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.

which has eigenvalues

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

and eigenvectors

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

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.

which are illustrated in Fi[illegible]. In Figure 2-1, the discontinuity at x = 0.5 represents a barrier, whi[illegible] two initial river heights and is removed at t = 0. Walls are present [illegible] and at x [illegible]ulting in reflection at both boundaries.

Note that, for t[illegible] problem, if $\frac{1}{\phi_0} > 7$[illegible] 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 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

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

and we have initial conditions

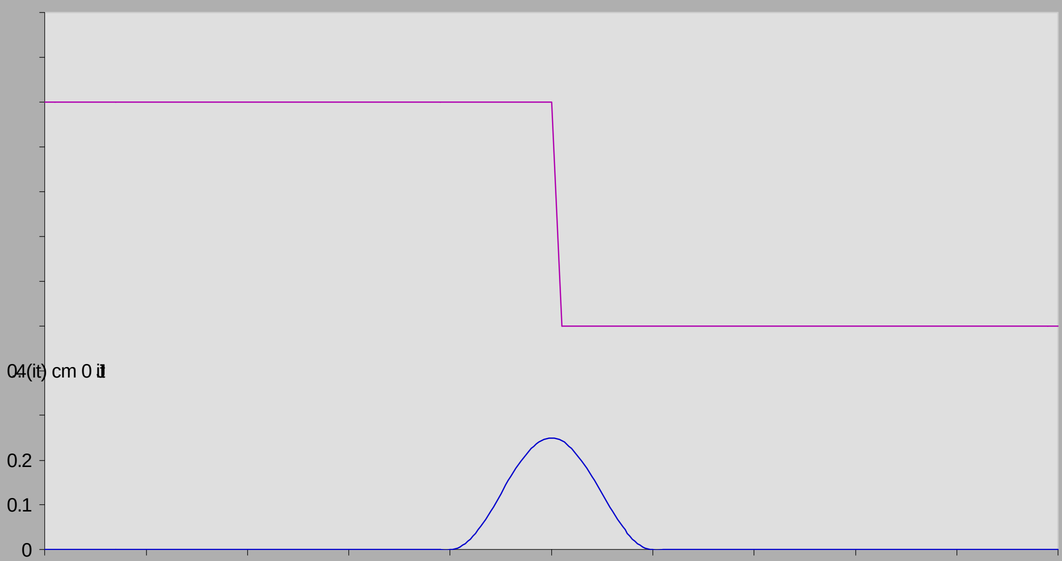
$$u(x,0) = 0 \quad \text{and} \quad h(x,0) = \begin{cases} 1 - B(x) & \text{if } 0 \leq x \leq \frac{1}{2} \\ \phi_0 - B(x) & \text{if } \frac{1}{2} < x \leq 1 \end{cases},$$

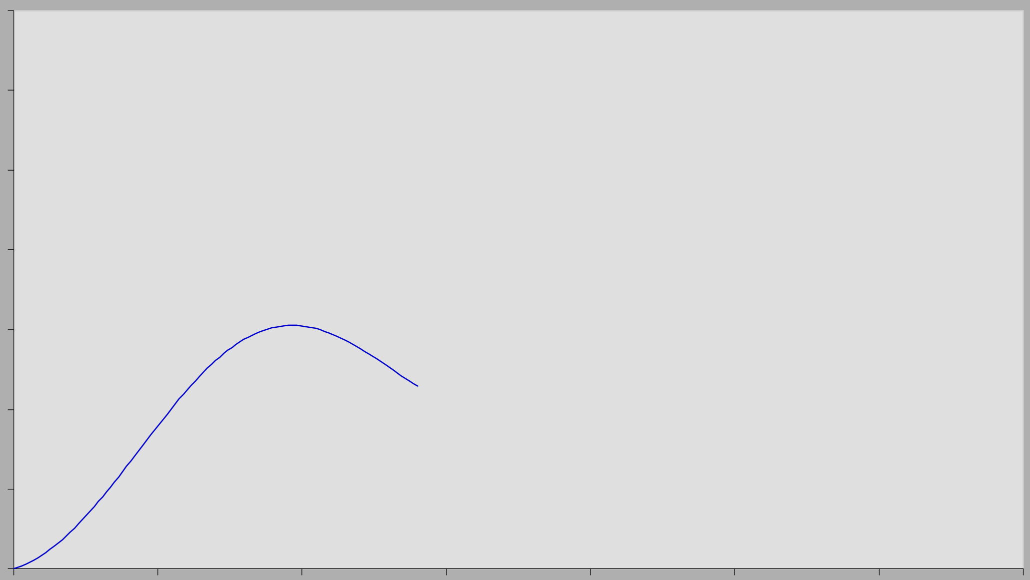
techniques for approximating the Shallow Water Equations so that we can apply the different numerical techniques to the four test problems.

0.4

0.5

Initial Conditions for Problem B





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.

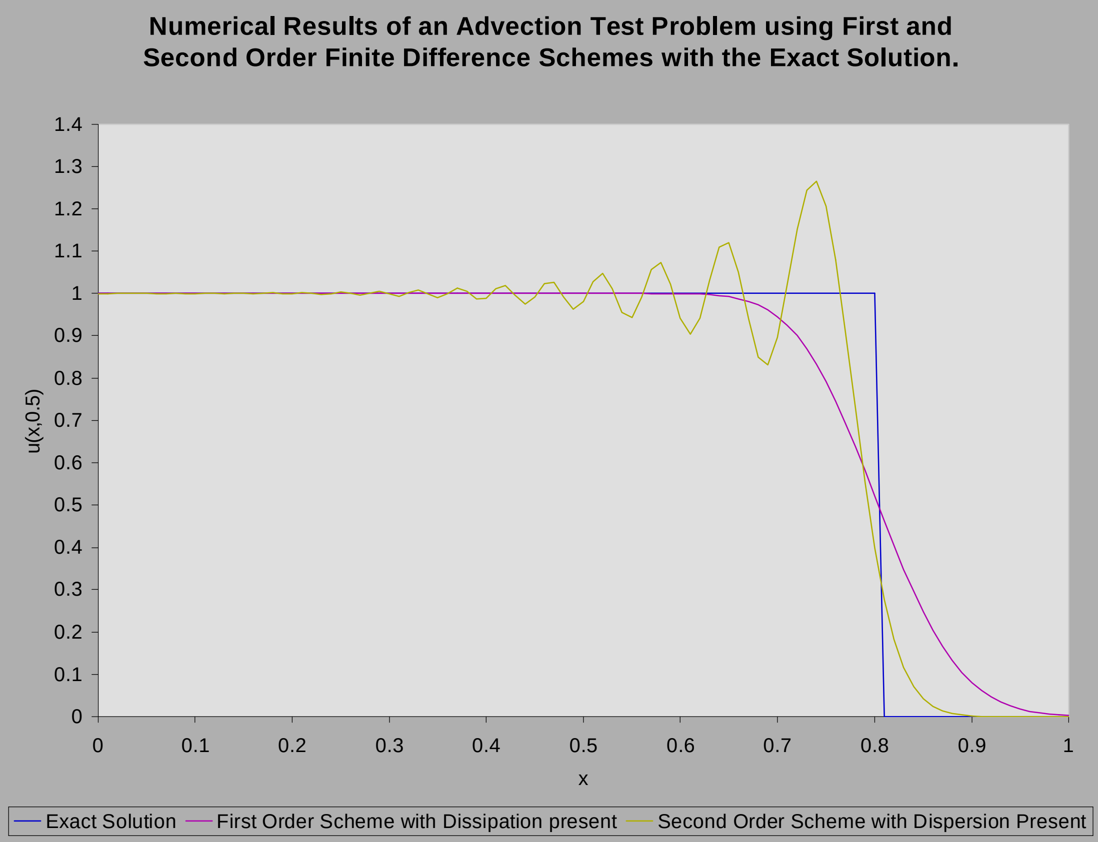


Figure 3-1: Illustration of Dispersion and Dissipation

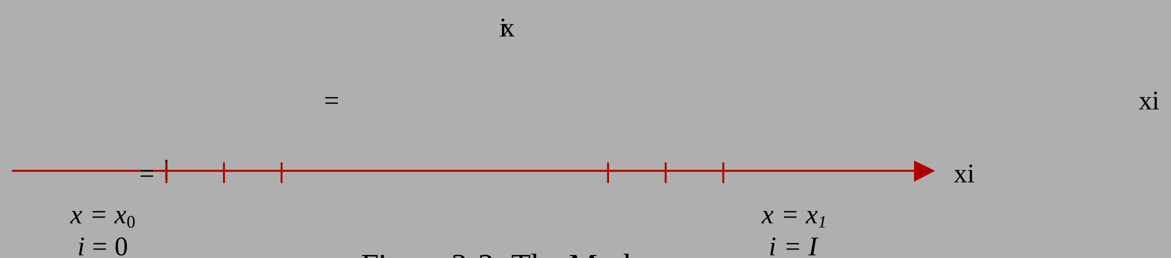


Figure 3-2: The Mesh

3.1

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 $\mathbf{F}^*$

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	$(\ )\ \frac{\vert\ \vert}{\vert\ \vert}$

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

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

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

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

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

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

where

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

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

and

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

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

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

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

and $\mathbf{Q}$ 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 $\mathbf{A}$ denotes the Jacobian matrix of $\mathbf{F}(\mathbf{w})$

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

which has eigenvalues

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

and eigenvectors

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

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

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

which has eigenvalues

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

and eigenvectors

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

where

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

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

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

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

ɉ0T*

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	-

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

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

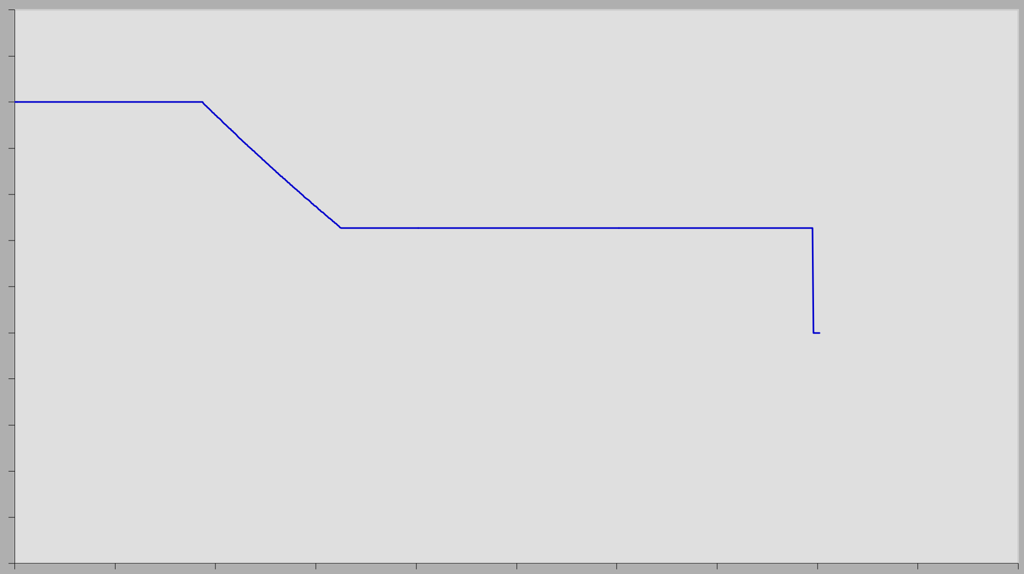
4.2.4 Numerical Results for Problem D

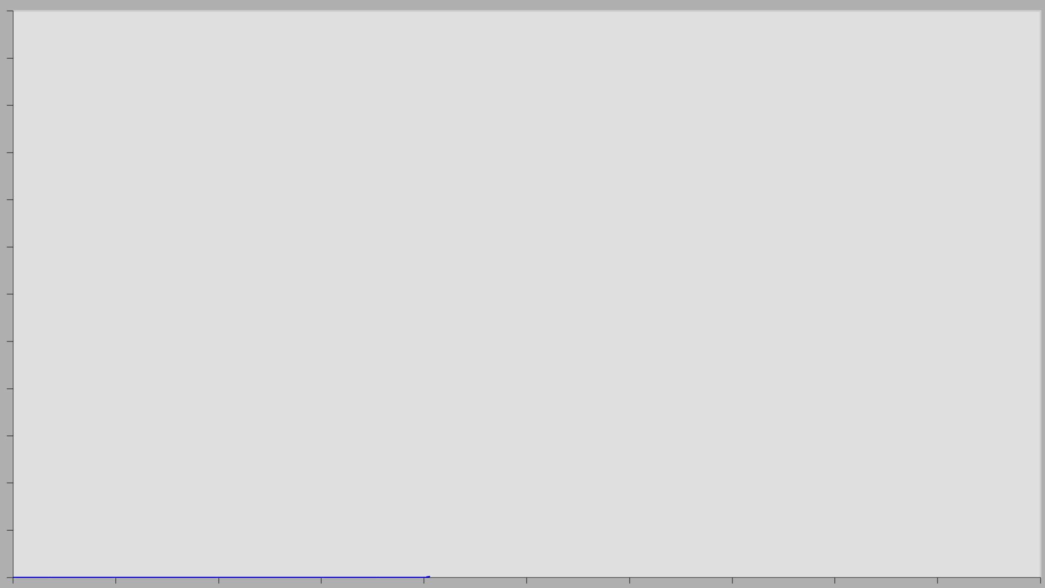
Now, by using approaches (S-2), (S-3a) and (S-4) to approximate Problem D, the results in Figure 4-20a to Figure 4-23b were obtained. Here, we can see that Hubbard's approach was the only approach that did not produce movement after x = 216,000*m* at t = 10,800*s*. Roe's scheme with source term added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

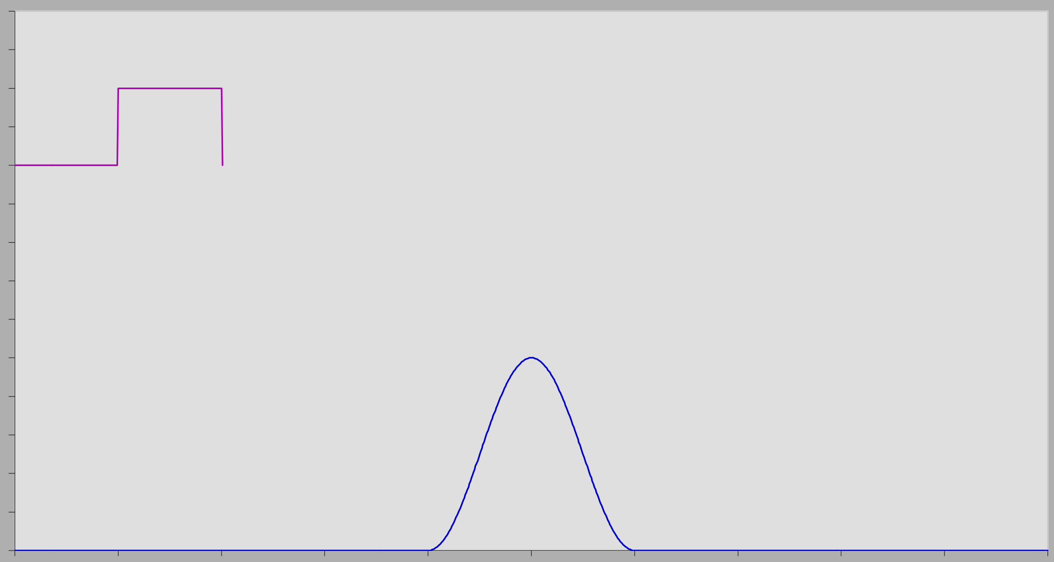
4.2.5 Overall Comparison of the Flux- Limited Second Order Approaches

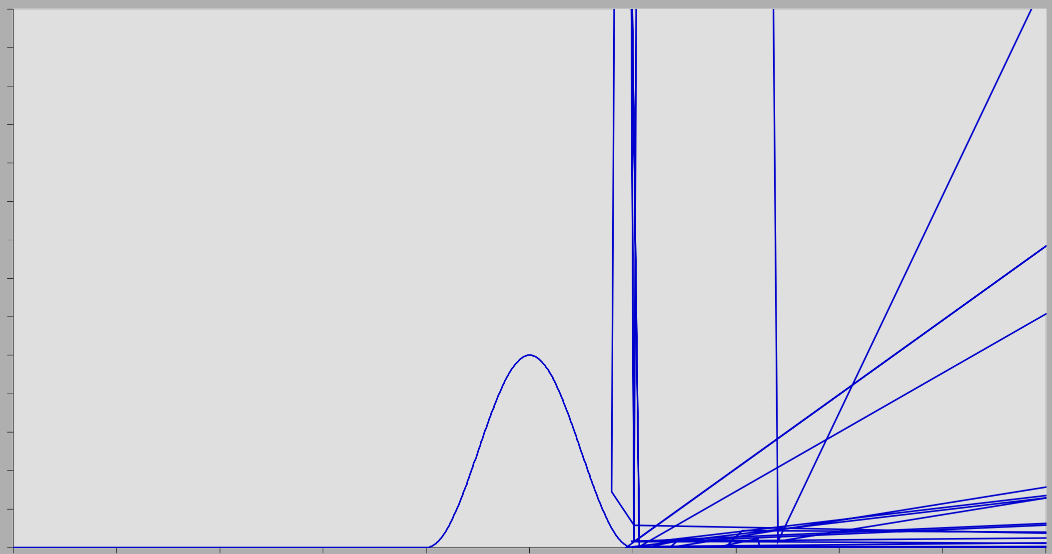
From the results of this sub-section, we have seen that Hubbard's approach has produced the most accurate results for all test problems producing very accurate results even for Problem D. LeVeque & Yee's MacCormack approach, Roe's scheme with source term added and Hubbard's approach all produced almost 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.

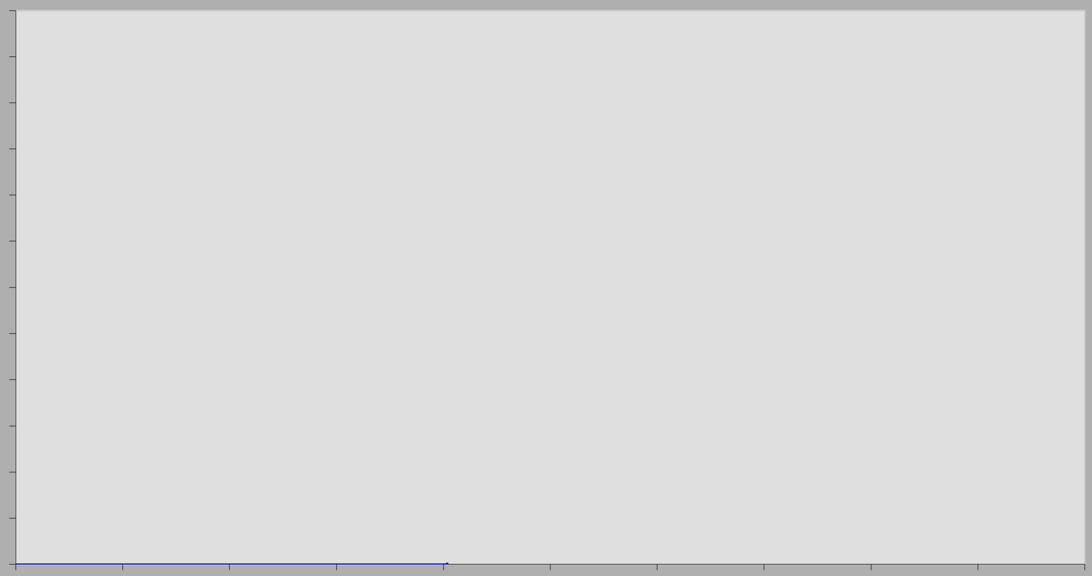
First Order Numerical Results:

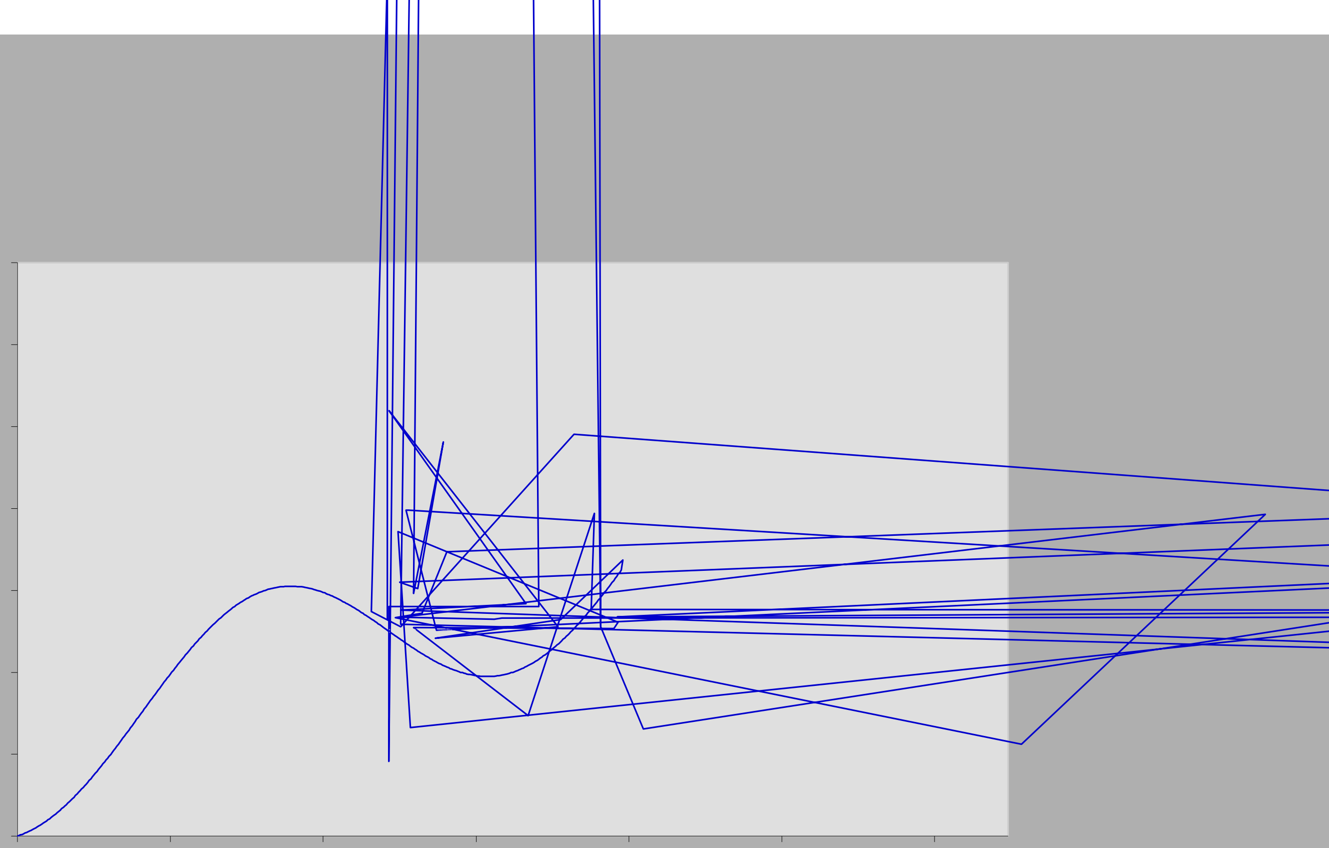




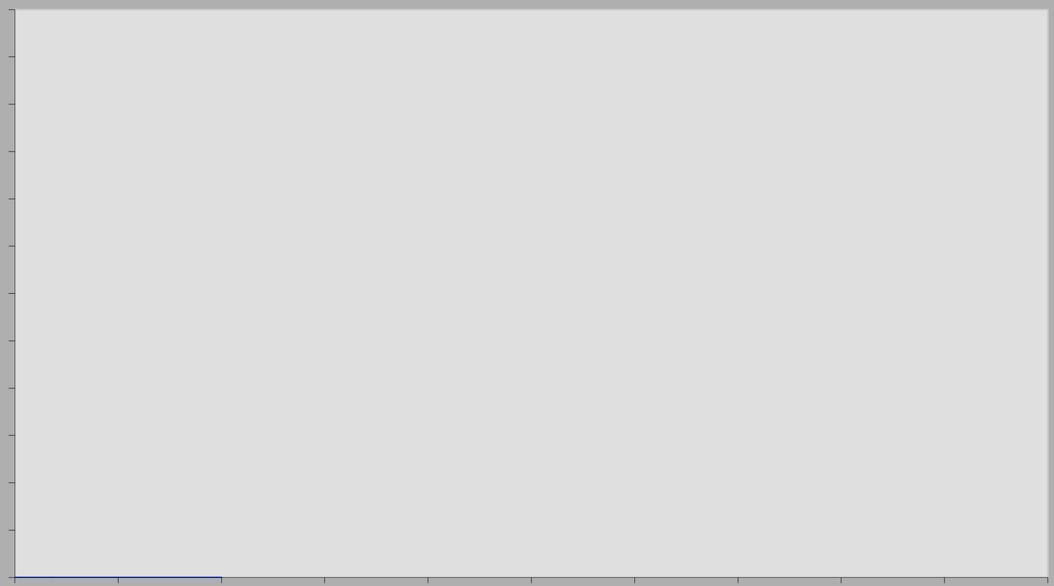


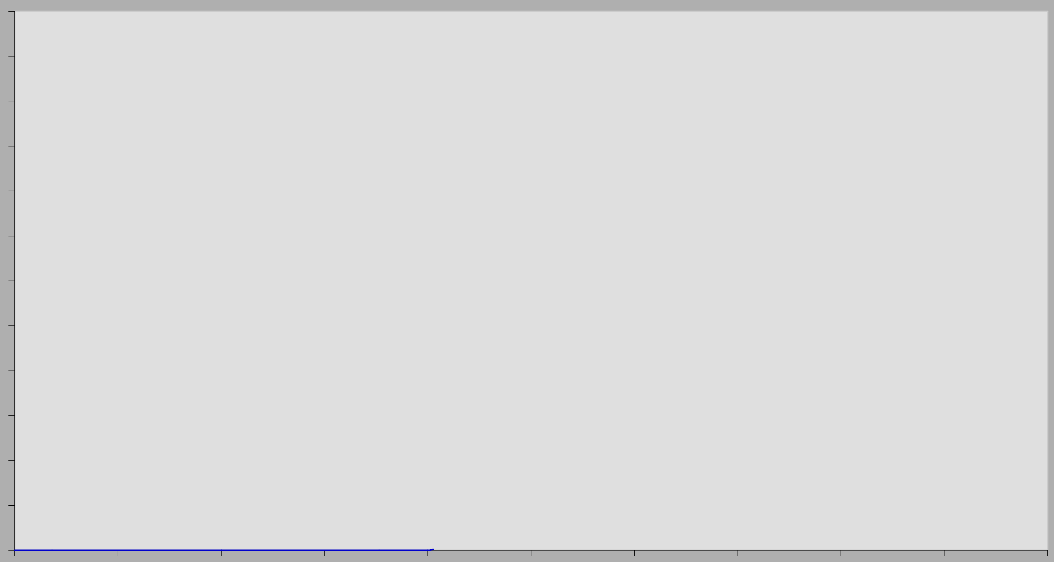


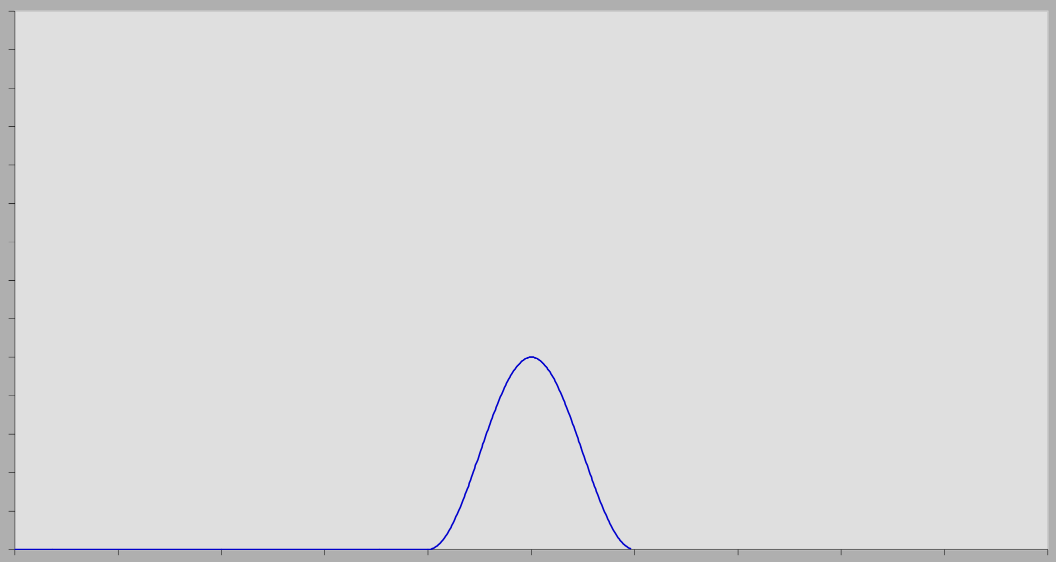


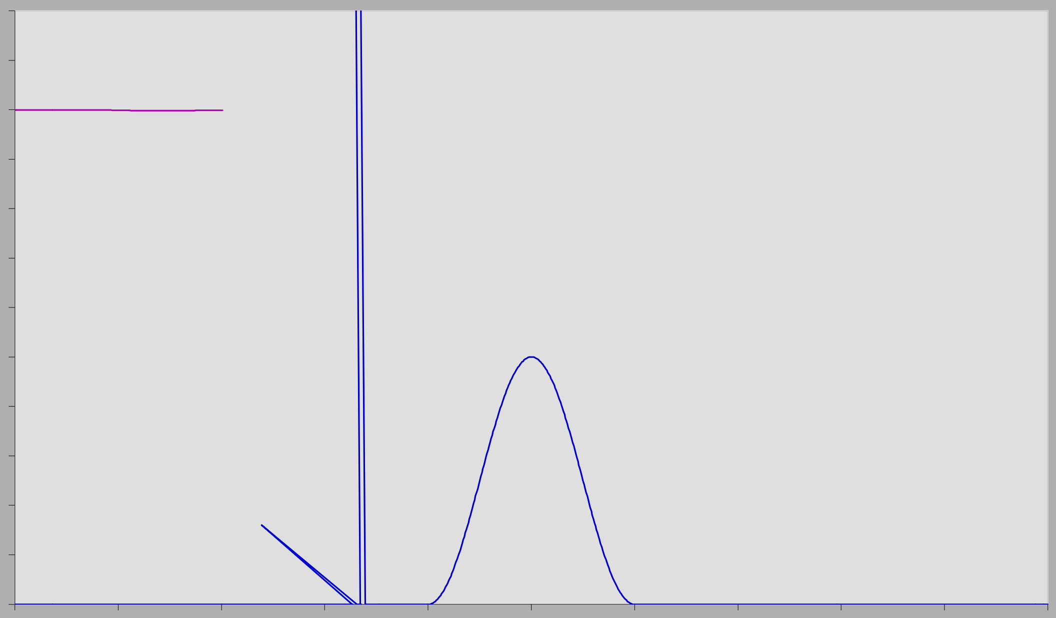


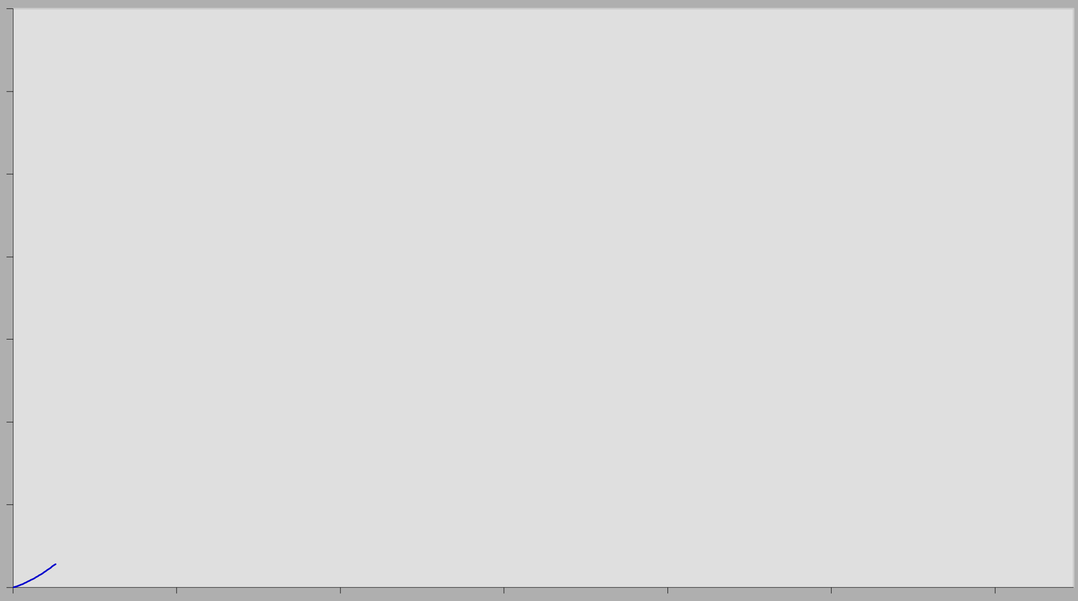
Flux-Limited Second Order Numerical Results:







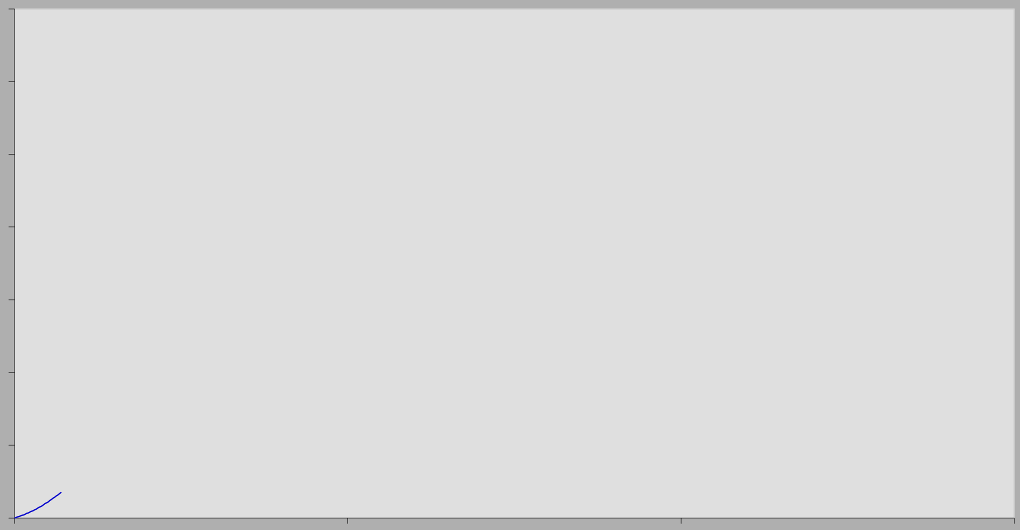




5 Conclusion

In Chapter 4, it was shown that Roe's scheme with source term decomposed and 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.



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*