ENGR30003
Numerical Programming for Engineers

Assignment #2 - Report

Goh Xing Yang
1001969
October 22, 2020

Contents

1 Shock-wave Root Finding

1.1 Analytical Solution

The half angle θ, shock oblique angle β and Mach number M for compressible flow at supersonic speeds is given by the equation:

$$tan(\theta) = 2\cot(\beta)\frac{M^2\sin^2(\beta) - 1}{M^2(\gamma + \cos(2\beta)) + 2} \tag{1}$$

For $\theta = 0$:

$$0 = 2\cot(\beta)\frac{M^2\sin^2(\beta) - 1}{M^2(\gamma + \cos(2\beta)) + 2}$$

Solve for β_u:

$$\begin{aligned} 2\cot(\beta_u) &= 0 \\ \beta_u &= \cot^{-1}(0) \\ \beta_u &= 90^\circ \end{aligned} \tag{2}$$

Solve for β_l:

$$\begin{aligned} \frac{M^2\sin^2(\beta_l) - 1}{M^2(\gamma + \cos(2\beta_l)) + 2} &= 0 \\ M^2\sin^2(\beta_l) - 1 &= 0 \\ \beta_l &= \sin^{-1}\left(\pm\frac{1}{M}\right) \end{aligned}$$

Since β is a non-negative value:

$$\beta_l = \sin^{-1}\left(\frac{1}{M}\right) \tag{3}$$

1.2 Graphical Solution

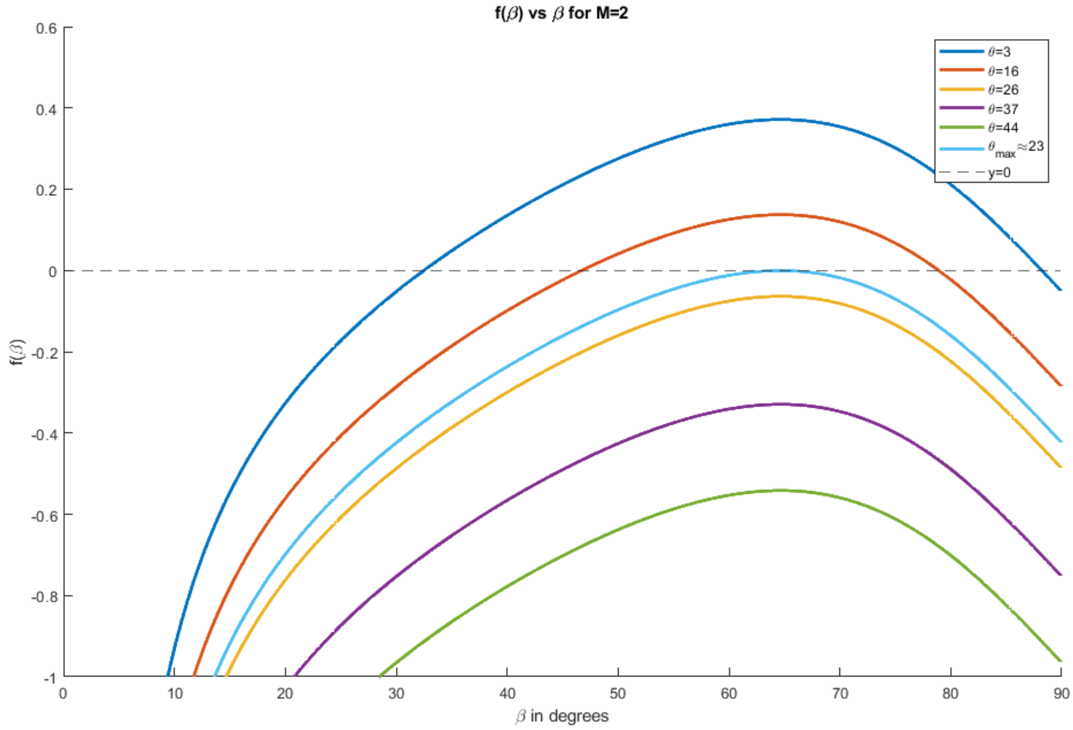


Figure 1: f(β) vs β for M=2

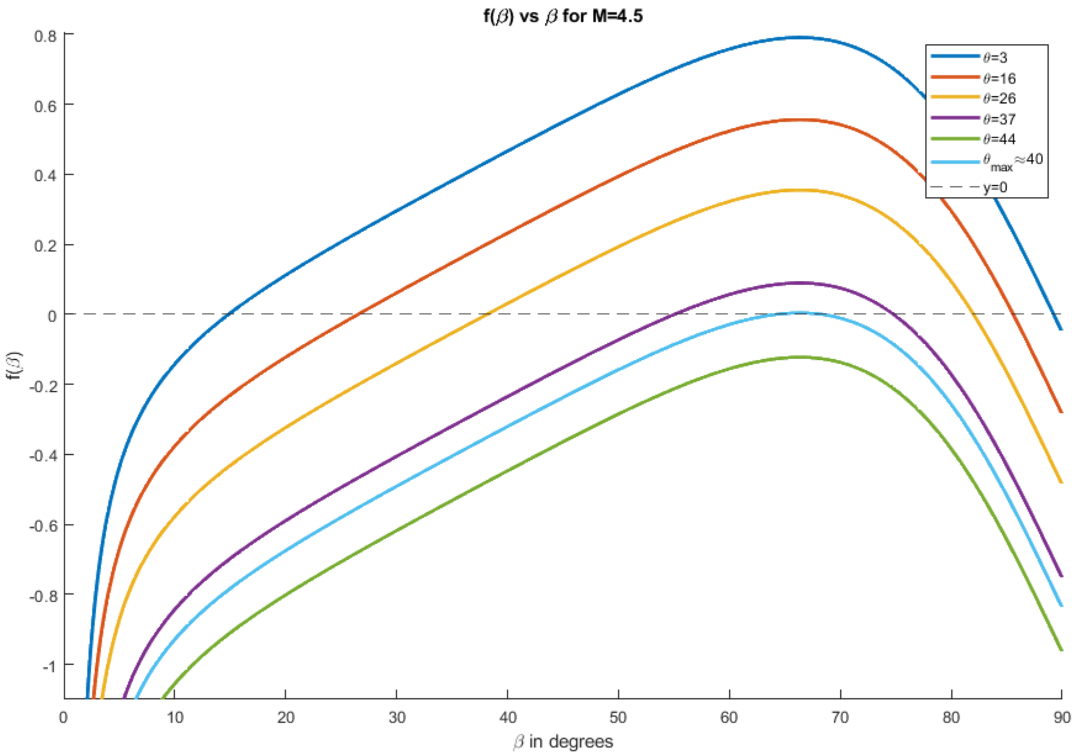


Figure 2: f(β) vs β for M=4.5

Plotting the f(β) vs β relationship for different M and theta values shows that increasing the theta value shifts the curve down without altering the shape of the function. Increasing the Mach number results in a steeper gradient for the function and increases the f(β) value.

For M=2, by visually observing the function curves as θ changes, θ_{max} is approximately 23° and for M=4.5, θ_{max} is approximately 40°. These θ_{max} approximations are plotted out in the figures.

1.3 Solving Shock-wave Equation

Using the Newton-Raphson root finding approach $x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$, the values of β can be identified for a given M and θ.

$f(\beta)$ will be represented as:

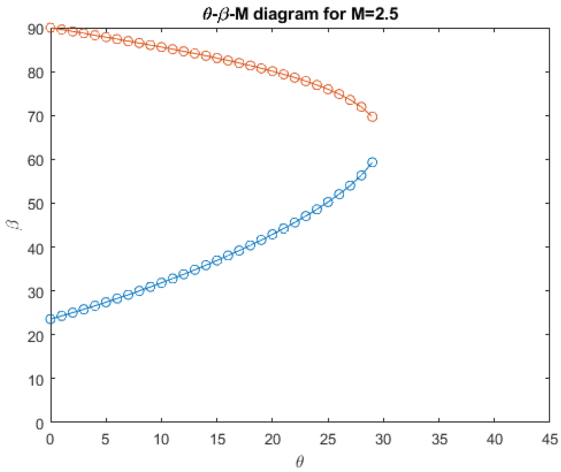
$$f(\beta) = 2\cot(\beta)\frac{M^2\sin^2(\beta) - 1}{M^2(\gamma + \cos(2\beta)) + 2} - \tan(\theta) \tag{4}$$

Where finding the roots of the equation will provide the solution of β.

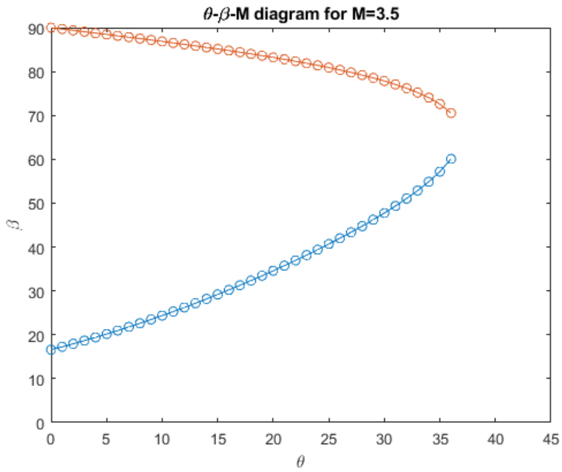
The derivative $f'(\beta)$ will also be evaluated to perform Newton-Raphson's root finding approach.

$$\begin{aligned} f'(\beta) = & \frac{4M^2\cos(\beta)\sin(\beta)}{\tan(\beta)((\gamma + \cos(2\beta)))M^2 + 2)} \\ & - \frac{2\left(\tan(\beta)^2 + 1\right)\left(M^2\sin(\beta)^2 - 1\right)}{\tan(\beta)^2\left((\gamma + \cos(2\beta))M^2 + 2\right)} \\ & + \frac{4M^2\sin(2\beta)\left(M^2\sin(\beta)^2 - 1\right)}{\tan(\beta)\left((\gamma + \cos(2\beta))M^2 + 2\right)^2} \end{aligned} \tag{5}$$

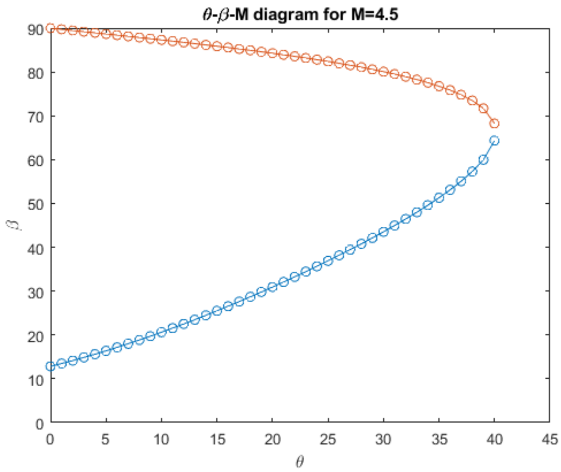
Since Newton-Raphson is an open method of root convergence, it requires an initial guess. Since $\beta_l = \sin^{-1}\left(\frac{1}{M}\right)$ and $\beta_u = 90°$ when $\theta = 0°$, these values will be used for the initial guess when $\theta = 1°$ and subsequent results of β_l and β_u are used for initial guesses as theta increments.



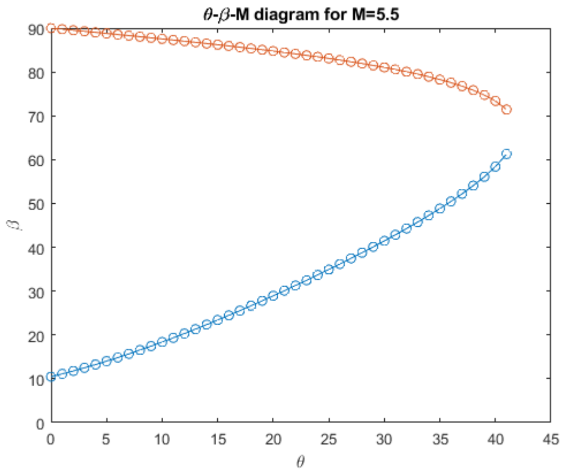
(a) Mach number = 2.5



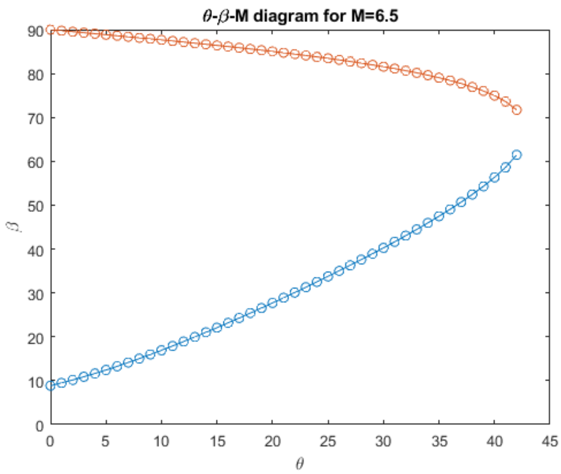
(b) Mach number = 3.5



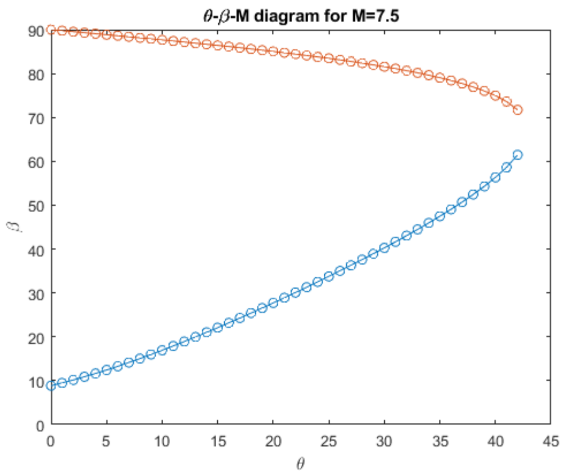
(c) Mach number = 4.5



(d) Mach number = 5.5



(e) Mach number = 6.5



(f) Mach number = 7.5

Figure 3: $\theta - \beta - M$ diagrams for various Mach numbers

2 Regression

The least squares problem for linear regression $y = ax + b$ can be expressed as:

$$\begin{bmatrix} \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} x_i & N \end{bmatrix} \left\{ \begin{array}{c} a \\ b \end{array} \right\} = \left\{ \begin{array}{c} \sum_{i=1}^{N} x_i y_i \\ \sum_{i=1}^{N} y_i \end{array} \right\} \tag{6}$$

where the mean of x_i is:

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N} \tag{7}$$

Expanding the matrix provides the following equations:

$$a \sum_{i=1}^{N} x_i^2 + b \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} x_i y_i \tag{8}$$

$$a \sum_{i=1}^{N} x_i + bN = \sum_{i=1}^{N} y_i \tag{9}$$

Rearranging equation (8) provides:

$$b = \frac{\sum_{i=1}^{N} y_i}{N} - \frac{a \sum_{i=1}^{N} x_i}{N} \tag{10}$$

Using the relationship in (7), this can be expressed as:

$$b = \bar{y} - a\bar{x} \tag{11}$$

Substituting equation (10) into equation (8)

$$a \sum_{i=1}^{N} x_i^2 + \left(\frac{\sum_{i=1}^{N} y_i}{N} - \frac{a \sum_{i=1}^{N} x_i}{N} \right) \left(\sum_{i=1}^{N} x_i \right) = \sum_{i=1}^{N} x_i y_i$$

$$a \sum_{i=1}^{N} x_i^2 + \frac{\sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{N} - \frac{a \sum_{i=1}^{N} x_i \sum_{i=1}^{N} x_i}{N} = \sum_{i=1}^{N} x_i y_i$$

$$a \left(\sum_{i=1}^{N} x_i^2 - \bar{x} \sum_{i=1}^{N} x_i \right) = \sum_{i=1}^{N} x_i y_i - \frac{\sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{N}$$

$$a \left(\sum_{i=1}^{N} x_i^2 - \bar{x} \sum_{i=1}^{N} x_i \right) = \sum_{i=1}^{N} x_i y_i - \bar{y} \sum_{i=1}^{N} x_i$$

Since $\bar{x}$ is the mean, $\sum_{i=1}^{N}(x_i - \bar{x}) = 0$, adding these terms to the equation does not change it.

$$a \left(\sum_{i=1}^{N} x_i^2 - \bar{x} \sum_{i=1}^{N} x_i - \bar{x} \sum_{i=1}^{N} (x_i - \bar{x}) \right) = \sum_{i=1}^{N} x_i y_i - \bar{y} \sum_{i=1}^{N} x_i - \bar{x} \sum_{i=1}^{N} (y_i - \bar{y})$$

$$a \sum_{i=1}^{N} (x_i^2 - \bar{x} x_i - \bar{x}(x_i - \bar{x})) = \sum_{i=1}^{N} (x_i y_i - \bar{y} x_i - \bar{x}(y_i - \bar{y}))$$

$$a \sum_{i=1}^{N} (x_i^2 - 2\bar{x} x_i + \bar{x}^2) = \sum_{i=1}^{N} (x_i y_i - \bar{y} x_i - \bar{x} y_i - \bar{x}\bar{y})$$

$$a\sum_{i=1}^{N}(x_i-\bar{x})^2=\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})$$

$$a=\frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{N}(x_i-\bar{x})^2} \tag{12}$$

From equation (12), the linear regression will fail when the denominator $\sum_{i=1}^{N}(x_i-\bar{x})^2=0$, which will occur when there is only a single data point available or if the function is a constant function as for each data point in these scenarios, $\bar{x}=x_i$.

3 Linear Algebraic Systems

For an augmented linear tri-diagonal matrix:

$$\left[\begin{array}{cccccc|c} b_1 & c_1 & 0 & 0 & \dots & 0 & r_1 \\ a_2 & b_2 & c_2 & 0 & \dots & 0 & r_2 \\ 0 & a_3 & b_3 & c_3 & \dots & \vdots & r_3 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & a_{N-1} & b_{N-1} & c_{N-1} & \vdots \\ 0 & \dots & 0 & 0 & a_N & b_N & r_N \end{array}\right] \tag{13}$$

Performing Gaussian elimination on (13), for rows n with indices from i=2 to i=N:

$$n_i^*=n_i-\frac{a_i n_{i-1}^*}{b_{i-1}^*}$$

This makes all the $a_i=0$ in system (13).
To illustrate this process, a generalised 4x4 matrix is used:

$$\left[\begin{array}{cccc|c} b_1 & c_1 & 0 & 0 & r_1 \\ a_2 & b_2 & c_2 & 0 & r_2 \\ 0 & a_3 & b_3 & c_3 & r_3 \\ 0 & 0 & a_4 & b_4 & r_4 \end{array}\right]$$

$$R2^*=R2-\frac{a_2 R1}{b_1}$$

$$\left[\begin{array}{cccc|c} b_1 & c_1 & 0 & 0 & r_1 \\ 0 & b_2-\frac{a_2c_1}{b_1} & c_2 & 0 & r_2-\frac{a_2r_1}{b_1} \\ 0 & a_3 & b_3 & c_3 & r_3 \\ 0 & 0 & a_4 & b_4 & r_4 \end{array}\right]=\left[\begin{array}{cccc|c} b_1 & c_1 & 0 & 0 & r_1 \\ 0 & b_2^* & c_2 & 0 & r_2^* \\ 0 & a_3 & b_3 & c_3 & r_3 \\ 0 & 0 & a_4 & b_4 & r_4 \end{array}\right]$$

$$R3^*=R3-\frac{a_3 R2^*}{b_2^*}$$

$$\left[\begin{array}{cccc|c} b_1 & c_1 & 0 & 0 & r_1 \\ 0 & b_2^* & c_2 & 0 & r_2^* \\ 0 & 0 & b_3-\frac{a_3c_2}{b_2^*} & c_3 & r_3-\frac{a_3r_2^*}{b_2^*} \\ 0 & 0 & a_4 & b_4 & r_4 \end{array}\right]=\left[\begin{array}{cccc|c} b_1 & c_1 & 0 & 0 & r_1 \\ 0 & b_2^* & c_2 & 0 & r_2^* \\ 0 & 0 & b_3^* & c_3 & r_3^* \\ 0 & 0 & a_4 & b_4 & r_4 \end{array}\right]$$

$$R4^*=R4-\frac{a_4 R3^*}{b_3^*}$$

$$\left[\begin{array}{cccc|c} b_1 & c_1 & 0 & 0 & r_1 \\ 0 & b_2^* & c_2 & 0 & r_2^* \\ 0 & 0 & b_3^* & c_3 & r_3^* \\ 0 & 0 & 0 & b_4 - \frac{a_4 c_3}{b_3^*} & r_4 - \frac{a_4 r_3^*}{b_3^*} \end{array}\right] = \left[\begin{array}{cccc|c} b_1 & c_1 & 0 & 0 & r_1 \\ 0 & b_2^* & c_2 & 0 & r_2^* \\ 0 & 0 & b_3^* & c_3 & r_3^* \\ 0 & 0 & 0 & b_4^* & r_4^* \end{array}\right]$$

Now, solving for x:

$$b_4^* x_4 = r_4^* \qquad \rightarrow \qquad x_4 = \frac{r_4^*}{b_4^*}$$

$$b_3^* x_3 + c_3 x_4 = r_3^* \qquad \rightarrow \qquad x_3 = \frac{r_3^* - c_3 x_4}{b_3^*}$$

$$b_2^* x_2 + c_2 x_3 = r_2^* \qquad \rightarrow \qquad x_2 = \frac{r_2^* - c_2 x_3}{b_2^*}$$

$$b_1 x_1 + c_1 x_2 = r_1 \qquad \rightarrow \qquad x_1 = \frac{r_1 - c_1 x_2}{b_1}$$

Generalising the process above, for a general tri-diagonal matrix system (13):

$$\begin{bmatrix} b_1^* & c_1 & 0 & 0 & \dots & 0 \\ 0 & b_2^* & c_2 & 0 & \dots & 0 \\ 0 & 0 & b_3^* & c_3 & \dots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & b_{N-1}^* & c_{N-1} \\ 0 & \dots & 0 & 0 & 0 & b_N^* \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ \vdots \\ x_N \end{Bmatrix} = \begin{Bmatrix} r_1^* \\ r_2^* \\ r_3^* \\ \vdots \\ \vdots \\ r_N^* \end{Bmatrix} \tag{14}$$

The tri-diagonal matrix system (13) can be rewritten as (14), where:

$$b_i^* = \begin{cases} b_i & \text{for} \quad i = 1 \\ b_i - a_i c_{i-1}/b_{i-1}^* & \text{for} \quad i = 2, 3, \dots, N \end{cases}$$

$$r_i^* = \begin{cases} r_i & \text{for} \quad i = 1 \\ r_i - a_i r_{i-1}^*/b_{i-1}^* & \text{for} \quad i = 2, 3, \dots, N \end{cases}$$

$$x_i = \begin{cases} r_i^*/b_i^* & \text{for} \quad i = N \\ \left(r_i^* - c_i x_{i+1}\right)/b_i^* & \text{for} \quad i = N-1, N-2, \dots, 1 \end{cases}$$

4 Interpolation

4.1 Graphs

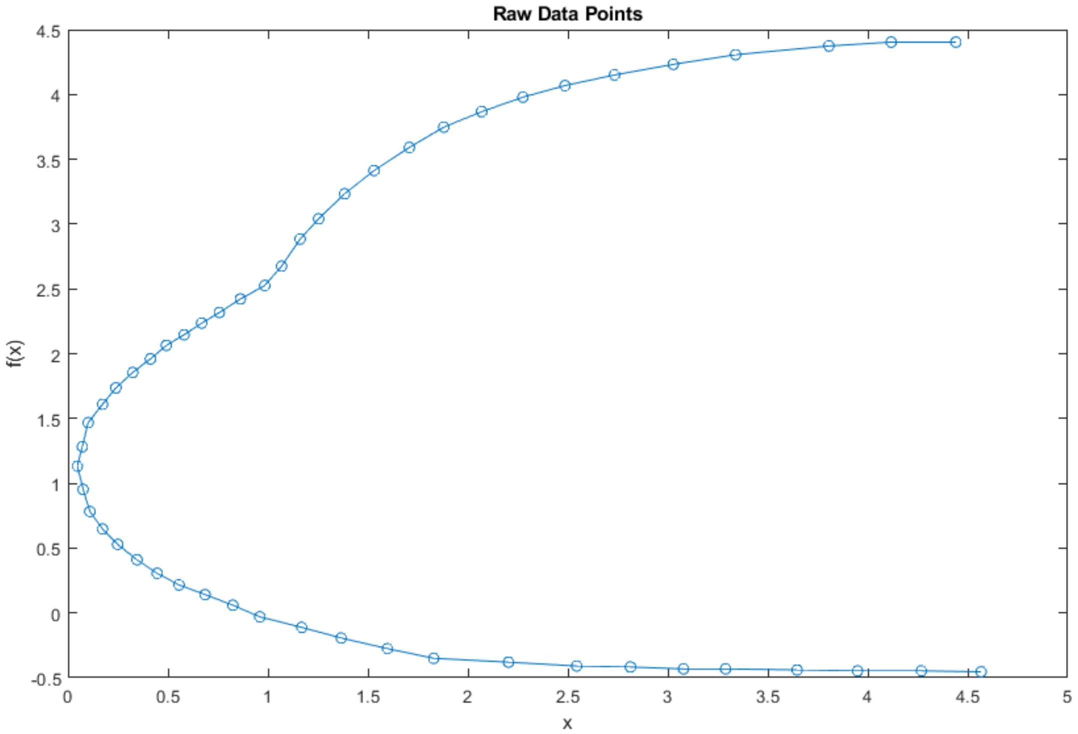


Figure 4: Provided data points

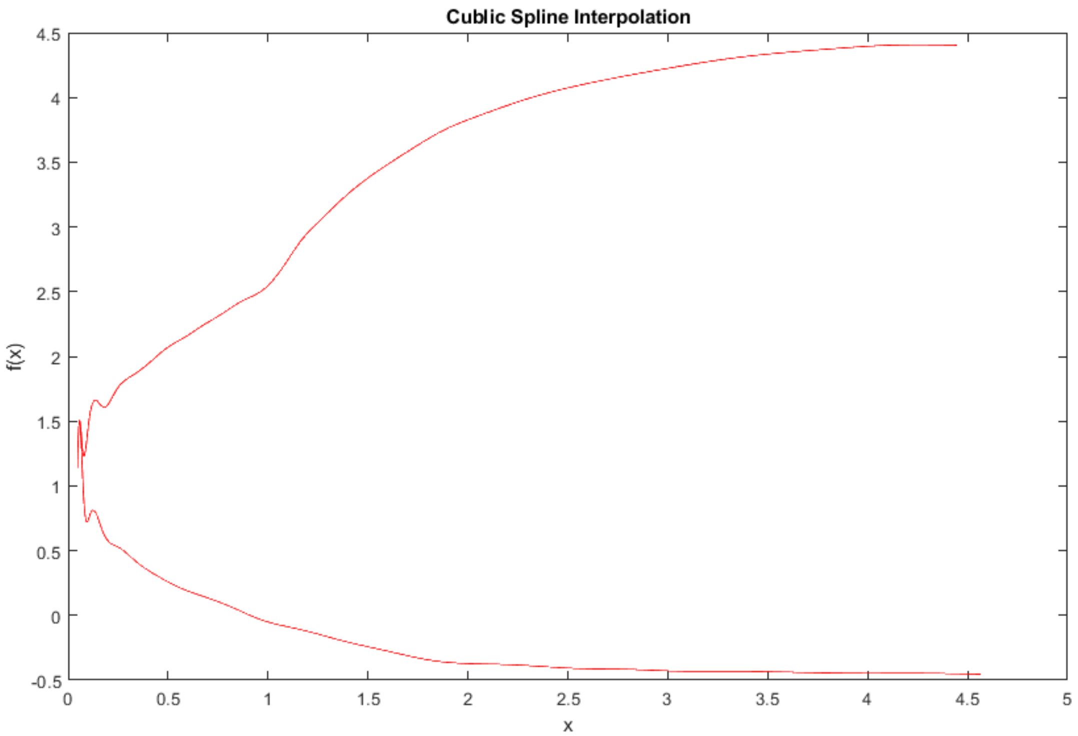


Figure 5: Cubic spline interpolation

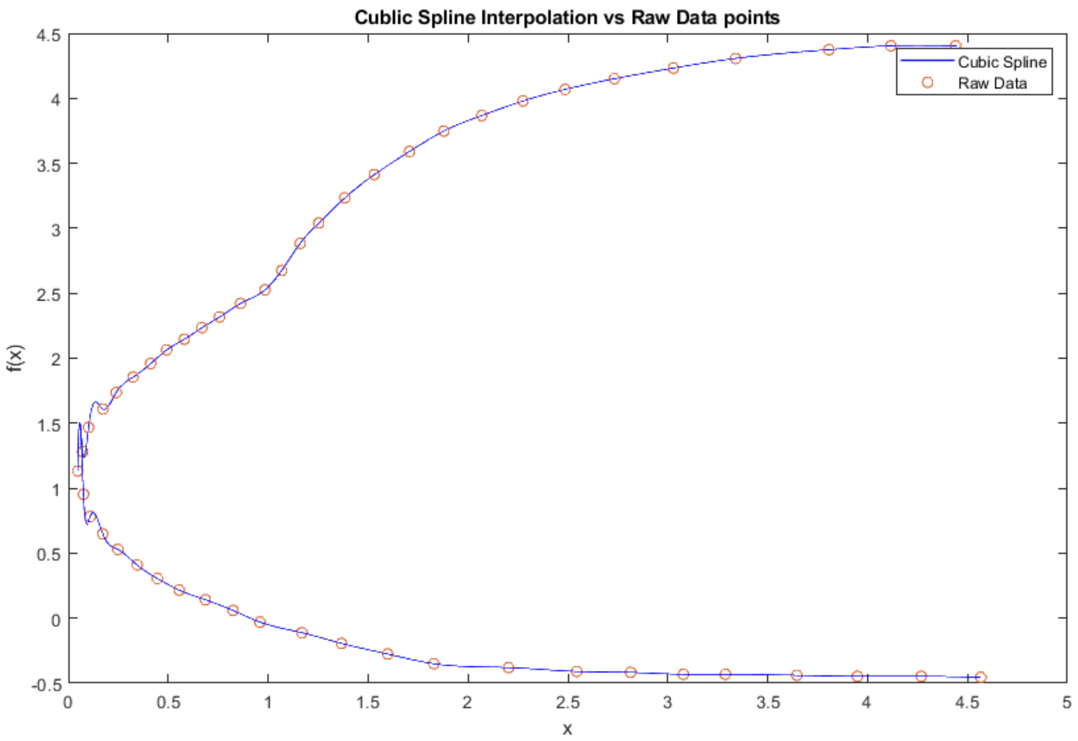


Figure 6: Cubic spline interpolation vs raw data points

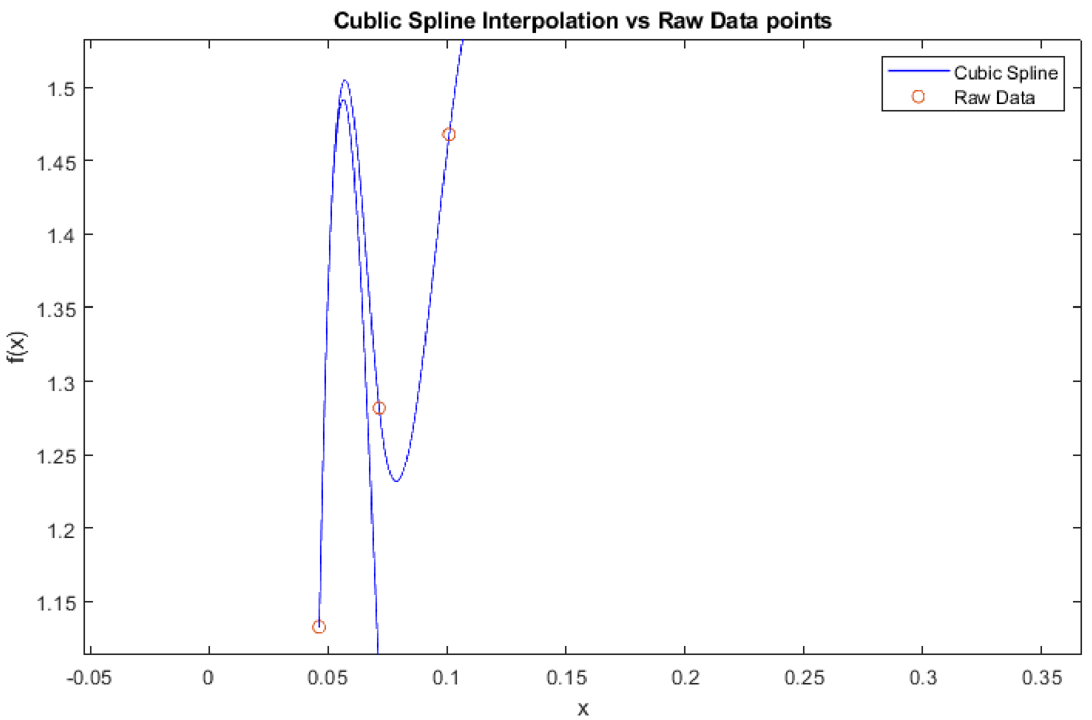


Figure 7: Cubic spline interpolation zoomed

4.2 Discussion

The cubic spline generally fits the given data well as seen in figure (6), however, closer inspection of the points in figure (7), there is a large error at the start of the interpolated graph. This is a result of the cubic spline equation bounding the first and second derivative to be continuous at the point where x is smallest ($x = 0.046$). This means that the spline equation between the points connected to the inflection point in the y-axis will have a large degree of error. Due to the graph of the raw data points failing the vertical line test, it is not a function and not suitable for cubic interpolation.

5 Advection Differential Equations

5.1 Graphs

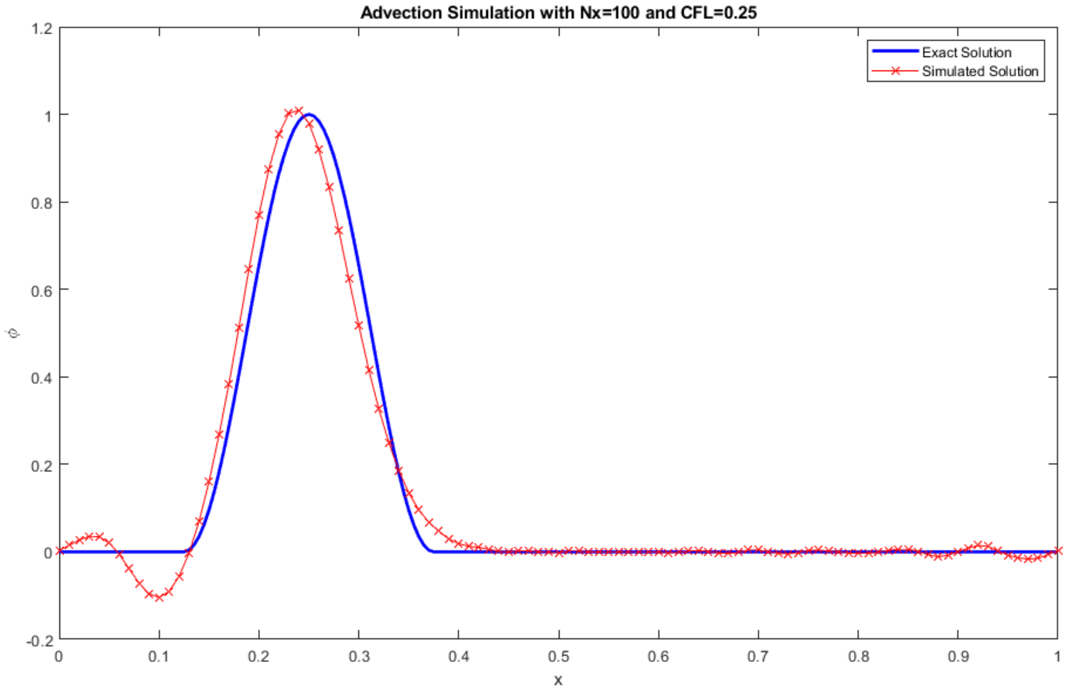


Figure 8: Provided data points

The simulated solution in figure 8 has a large amplitude oscillation from $x = 0$ to $x = 0.14$ that deviates from the exact solution. The phi values seems to match the general shape of the exact solution from $x = 0.14$ to $x = 0.35$ where the wave occurs, however, there is a noticeable phase shift in the simulated solution to the left and a slightly larger peak value when comparing to the exact solution. There is also a error that occurs when the wave ends at $x = 0.34$ to $x = 0.42$. The simulated solution is relatively accurate from $x = 0.43$ to $x = 1$, however, there are oscillations that gradually grow in amplitude. Overall this simulation with Nx=100 and CFL=0.25 is not an accurate one.

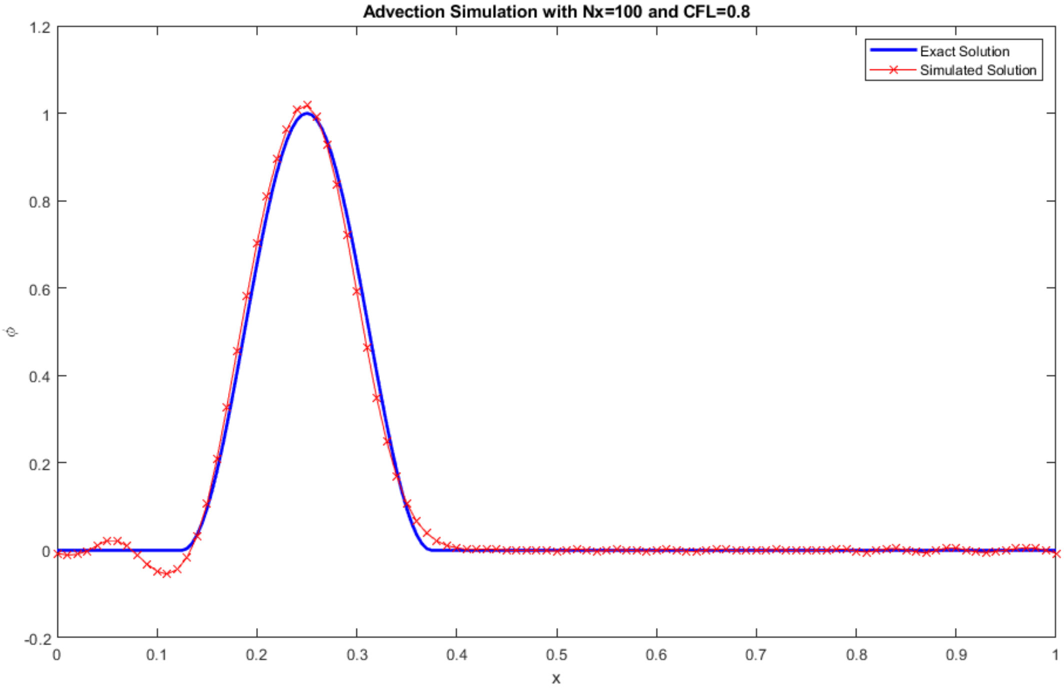


Figure 9: Provided data points

The simulated solution in figure 9 still has substantial oscillation from $x = 0$ to $x = 0.14$ that deviates from the exact solution, however, the amplitudes are lower than the ones observed in figure 8 in this region. The phi values seems to match the general shape of the exact solution from $x = 0.14$ to $x = 0.35$ where the wave occurs, however, there is a small phase shift in the simulated solution to the left when comparing to the exact solution, albeit much less distinct than the phase shift in figure 8. There is also a error that occurs when the wave ends at $x = 0.34$ to $x = 42$ that is much less pronounced than in figure 8. Also, the peak value of the simulated solution is slightly larger than the exact solution but it is again less pronounced than in figure 8. The simulated solution is accurate from $x = 0.43$ to $x = 1$with minor oscillations that slowly grows in amplitude. Overall this simulation with Nx=100 and CFL=0.8 is not very accurate and the features are very similar to the one in figure 8, but with a slightly better accuracy.

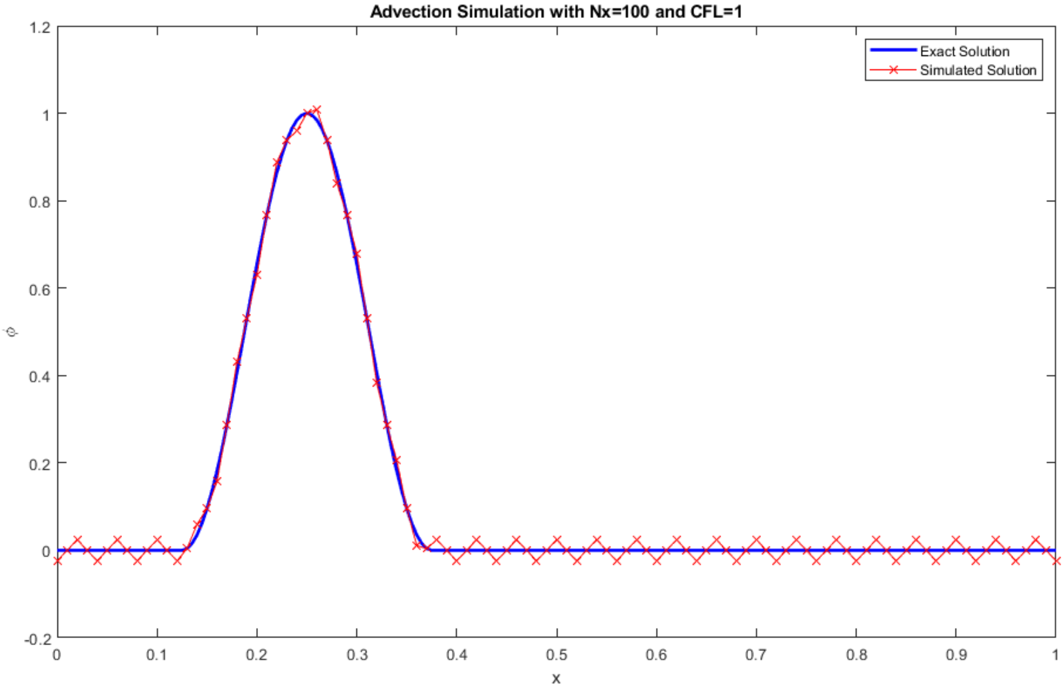


Figure 10: Provided data points

The simulated solution in figure 10 create a very accurate simulation of the wave from $x = 0.14$ to $x = 0.36$, with only a slight error at the peak of the wave. However, there are substantial oscillations of constant amplitude and frequency outside this range. Overall, this simulation with Nx=100 and CFL=1 provides an accurate representation of the wave but has the worst errors elsewhere where the function is constant.

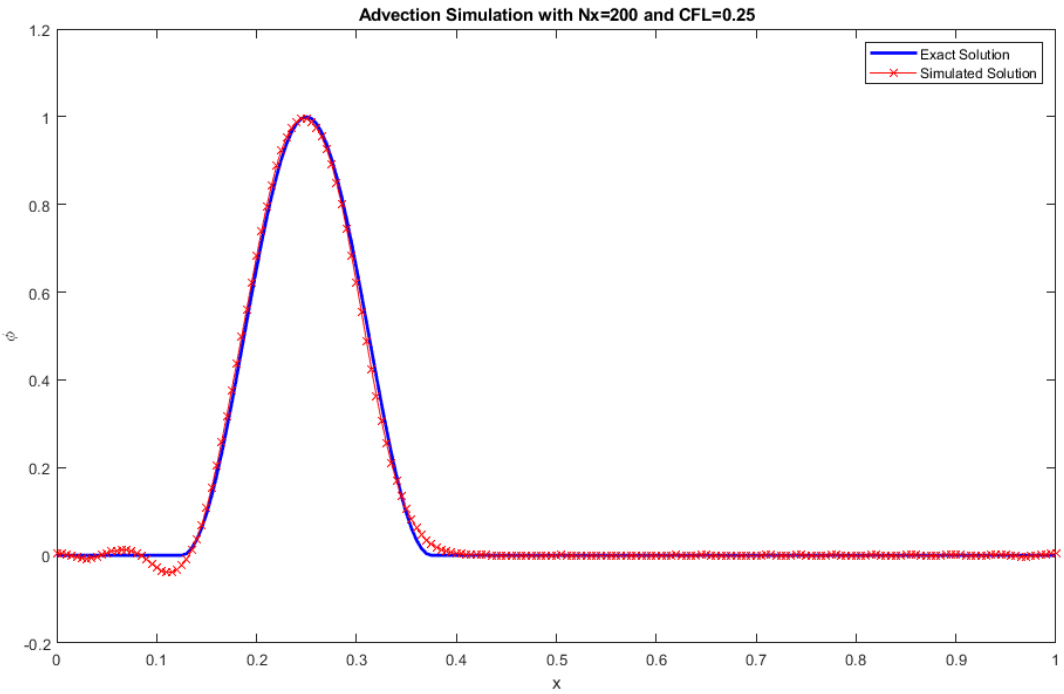


Figure 11: Provided data points

The simulated solution in figure 11 has growing oscillation from $x = 0$ to $x = 0.14$ that deviates from the exact solution and slight errors after the wave ends from $x = 0.34$ to $x = 0.38$. However, the simulated solution is very accurate elsewhere. This simulation with Nx=200 and CFL=0.25 only has slight errors and generally more accurate than the Nx=100 simulations.

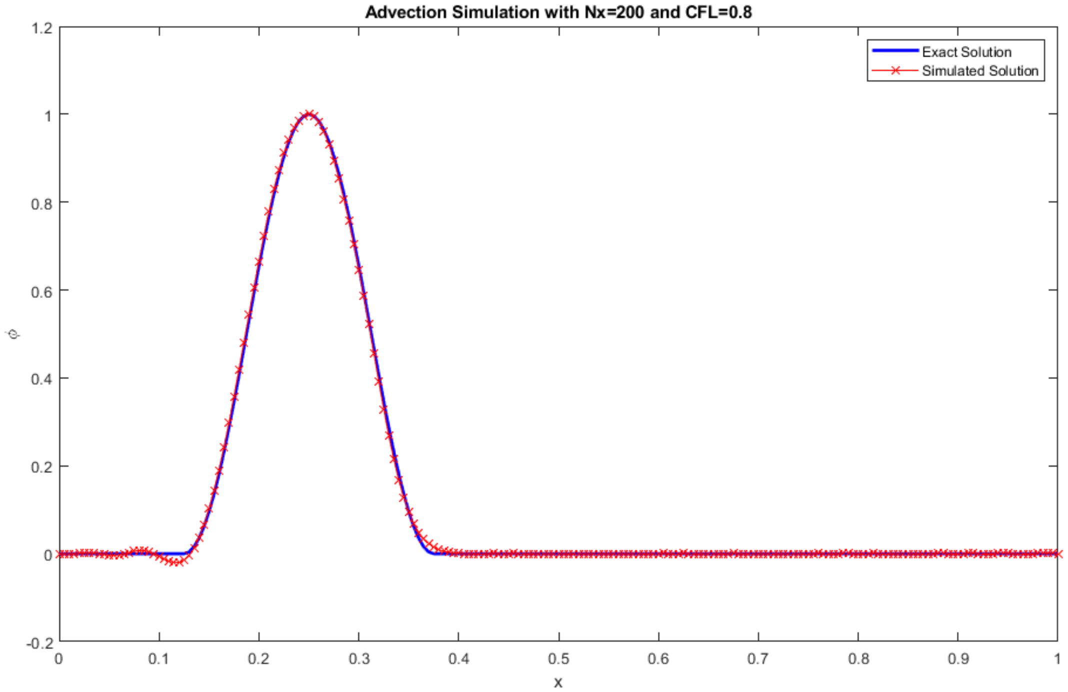


Figure 12: Provided data points

The simulated solution in figure 12 has small errors from $x = 0$ to $x = 0.14$ with oscillations that deviates from the exact solution, and very slight errors after the wave ends from $x = 0.36$ to $x = 0.38$. However, the simulated solution is very accurate elsewhere. This simulation with Nx=200 and CFL=0.8 only has slight errors and is the most accurate solution yet with a very accurate representation of the peak and only slight inaccuracies at the start and the end of the wave.

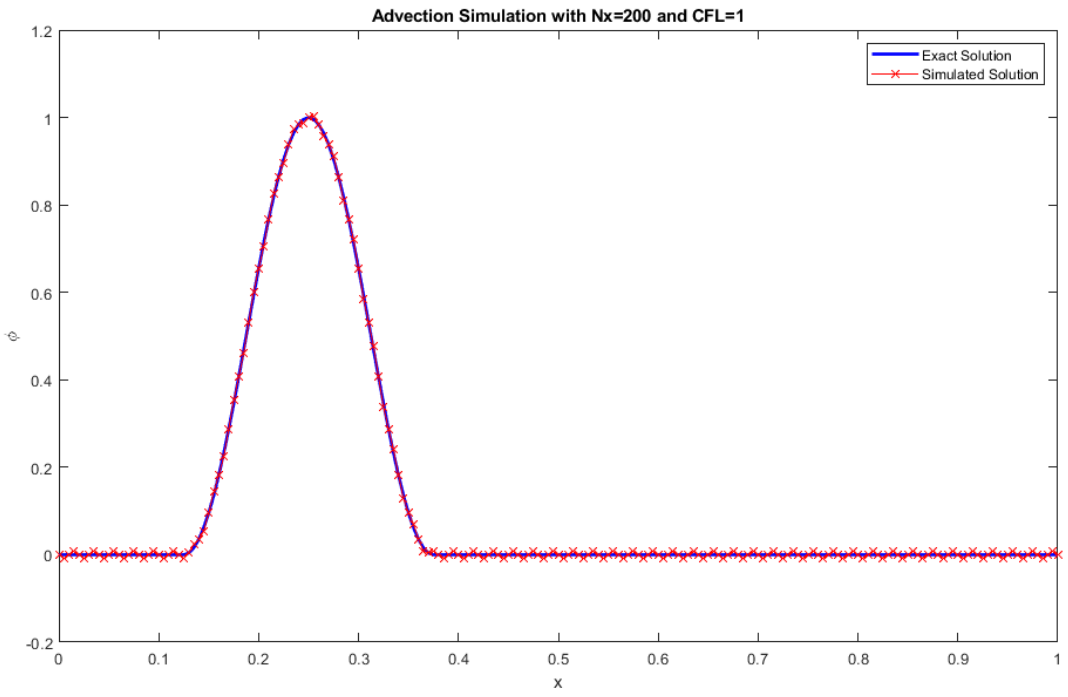


Figure 13: Provided data points

The simulated solution in figure 13 almost perfectly matches the exact solution, however, the oscillations of constant amplitude and frequency where the function is constant still occurs much like figure 10, which shows a pattern that occurs when using a CFL value of 1. These oscillations have a much lower amplitude than figure 10. Overall, the simulation with Nx=200 and CFL=1 is very accurate with the only issue being it's low amplitude, high frequency oscillations.

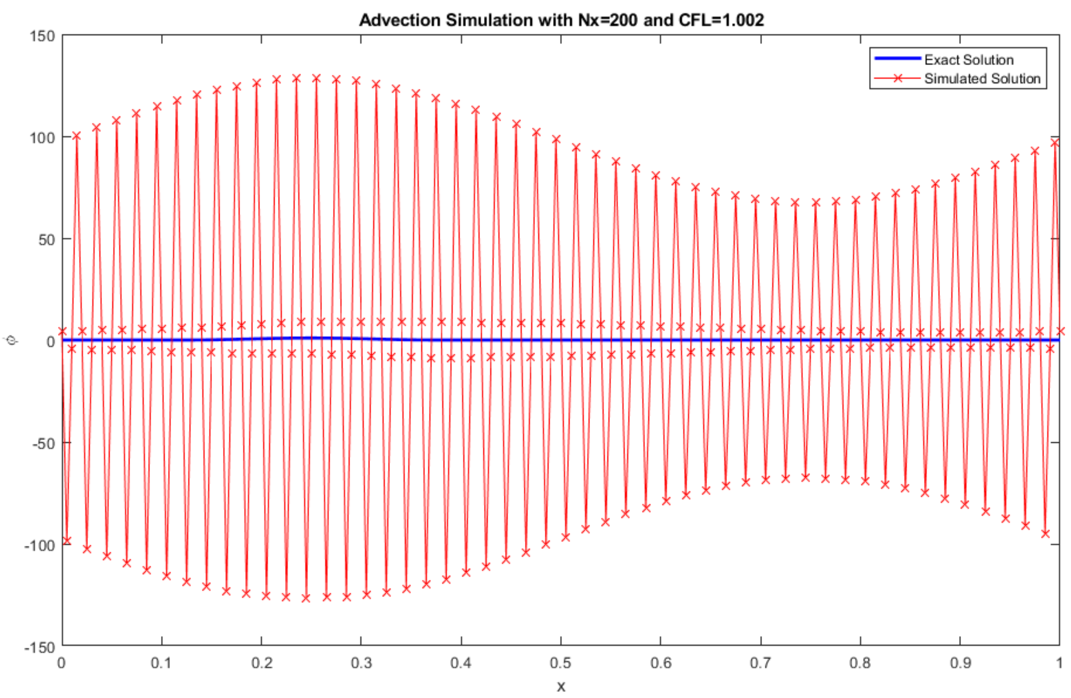


Figure 14: Provided data points

With a CFL=1.002 > 1, the simulation is not stable and diverges as the discrete time steps Δt becomes greater than the time taken for the wave to travel to adjacent grid points. This leads to a simulated solution that has oscillatory behaviour common to CFL=1, however, these oscillation have amplitudes that are much greater (factor>100) and occur throughout the entire domain $x = 0$ to $x = 1$ at a low frequency.

5.2 Discussion

To summarise, increasing the number of sampled x points (reducing Δx) makes the simulation much more accurate especially when simulating the wave. The CFL value effects the oscillation properties of the simulation, with CFL=1 creating oscillations that are constant in frequency and amplitude that occur where the solution should be constant. A CFL<1 creates gradually increasing oscillations, with a low CFL like 0.25 having the greatest growth of oscillations and the greatest deviation error from the exact solution before the wave. The frequency and amplitude of these oscillations changes with the number of sampled x points, with a lower Δx reducing the amplitude and frequency and subsequently, reducing the error. When the CFL>1 as seen in figure 14, the simulation is no longer stable and the simulation oscillates at very high amplitudes and at a low frequency.