = \begin{cases} 64.5 + 4\sin\left(\pi\left(\dfrac{4t}{86400} - \dfrac{1}{2}\right)\right) & \text{if} \quad t \le 43200 \\ 60.5 & \text{if} \quad t > 43200 \end{cases},$$

representing an incoming wave which is illustrated in Figure 2-7, and $u(L,t) = 0$. For this test problem, the reference depth D is taken as $D = 60.5$.

This problem represents a tidal wave propagating on a variable depth riverbed. Here, $h(0,t)$ represents a tidal wave of $4m$ amplitude and in Figure 2-7, we can see that at t = 21,600s, the tidal wave has reached it's full height of $8m$ at inflow whereas at t = 43,200s, the tidal wave has disappeared. Also, since the waves propagate at speed $\sqrt{gh}$, the tidal wave should only reach as far as 216,000m at t = 10,800s, when L = 648,000m.

Throughout this chapter, we have discussed four different test problems for the Shallow Water Equations. In the next chapter, we will discuss some numerical

techniques for approximating the Shallow Water Equations so that we can apply the different numerical techniques to the four test problems.

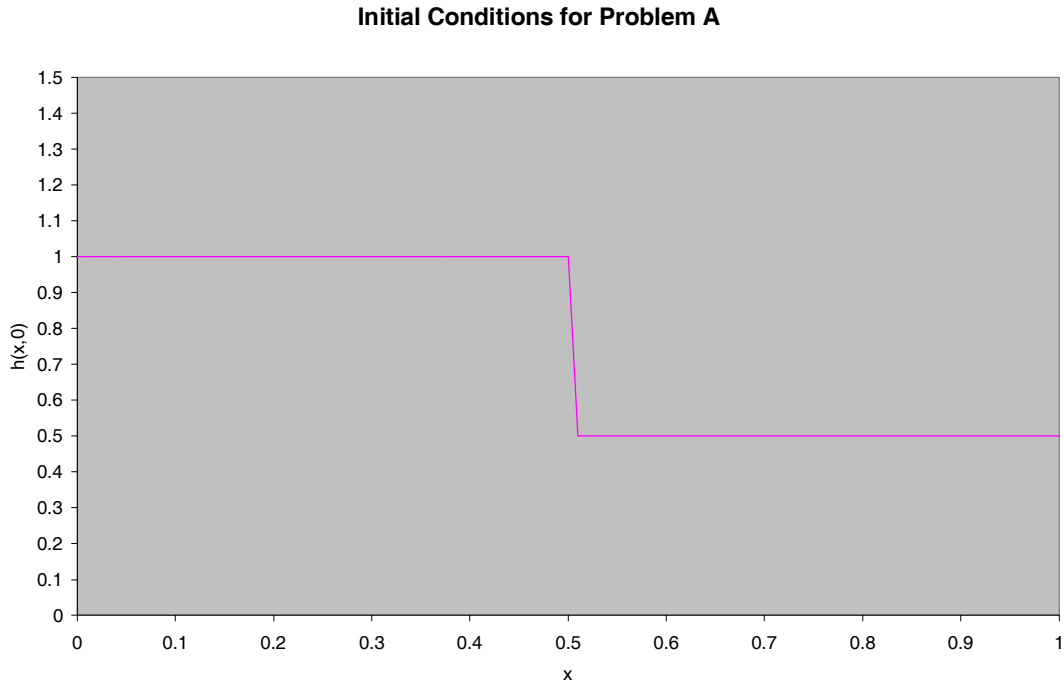

Figure 2-1: Initial condition $h(x,0)$ for Problem A

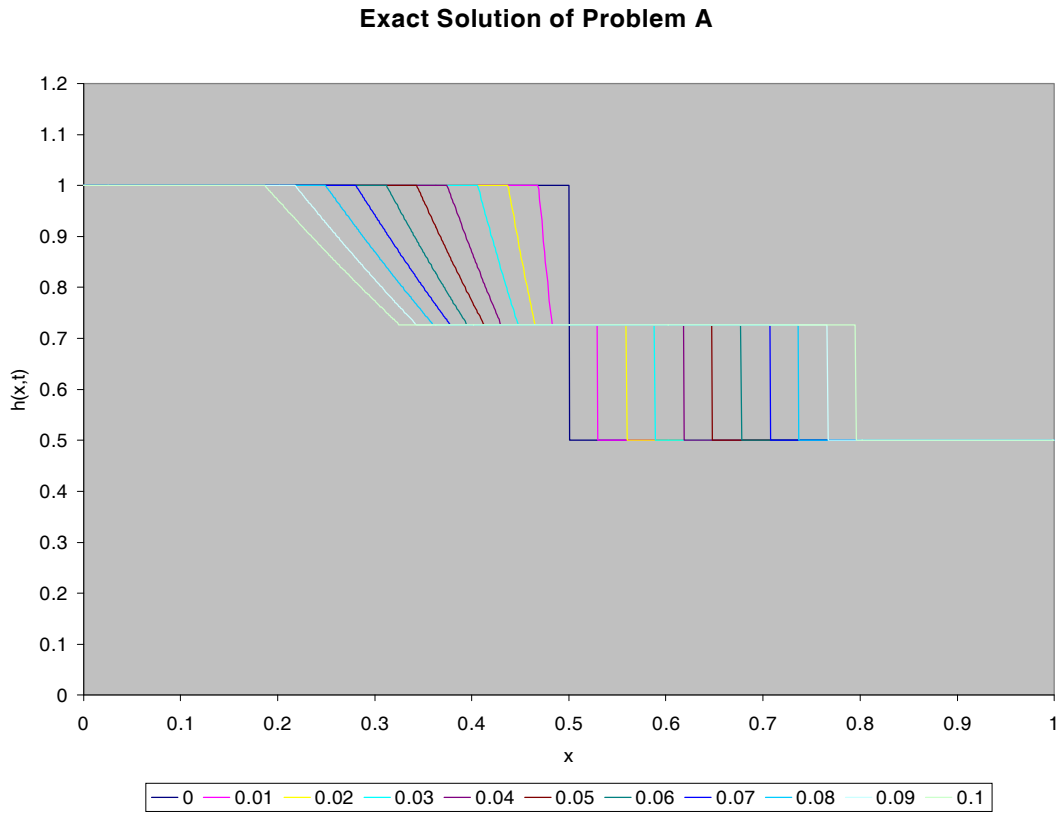

Figure 2-2: Illustration of the exact solution of Problem A

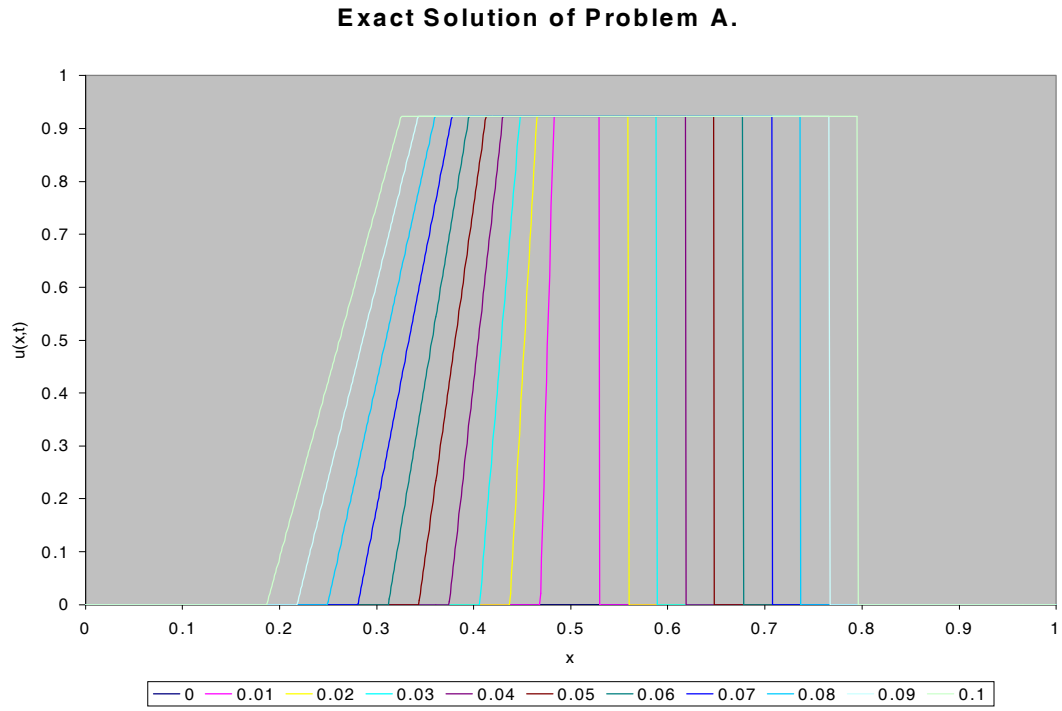

Figure 2-3: Illustration of the exact solution of Problem A

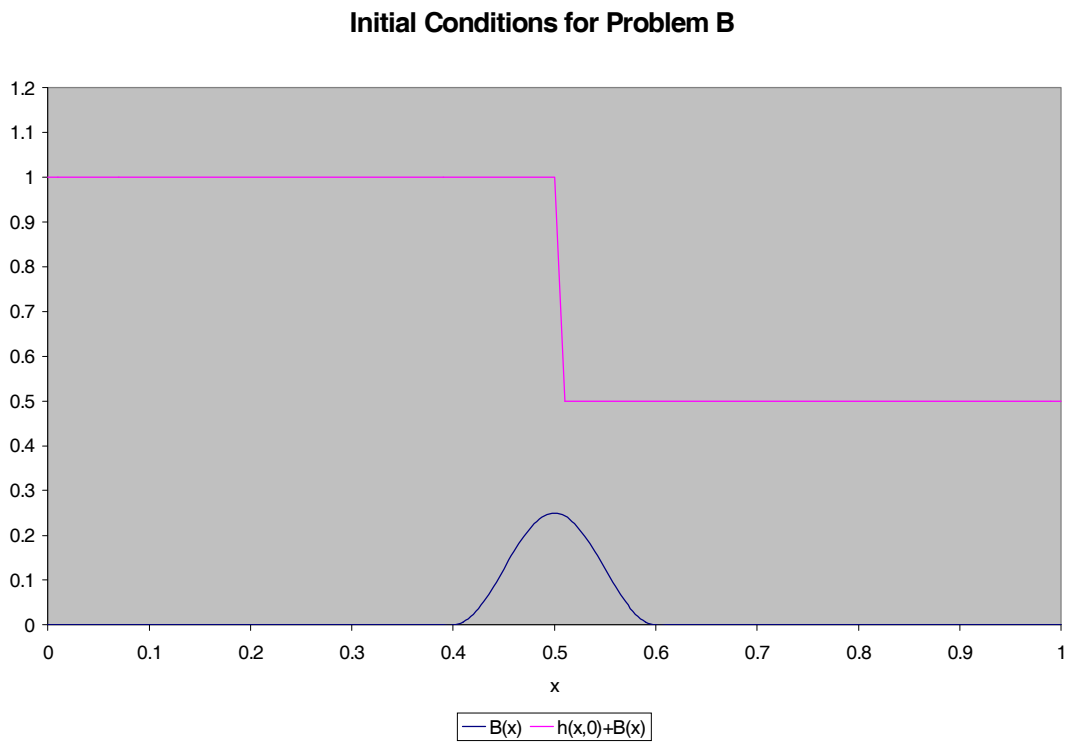

Figure 2-4: Initial condition $h(x,0)$ and $B(x)$ for Problem B.

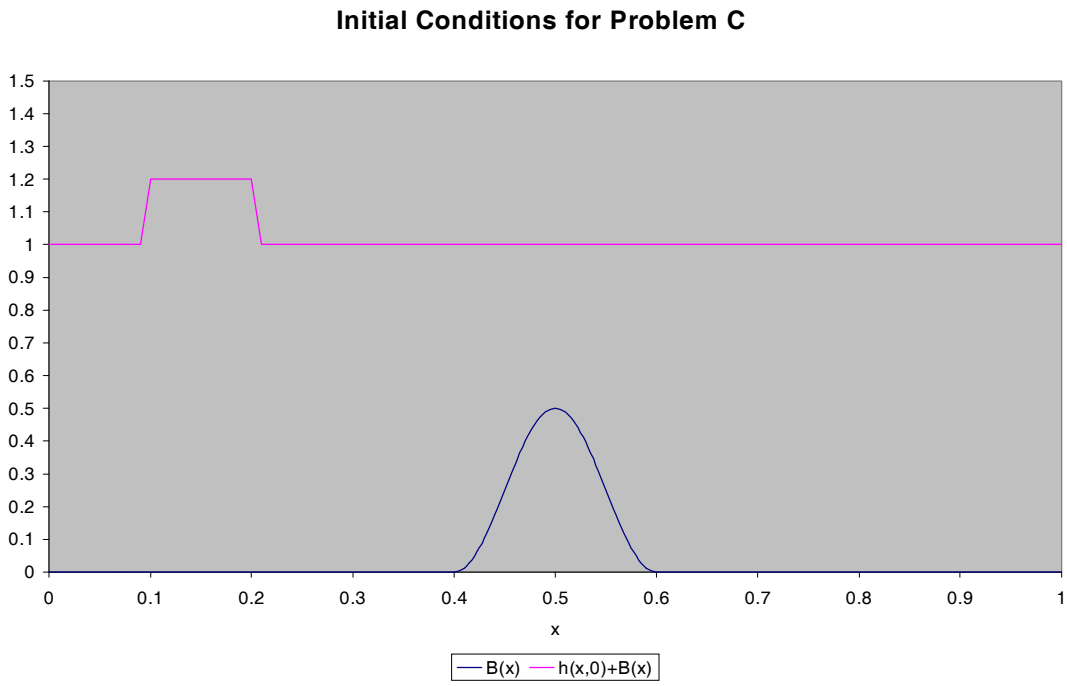

Figure 2-5: Initial condition $h(x,0)$ and $B(x)$ for Problem C.

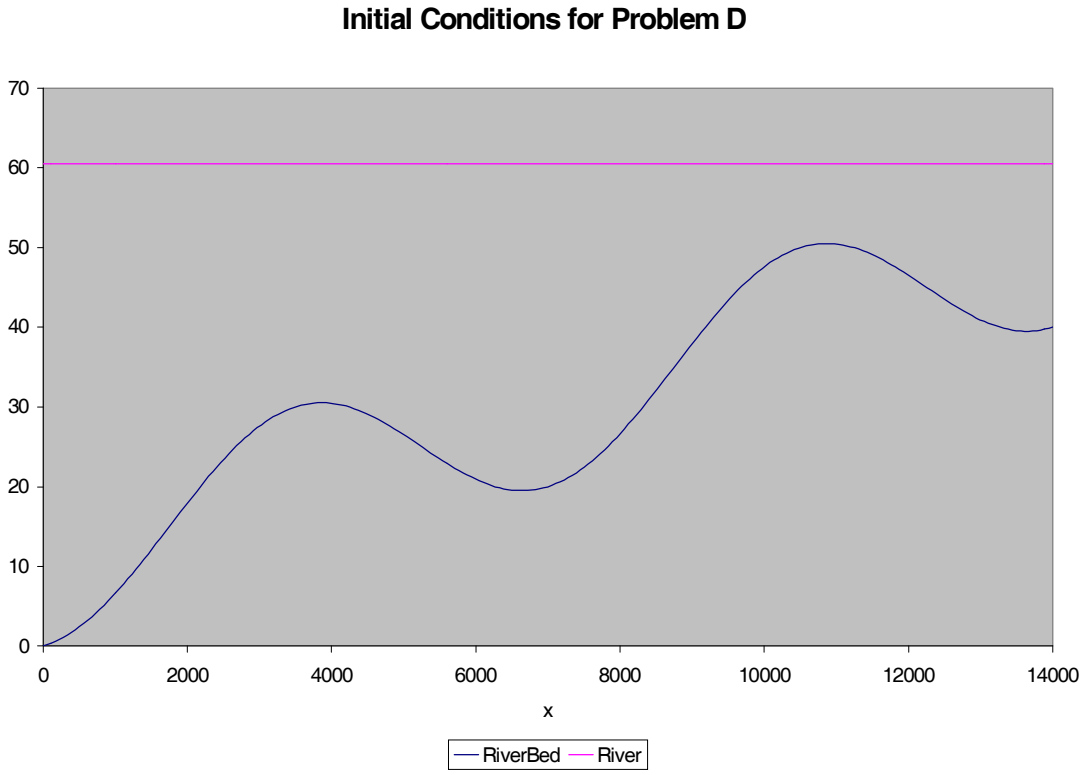

Figure 2-6: Initial condition $h(x,0)$ and $B(x)$ for Problem D.

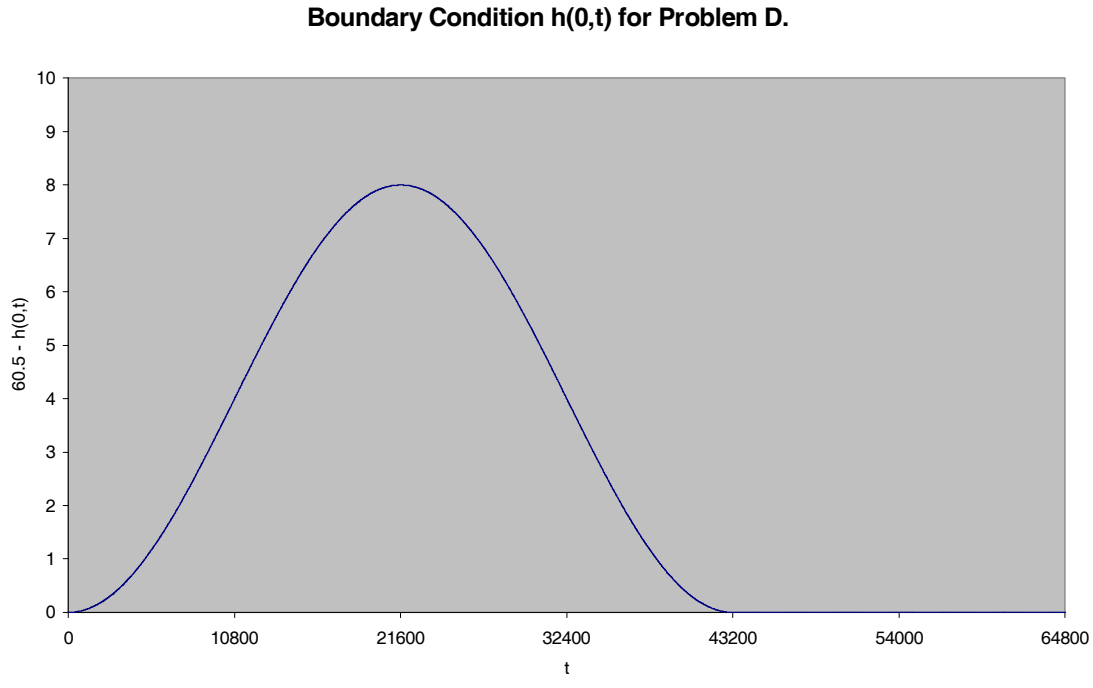

Figure 2-7: Boundary condition $h(0,t)$ for Problem D.

3 Numerical Schemes

There are a variety of numerical techniques for approximating (1.1), e.g. finite element methods, finite volume methods, etc. In this report, we will discuss the finite difference approach (see LeVeque[10] and Kroner[8]), adaptations of Roe's approximate Riemann solver (see Glaister[4], Hubbard[6] and Roe[12]) and the Q-Schemes of Bermudez & Vazquez[1].

These approaches will be used to derive first order and second order numerical schemes where if a numerical scheme is first order then the scheme is dissipative and if a numerical scheme is second order then the scheme is dispersive (see Figure 3-1). Dissipation occurs when the travelling wave's amplitude decreases resulting in the numerical solution being smeared. Dispersion occurs when waves travel at different wave speeds and results in oscillations being present in the numerical results. Both dissipation and dispersion can cause very significant errors in the numerical results, see Figure 3-1, and can sometimes give completely inaccurate numerical results.

One way to minimise dissipation and dispersion is to use a numerical method which satisfies the Total Variational Diminishing property (see Sweby[14] and Harten[5]). Flux-limiter methods satisfy the TVD property and switch between a second order approximation when the region is smooth and a first order approximation when near a discontinuity. Flux-limiter methods will also be applied to the different numerical approaches so that oscillations present in the numerical solution can be minimised.

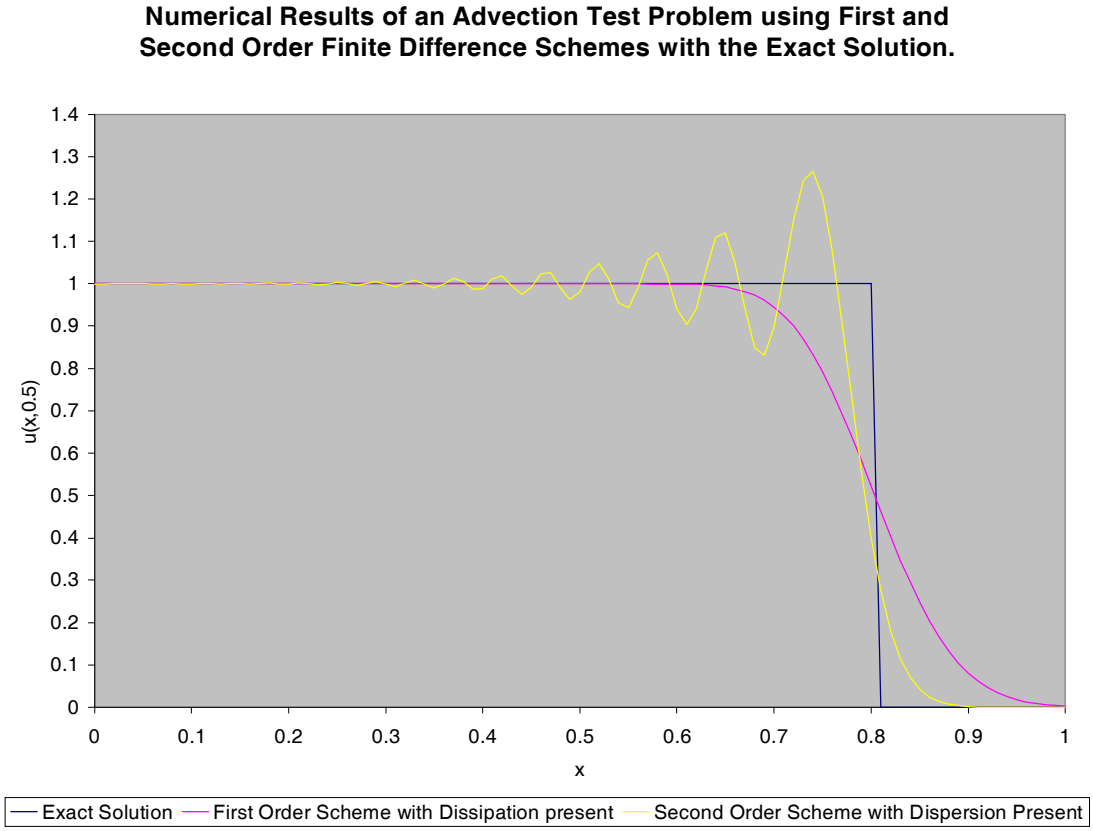

Figure 3-1: Illustration of Dispersion and Dissipation

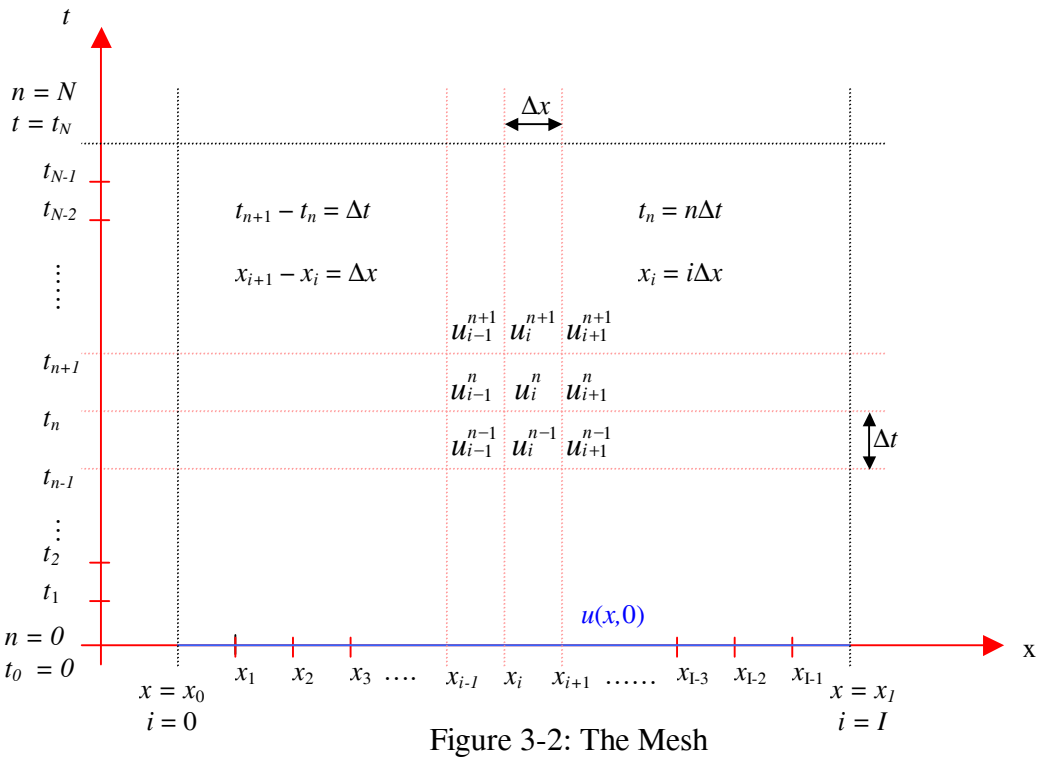

Figure 3-2: The Mesh

3.1 Mesh & Boundary Conditions

Before we discuss the various numerical approaches for approximating the Shallow Water Equations (1.1), we must define the mesh and then look at the numerical boundary conditions required to implement the numerical approaches correctly.

In this report, we will use a fixed mesh over the finite region $x_o \le x \le x_I$ and $0 \le t \le t_N$, which is illustrated in Figure 3-2. Here, the numerical solution is denoted by $u_i^n \approx u(i\Delta x, n\Delta t)$ where $\Delta x = x_i - x_{i-1}$ and $\Delta t = t_n - t_{n-1}$ for all i and n.

Numerical boundary conditions are required at $x = x_0$ and $x = x_I$ and are derived separately for each test problem. For

i) Problem A and Problem B, walls are present at the upstream boundary, $x = 0$, and the downstream boundary, $x = 1$, so we will need to reflect the velocity at the two boundaries

$$u_{-i}^n = -u_i^n \text{ and } u_{I+i}^n = -u_{I-i}^n$$

where $i = 1,2$. For the water depth, we will assume that the depth is constant at the boundaries

$$h_{-i}^n = h_0^n \text{ and } h_{I+i}^n = h_I^n$$

where $i = 1, 2$.

ii) Problem C we will assume that the velocity and water depth are constant

$$h_{-i}^n = h_0^n, \quad h_{I+i}^n = h_I^n, \quad u_{-i}^n = u_0^n \text{ and } u_{I+i}^n = u_I^n$$

where $i = 1, 2$.

iii) Problem D we use the analytical boundary conditions to obtain

$$h_{-i}^n = h(0,t) \text{ and } u_{I+i}^n = 0$$

where $i = 0, 1, 2$. For the second boundary condition, we will use

$$h_{I+i}^n = h_{I+i}^0 \quad \text{and} \quad u_{-i}^n = u_0^n$$

where $i = 1, 2$.

3.2 Finite Difference Method

One approach widely used to numerically approximate (1.1) is the finite difference method. This method involves replacing the derivatives of (1.1) with finite difference approximations, e.g.

$$\frac{\partial \mathbf{w}}{\partial t} = \frac{\mathbf{w}_i^{n+1} - \mathbf{w}_i^n}{\Delta t}$$

which is a forward difference approximation in time, to obtain a finite difference scheme. Great care must be taken when using finite differences to construct a finite difference scheme as we need to ensure that the scheme is conservative. A finite difference scheme that is not conservative may propagate discontinuities at the wrong wave speed, if at all, giving inaccurate numerical results. To ensure we obtain a conservative scheme, we only construct finite difference schemes of the form

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - \frac{\Delta t}{\Delta x}\left[\mathbf{F}^*_{i+1/2} - \mathbf{F}^*_{i-1/2}\right], \tag{3.1}$$

where $\mathbf{F}^*$ is called the numerical flux function. Notice that (3.1) approximates the homogeneous problem (2.1). Obtaining a conservative scheme when a source term is present can be very difficult, especially when the source term is stiff. However, there are a variety of numerical techniques which approximate the source term and they generally fall under two categories: a pointwise approach, where we add a source term approximation to (3.1), i.e.

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - \frac{\Delta t}{\Delta x}\left[\mathbf{F}^*_{i+1/2} - \mathbf{F}^*_{i-1/2}\right] + \Delta t \mathbf{R}_i^n \tag{3.2}$$

or a physical approach, which involves taking an average value of the source term and upwinding (see Sweby[15] for more details), i.e.

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - \frac{\Delta t}{\Delta x}\left[\mathbf{F}^*_{i+1/2} - \mathbf{F}^*_{i-1/2}\right] + \Delta t \mathbf{R}_i^* . \tag{3.3}$$

We will now look at specific schemes which numerically approximate the Shallow Water Equations7.

3.2.1 First Order Lax- Friedrichs Approach

We can use the Lax-Friedrichs scheme with a source term approximation 'added' to approximate (1.1), i.e.

$$\mathbf{w}_i^{n+1} = \frac{\mathbf{w}_{i+1}^n + \mathbf{w}_{i-1}^n}{2} - \frac{s}{2}\left(\mathbf{F}_{i+1}^n - \mathbf{F}_{i-1}^n\right) + s\mathbf{R}_i^n \qquad \text{(S-1)}$$

where

$$s = \frac{\Delta t}{\Delta x},\ \mathbf{w}_i^n = \begin{bmatrix} h_i^n \\ h_i^n u_i^n \end{bmatrix},\ \mathbf{F}_i^n = \begin{bmatrix} h_i^n u_i^n \\ h_i^n \left(u_i^n\right)^2 + \frac{1}{2} g \left(h_i^n\right)^2 \end{bmatrix} \text{ and } \mathbf{R}_i^n = \begin{bmatrix} 0 \\ -g\, h_i^n \left(B_i - B_{i-1}\right) \end{bmatrix}.$$

Notice that this scheme is in a similar form of (3.2) and that the Jacobian matrix of F(w) associated with the system (1.1) does not need to be approximated. However, the Lax-Friedrichs scheme can suffer from oscillations but sometimes these oscillations can be minimised by using sufficiently small step-sizes.

3.2.2 Explicit MacCormack Approach

We can also use an approach discussed by LeVeque & Yee[9] to approximate (1.1), which involves modifying the explicit MacCormack scheme, i.e.

$$\mathbf{w}_i^{n+1} = \frac{1}{2}\left(\mathbf{w}_i^n + \mathbf{w}_i^{(1)}\right) - \frac{s}{2}\left(\mathbf{F}_i^{(1)} - \mathbf{F}_{i-1}^{(1)}\right) + \frac{s}{2}\mathbf{R}_i^{(1)} \qquad \text{(S-2a)}$$

where

$$\mathbf{w}_i^{(1)} = \mathbf{w}_i^n - s\left(\mathbf{F}_{i+1}^n - \mathbf{F}_i^n\right) + s\mathbf{R}_i^n,$$

$$s = \frac{\Delta t}{\Delta x},\ \mathbf{w}_i^n = \begin{bmatrix} h_i^n \\ h_i^n u_i^n \end{bmatrix},\ \mathbf{F}_i^n = \begin{bmatrix} h_i^n u_i^n \\ h_i^n \left(u_i^n\right)^2 + \frac{1}{2} g \left(h_i^n\right)^2 \end{bmatrix} \text{ and } \mathbf{R}_i^n = \begin{bmatrix} 0 \\ -g\, h_i^n \left(B_i - B_{i-1}\right) \end{bmatrix}.$$

Unfortunately, this scheme suffers from dispersion and results in oscillations being present in the numerical approximation. However, LeVeque & Yee[9] eliminated

these oscillations by adapting (S-2a) so that the scheme satisfies the TVD property and obtained

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^{(2)} + \left(\mathbf{X}_{i+1/2}\boldsymbol{\varphi}_{i+1/2} - \mathbf{X}_{i-1/2}\boldsymbol{\varphi}_{i-1/2}\right) \tag{S-2b}$$

where $\mathbf{X}$ represents a matrix containing the right eigenvectors $\mathbf{e}$, $\mathbf{w}_i^{(2)}$ is the numerical approximation derived from scheme (S-2a), i.e.

$$\mathbf{w}_i^{(2)} = \frac{1}{2}\left(\mathbf{w}_i^n + \mathbf{w}_i^{(1)}\right) - \frac{s}{2}\left(\mathbf{F}_i^{(1)} - \mathbf{F}_{i-1}^{(1)}\right) + \frac{s}{2}\mathbf{R}_i^{(1)},$$

and

$$\boldsymbol{\varphi}_{i+1/2} = \begin{bmatrix} \phi_{i+1/2}^1 \\ \phi_{i+1/2}^2 \end{bmatrix}$$

where

$$\phi_{i+1/2}^k = \frac{1}{2}\left(\left|v_{i+1/2}^k\right| - \left(v_{i+1/2}^k\right)^2\right)\left(\alpha_{i+1/2}^k - Q_{i+1/2}^k\right),$$

$$Q_{i+1/2}^k = \min\text{mod}\left(\alpha_{i-1/2}^k, \alpha_{i+1/2}^k, \alpha_{i+3/2}^k\right),$$

$$\min\text{mod}(a,b,c) = \begin{cases} d\min\left(|a|,|b|,|c|\right) & \text{if } d = \text{sgn}(a) = \text{sgn}(b) = \text{sgn}(c) \\ 0 & \text{otherwise} \end{cases},$$

$$\begin{bmatrix} \alpha_{i+1/2}^1 \\ \alpha_{i+1/2}^2 \end{bmatrix} = \mathbf{X}_{i+1/2}^{-1}\left(\mathbf{w}_{i+1}^n - \mathbf{w}_i^n\right) \text{ and } v_{i+1/2}^k = \frac{\Delta t}{\Delta x}\lambda_{i+1/2}^k.$$

Here, k represents the k^{th} component of the vector and λ, $\mathbf{e}$ and α represents the eigenvalues, eigenvectors and wave strengths associated with the system (1.1) respectively, which can be obtained by the decomposition

$$\Delta\mathbf{F} = \sum_{k=1}^{2} \tilde{\alpha}_k \tilde{\lambda}_k \tilde{\mathbf{e}}_k = \tilde{\mathbf{A}}\Delta\mathbf{w}.$$

Variables which have $\tilde{}$ represent the Roe average and $\tilde{\mathbf{A}}$ is the Jacobian matrix evaluated at Roe's average state. We will discuss the method of decomposition in Section 3.3.

3.2.3 Semi-Implicit MacCormack Approach

LeVeque & Yee[9] also obtained a semi-implicit approach to approximate (1.1), which involves modifying the explicit MacCormack scheme so that

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^{(2)} + \left(\mathbf{X}_{i+1/2}\boldsymbol{\varphi}_{i+1/2} - \mathbf{X}_{i-1/2}\boldsymbol{\varphi}_{i-1/2}\right) \tag{3.4}$$

where

$$\mathbf{w}_i^{(2)} = \frac{1}{2}\left(\mathbf{w}_i^n + \mathbf{w}_i^{(1)}\right) - \frac{s}{2}\left[\mathbf{I} - \frac{s}{2}\left[\frac{\partial \mathbf{R}}{\partial \mathbf{w}}\right]\right]^{-1}\left(\mathbf{F}_i^{(1)} - \mathbf{F}_{i-1}^{(1)}\right) + \frac{s}{2}\left[\mathbf{I} - \frac{s}{2}\left[\frac{\partial \mathbf{R}}{\partial \mathbf{w}}\right]\right]^{-1}\mathbf{R}_i^{(1)},$$

$$\mathbf{w}_i^{(1)} = \mathbf{w}_i^n - s\left[\mathbf{I} - \frac{s}{2}\left[\frac{\partial \mathbf{R}}{\partial \mathbf{w}}\right]\right]^{-1}\left(\mathbf{F}_{i+1}^n - \mathbf{F}_i^n\right) + s\left[\mathbf{I} - \frac{s}{2}\left[\frac{\partial \mathbf{R}}{\partial \mathbf{w}}\right]\right]^{-1}\mathbf{R}_i^n$$

and

$$\left[\frac{\partial \mathbf{R}}{\partial \mathbf{w}}\right]_i^n = \begin{bmatrix} 0 & 0 \\ -g\left(B_i - B_{i-1}\right) & 0 \end{bmatrix}$$

which implies that

$$\left[\mathbf{I} - \frac{s}{2}\left[\frac{\partial \mathbf{R}}{\partial \mathbf{w}}\right]\right]^{-1} = \begin{bmatrix} 1 & 0 \\ -\frac{s}{2}g\left(B_i - B_{i-1}\right) & 1 \end{bmatrix}.$$

Also, notice that for the Shallow Water Equations

$$\left[\mathrm{I} - \frac{s}{2}\left[\frac{\partial \mathbf{R}}{\partial \mathbf{w}}\right]\right]^{-1}\mathbf{R}_i^n = \begin{bmatrix} 1 & 0 \\ -\frac{s}{2}g\left(B_i - B_{i-1}\right) & 1 \end{bmatrix}\begin{bmatrix} 0 \\ -g\,h_i^n\left(B_i - B_{i-1}\right) \end{bmatrix} = \begin{bmatrix} 0 \\ -g\,h_i^n\left(B_i - B_{i-1}\right) \end{bmatrix} = \mathbf{R}_i^n. \tag{3.5}$$

Hence, we can re-write (3.4) as

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^{(2)} + \left(\mathbf{X}_{i+1/2}\boldsymbol{\varphi}_{i+1/2} - \mathbf{X}_{i-1/2}\boldsymbol{\varphi}_{i-1/2}\right) \tag{S-2c}$$

where

$$\mathbf{w}_i^{(2)} = \frac{1}{2}\left(\mathbf{w}_i^n + \mathbf{w}_i^{(1)}\right) - \frac{s}{2}\left[\mathbf{I} - \frac{s}{2}\left[\frac{\partial \mathbf{R}}{\partial \mathbf{w}}\right]\right]^{-1}\left(\mathbf{F}_i^{(1)} - \mathbf{F}_{i-1}^{(1)}\right) + \frac{s}{2}\mathbf{R}_i^{(1)}$$

and

$$\mathbf{w}_i^{(1)} = \mathbf{w}_i^n - s\left[\mathbf{I} - \frac{s}{2}\left[\frac{\partial \mathbf{R}}{\partial \mathbf{w}}\right]\right]^{-1}\left(\mathbf{F}_{i+1}^n - \mathbf{F}_i^n\right) + s\,\mathbf{R}_i^n.$$

3.3 Approximate Riemann Solvers

Roe [12] derived an approach which approximates systems of conservation laws by using a piecewise constant approximation

$$\mathbf{w}(x,t_n)=\begin{cases}\mathbf{w}_L & \text{if } x_L-\dfrac{\Delta x}{2}<x<x_L+\dfrac{\Delta x}{2}\\ \mathbf{w}_R & \text{if } x_R-\dfrac{\Delta x}{2}<x<x_R+\dfrac{\Delta x}{2}\end{cases},$$

where $\mathbf{w}_L$ and $\mathbf{w}_R$ represent the piecewise constant states at t_n, and determining the exact solution of a linearised Riemann problem which is related to (2.1)

$$\frac{\partial \mathbf{w}}{\partial t}+\tilde{\mathbf{A}}(\mathbf{w}_L,\mathbf{w}_R)\frac{\partial \mathbf{w}}{\partial x}=0, \tag{3.6}$$

where $\tilde{\mathbf{A}}(\mathbf{w}_L,\mathbf{w}_R)\approx\dfrac{\partial \mathbf{F}}{\partial \mathbf{w}}$ is the linearised Jacobian matrix and $\tilde{}$ is called the Roe average. The eigenvalues and eigenvectors of $\tilde{\mathbf{A}}$ are $\tilde{\lambda}$ and $\tilde{\mathbf{e}}$ respectively and are determined from the decomposition

$$\Delta\mathbf{F}=\sum_{k=1}^{M}\tilde{\alpha}_k\tilde{\lambda}_k\tilde{\mathbf{e}}_k=\tilde{\mathbf{A}}\Delta\mathbf{w}$$

where $\Delta\mathbf{w}=\mathbf{w}_R-\mathbf{w}_L$, M is the number of waves present in the system and $\tilde{\alpha}$ represents the wave strengths, i.e. $\tilde{\alpha}_k=\Delta\mathbf{w}_k$. Once the eigenvalues, eigenvectors and wave strengths associated with the linearised Riemann problem have been obtained, Roe's scheme [12] can be used

$$\mathbf{w}_i^{n+1}=\mathbf{w}_i^n-s\left(\mathbf{F}^*_{i+1/2}-\mathbf{F}^*_{i-1/2}\right), \tag{3.7}$$

where

$$\mathbf{F}^*_{i+1/2}=\frac{1}{2}\left(\mathbf{F}^n_{i+1}+\mathbf{F}^n_i\right)-\frac{1}{2}\left(\sum_{k=1}^{M}\tilde{\alpha}_k|\tilde{\lambda}_k|\left(1-\phi_k\left(1-|\nu_k|\right)\right)\tilde{\mathbf{e}}_k\right)_{i+1/2},$$

$$\nu_k=\frac{\Delta t}{\Delta x}\tilde{\lambda}_k,\ \theta_k=\frac{(\tilde{\alpha}_k)_{I\pm1/2}}{(\tilde{\alpha}_k)_{i\pm1/2}},\ I=i-\mathrm{sgn}\left((\nu_k)_{i\pm1/2}\right)$$

and ϕ_k can be any of the flux-limiters listed in Table 3-1.

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	$\phi(\theta) = \dfrac{\lvert\theta\rvert + \theta}{1 + \lvert\theta\rvert}$
van Albada	$\phi(\theta) = \dfrac{\theta^2 + \theta}{1 + \theta^2}$

Table 3-1: Some second order flux-limiters.

3.3.1 Roe's Scheme with Source Term Added

As presented so far, Roe's scheme approximates homogeneous systems of conservation laws. We now extend (3.7) to numerically approximate the Shallow Water Equations, which has a source term present.

One approach is to add a source term approximation to (3.7)

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - s\left(\mathbf{F}_{i+1/2}^* - \mathbf{F}_{i-1/2}^*\right) + s\mathbf{R}_i^n \qquad \text{(S-3a)}$$

where

$$\mathbf{F}_{i+1/2}^* = \frac{1}{2}\left(\mathbf{F}_{i+1}^n + \mathbf{F}_i^n\right) - \frac{1}{2}\left(\sum_{k=1}^{M} \tilde{\alpha}_k \lvert\tilde{\lambda}_k\rvert \left(1 - \phi_k\left(1 - \lvert\nu_k\rvert\right)\right)\tilde{\mathbf{e}}_k\right)_{i+1/2},$$

$$\nu_k = \frac{\Delta t}{\Delta x}\tilde{\lambda}_k\,,\ \theta_k = \frac{(\tilde{\alpha}_k)_{I\pm1/2}}{(\tilde{\alpha}_k)_{i\pm1/2}},\ I = i - \operatorname{sgn}\left((\nu_k)_{i\pm1/2}\right).$$

and

$$\mathbf{R}_i^n = \begin{bmatrix} 0 \\ -g\,h_i^n\left(B_i - B_{i-1}\right) \end{bmatrix}.$$

We now need to obtain the eigenvalues, eigenvectors and wave strengths associated with the linearised Riemann problem from the decomposition. For the Shallow Water Equations, we may obtain eigenvalues, eigenvectors and wave strengths (see Glaister[3] and Hubbard[6])

$$\tilde{\lambda}_1 = \tilde{u} + \tilde{c}\,,\quad \tilde{\lambda}_2 = \tilde{u} - \tilde{c}\,,$$

$$\tilde{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ \tilde{u} + \tilde{c} \end{bmatrix}, \quad \tilde{\mathbf{e}}_2 = \begin{bmatrix} 1 \\ \tilde{u} - \tilde{c} \end{bmatrix},$$

$$\tilde{\alpha}_1 = \frac{1}{2}\Delta h + \frac{1}{2\tilde{c}}\left(\Delta(hu) - \tilde{u}\Delta h\right)$$

and

$$\tilde{\alpha}_2 = \frac{1}{2}\Delta h - \frac{1}{2\tilde{c}}\left(\Delta(hu) - \tilde{u}\Delta h\right),$$

where

$$\tilde{u} = \frac{\sqrt{h_R}\,u_R + \sqrt{h_L}\,u_L}{\sqrt{h_L} + \sqrt{h_L}} \quad \text{and} \quad \tilde{c} = \sqrt{g\frac{h_R + h_L}{2}}\ .$$

A semi-implicit approach of (S-3a) can be obtained by approximating the source term at t_{n+1} instead of at t_n, i.e.

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - s\left(\mathbf{F}_{i+1/2}^* - \mathbf{F}_{i-1/2}^*\right) + s\mathbf{R}_i^{n+1}. \tag{3.8}$$

Now, by using Taylor's theorem, we may obtain

$$\mathbf{R}_i^{n+1} \approx \mathbf{R}_i^n + \left(\mathbf{w}_i^{n+1} - \mathbf{w}_i^n\right)\left[\frac{\partial \mathbf{R}}{\partial \mathbf{w}}\right]$$

and by substituting into (3.8) and using (3.5), we may obtain

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - s\left[\mathbf{I} - \frac{s}{2}\left[\frac{\partial \mathbf{R}}{\partial \mathbf{w}}\right]\right]^{-1}\left(\mathbf{F}_{i+1/2}^* - \mathbf{F}_{i-1/2}^*\right) + s\,\mathbf{R}_i^n\,. \tag{S-3b}$$

where

$$\mathbf{F}_{i+1/2}^* = \frac{1}{2}\left(\mathbf{F}_{i+1}^n + \mathbf{F}_i^n\right) - \frac{1}{2}\left(\sum_{k=1}^{M}\tilde{\alpha}_k\left|\tilde{\lambda}_k\right|\left(1 - \phi_k\left(1 - |\nu_k|\right)\right)\tilde{\mathbf{e}}_k\right)_{i+1/2},$$

$$\nu_k = \frac{\Delta t}{\Delta x}\tilde{\lambda}_k\,,\ \theta_k = \frac{\left(\tilde{\alpha}_k\right)_{I\pm 1/2}}{\left(\tilde{\alpha}_k\right)_{i\pm 1/2}}\ ,\ I = i - \mathrm{sgn}\left(\left(\nu_k\right)_{i\pm 1/2}\right)$$

and

$$\mathbf{R}_i^n = \begin{bmatrix} 0 \\ -g\,h_i^n\left(B_i - B_{i-1}\right) \end{bmatrix}.$$

3.3.2 Roe's Scheme with Source Term Decomposed

Instead of adding a source term approximation to Roe's scheme [12], we can use an approach discussed by Glaister[3]. This method approximates the source term by decomposing the source term in a similar way that we did for the flux terms, i.e.

$$\frac{1}{\Delta x}\sum_{k=1}^{M}\tilde{\beta}_k\,\tilde{\mathbf{e}}_k = \tilde{\mathbf{R}}\,.$$

Here, $\tilde{\beta}_k$ are the coefficients of the decomposition of the source term onto the eigenvectors of the characteristic decomposition (see Glaister[3] and Hubbard[6]).

For the Shallow Water Equations, we may obtain

$$\tilde{\beta}_1 = \frac{\tilde{c}\,\Delta H}{2} \quad \text{and} \quad \tilde{\beta}_2 = -\frac{\tilde{c}\,\Delta H}{2}\,. \tag{3.9}$$

Alternatively, we can use

$$\tilde{\beta}_1 = -\frac{\tilde{c}\,\Delta B}{2} \quad \text{and} \quad \tilde{\beta}_2 = \frac{\tilde{c}\,\Delta B}{2}$$

for the Shallow Water equations. Once the values of $\tilde{\beta}_k$ have been obtained, we can approximate the source term by using

$$\mathbf{R}_i^{*} = \frac{1}{\Delta x}\left(\mathbf{R}_{i+1/2}^{-} + \mathbf{R}_{i-1/2}^{+}\right)$$

where

$$\mathbf{R}_{i+1/2}^{\pm} = \frac{1}{2}\left(\sum_{k=1}^{2}\tilde{\beta}_k\,\tilde{\mathbf{e}}_k\left(1\pm\operatorname{sgn}(\tilde{\lambda}_k)\right)\right)_{i+1/2},$$

which is a first order approximation of the source term. Hence, Roe's first order scheme with the source term decomposed may be written

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^{n} - s\left(\mathbf{F}_{i+1/2}^{*} - \mathbf{F}_{i-1/2}^{*}\right) + s\left(\mathbf{R}_{i+1/2}^{-} + \mathbf{R}_{i-1/2}^{+}\right) \tag{S-3c}$$

where

$$\mathbf{F}_{i+1/2}^{*} = \frac{1}{2}\left(\mathbf{F}_{i+1}^{n} + \mathbf{F}_{i}^{n}\right) - \frac{1}{2}\left(\sum_{k=1}^{2}\tilde{\alpha}_k\left|\tilde{\lambda}_k\right|\tilde{\mathbf{e}}_k\right)_{i+1/2}$$

and

$$\mathbf{R}^{\pm}_{i+1/2} = \frac{1}{2}\left(\sum_{k=1}^{2} \tilde{\beta}_k \tilde{\mathbf{e}}_k \left(1 \pm \mathrm{sgn}(\tilde{\lambda}_k)\right) \right)_{i+1/2}.$$

We can also obtain a second order accurate approximation of the source term and thus obtain Roe's second order scheme with source term decomposed

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - s\left(\mathbf{F}^*_{i+1/2} - \mathbf{F}^*_{i-1/2}\right) + s\left(\mathbf{R}^-_{i+1/2} + \mathbf{R}^+_{i-1/2}\right) \qquad \text{(S-3d)}$$

where

$$\mathbf{F}^*_{i+1/2} = \frac{1}{2}\left(\mathbf{F}^n_{i+1} + \mathbf{F}^n_i\right) - \frac{s}{2}\left(\sum_{k=1}^{2} \tilde{\alpha}_k (\tilde{\lambda}_k)^2 \tilde{\mathbf{e}}_k \right)_{i+1/2}$$

and

$$\mathbf{R}^{\pm}_{i+1/2} = \frac{s}{2}\left[\sum_{k=1}^{2} \tilde{\beta}_k \tilde{\mathbf{e}}_k |\tilde{\lambda}_k| \left(1 \pm \mathrm{sgn}(\tilde{\lambda}_k)\right) \right]_{i+1/2}.$$

In addition, we can also use an approach proposed by Hubbard[6] which involves applying TVD to the source term approximation as well as the conservation law approximation, i.e.

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - s\left(\mathbf{F}^*_{i+1/2} - \mathbf{F}^*_{i-1/2}\right) + s\left(\mathbf{R}^-_{i+1/2} + \mathbf{R}^+_{i-1/2}\right) \qquad \text{(S-4)}$$

where

$$\mathbf{F}^*_{i+1/2} = \frac{1}{2}\left(\mathbf{F}^n_{i+1} + \mathbf{F}^n_i\right) - \frac{1}{2}\left(\sum_{k=1}^{M} \tilde{\alpha}_k |\tilde{\lambda}_k| \left(1 - \phi_k\left(1 - |\nu_k|\right)\right) \tilde{\mathbf{e}}_k \right)_{i+1/2},$$

$$\mathbf{R}^{\pm}_{i+1/2} = \frac{1}{2}\left[\sum_{k=1}^{2} \tilde{\beta}_k \tilde{\mathbf{e}}_k \left(1 \pm \mathrm{sgn}(\tilde{\lambda}_k)\right)\left(1 - \phi_k\left(1 - |\nu_k|\right)\right) \right]_{i+1/2},$$

$$\nu_k = \frac{\Delta t}{\Delta x}\tilde{\lambda}_k, \quad \theta_k = \frac{(\tilde{\alpha}_k)_{I\pm 1/2}}{(\tilde{\alpha}_k)_{i\pm 1/2}}, \quad I = i - \mathrm{sgn}\left((\nu_k)_{i\pm 1/2}\right)$$

and ϕ_k can be any of the flux-limiter functions listed in Table 3-1.

3.3.3 The C – Property of Bermudez & Vazquez

The advantage of decomposing the source term as well as the conservation law is that we obtain a numerical scheme that satisfies the approximate C-property. Bermudez & Vazquez[1] state that in order for a numerical scheme to satisfy

i) The approximate C-property, the numerical scheme must be at least second order accurate when applied to the quiescent flow case, i.e. $u \equiv 0$ and $h \equiv H$.

ii) The exact C-property, the numerical scheme must be exact when applied to the quiescent flow case, i.e. $u \equiv 0$ and $h \equiv H$.

Hubbard's approach and Roe's scheme with source term decomposed all satisfy the exact C-property and should produce very accurate numerical results. However, Roe's scheme with source term added does not even satisfy the approximate C-property and may give misleading results as the source term becomes significant.

3.4 Q-Schemes of Bermudez & Vazquez

Bermudez & Vazquez[1] discussed a variety of Q-Schemes, which numerically approximate (1.1). All of the Q-Schemes discussed were used with the following first order equation

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - s\left(\mathbf{F}_{i+1/2}^* - \mathbf{F}_{i-1/2}^*\right) + \Delta t \mathbf{R}_i^n \tag{S-5}$$

where

$$\mathbf{F}_{i+1/2}^* = \frac{1}{2}\left(\mathbf{F}_{i+1}^n + \mathbf{F}_i^n\right) - \frac{1}{2}\left|\mathbf{Q}\left(\mathbf{w}_i^n, \mathbf{w}_{i+1}^n\right)\right|\left(\mathbf{w}_{i+1}^n - \mathbf{w}_i^n\right),$$

$$\mathbf{R}_i^n = \mathbf{R}_L\left(x_{i-1}, x_i, \mathbf{w}_{i-1}^n, \mathbf{w}_i^n\right) + \mathbf{R}_R\left(x_i, x_{i+1}, \mathbf{w}_i^n, \mathbf{w}_{i+1}^n\right),$$

$$\mathbf{R}_L\left(x_{i-1}, x_i, \mathbf{w}_{i-1}^n, \mathbf{w}_i^n\right) = \frac{1}{2}\left[\mathbf{I} + \left|\mathbf{Q}\left(\mathbf{w}_{i-1}^n, \mathbf{w}_i^n\right)\right|\mathbf{Q}^{-1}\left(\mathbf{w}_{i-1}^n, \mathbf{w}_i^n\right)\right]\hat{\mathbf{R}}\left(x_{i-1}, x_i, \mathbf{w}_{i-1}^n, \mathbf{w}_i^n\right),$$

$$\mathbf{R}_R\left(x_i, x_{i+1}, \mathbf{w}_i^n, \mathbf{w}_{i+1}^n\right) = \frac{1}{2}\left[\mathbf{I} - \left|\mathbf{Q}\left(\mathbf{w}_i^n, \mathbf{w}_{i+1}^n\right)\right|\mathbf{Q}^{-1}\left(\mathbf{w}_i^n, \mathbf{w}_{i+1}^n\right)\right]\hat{\mathbf{R}}\left(x_i, x_{i+1}, \mathbf{w}_i^n, \mathbf{w}_{i+1}^n\right),$$

$$\hat{\mathbf{R}}\left(x_i, x_{i+1}, \mathbf{w}_i^n, \mathbf{w}_{i+1}^n\right) = \begin{bmatrix} 0 \\ g\left(\frac{h_i^n + h_{i+1}^n}{2}\right)\left(\frac{H(x_{i+1}) - H(x_i)}{\Delta x}\right) \end{bmatrix}$$

and **Q** is a matrix calculated by using a certain Q-Scheme. Bermudez & Vazquez[1] discussed a variety of Q-Schemes but generally concentrated on the Q-scheme of van Leer and a Q-Scheme which is equivalent to Roe's first order scheme (S-3c).

The Q-Scheme of van Leer is

$$\mathbf{Q}\left(\mathbf{w}_i^n, \mathbf{w}_{i+1}^n\right) = \mathbf{A}\left(\frac{\mathbf{w}_i^n + \mathbf{w}_{i+1}^n}{2}\right),$$

where **A** denotes the Jacobian matrix of $\mathbf{F}(\mathbf{w})$

$$\mathbf{A}(\mathbf{w}) = \begin{bmatrix} 0 & 1 \\ gh - u^2 & 2u \end{bmatrix},$$

which has eigenvalues

$$\lambda_1 = u + \sqrt{gh} \quad \text{and} \quad \lambda_2 = u - \sqrt{gh}$$

and eigenvectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ u + \sqrt{gh} \end{bmatrix} \quad \text{and} \quad \mathbf{e}_2 = \begin{bmatrix} 1 \\ u - \sqrt{gh} \end{bmatrix}.$$

The Q-Scheme that is equivalent to Roe's first order scheme is

$$\mathbf{Q}\left(\mathbf{w}_i^n, \mathbf{w}_{i+1}^n\right) = \mathbf{A}\left(\tilde{\mathbf{w}}\right) = \begin{bmatrix} 0 & 1 \\ \tilde{c}^2 - \tilde{u}^2 & 2\tilde{u} \end{bmatrix},$$

which has eigenvalues

$$\tilde{\lambda}_1 = \tilde{u} + \tilde{c}\,, \quad \tilde{\lambda}_2 = \tilde{u} - \tilde{c}$$

and eigenvectors

$$\tilde{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ \tilde{u} + \tilde{c} \end{bmatrix}, \quad \tilde{\mathbf{e}}_2 = \begin{bmatrix} 1 \\ \tilde{u} - \tilde{c} \end{bmatrix},$$

where

$$\tilde{u} = \frac{\sqrt{h_R}\,u_R + \sqrt{h_L}\,u_L}{\sqrt{h_L} + \sqrt{h_L}} \quad \text{and} \quad \tilde{c} = \sqrt{g\frac{h_R + h_L}{2}}\,.$$

For both Q-Schemes, the modulus of matrix **A** can be obtained by using

$$|\mathbf{A}| = \mathbf{X}^{-1}|\mathbf{\Lambda}|\mathbf{X}$$

where $\mathbf{\Lambda}$ represents a matrix whose diagonal elements are the eigenvalues of $\mathbf{A}$,

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

and $\mathbf{X}$ denotes a matrix containing the right eigenvectors,

$$\mathbf{X}(\mathbf{w}) = \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix}.$$

Hence, for both Q-Schemes we may obtain

$$|\mathbf{A}| = \frac{1}{\lambda_2 - \lambda_1}\begin{bmatrix} |\lambda_1|\lambda_2 - \lambda_1|\lambda_2| & |\lambda_2| - |\lambda_1| \\ \lambda_2\lambda_1(|\lambda_1| - |\lambda_2|) & \lambda_2|\lambda_2| - \lambda_1|\lambda_1| \end{bmatrix},$$

and for the source term approximation we may obtain

$$\hat{\mathbf{R}}_L(x_{i-1}, x_i, \mathbf{w}_{i-1}^n, \mathbf{w}_i^n) = \frac{\hat{\mathbf{R}}_2(x_{i-1}, x_i, \mathbf{w}_{i-1}^n, \mathbf{w}_i^n)}{((\lambda_2 - \lambda_1)\lambda_2\lambda_1)_{i-1/2}}\begin{bmatrix} (\lambda_1|\lambda_2| - \lambda_2|\lambda_1|) \\ \lambda_2\lambda_1(\lambda_2 - \lambda_1 + |\lambda_2| - |\lambda_1|) \end{bmatrix}_{i-1/2}$$

and

$$\hat{\mathbf{R}}_R(x_i, x_{i+1}, \mathbf{w}_i^n, \mathbf{w}_{i+1}^n) = \frac{\hat{\mathbf{R}}_2(x_i, x_{i+1}, \mathbf{w}_i^n, \mathbf{w}_{i+1}^n)}{((\lambda_2 - \lambda_1)\lambda_2\lambda_1)_{i+1/2}}\begin{bmatrix} -(\lambda_1|\lambda_2| - \lambda_2|\lambda_1|) \\ \lambda_2\lambda_1(\lambda_2 - \lambda_1 - |\lambda_2| + |\lambda_1|) \end{bmatrix}_{i+1/2}$$

where

$$\hat{\mathbf{R}}_2(x_i, x_{i+1}, \mathbf{w}_i^n, \mathbf{w}_{i+1}^n) = \frac{g}{2\Delta x}(h_i^n + h_{i+1}^n)(H(x_{i+1}) - H(x_i)).$$

Again, notice that (S-5) with either Q-Scheme satisfies the exact C-property.

In this chapter, we have discussed a variety of numerical schemes which approximate (1.1). In the next chapter, we will apply these schemes to the four test problems discussed in Chapter 2 to find out which approach is the most accurate.

4 Numerical Results

In this chapter, we will apply the schemes discussed in Chapter 3 and listed in Table 4-1 to the four test problems discussed in Chapter 2 to find out which approach produces the most accurate results. We will not discuss the results of the semi-implicit approaches as they produced almost identical results to the explicit approaches. Also, the numerical results of Bermudez & Vazquez's Q-Schemes will not be discussed as the two Q-Schemes produced almost identical results to Roe's first order scheme with source term decomposed, i.e. (S-3b).

For the first three test problems, step-sizes $\Delta x = 0.001$ and $\Delta t = 0.0001$ will be used with a final time of $t = 0.1$. A comparison will be made at $t = 0.1$ and numerical results will also be shown for $t = 0.01m$, where $m = 0$ to 10.

For test Problem D, step-sizes $\Delta x = 2800$ and $\Delta t = 1$ will be used with a bed length of $L = 648{,}000m$ and a final time of $t = 10{,}800s$. A comparison will be made at $t = 10{,}800s$ and numerical results will also be shown for $t = 1080m$, where $m = 0$ to 10.

Name Of Approach	Reference No.	Order	Paper
Lax-Friedrichs	(S-1)	1	-
MacCormack	(S-2b)	2	Yee[16], LeVeque & Yee[9],
Roe's Scheme with Source Term 'added'	(S-3a)	1 / 2	Hubbard[6], Glaister[3], Roe[12]
Roe's Scheme with Source Term Decomposed	(S-3b)	1	
Hubbard's Approach	(S-4)	2	Hubbard[6]

Table 4-1: Different approaches for numerically approximating (1.1).

Only first order and flux-limited second order numerical results will be discussed and the Minmod flux-limiter will be used with all flux-limited second order approaches.

4.1 First Order Comparison

4.1.1 Numerical Results for Problem A

For this test problem, no source term is present so schemes (S-3a) and (S-3b) are the same. Now, by using schemes (S-1) and (S-3b) to approximate Problem A and comparing with the exact solution, we may obtain the numerical results in Figure 4-1a and Figure 4-1b. Here, we can see that Roe's scheme is more accurate than the Lax-Friedrichs approach since the Lax-Friedrichs approach is more dissipative than Roe's scheme. Also, the Lax-Friedrichs scheme suffered badly from oscillations if larger step-sizes were used whereas Roe's scheme remained accurate but became more dissipative.

4.1.2 Numerical Results for Problem B

For this test problem, a source term is now present so approaches (S-3a) and (S-3b) are no longer equivalent. By using schemes (S-1), (S-3a) and (S-3b) to approximate Problem B, the results in Figure 4-2a to Figure 4-3b were obtained. Here we can see that Roe's scheme with source term decomposed has produced the most accurate results. Roe's scheme with source term added produced almost identical results to Roe's scheme with source term decomposed showing that, for Problem B, adding a source term approximation to Roe's scheme produces sufficiently accurate results. The Lax-Friedrichs approach produced the least accurate results suffering badly from dissipation and the approach also misplaced the disturbance caused by the riverbed.

4.1.3 Numerical Results for Problem C

For this test problem, the source term is becoming more significant, i.e. the variation in the riverbed is becoming more pronounced, which may cause some schemes to produce inaccurate results. By using approaches (S-1), (S-3a) and (S-3b) to approximate Problem C, the results in Figure 4-4a to Figure 4-7b were obtained. Here, we can see that Roe's scheme with source term decomposed has produced the most accurate results but the results are no longer almost identical to Roe's scheme with source term added. Adding the source term in this case has produced movement for $0.3 > x > 0.55$ whereas decomposing the source term has produced no movement for $0.3 > x > 0.55$. This is because Roe's scheme with source term decomposed satisfies the exact C-property whereas Roe's scheme with source term added does not. The Lax-Friedrichs approach has produced the least accurate results due to the scheme suffering badly from dissipation and the scheme has also produced more movement than Roe's scheme with source term added for $0.3 > x > 0.55$. Also, from Figure 4-7b we can see that the Lax-Friedrichs approach has started producing oscillations at the peak of the pulse even though small step-sizes have been used.

4.1.4 Numerical Results for Problem D

For this test problem, the source term is very difficult to approximate accurately which may cause some approaches to produce very inaccurate numerical results. By applying schemes (S-1), (S-3a) and (S-3b) to Problem D, the results in Figure 4-8a to Figure 4-11b were obtained. Here, we can see that Roe's scheme with source term decomposed has produced the most accurate results since it was the only approach not to produce movement after $x = 216{,}000m$ at $t = 10{,}800s$. This is because the numerical scheme satisfies the exact C-property whereas the other approaches do not even satisfy the approximate C-property. Roe's scheme with source term added was

the second most accurate due to the approach producing movement after x = 216,000 at t = 10,800s. The Lax-Friedrichs approach failed to accurately approximate Problem D at all due to the scheme producing oscillations over the whole domain.

4.1.5 Overall Comparison of First Order Approaches

From the results of this sub-section, we have seen that Roe's scheme with source term decomposed has produced very accurate results for all test problems suffering only slightly from dissipation. Roe's scheme with source term added produced accurate numerical results for the first two test problems, but as the source term became more significant, the numerical scheme started to produce less accurate results. In Problem D, Roe's scheme with source term added produced movement after x = 216,000 at t = 10,800s making the scheme very inaccurate. The Lax-Friedrichs approach is accurate for the most basic test problems but only when sufficiently small step-sizes are used otherwise the scheme suffers from oscillations. Also, as the source term became significant the Lax-Friedrichs approach became impractical suffering badly from oscillations even when small step-sizes were used. Hence, Roe's scheme with source term decomposed produced the most accurate results for all test problems in the first order case since it was the only approach to satisfy the exact C-property.

4.2 Flux-Limited Second Order Comparison

4.2.1 Numerical Results for Problem A

Now, by using approaches (S-2b) and (S-4) to approximate Problem A, the results in Figure 4-12a and Figure 4-12b were obtained. Here, we can see that Roe's scheme and the MacCormack approach have produced almost identical results. Both approaches have produced extremely accurate numerical results but Roe's scheme was the most accurate.

4.2.2 Numerical Results for Problem B

By using approaches (S-2), (S-3a) and (S-4) to approximate Problem B, the results in Figure 4-13a to Figure 4-14b were obtained. Here, we can see that Roe's scheme with source term added, Hubbard's approach and the MacCormack approach have all produced almost identical results. All three approaches have produced very accurate results.

4.2.3 Numerical Results for Problem C

By using approaches (S-2), (S-3a) and (S-4) to approximate Problem C, the results in Figure 4-15a to Figure 4-18b were obtained. Here, we can see that Hubbard's approach has produced the most accurate results. The other two approaches have produced movement in the region $0.3 > x > 0.55$ with the MacCormack approach producing the most amount of movement of the three approaches. If we use a larger step-size of $\Delta x = 0.01$ and $\Delta t = 0.001$ with approaches (S-2), (S-3a) and (S-4), we may obtain the results in Figure 4-19a and Figure 4-19b. Here we can see that Hubbard's approach has again produced the most accurate numerical results and has not produced movement in the region $0.3 > x > 0.55$. However, the MacCormack approach and Roe's scheme with source term added have both produced very inaccurate numerical results. Roe's scheme with source term added has produced inaccurate numerical results due to the approach producing a considerable amount of movement in the region $0.3 > x > 0.55$ and the approach has slightly smeared the pulse at $x = 0.5$. The MacCormack approach has produced less movement in the region $0.3 > x > 0.55$ than Roe's scheme with source term added but the approach has smeared the pulse at $x = 0.5$ considerably more than Roe's scheme with source term added. Hence, Hubbard's approach is still very accurate when the step-size is increased but the other two approaches are becoming impractical.

4.2.4 Numerical Results for Problem D

Now, by using approaches (S-2), (S-3a) and (S-4) to approximate Problem D, the results in Figure 4-20a to Figure 4-23b were obtained. Here, we can see that Hubbard's approach was the only approach that did not produce movement after x = 216,000m at t = 10,800s. Roe's scheme with source term added and LeVeque & Yee's MacCormack approach produced very similar results but both produced movement after x = 216,000m at t = 10,800s. Hence, Hubbard's produced the most accurate results.

4.2.5 Overall Comparison of the Flux- Limited Second Order Approaches

From the results of this sub-section, we have seen that Hubbard's approach has produced the most accurate results for all test problems producing very accurate results even for Problem D. LeVeque & Yee's MacCormack approach, Roe's scheme with source term added and Hubbard's approach all produced almost identical results for the first two test problems. However, as the source term became more significant, LeVeque & Yee's MacCormack approach and Roe's scheme with source term added produced less accurate results but still produced very similar results due to both approaches adding a source term approximation on. When the source term became significant and a larger step-size was used, the MacCormack approach and Roe's scheme with source term added both became impractical but Hubbard's approach still produced very accurate numerical results.

First Order Numerical Results:

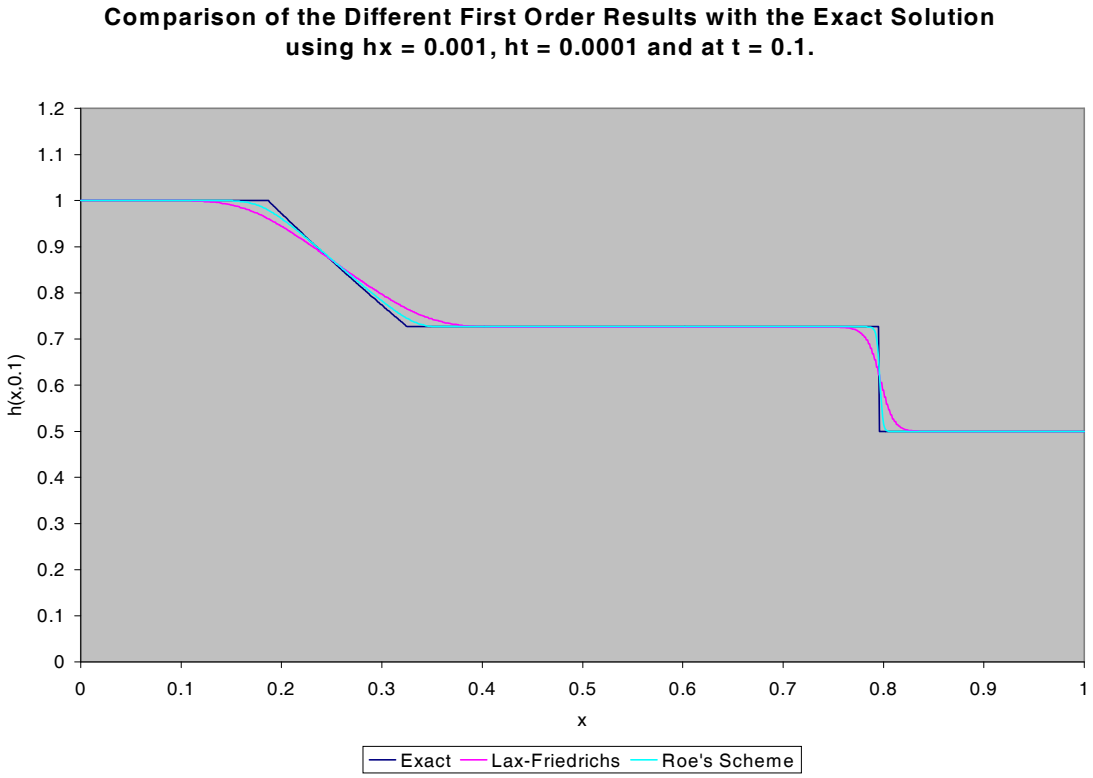

Figure 4-1a: Comparison of the first order results for Problem A

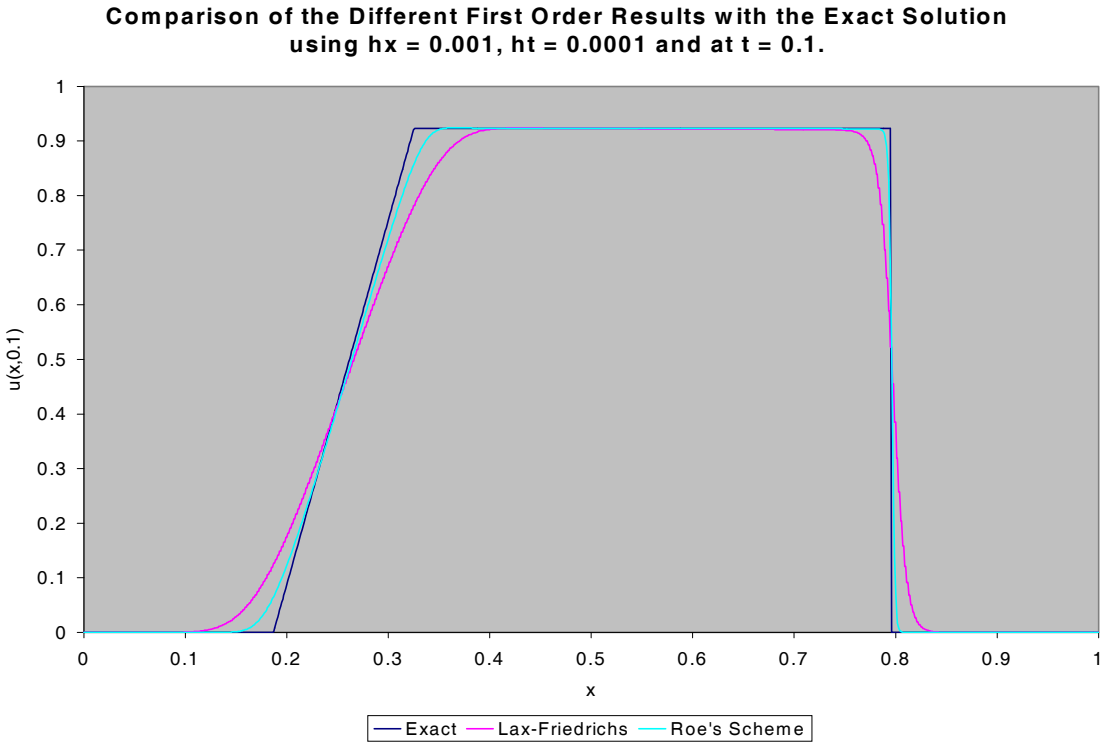

Figure 4-1b: Comparison of the results for Problem A

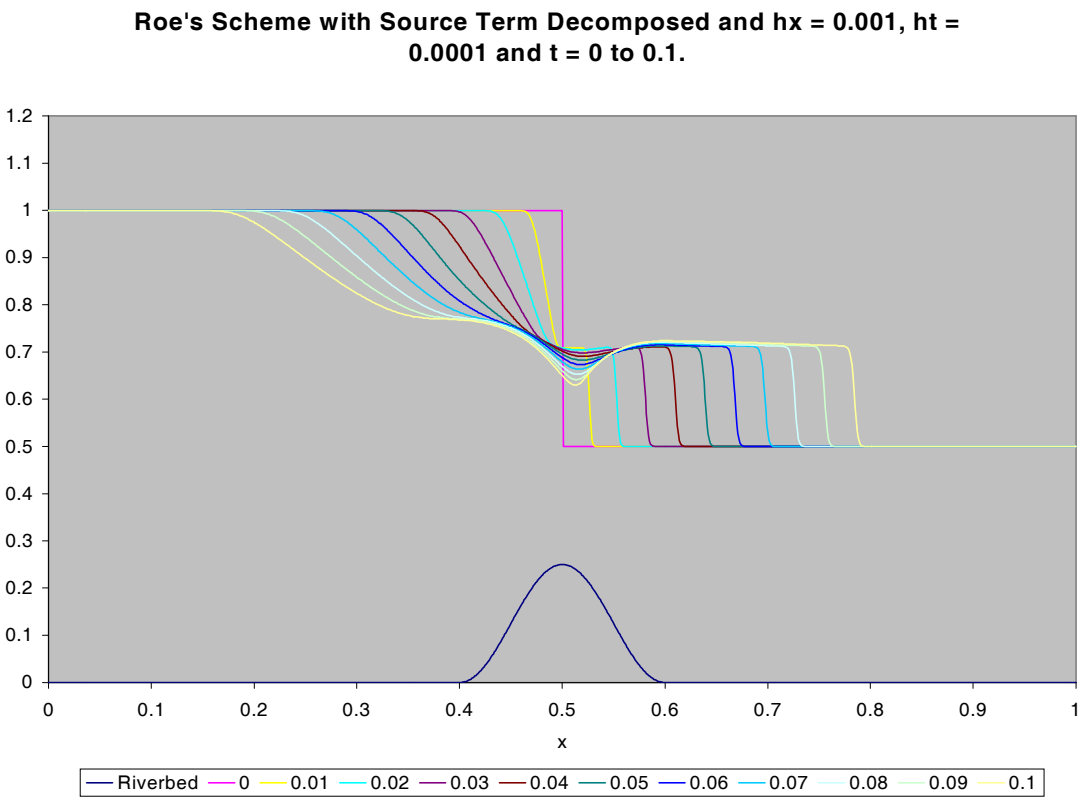

Figure 4-2a: Results of Roe's scheme with source term decomposed for Problem B

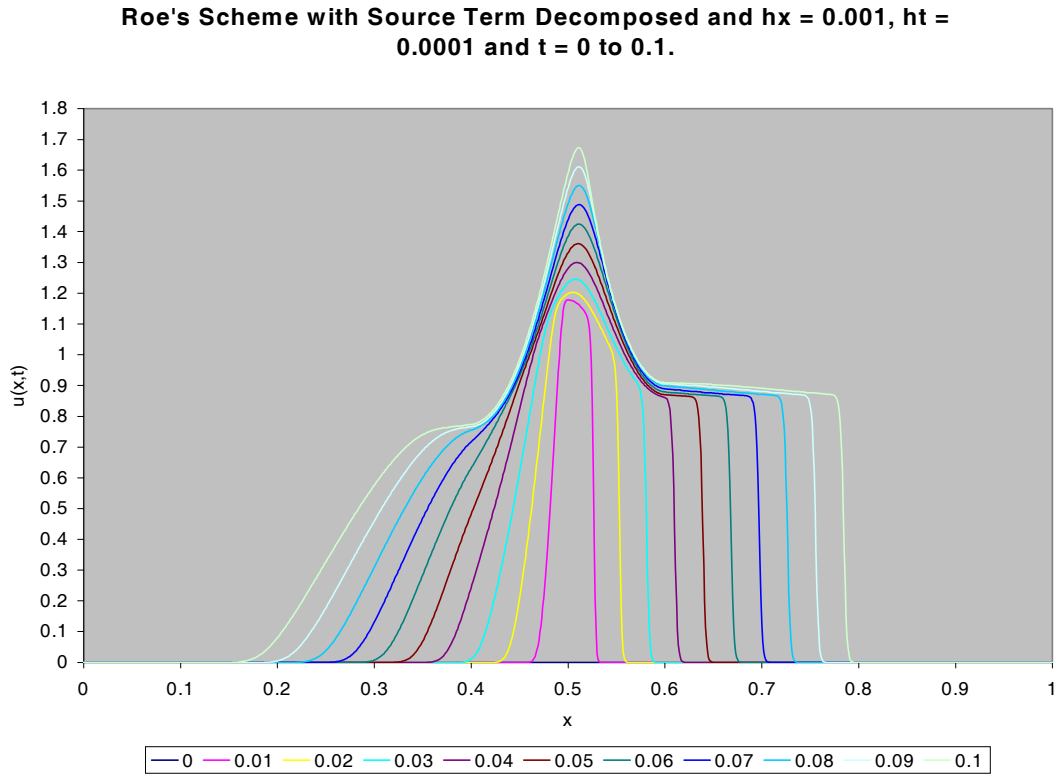

Figure 4-2b: Results of Roe's scheme with source term decomposed for Problem B

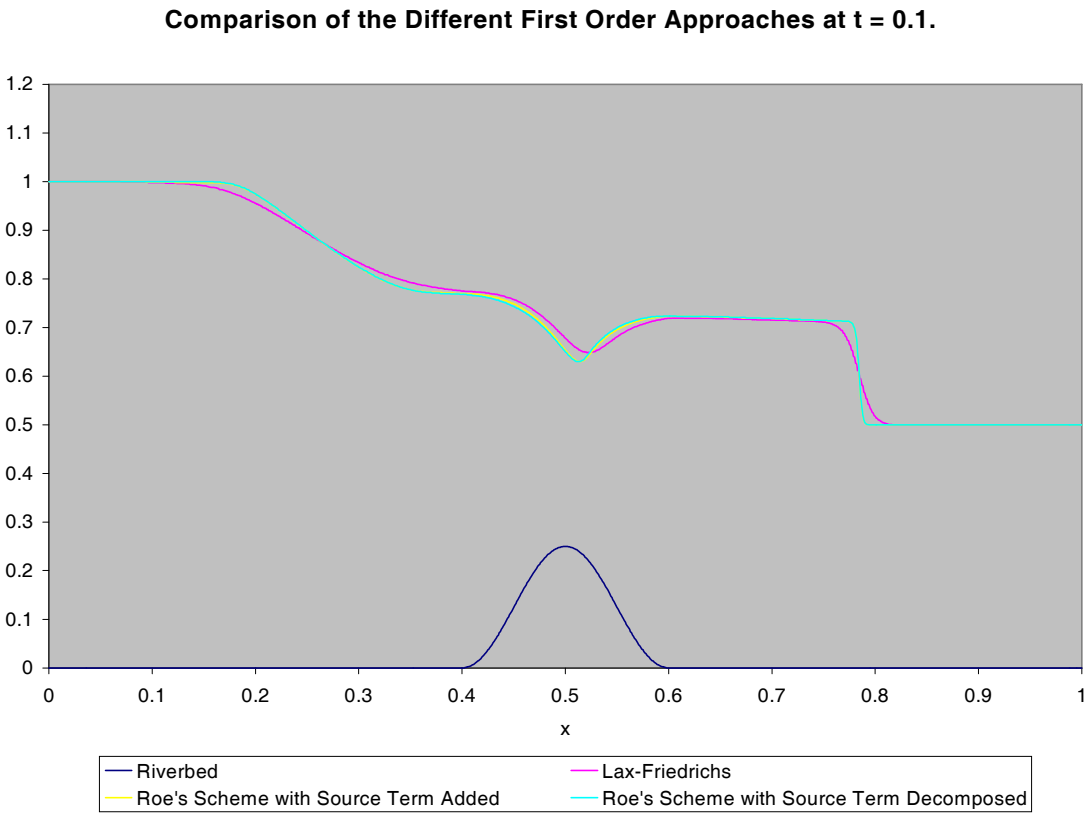

Figure 4-3a: Comparison of the different first order approaches for Problem B

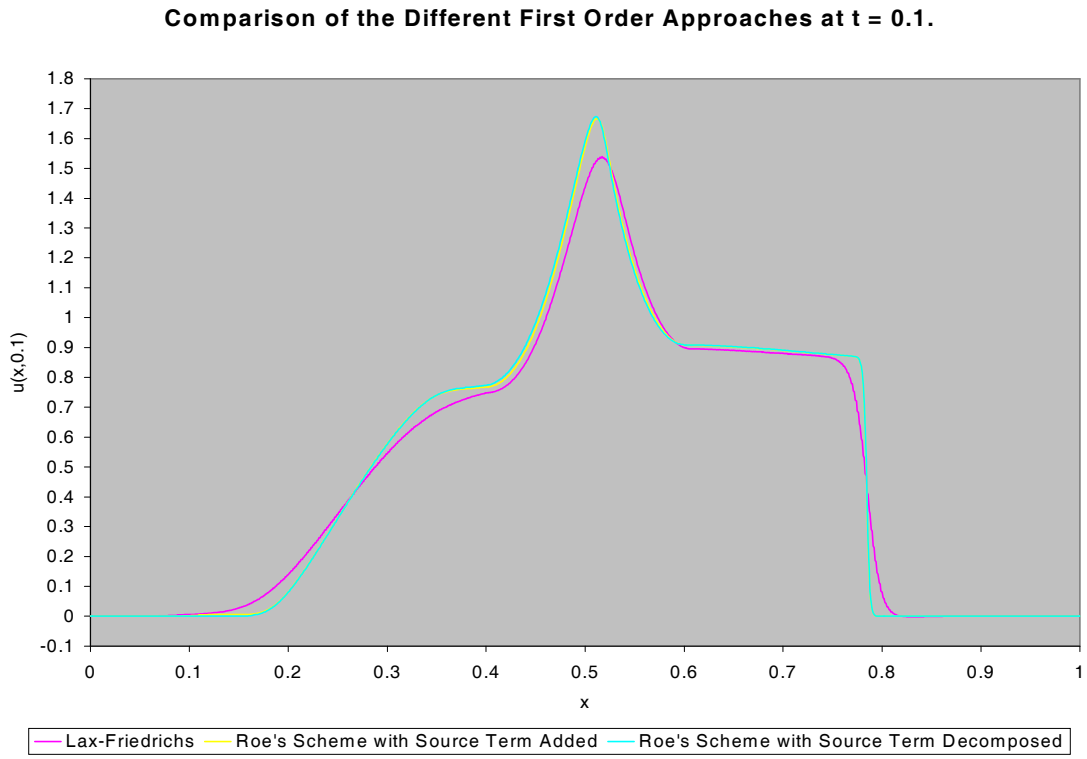

Figure 4-3b: Comparison of the different first order approaches for Problem B

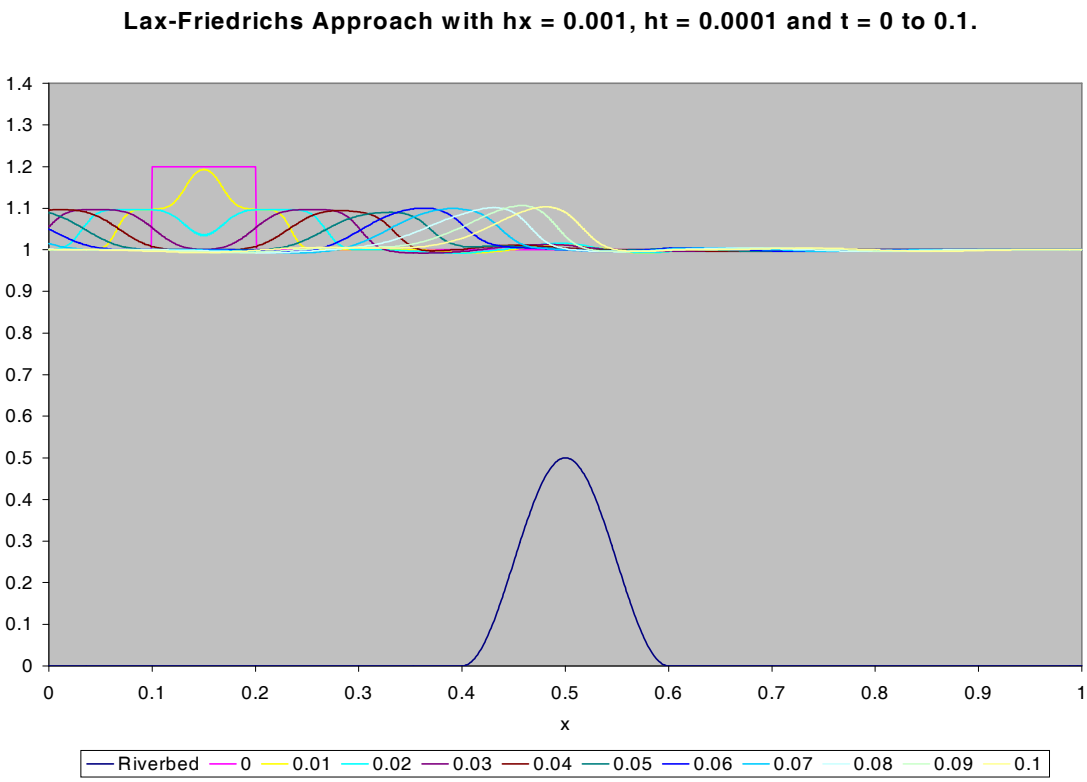

Figure 4-4a: Results of Lax-Friedrichs approach for Problem C

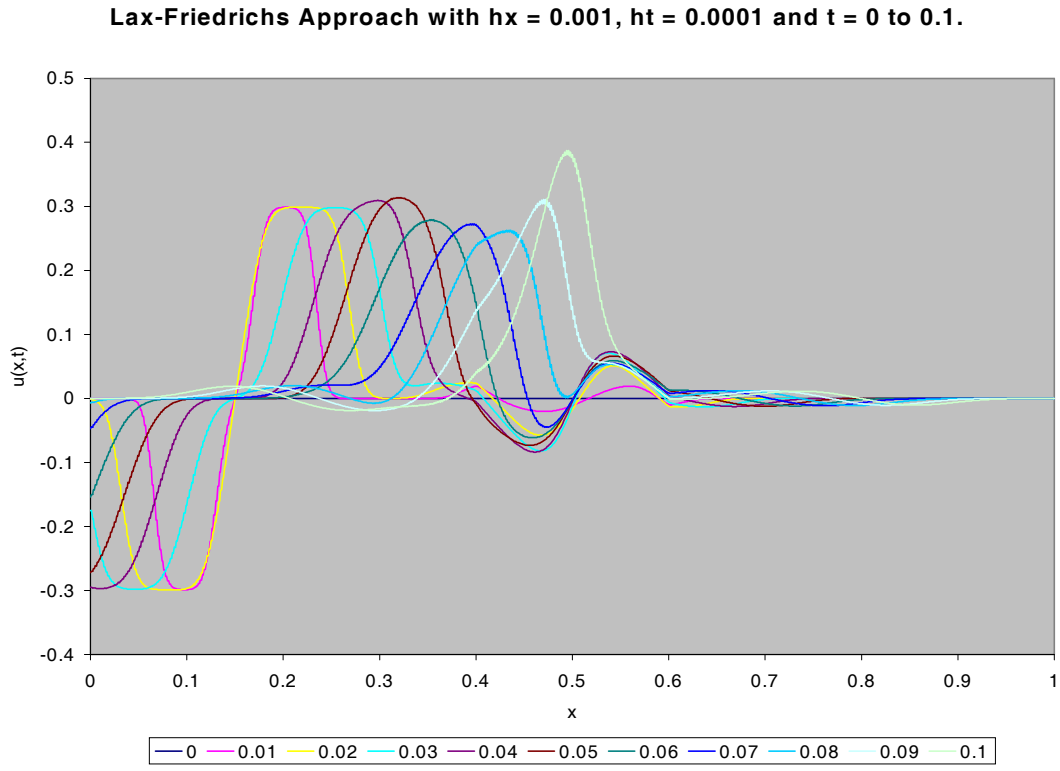

Figure 4-4b: Results of Lax-Friedrichs approach for Problem C

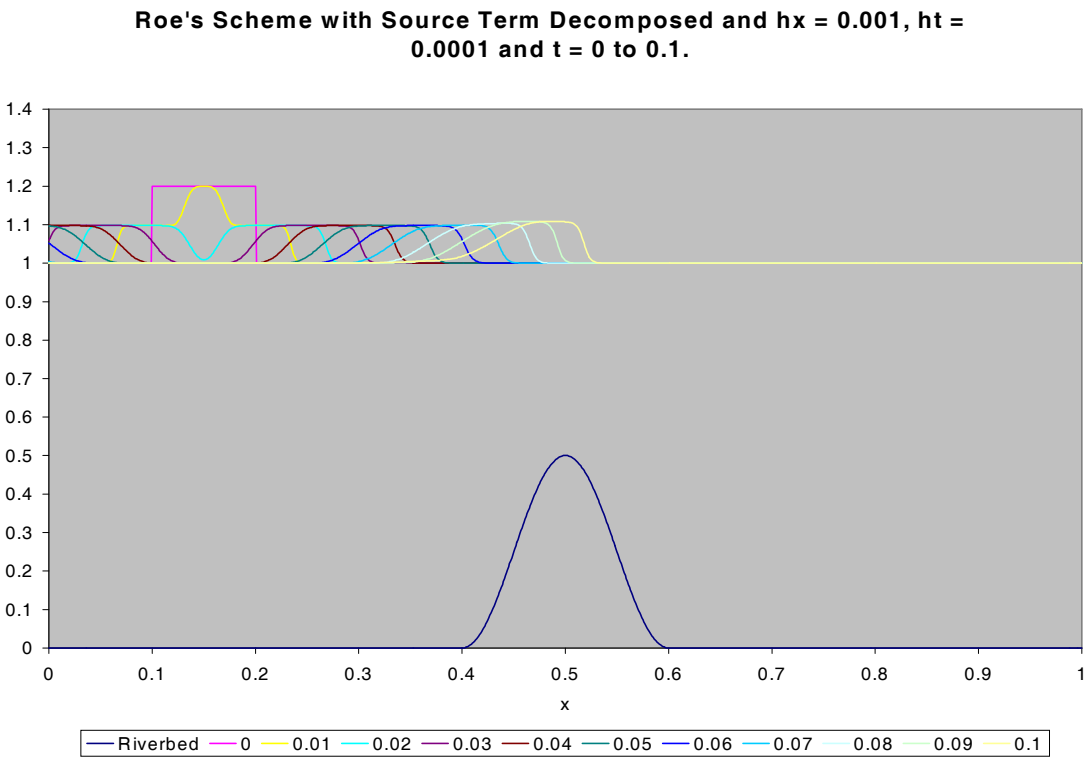

Figure 4-5a: Results of Roe's scheme with source term decomposed for Problem C

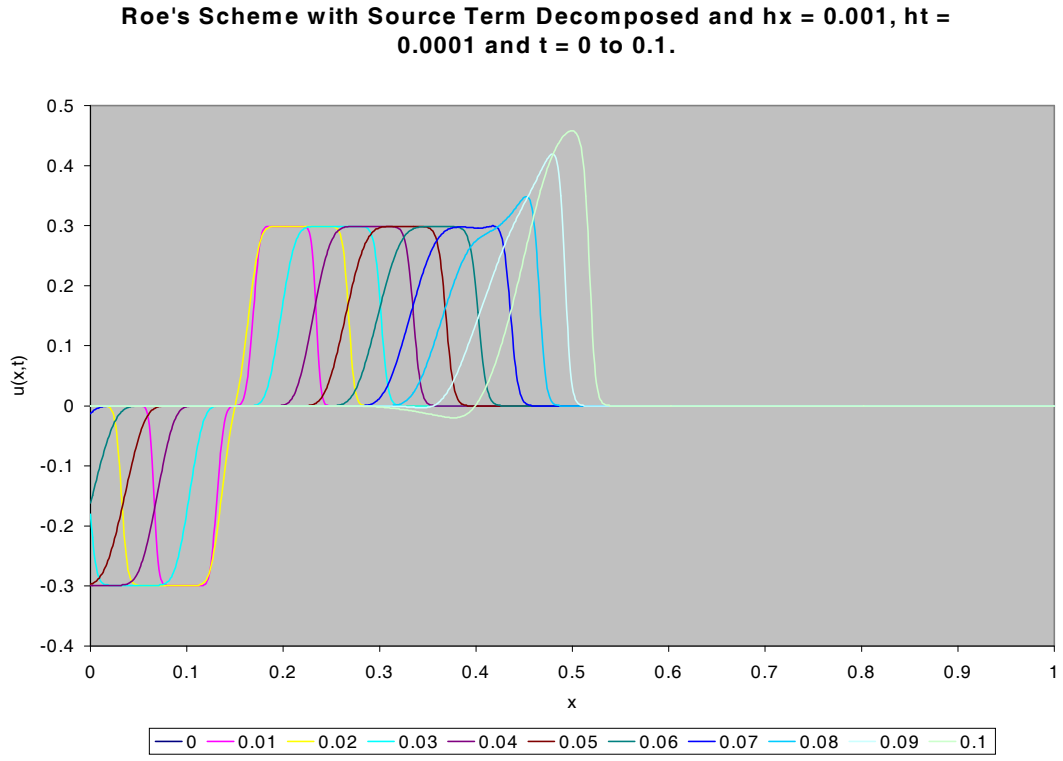

Figure 4-5b: Results of Roe's scheme with source term decomposed for Problem C

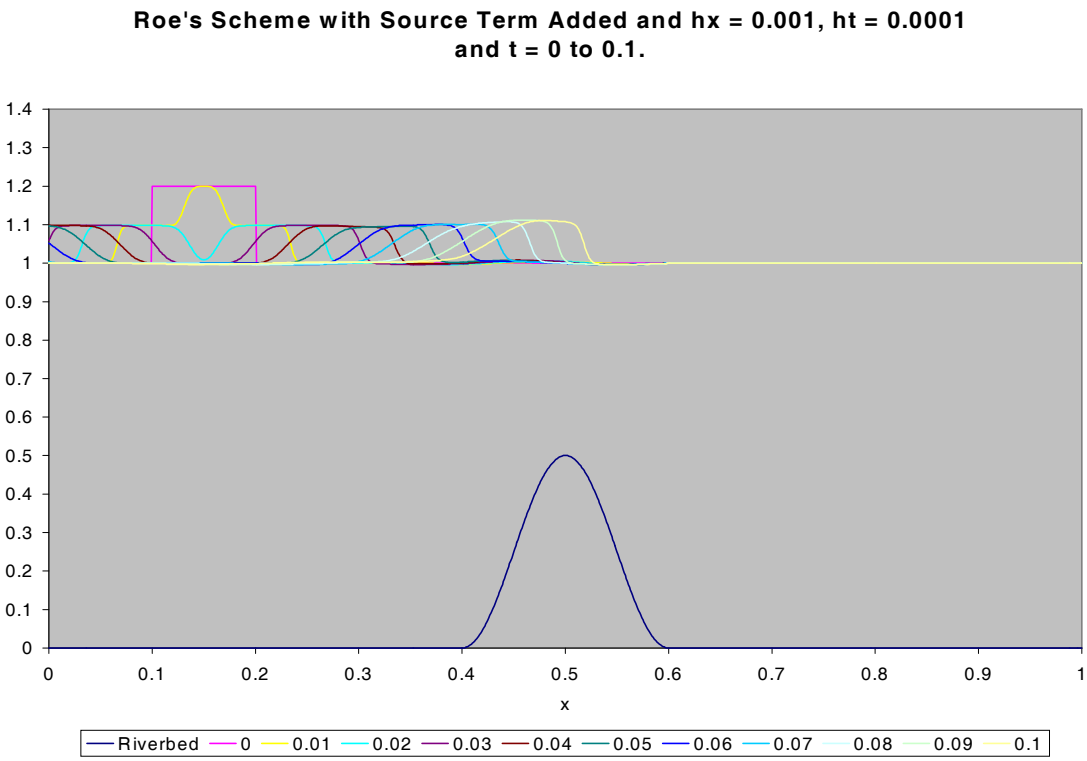

Figure 4-6a: Results of Roe's scheme with source term added for Problem C

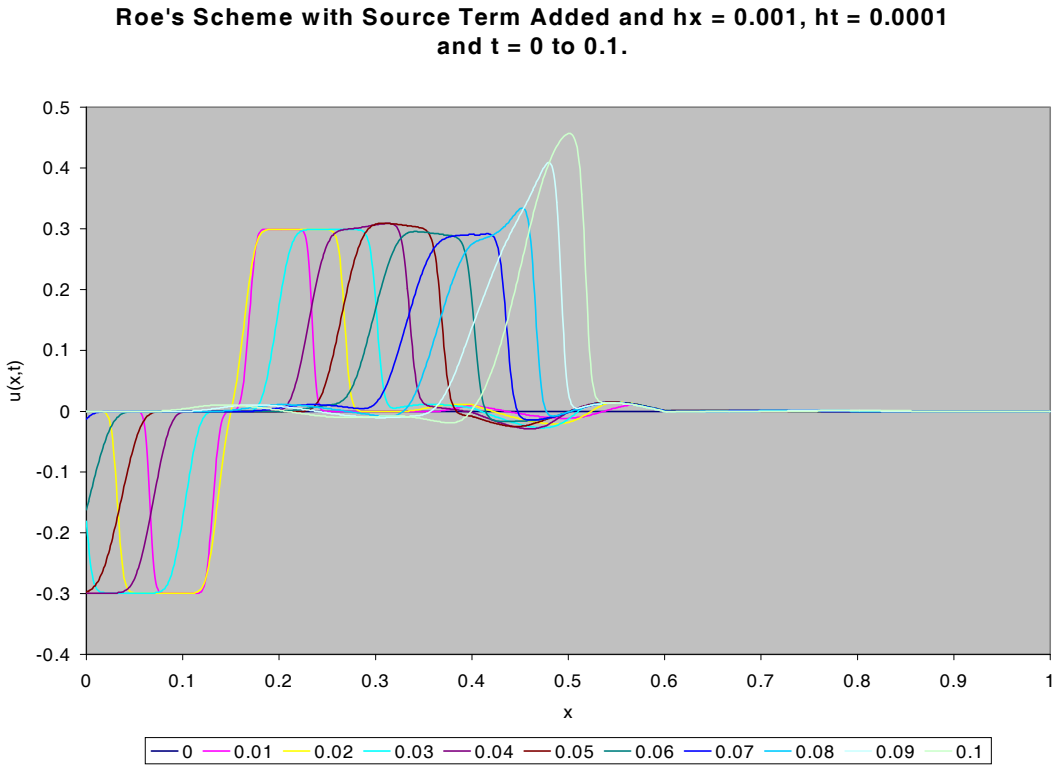

Figure 4-6b: Results of Roe's scheme with source term added for Problem C

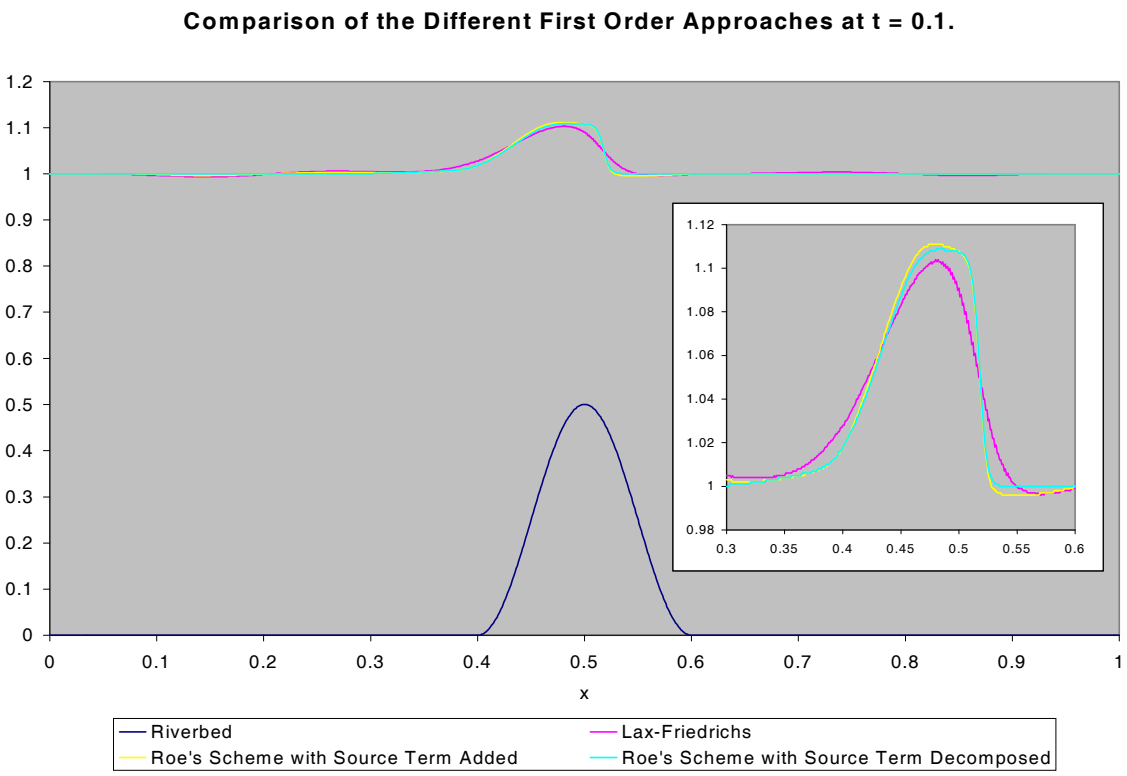

Figure 4-7a: Comparison of the different first order approaches for Problem C

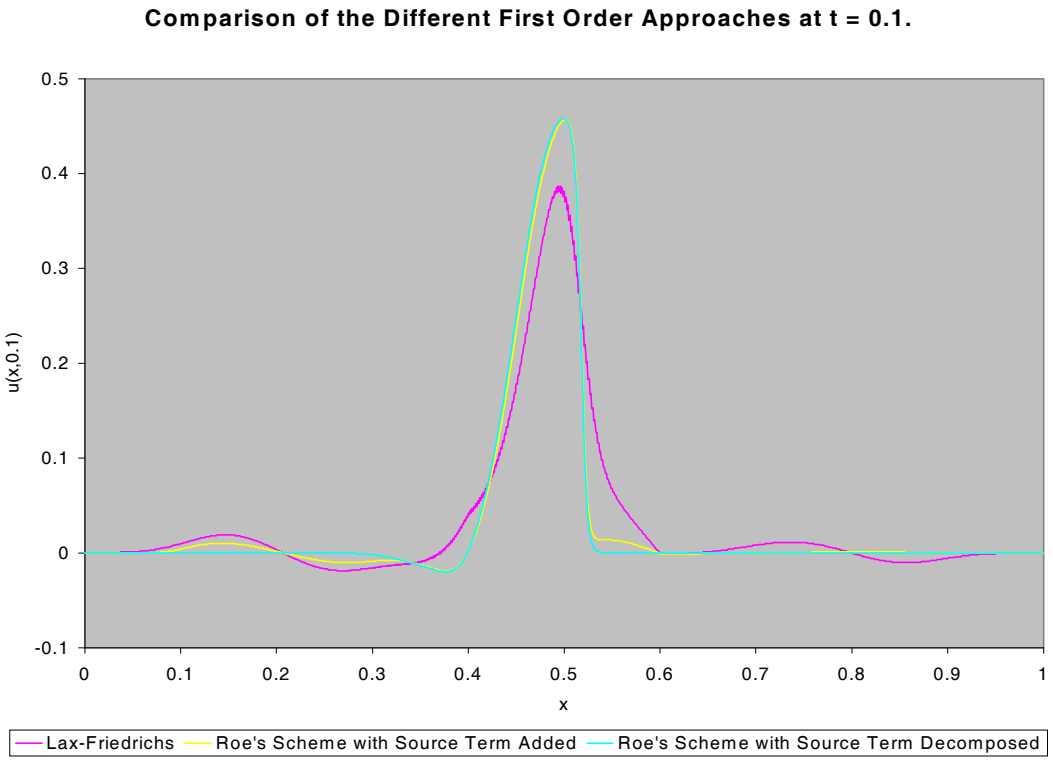

Figure 4-7b: Comparison of the different first order approaches for Problem C

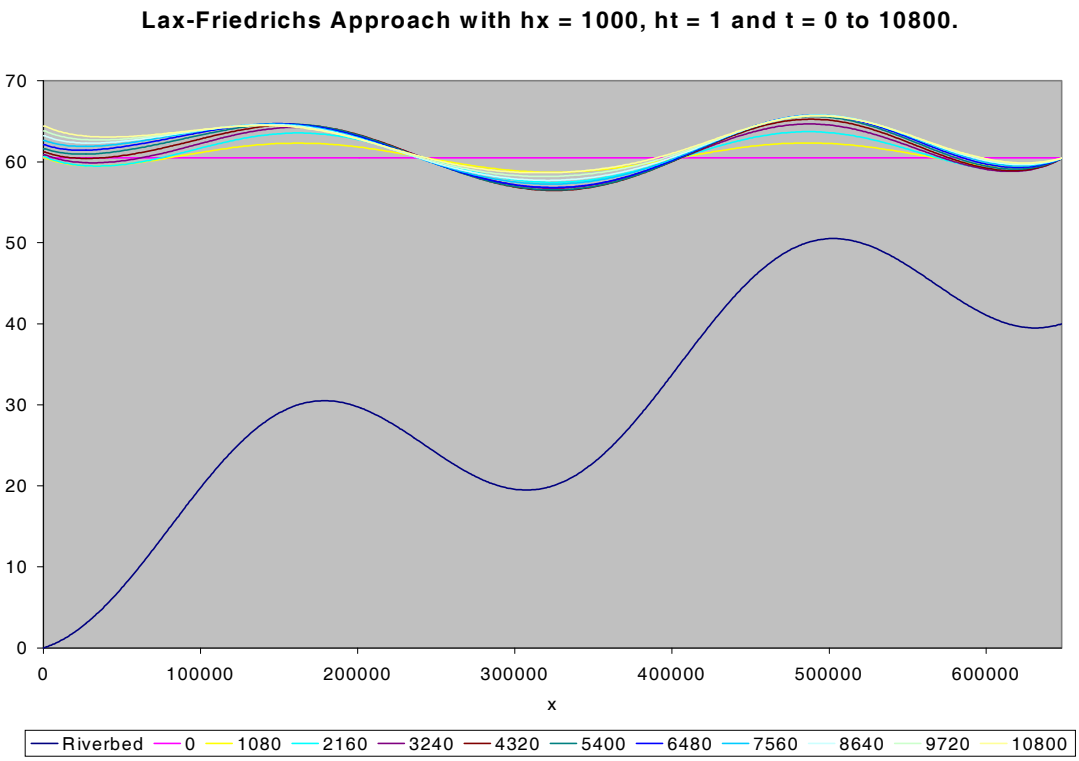

Figure 4-8a: Results of Lax-Friedrichs approach for Problem D

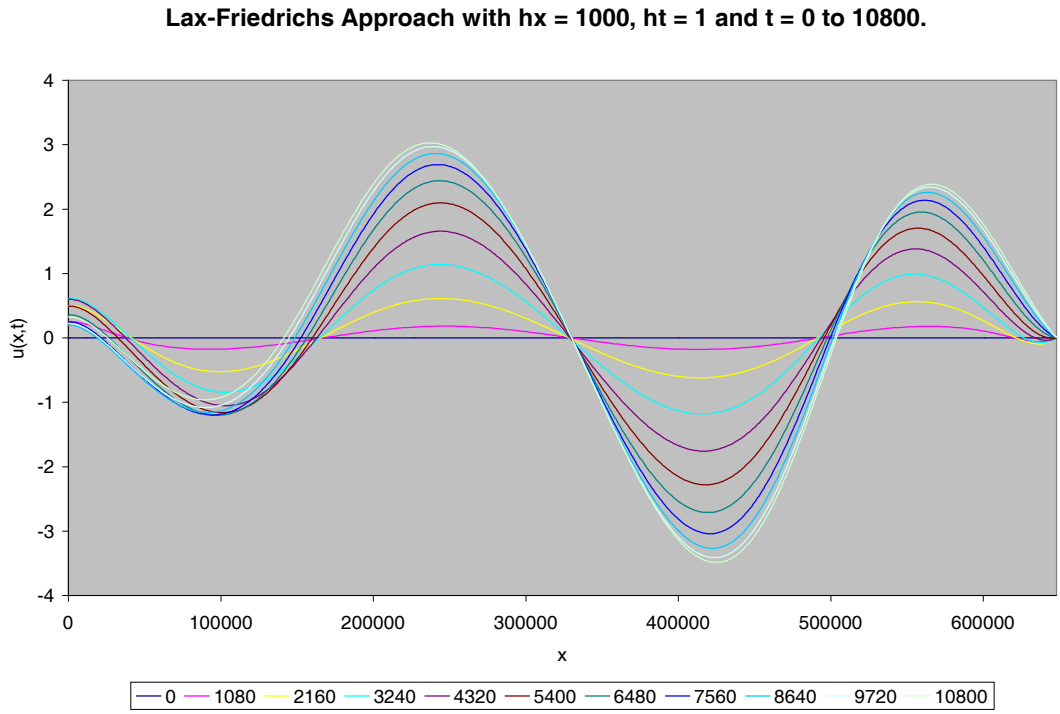

Figure 4-8b: Results of Lax-Friedrichs approach for Problem D

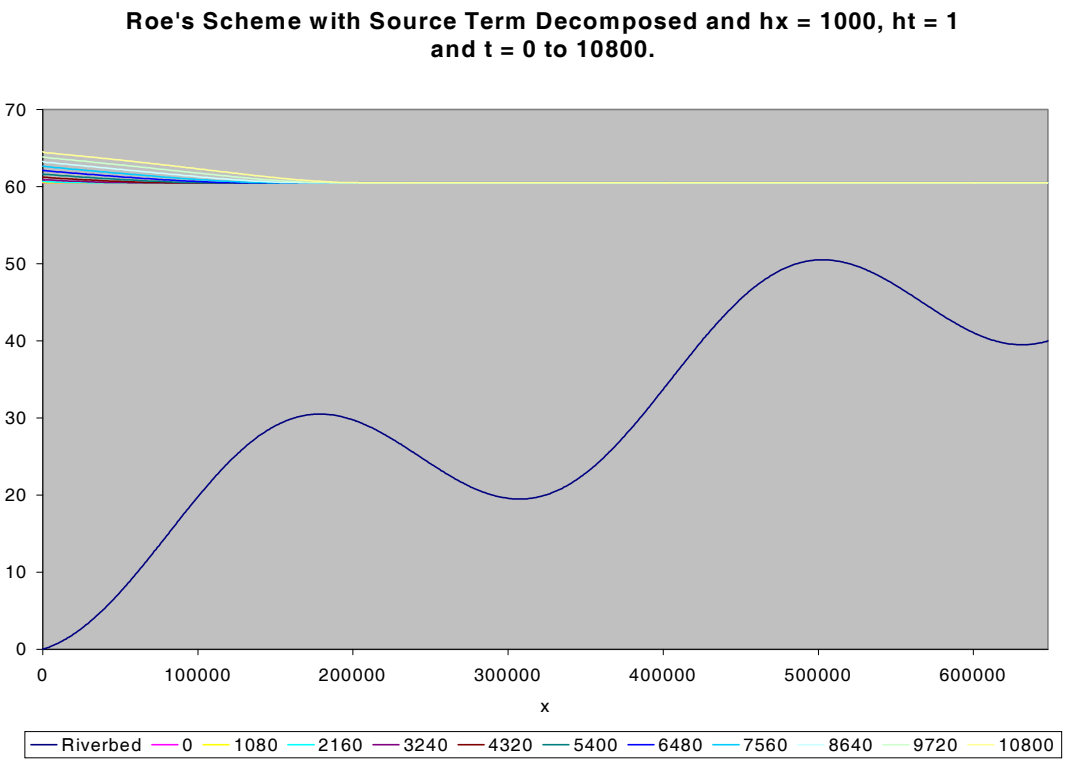

Figure 4-9a: Results of Roe's scheme with source term decomposed for Problem D

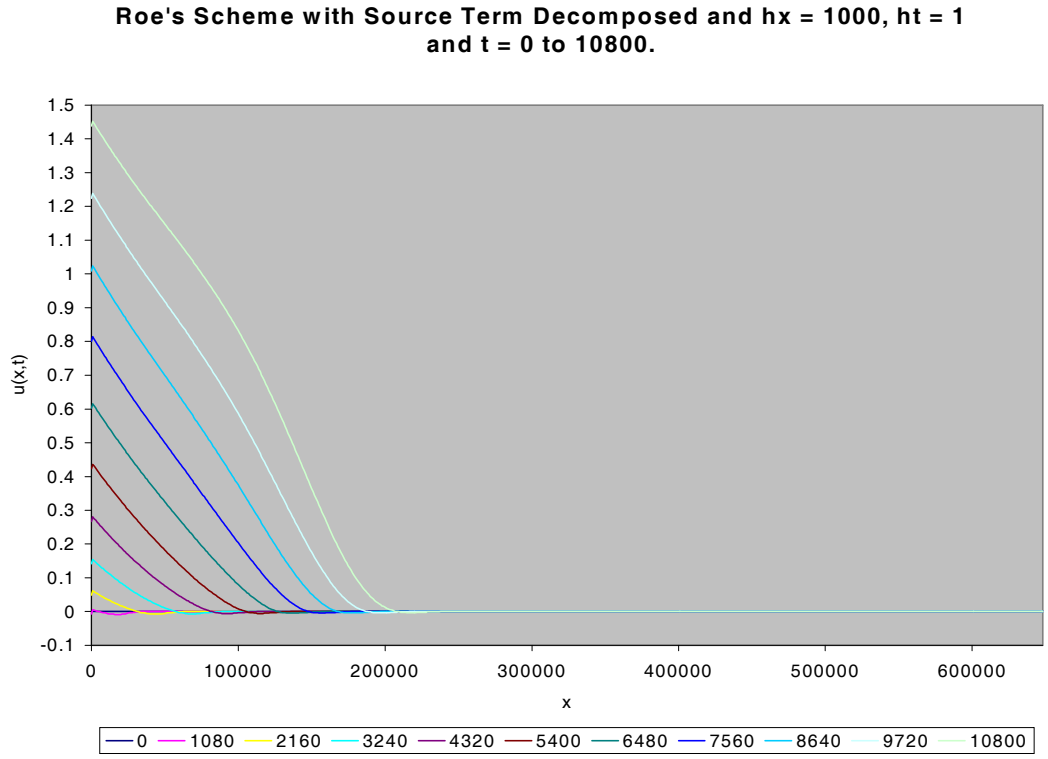

Figure 4-9b:Results of Roe's scheme with source term decomposed for Problem D

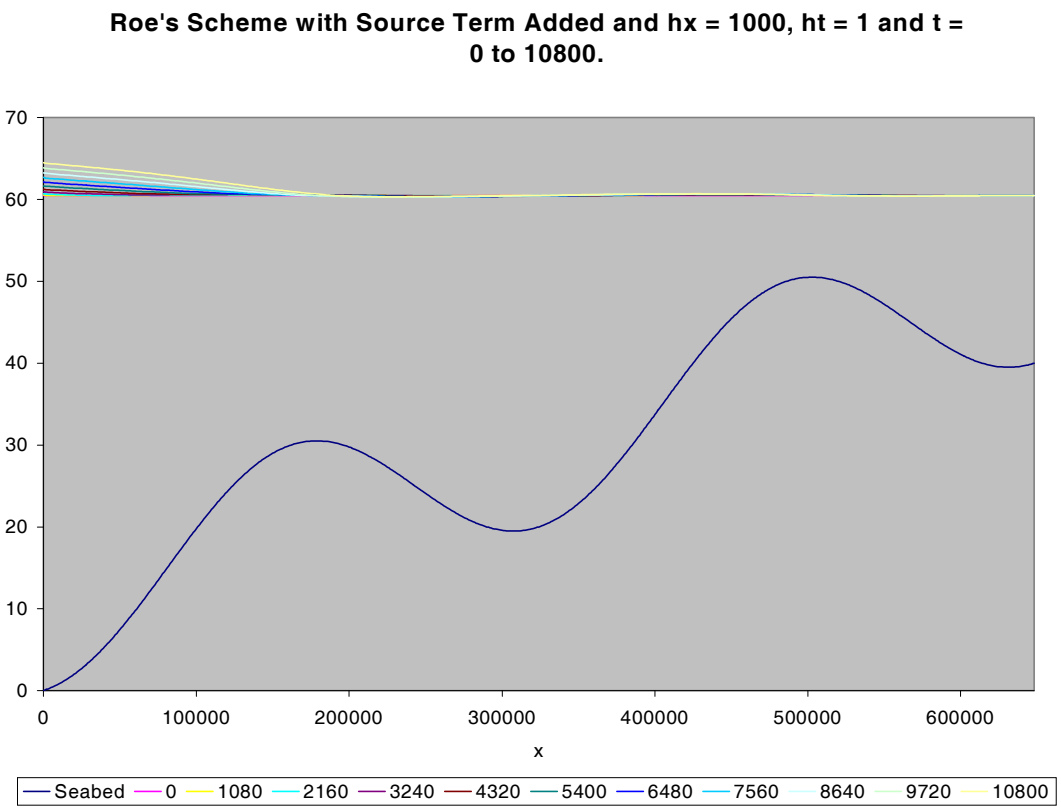

Figure 4-10a: Results of Roe's scheme with source term added for Problem D

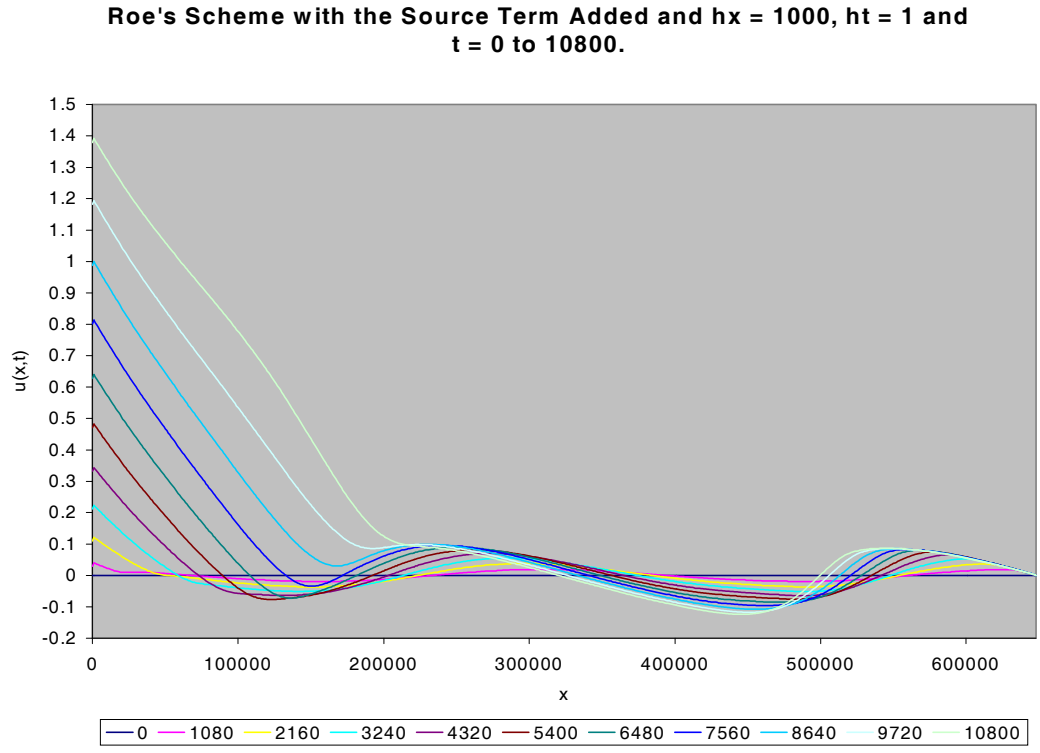

Figure 4-10b: Results of Roe's scheme with source term added for Problem D

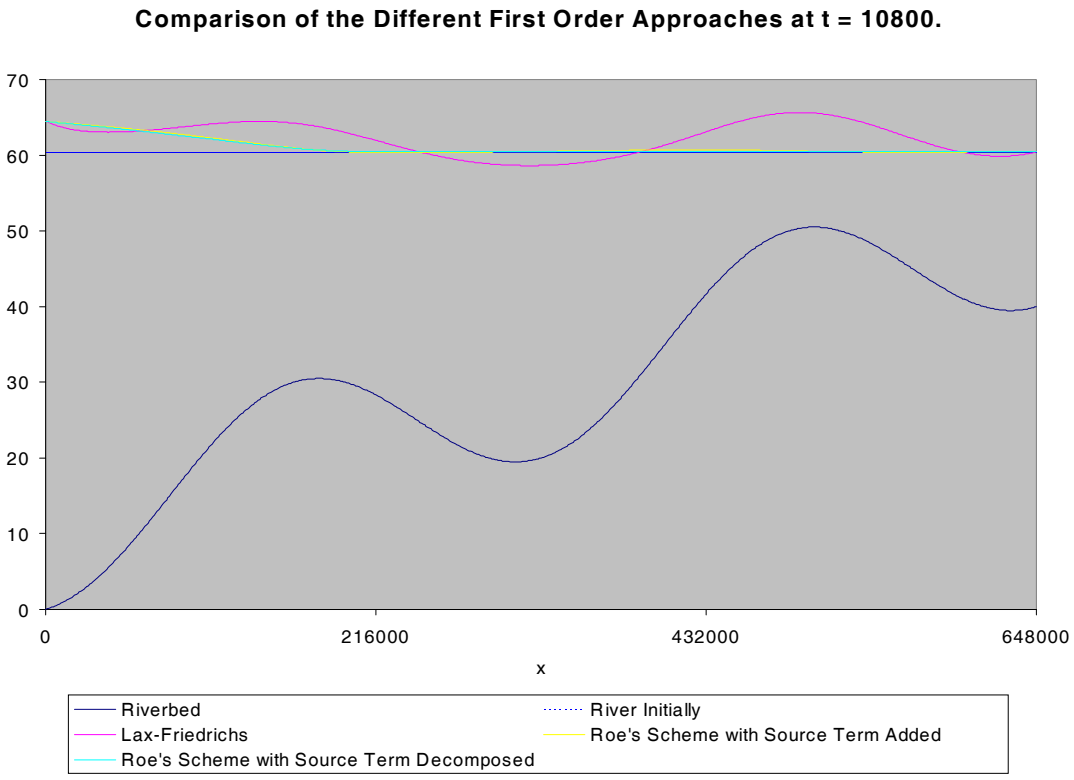

Figure 4-11a: Comparison of the different first order approaches for Problem D

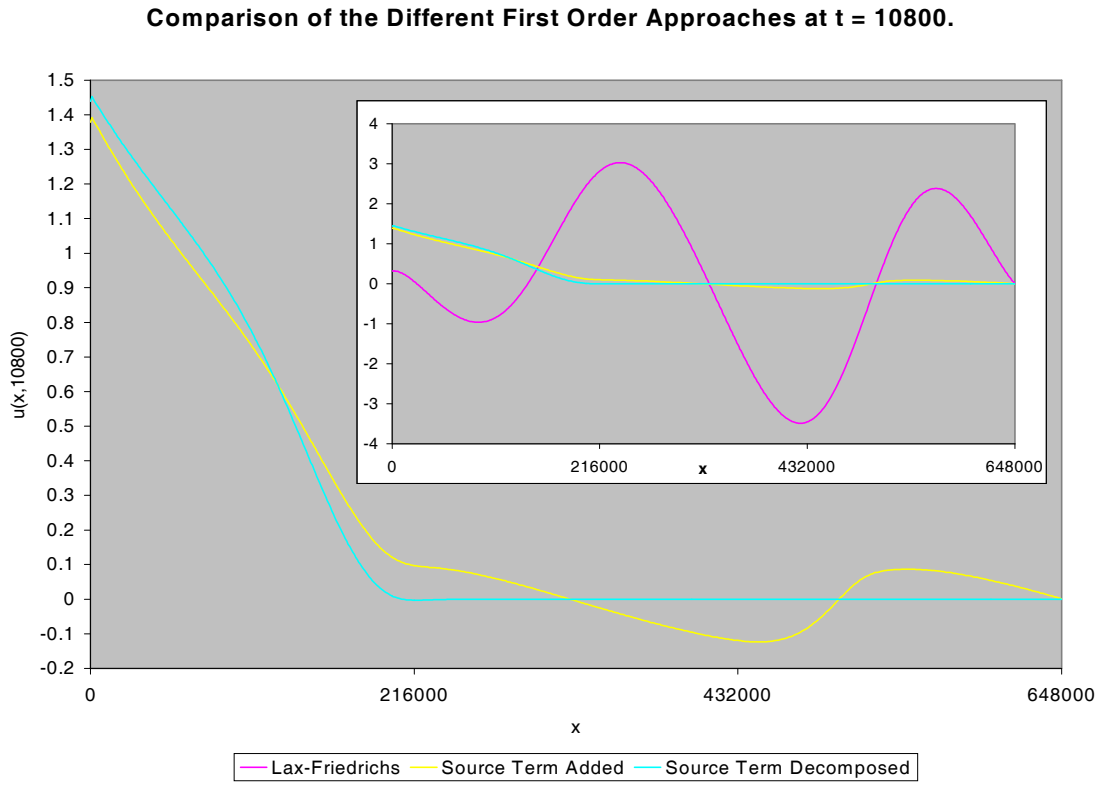

Figure 4-11b: Comparison of the different first order approaches for Problem D

Flux-Limited Second Order Numerical Results:

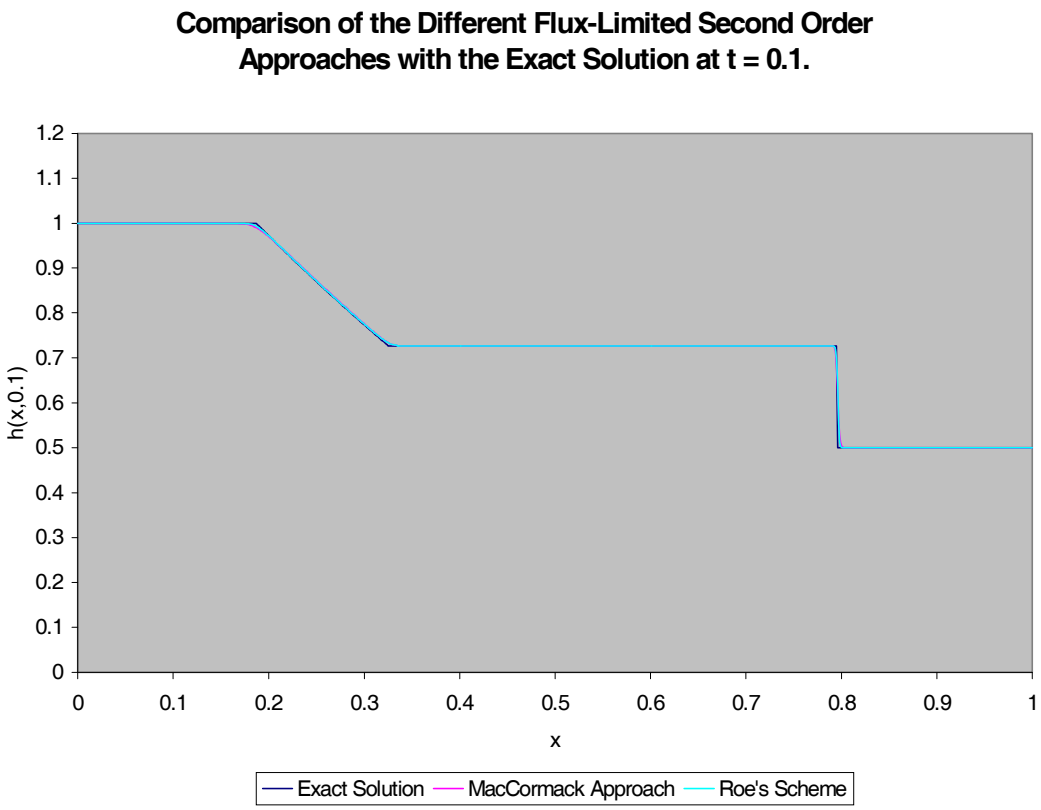

Figure 4-12a: Comparison of the flux-limited second order results for Problem A

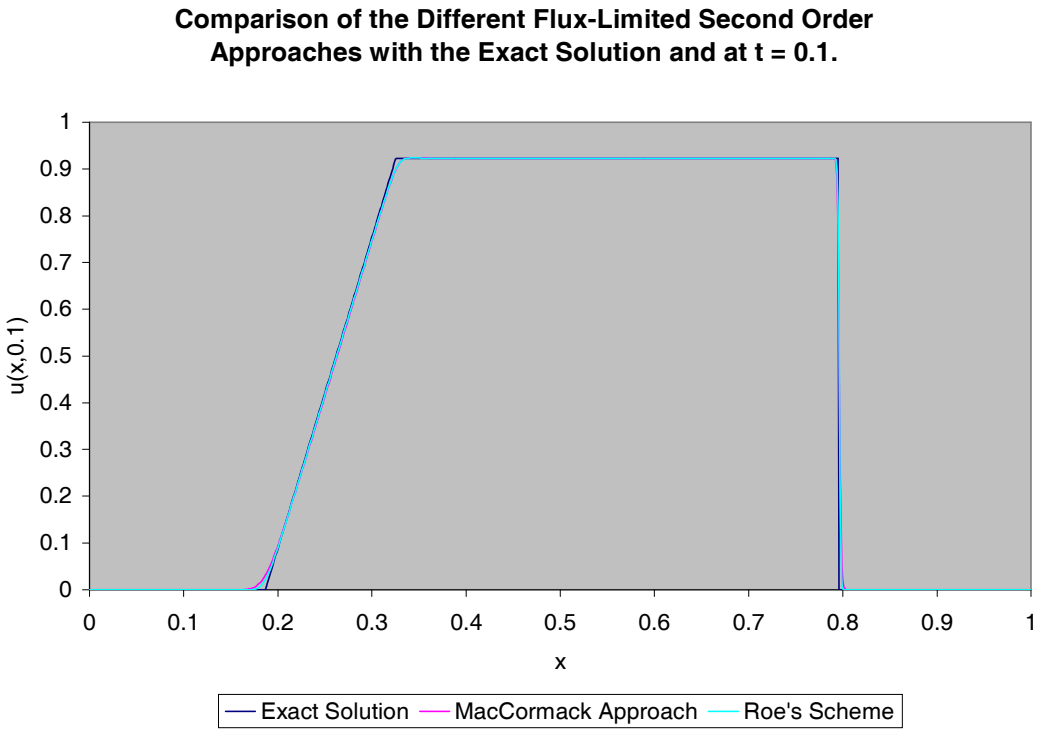

Figure 4-12b: Comparison of the flux-limited second order results for Problem A

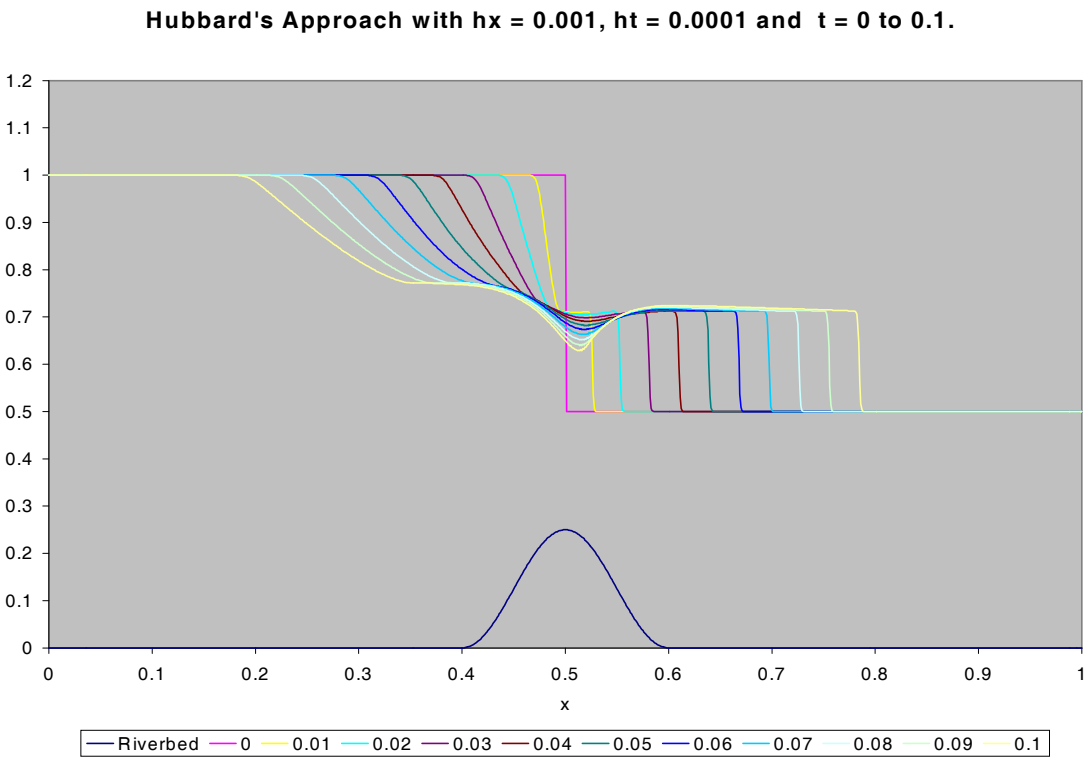

Figure 4-13a: Results of Hubbard's approach for Problem B

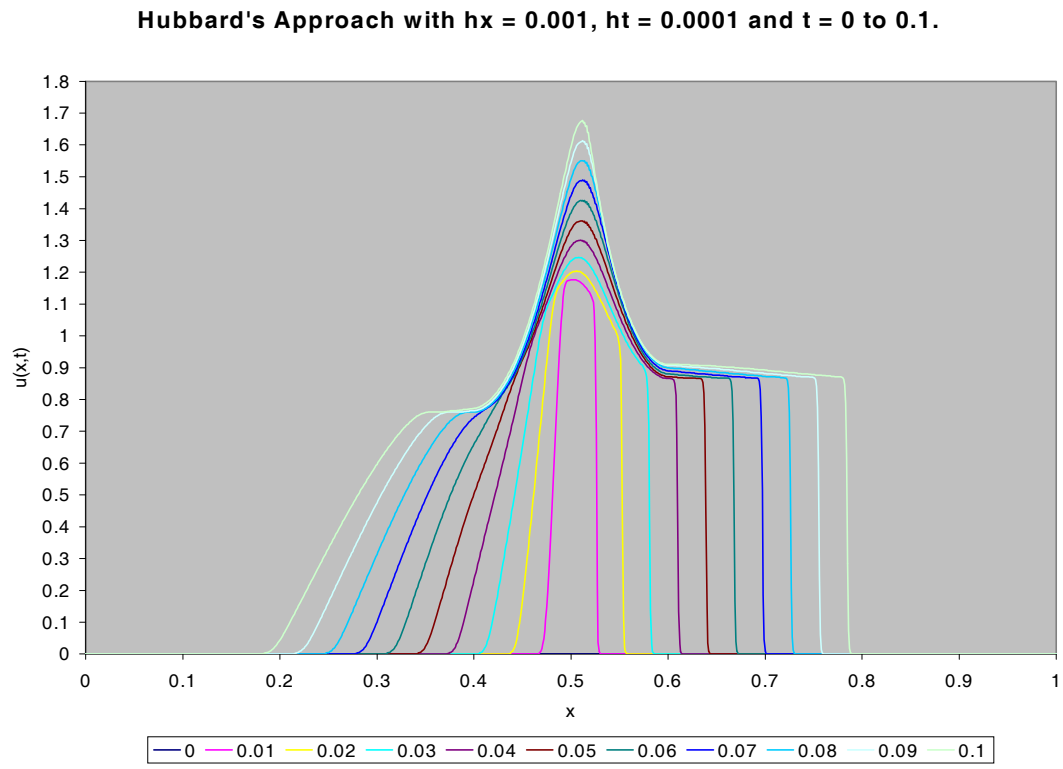

Figure 4-13b: Results of Hubbard's approach for Problem B

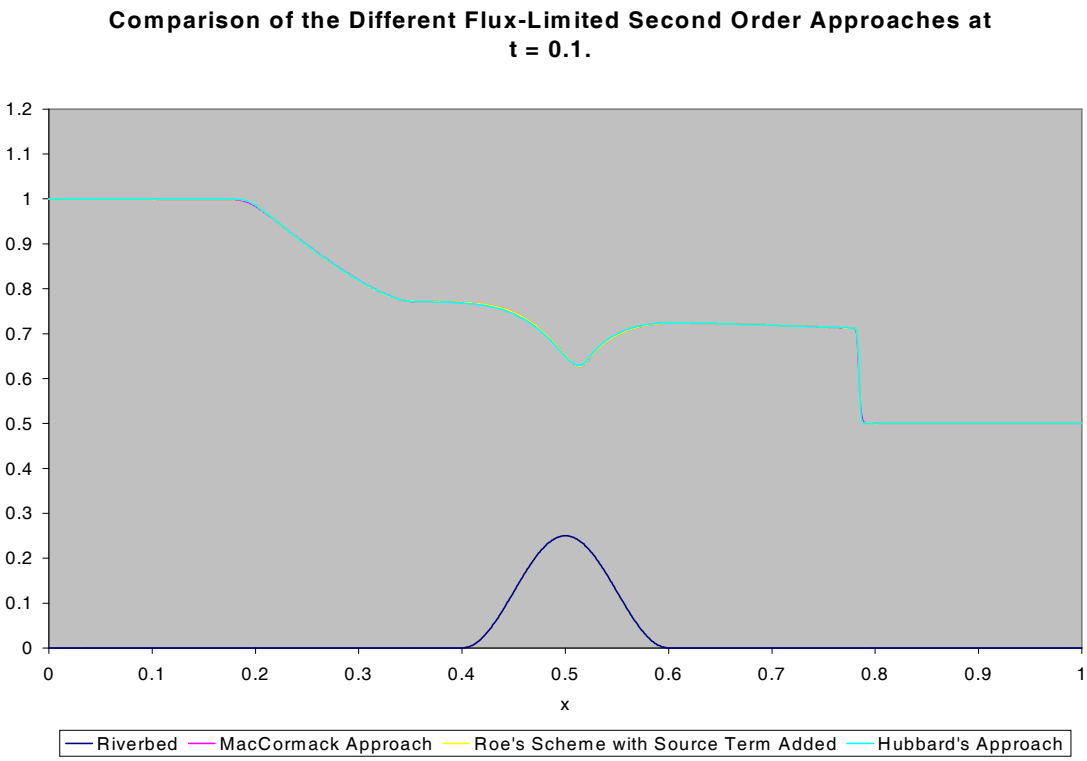

Figure 4-14a: Comparison of the flux-limited second order results for Problem B

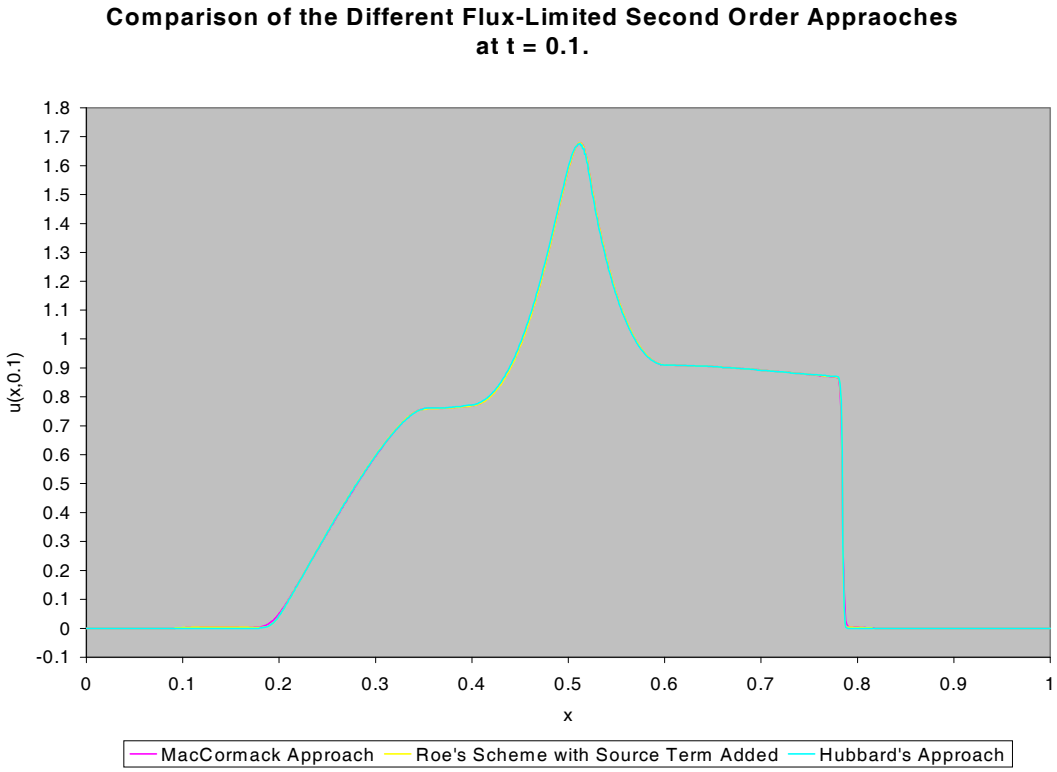

Figure 4-14b: Comparison of the flux-limited second order results for Problem B

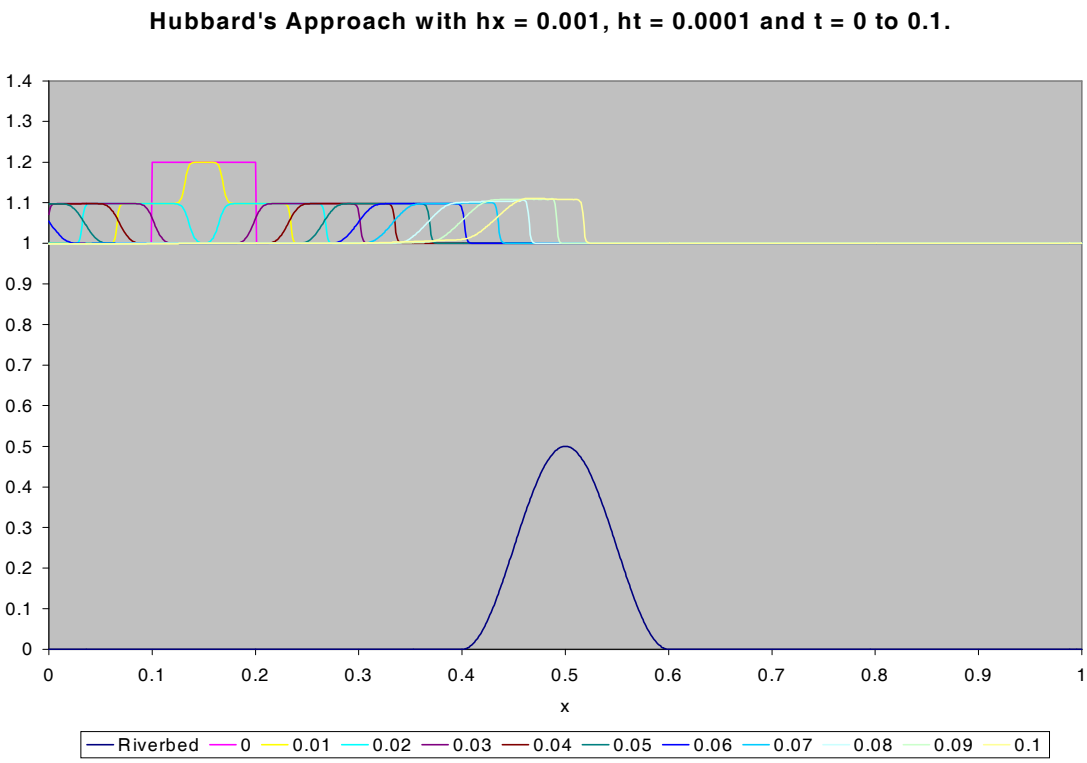

Figure 4-15a: Results of Hubbard's approach for Problem C

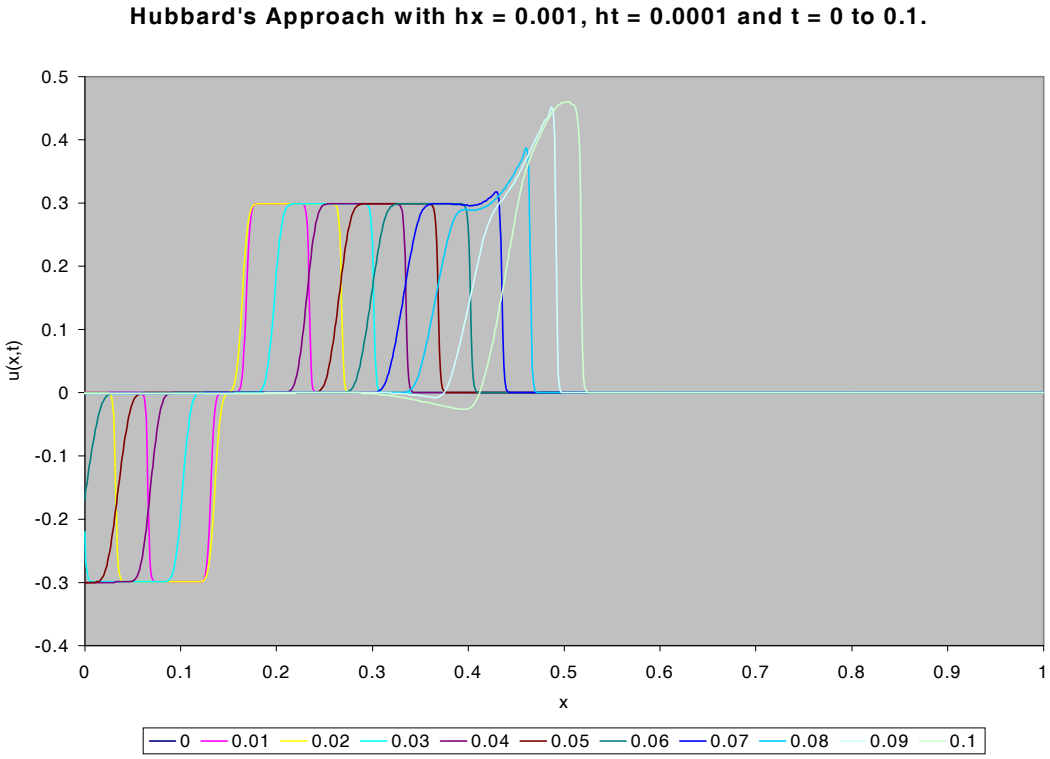

Figure 4-15b: Results of Hubbard's approach for Problem C

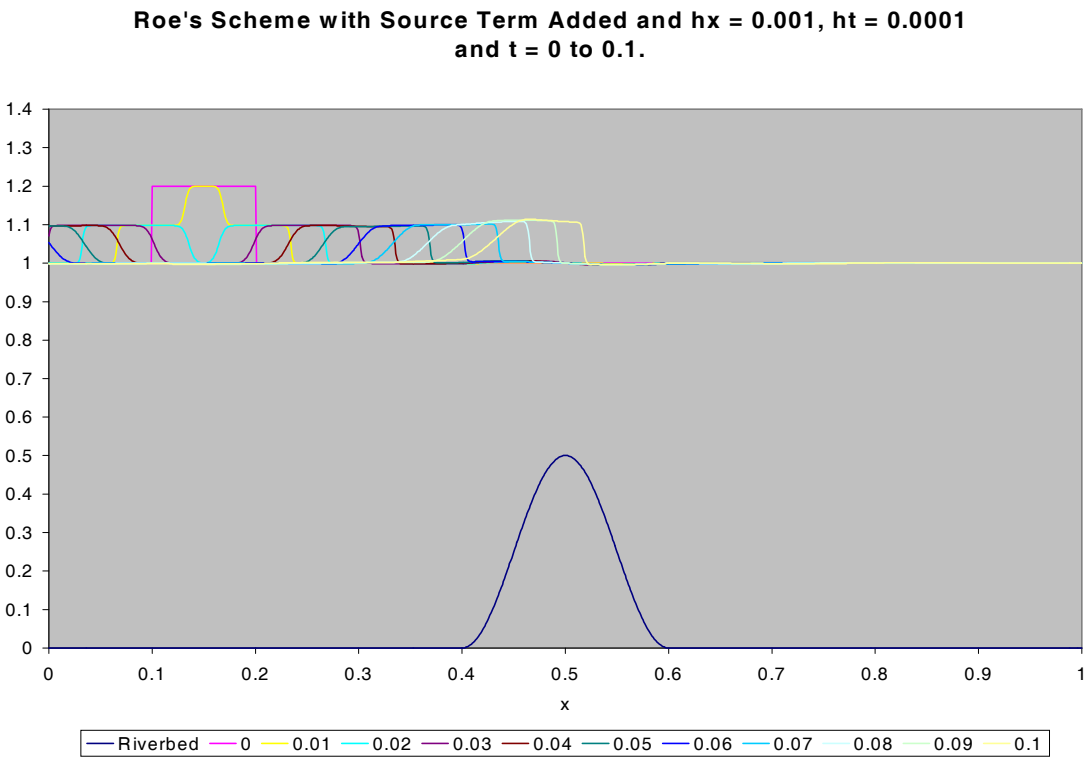

Figure 4-16a: Results of Roe's scheme with source term added for Problem C

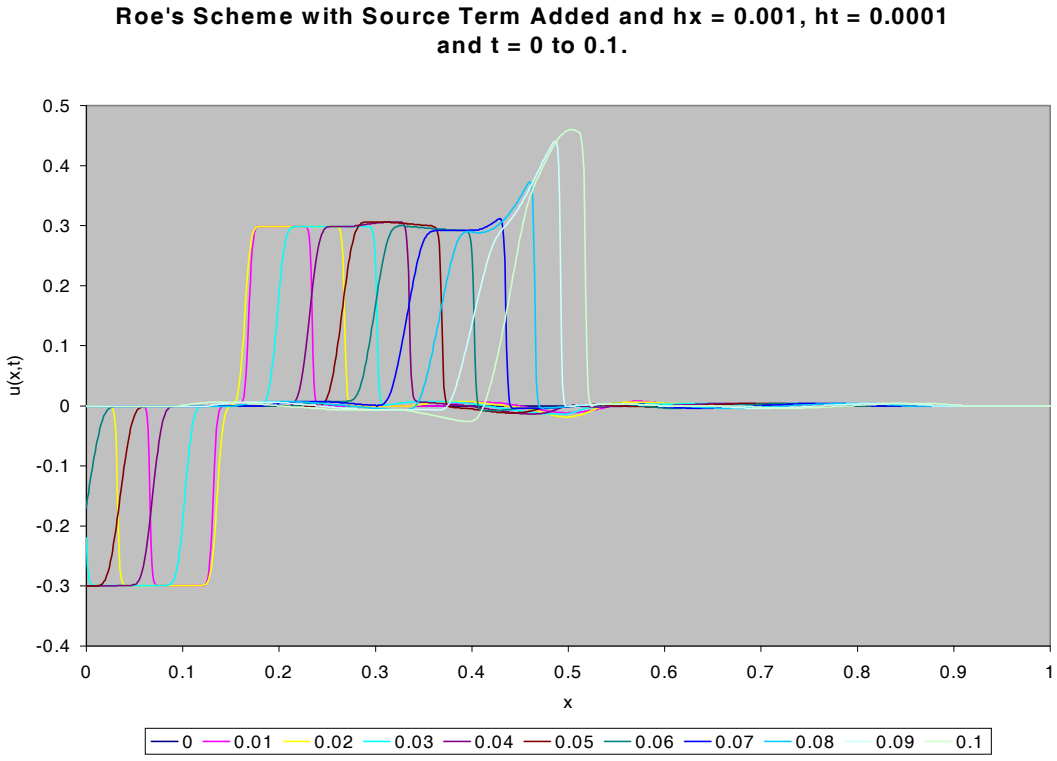

Figure 4-16b: Results of Roe's scheme with source term added for Problem C

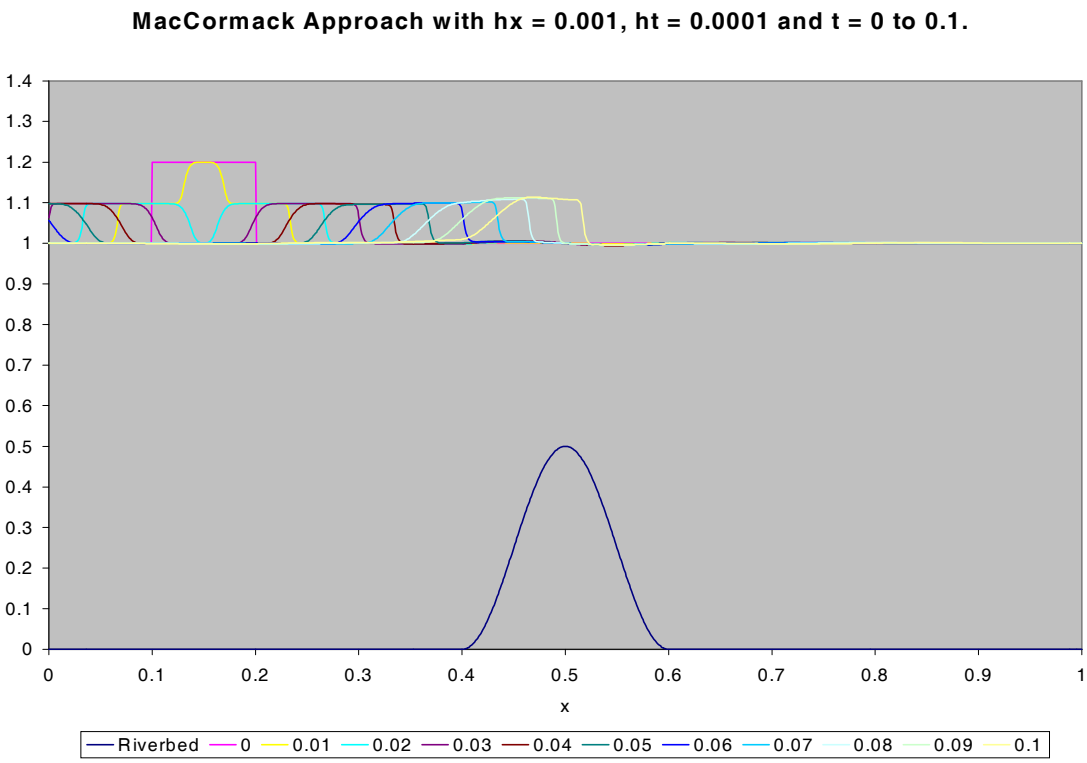

Figure 4-17a: Results of LeVeque & Yee's MacCormack approach for Problem C

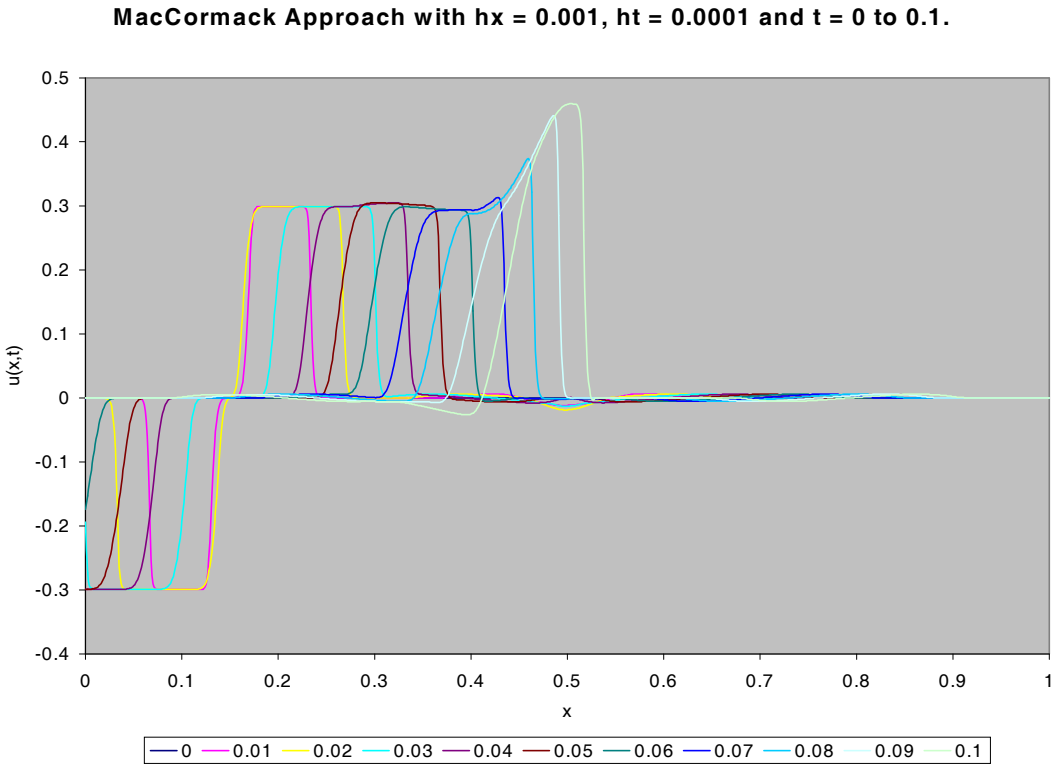

Figure 4-17b: Results of LeVeque & Yee's MacCormack approach for Problem C

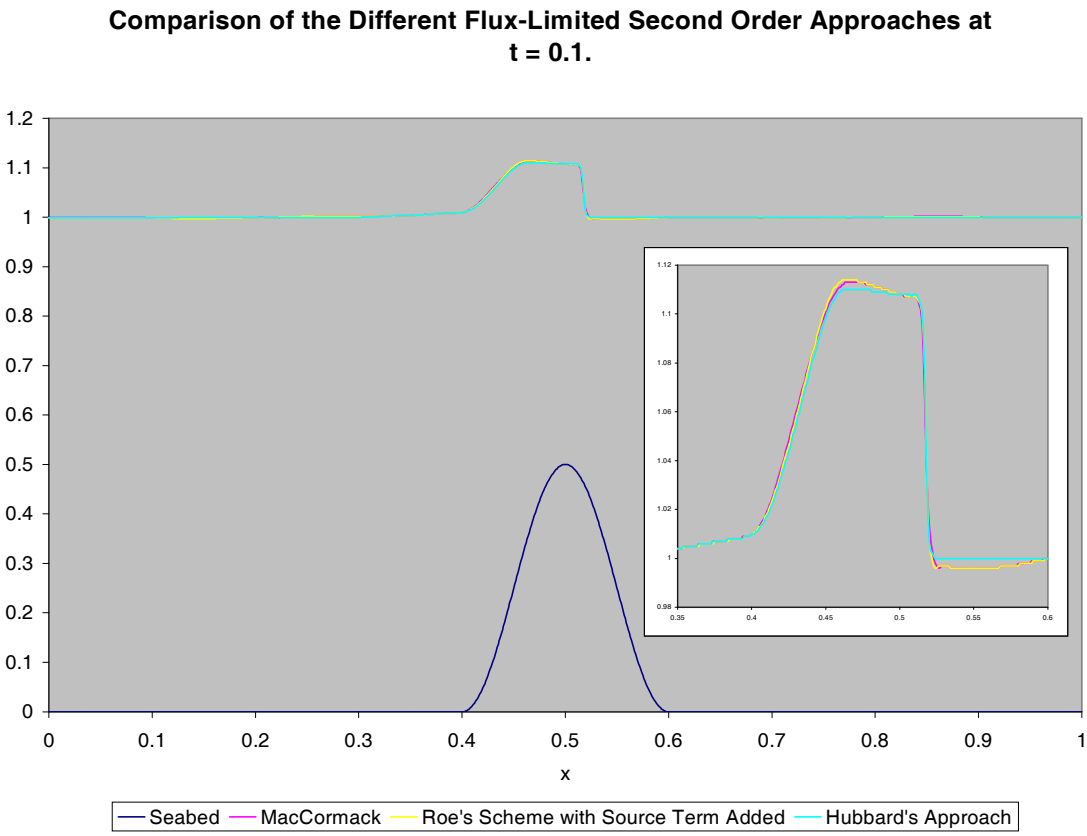

Figure 4-18a: Comparison of the flux-limited second order results for Problem C

Figure 4-18b: Comparison of the flux-limited second order results for Problem C

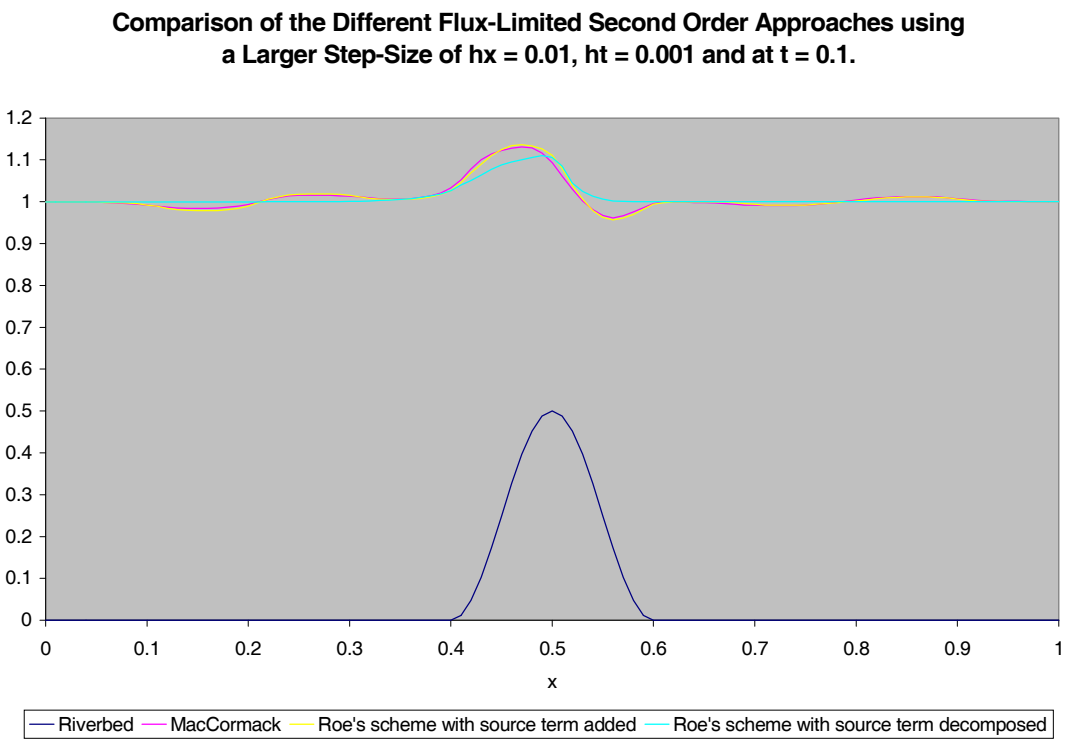

Figure 4-19a: Comparison of the flux-limited second order results for Problem C

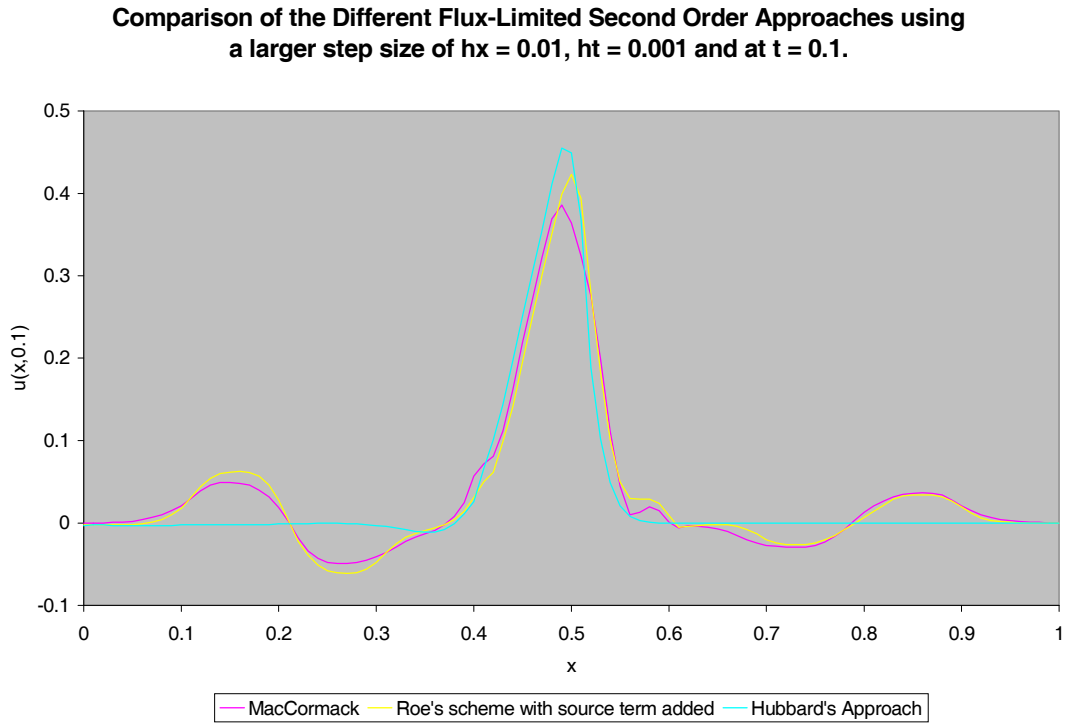

Figure 4-19b: Comparison of the flux-limited second order results for Problem C

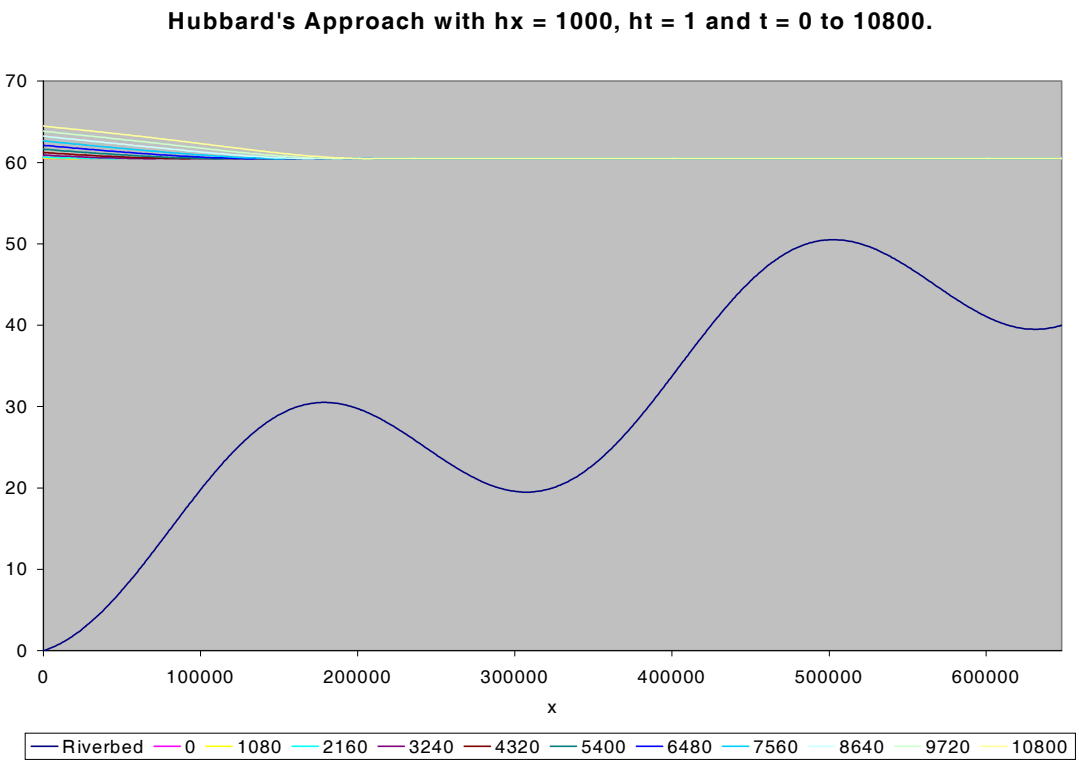

Figure 4-20a: Results of Hubbard's approach for Problem D

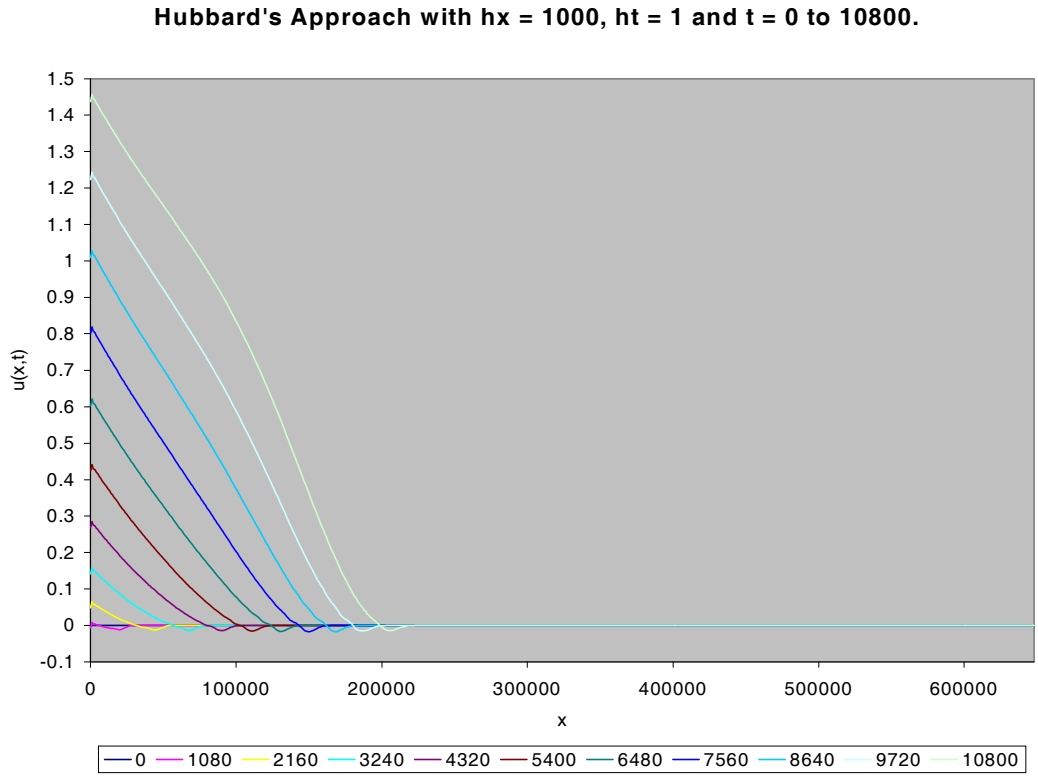

Figure 4-20b: Results of Hubbard's approach for Problem D

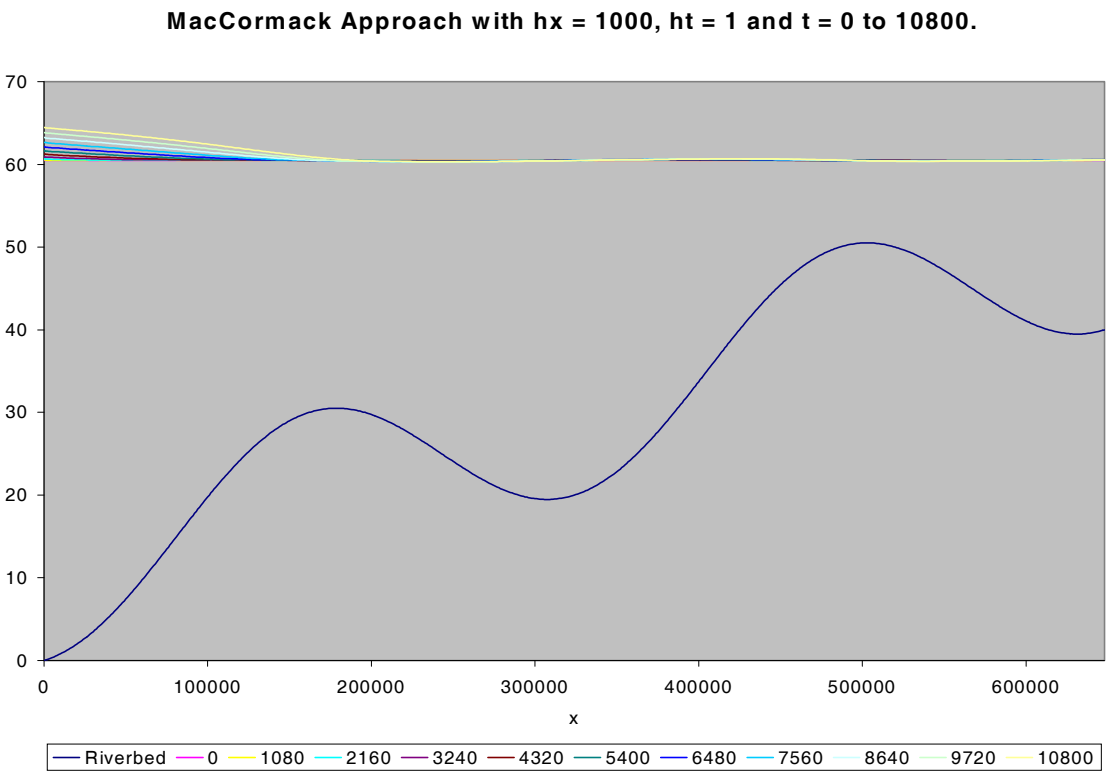

Figure 4-21a: Results of LeVeque & Yee's MacCormack approach for Problem D

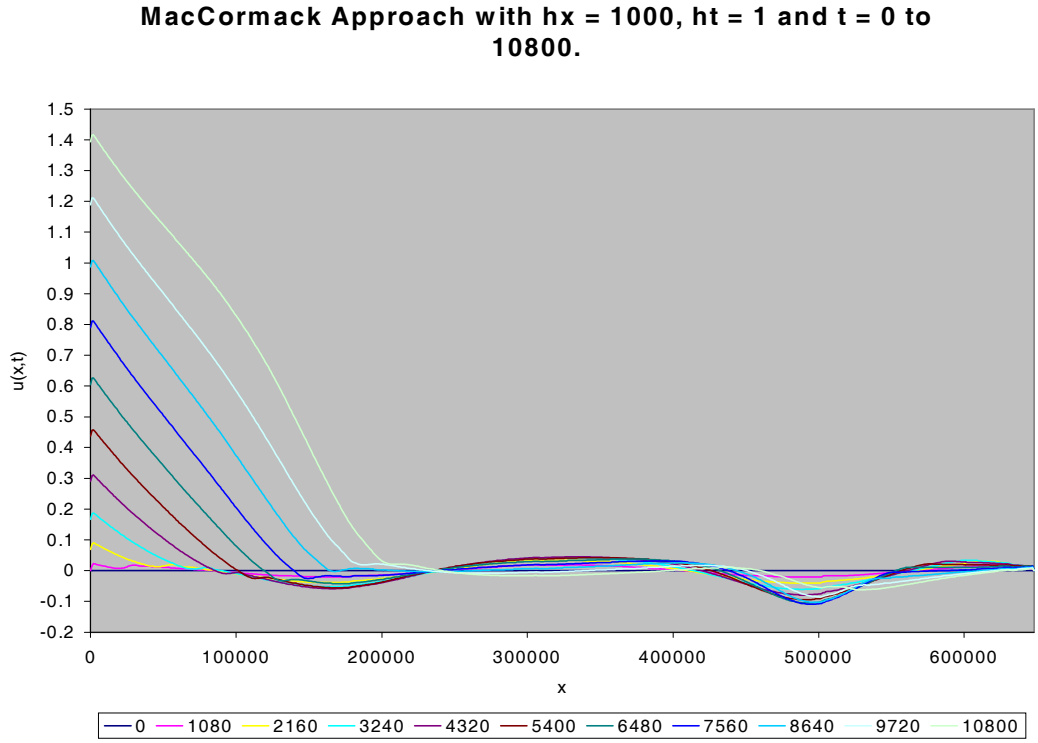

Figure 4-21b: Results of LeVeque & Yee's MacCormack approach for Problem D

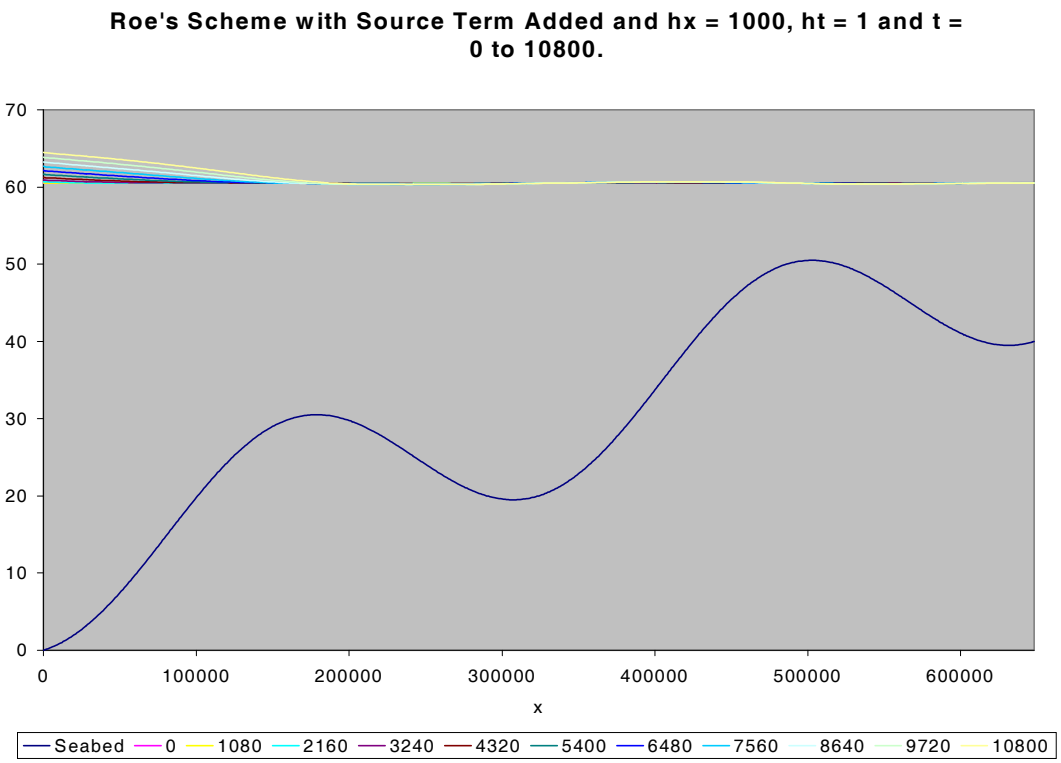

Figure 4-22a: Results of Roe's scheme with source term added for Problem D

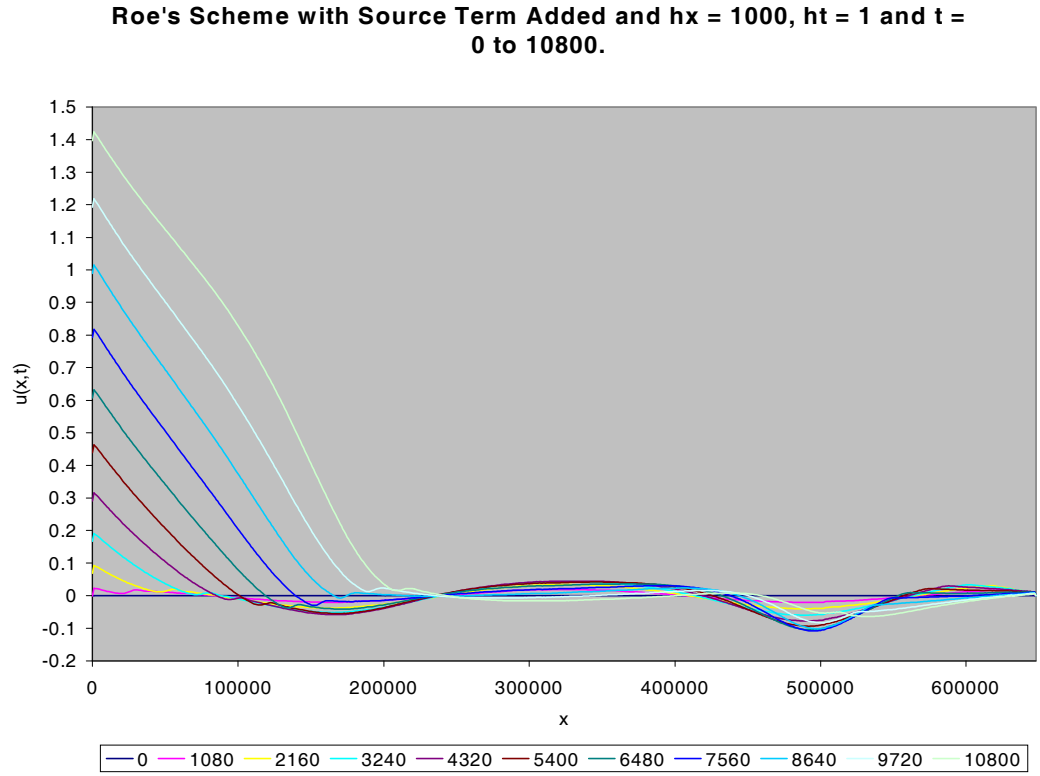

Figure 4-22b: Results of Roe's scheme with source term added for Problem D

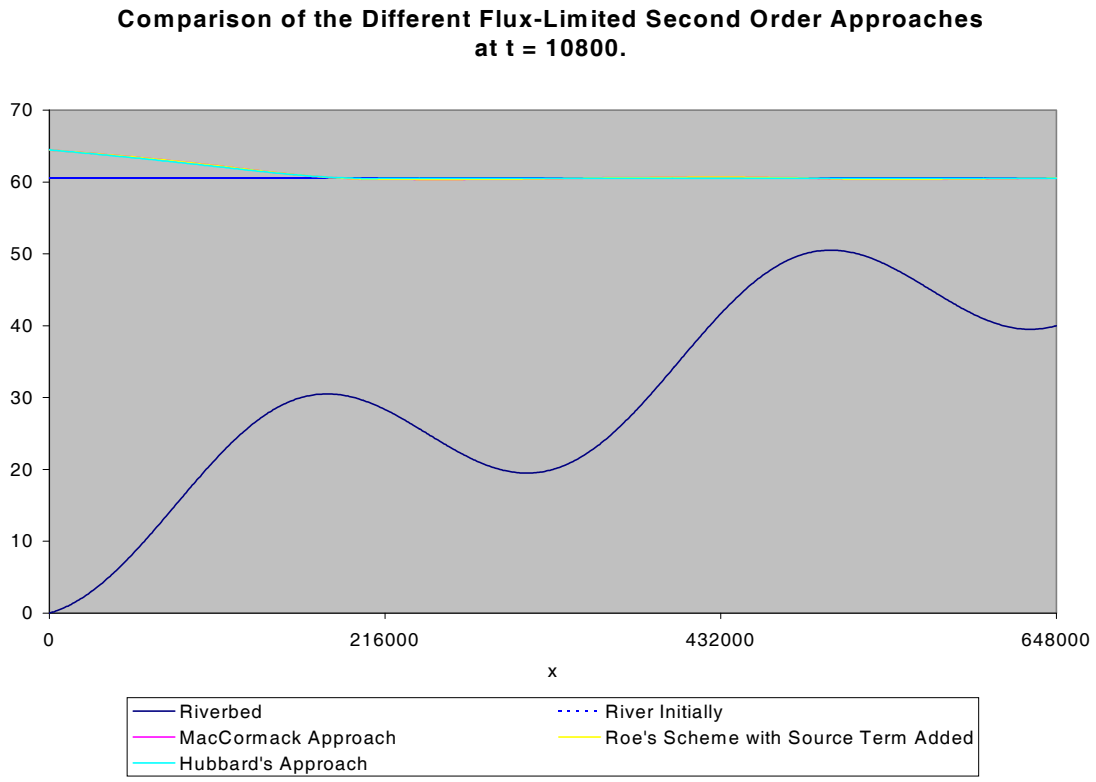

Figure 4-23a: Comparison of the flux-limited second order results for Problem D

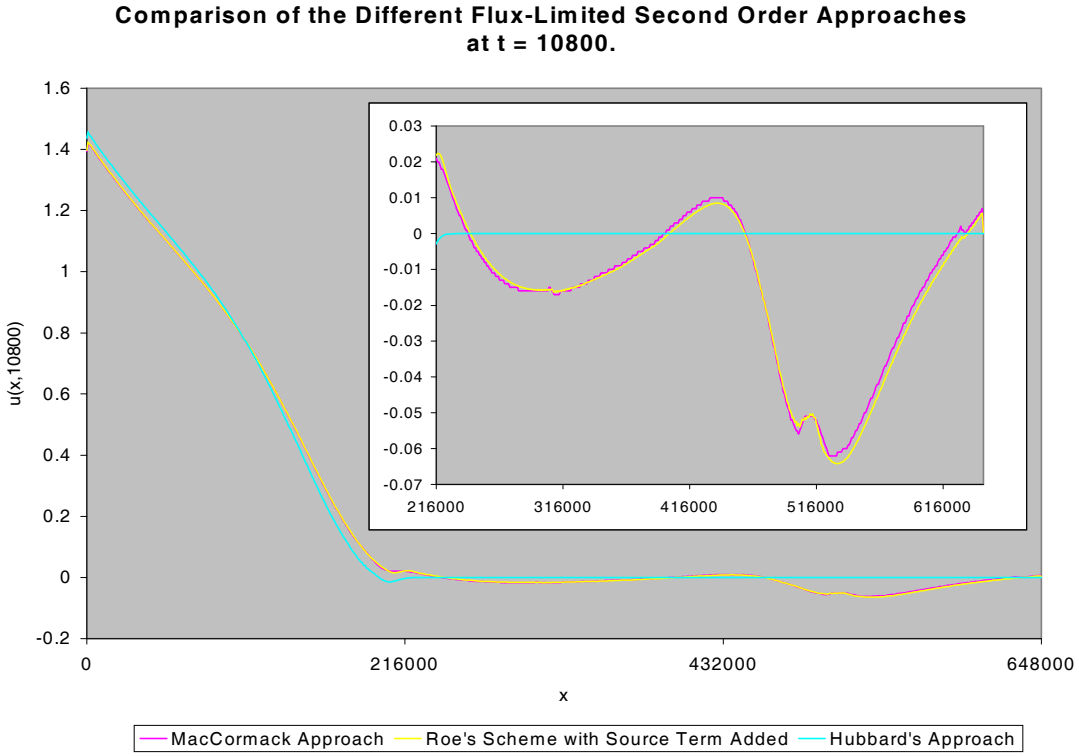

Figure 4-23b: Comparison of the flux-limited second order results for Problem D

5 Conclusion

In Chapter 4, it was shown that Roe's scheme with source term decomposed and Hubbard's approach produced the most accurate numerical results for all four test problems. All the other approaches were accurate for the first two test problems, but as the source term became more significant, the other approaches became inaccurate. In Problem D we saw that Roe's scheme with source term decomposed and Hubbard's approach was the only approach that did not produce movement after $x = 216{,}000m$ at $t = 10{,}800s$. This is because the other approaches did not satisfy the C-property of Bermudez & Vazquez[1].

Overall, Hubbard's approach, which is Roe's scheme with source term decomposed with a flux-limiter applied to both the conservation law and the source term, produced the most accurate numerical results. However, if we use the Superbee flux-limiter instead of the Minmod flux-limiter with Hubbard's approach then the approach produces inaccurate numerical results. If we use (S-4) with Roe's Superbee flux-limiter, the Minmod flux-limiter and van Leer's flux-limiter and use $\Delta x = 1000$, $\Delta t = 1$ and $t = 0$ to $10{,}800s$, we may obtain Figure 5-1a to Figure 5-4b. Figure 5-1a to Figure 5-2b show that by using either Roe's Superbee flux-limiter or van Leer's flux-limiter with Hubbard's approach, oscillations occur in the numerical results. Figure 5-3a and Figure 5-3b show that using the Minmod flux-limiter with Hubbard's approach produces no oscillations in the numerical results. Figure 5-4a and Figure 5-4b compare the results obtained by using the three different flux-limiters with Hubbard's approach at $t = 10{,}800s$. Here, we can see that Roe's Superbee flux-limiter has

produced the most oscillations, followed by van Leer's flux-limiter and the Minmod flux-limiter has produced no oscillations. This suggests that applying a flux-limiter to the source term approximation when the source term is significant can create oscillations in the numerical results. However, if we do not apply a flux-limiter to the source term approximation but we do to the conservation law then we will no longer be able to obtain a numerical scheme which satisfies the C-property of Bermudez & Vazquez[1]. The only solution at present is to use the Minmod flux-limiter with Hubbard's approach or to use Roe's first order scheme with source term decomposed.

Throughout this report, we have seen that adding a source term approximation can produce accurate numerical results but as the source term becomes significant, adding a source term approximation can give very inaccurate numerical results. We have also shown that by decomposing the source term as well as the conservation law, we may obtain very accurate numerical results for the first order case. However, when applying flux-limiters to the source term as well as the conservation law to ensure the numerical scheme satisfies the exact C-property, oscillations can occur in the numerical solution depending on which flux-limiter is used.

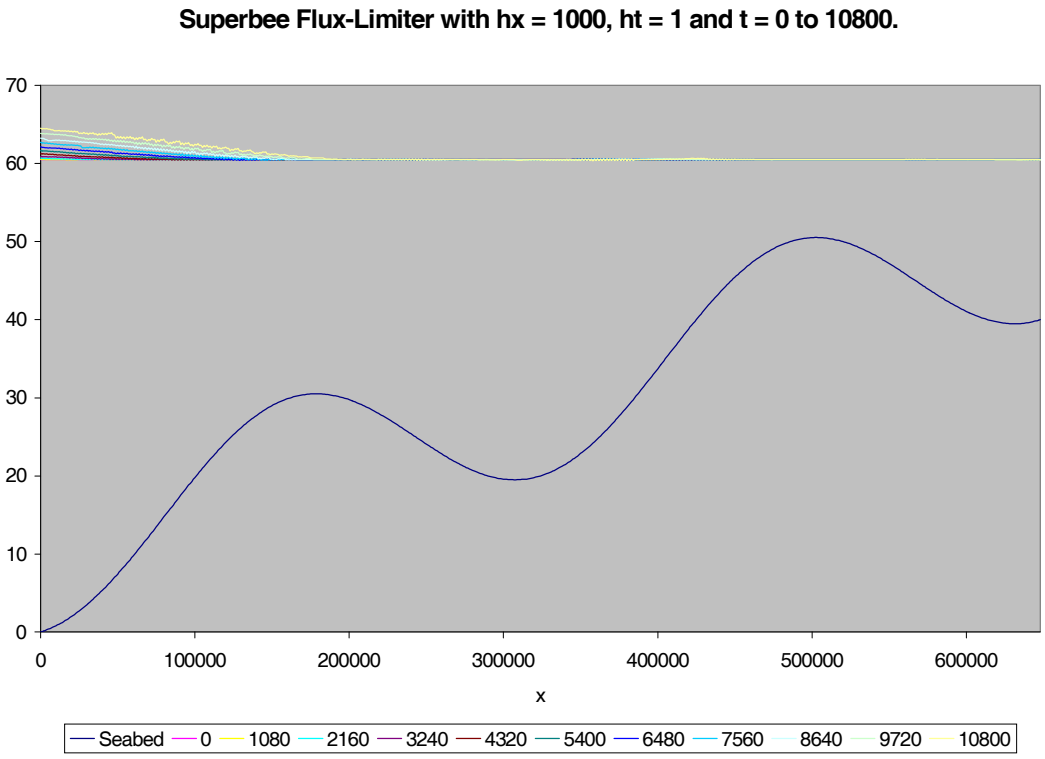

Figure 5-1a: Superbee Flux-Limiter results for Problem D

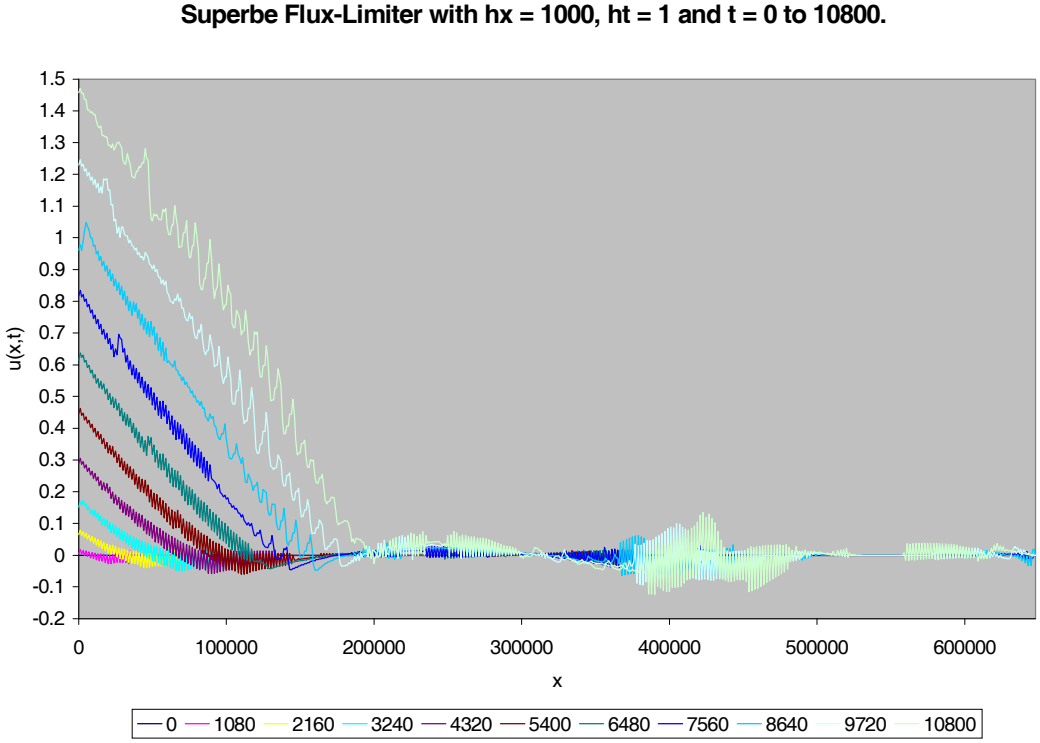

Figure 5-1b: Superbee Flux-Limiter results for Problem D

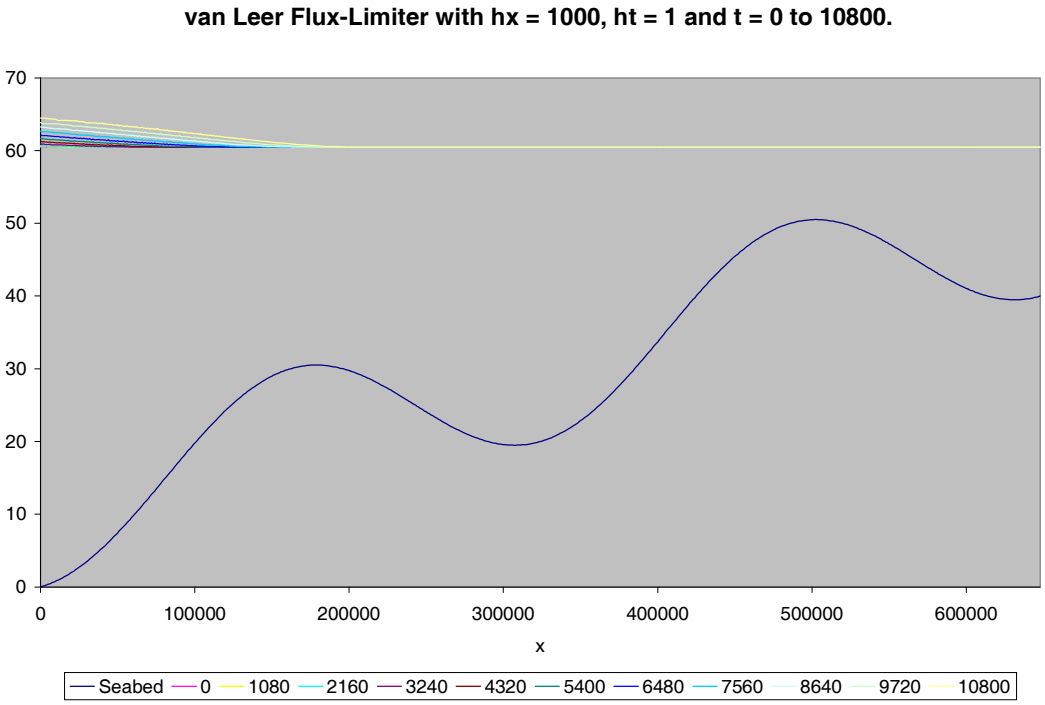

Figure 5-2a: van Leer Flux-Limiter results for Problem D

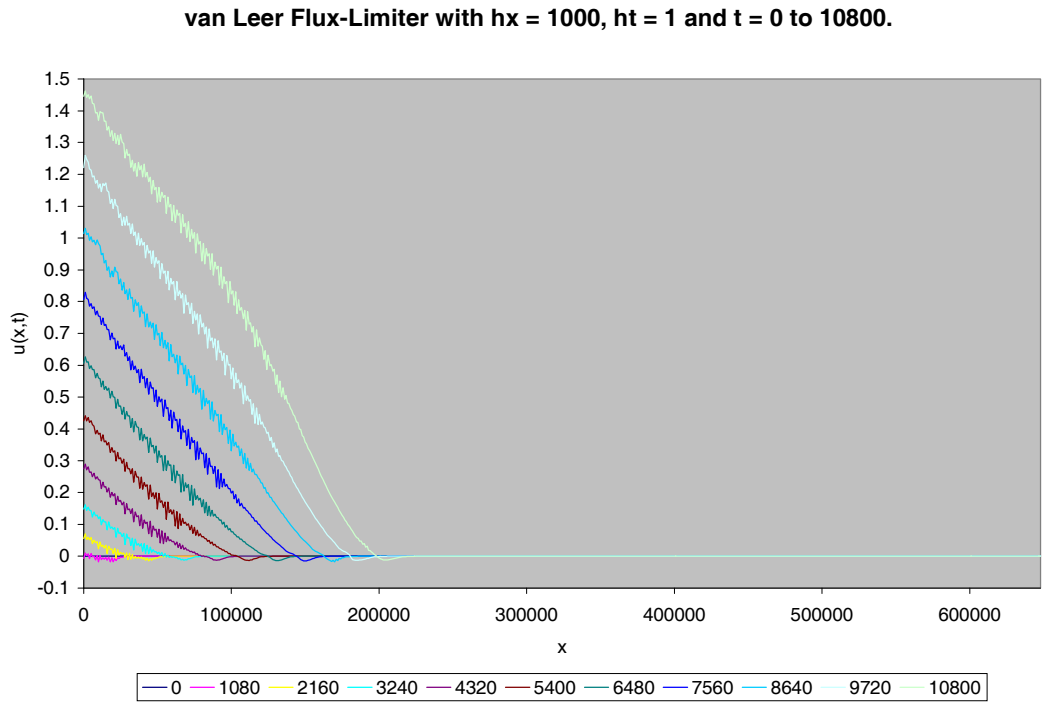

Figure 5-2b: van Leer Flux-Limiter results for Problem D

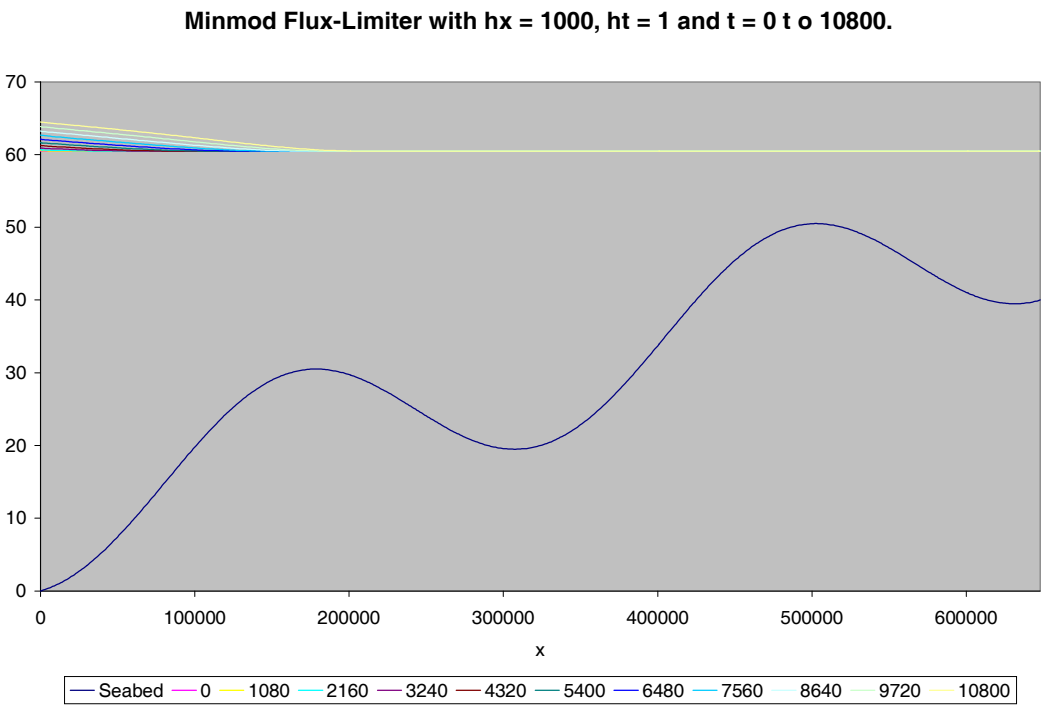

Figure 5-3a: Minmod Flux-Limiter results for Problem D

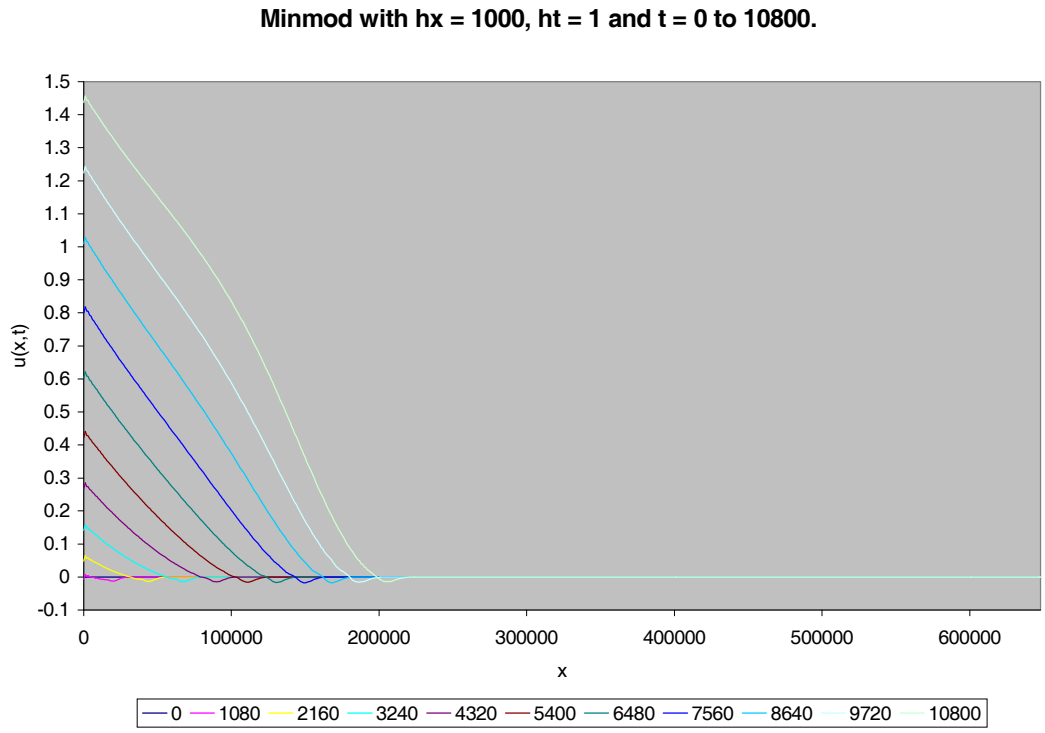

Figure 5-3b: Minmod Flux-Limiter results for Problem D

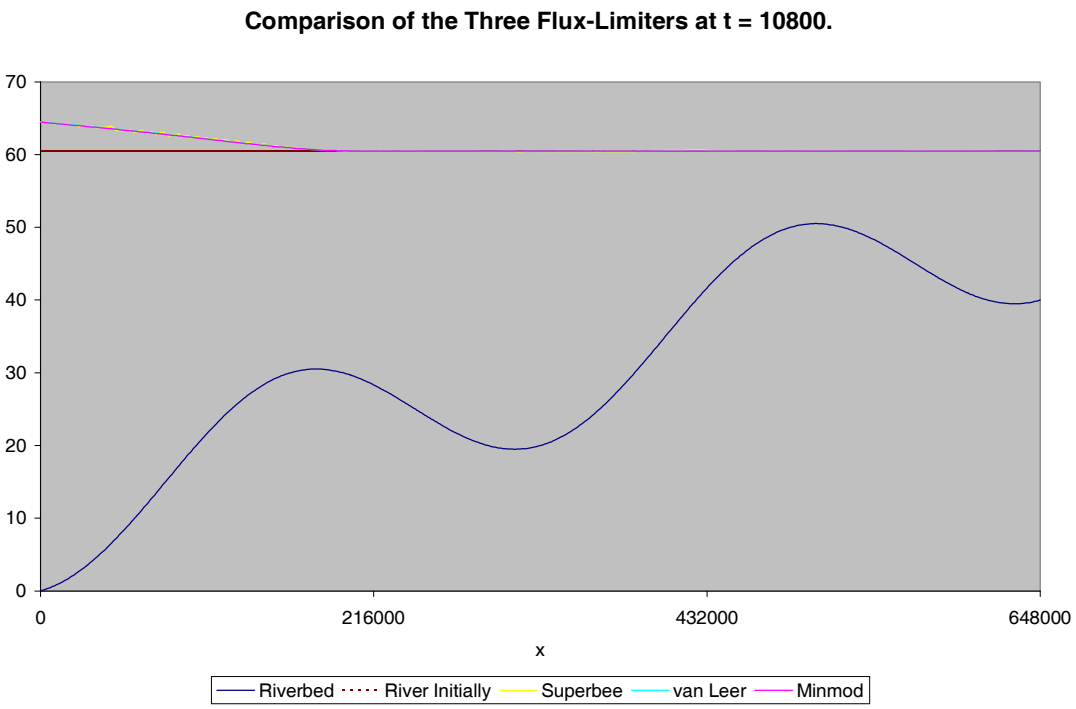

Figure 5-4a: Comparison of the Flux-Limiter results for Problem D

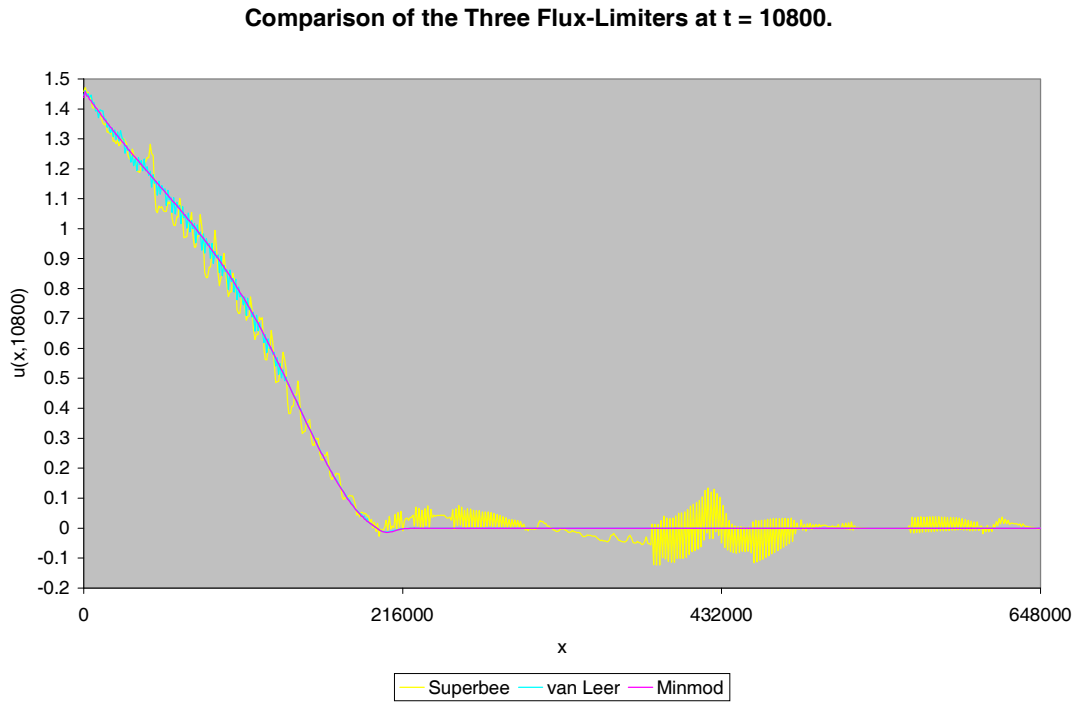

Figure 5-4b: Comparison of the Flux-Limiter results for Problem D

Acknowledgements

I would like to thank my supervisor Dr. P.K. Sweby and Dr. M.E. Hubbard for their guidance and support. In addition, I would like to thank the EPSRC and HR Wallingford for their funding.

References

1. 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).

2. P. Garcia-Navarro, Some Considerations and Improvements on the Performance of Roe's scheme for 1D Irregular Geometries, *Internal Report* **23** (1997).

3. P. Glaister, Difference Schemes for the Shallow Water Equations, *Numerical Analysis Report* **9/87**, *University Of Reading* (1987).

4. P. Glaister, Approximate Riemann Solutions of the Shallow Water Equations, *J. Hydr. Research* **Vol. 26, No.3** (1988).

5. A. Harten, High Resolution Schemes for Conservation Laws, *J. Comput. Phys.* **49** (1983).

6. M.E. Hubbard, Private Communication (University of Reading).

7. J. Hudson, Numerical Techniques for Conservation Laws with Source Terms, *MSc Dissertation*, *University of Reading* (1998).

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

9. 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).

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

11. R.J. LeVeque, Balancing Source Terms and Flux Gradients in High-Resolution Godunov Methods: The Quasi-Steady Wave-Propagation Algorithm, *J. Comput. Phys.* (1998).

12. P.L. Roe, Approximate Riemann Solvers, Parameter Vectors and Difference Schemes, *J. Comput. Phys.* **43**, 357 – 372 (1981).

13. J.J Stoker, Water Waves, *Wiley Interscience* (1957).

14. P.K. Sweby, High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws, *SIAM J. Num. Anal.* **21**, 995 (1984).

15. P.K. Sweby, Source Terms and Conservation Laws: A Preliminary Discussion, *Numerical Analysis Report* **6/89**, *University Of Reading* (1989)

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