Exploring the numerical stability and accuracy of several computational methods for solving the advection and inviscid Burgers' equations

James Portier

Date: April 2024

Abstract

This project focuses on the numerical analysis of the advection and inviscid Burgers' equations, fundamental to modeling transport phenomena in various scientific and engineering fields. Through numerical simulations, we explored the stability and accuracy of three distinct computational methods for the advection equation: the First Upwinding Differencing Method, the Lax-Wendroff Method, and the Euler BTCS Method. Each method was assessed under varying Courant numbers to understand its impact on the solution's stability and precision. We also investigated the advantages of adjusting the time and space step sizes to maintain a Courant number equal to unity, which significantly enhanced the accuracy of the Euler BTCS method. Additionally, we examined the Lax and MacCormack methods in solving the nonlinear inviscid Burgers' equation, highlighting the challenges in capturing shock formations and evaluating method efficacy in handling steep gradients and discontinuities. The findings underscore the significance of method selection and parameter tuning, particularly the synchronization of step sizes to achieve a unit Courant number, in obtaining accurate numerical solutions and offer insights into the behavior of linear versus nonlinear PDEs under different numerical schemes.

Keywords: Advection Equation, Inviscid Burgers' Equation, Courant Number, First Upwinding Method, Lax-Wendroff Method, Euler BTCS Method, Lax Method, MacCormack Method, Stability Analysis.

Contents

1 Introduction

In this section, we outline the problems to be solved in this project, some applications of them, and the numerical methods that will be used to solve them.

1.1 Problems to be solved

1.1.1 Advection equation (Linear problem)

Our first problem is the advection equation, which is given by:

$$\frac{\partial u}{\partial t} = -a\frac{\partial u}{\partial x} \tag{1}$$

with $a = 300\,\text{m/s}$ here. We'll take as initial conditions

$$u(x,0) = \begin{cases} 0 & \text{for } 0 \leq x < 50, \\ 100\sin\left(\frac{\pi(x-50)}{60}\right) & \text{for } 50 \leq x < 110, \\ 0 & \text{for } 110 \leq x \leq 300, \end{cases}$$

and boundary conditions,

$$u(0,t) = 0, \quad u(300,t) = 0.$$

There exists an *exact solution* for this advection equation, given by:

$$u(x,t) = u_0(x - at) \tag{2}$$

with $u_0(x) = u(x,0)$ and $a = 300\,\text{m/s}$. Our Python implementation will include a function for this exact solution, and will be used to quantify the errors of our numerical methods.

Applications of the Advection Equation

The advection equation, as formulated here, has numerous applications across various fields of science and engineering. One of the primary areas of application is in fluid dynamics, where it is used to model the transport of scalar quantities such as temperature, pollutants, or suspended particles in a flowing fluid. This is particularly useful in environmental engineering for predicting the dispersion of pollutants in air or water bodies. Meteorology is another field where the advection equation plays a critical role. It is used to predict the movement of heat, moisture, and pollutants in the atmosphere, which are essential for weather forecasting and climate models. Additionally, the equation is applicable in astrophysics to model the transport of heat and matter in different scales of the universe. In medical imaging, advection equations help in modeling the spread of substances through various tissues, which is crucial for techniques like dynamic contrast-enhanced imaging. Each of these applications demonstrates the importance of understanding and solving the advection equation to provide accurate predictions that are crucial for planning, safety, and efficient design in engineering and scientific endeavors.

1.1.2 Inviscid Burgers' equation (Non-linear problem)

In this report, we also consider the inviscid Burgers' equation, given by:

$$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x} \tag{3}$$

and using the initial conditions

$$u(x,0) = \begin{cases} 1 & \text{for } 0 \leq x < 2, \\ 0 & \text{for } 2 \leq x \leq 4, \end{cases}$$

and the boundary conditions

$$u(0,t) = 1, \quad u(4,t) = 0,$$

where $0 \leq x \leq 4$. It is useful at this point to introduce the inviscid Burgers' equation in conservative form:

$$\frac{\partial u}{\partial t} = -\frac{\partial}{\partial x}\left(\frac{u^2}{2}\right) = -\frac{\partial E}{\partial x} \tag{4}$$

where $E = \frac{u^2}{2}$. Burgers' equation (3) can be interpreted as the propagation of a wave with each point on the wave potentially having a different velocity and eventually therefore forming a discontinuity in the computational domain. The conservative form of the Burgers' equation, as presented in Equation (4), is particularly advantageous for numerical simulations involving shock waves or discontinuities. By expressing the equation in terms of the flux function E, we ensure the conservation of integral quantities across shock waves, which is a key physical property in systems governed by conservation laws. This form is adept at handling the abrupt changes in the solution and is essential for the accurate numerical resolution of shocks, thereby enabling a more stable and physically consistent computational approach when discontinuities are present.

Applications of the Inviscid Burgers' Equation

The inviscid Burgers' equation, despite its simplicity, serves as a fundamental model for understanding more complex phenomena in fluid dynamics and nonlinear wave theory. In *fluid dynamics*, the inviscid Burgers' equation is particularly useful for modeling the formation and propagation of *shock waves* in compressible fluids. The equation captures the essential nonlinear aspect of shock formation, where waves steepen and eventually merge into discontinuities, a critical behavior in the analysis of supersonic jet flows and explosions. This equation also serves as a basic model for certain types of *traffic flow analysis* where traffic is treated as a continuous fluid. Here, the velocity field represents the traffic speed, and the equation models the evolution of traffic density. This application is especially useful in understanding how traffic jams form, evolve, and how patterns of congestion can emerge from initially smooth traffic flow conditions. These applications underscore the importance of the inviscid Burgers' equation as a prototypical model for various complex systems involving nonlinear wave propagation and discontinuities.

1.2 Methods of solution for the Advection equation

Before introducing the numerical methods we'll use to solve our two problems, we introduce the *Courant number*, c, a dimensionless parameter critical in the analysis of numerical stability and accuracy. It is defined by:

$$c = \frac{a\Delta t}{\Delta x} \tag{5}$$

The Courant number essentially governs the time step size relative to the spatial step size to ensure stable and accurate numerical simulations.

1.2.1 The Lax-Wendroff method (explicit)

The *Lax-Wendroff* method, an explicit second-order numerical technique, is designed to provide improved accuracy for solving hyperbolic partial differential equations. It effectively captures wave prop-

agation phenomena by incorporating higher-order spatial and temporal derivatives. The method is described by the equation:

$$U_i^{n+1} = U_i^n - \frac{c}{2}(U_{i+1}^n - U_{i-1}^n) + \frac{c^2}{2}(U_{i-1}^n - 2U_i^n + U_{i+1}^n). \tag{6}$$

This method is of order $O((\Delta t)^2, (\Delta x)^2)$. It can be shown that the method is stable for $c \leq 1$ - in fact, we will show this later in Section 2. This method will be used to solve the Advection equation problem (1.1.1).

1.2.2 The First Upwinding Differencing Method (explicit)

The First Upwinding Differencing Method is a straightforward, explicit numerical scheme utilised primarily for its simplicity and robustness in handling advection problems. It is of first order in both space and time ($O(\Delta x, \Delta t)$) and employs a backward difference to approximate the spatial derivative. It's described by the following equation:

$$U_i^{n+1} = U_i^n - c(U_i^n - U_{i-1}^n). \tag{7}$$

Similar to the Lax-Wendroff method, it can be shown that this method is also stable for $c \leq 1$.

1.2.3 The Euler BTCS method (implicit)

The *Euler Backward Time, Central Space* (BTCS) method is an implicit numerical technique distinguished by its robust stability characteristics and ability to handle 'stiff' problems. It employs backward differencing for temporal discretisation and central differencing for spatial derivatives, offering a blend of accuracy and stability. This method, when applied to the model equation (1), is represented as follows:

$$\frac{c}{2}U_{i-1}^{n+1} - U_i^{n+1} - \frac{c}{2}U_{i+1}^{n+1} = -U_i^n \tag{8}$$

This method is *unconditionally stable* and is of order $O(\Delta t, \Delta x^2)$. Due to its implicit nature, this equation is applied to all grid points at each time step to yield a linear system of equations to solve. In this case, the coefficient matrix of the linear system is in tri-diagonal form. In order to efficiently solve this within our implementation, we use the `solve_banded` function.

Banded Solver

The `solve_banded` function from SciPy is optimised for solving linear systems with banded coefficient matrices and offers several advantages over the general `solve` function. It provides memory efficiency by requiring less memory, as only the non-zero diagonal bands need storage. In terms of computational speed, algorithms for banded matrices exploit the sparse structure for faster computations. Algorithmic optimisation is achieved through specialised methods, such as the LU decomposition for banded matrices, which are simpler and more efficient. Additionally, improved stability is ensured as tailored algorithms can offer better numerical stability by minimising operations prone to round-off errors. Since the Euler BTCS method involves a tri-diagonal system, we utilise the solve banded function for this method to capitalise on its enhanced performance and efficiency.

1.3 Methods of solution for the Inviscid Burgers' equation

1.3.1 The Lax method (explicit)

If a spatial average value of U_i^n is used in a typical FTCS method, i.e.,

$$U_i^n = \frac{U_{i-1}^n + U_{i+1}^n}{2}.$$

The Finite Difference Equation (FDE) then takes the form:

$$U_i^{n+1} = \frac{1}{2}(U_{i-1}^n + U_{i+1}^n) - \frac{c}{2}(U_{i+1}^n - U_{i-1}^n)$$

which is known as the *Lax method*. For the conservative form in Equation (4), this becomes:

$$U_i^{n+1} = \frac{1}{2}(U_{i+1}^n + U_{i-1}^n) - \frac{\Delta t}{2\Delta x}(E_{i+1}^n - E_{i-1}^n) \tag{9}$$

where $E_i^n = \frac{(U_i^n)^2}{2}$ represents the energy term at the grid points.

The Lax method is an explicit numerical scheme commonly used to stabilize numerical solutions by averaging adjacent spatial points, and is first-order accurate in time and second-order accurate in space. A Von Neumann stability analysis can show this method to be stable when $c \leq 1$.

1.3.2 The MacCormack method (explicit)

Some methods, including the ones we introduced previously, may struggle when it comes to solving non-linear problems such as the Inviscid Burgers' equation. There is a class of methods that make use of the finite difference equations at various points between t^n and t^{n+1}. These methods are known to work well with nonlinear hyperbolic equations, and are referred to as *predictor-corrector* methods. The *MacCormack* method is one such method, and is formulated as:

Predictor step:

$$U_i^* = U_i^n - c(U_{i+1}^n - U_i^n) \tag{10}$$

Corrector step:

$$U_i^{n+1} = \frac{1}{2}(U_i^n + U_i^*) - c(U_i^* - U_{i-1}^*) \tag{11}$$

For the conservative form in Equation (4) of the Inviscid Burgers' equation, this becomes:

Predictor step:

$$U_i^* = U_i^n - \frac{\Delta t}{\Delta x}(E_{i+1}^n - E_i^n) \tag{12}$$

Corrector step:

$$U_i^{n+1} = \frac{1}{2}\left(U_i^n + U_i^*\right) - \frac{\Delta t}{\Delta x}(E_i^* - E_{i-1}^*) \tag{13}$$

where, $E_i^n = \frac{(U_i^n)^2}{2}$ again, and for an intermediate state U^*, $E_i^* = \frac{(U_i^*)^2}{2}$, representing the energy terms at the grid points. The method is second order accurate in both time and space, and again, the stability requirement for this method is that $c \leq 1$ must be satisfied. Being a predictor-corrector method, it is well suited for nonlinear equations and is therefore fairly popular in computational fluid mechanics. Additionally, the order of differencing can be reversed at each time step, i.e., forward/backward followed by backward/forward, to enhance stability and accuracy in certain scenarios.

2 Stability Analysis

We now perform a technique called *Von Neumann Stability Analysis* in order to get constraints for stability for two of our methods - *Lax-Wendroff* and *MacCormack*. With the Von Neumann technique, the solution of the finite difference equation is expanded in a Fourier series. The decay or growth of the *amplification factor* obtained from this expansion indicates whether or not the numerical algorithm is stable.

2.1 Lax-Wendroff method

Recall the Lax-Wendroff method formulation:

$$U_i^{n+1} = U_i^n - \frac{c}{2}(U_{i+1}^n - U_{i-1}^n) + \frac{c^2}{2}(U_{i-1}^n - 2U_i^n + U_{i+1}^n).$$

where c is the Courant number. We next set $U_i^n \to e_i^n := \xi^n e^{l\theta i}$, where l is the imaginary unit. Assuming a perturbation of the form $e_i^n = \xi^n e^{l\theta i}$, and substituting this into the equation above, we get:

$$\begin{aligned}
\Leftrightarrow \xi^{n+1}e^{l\theta i} &= \xi^n e^{l\theta i} - \frac{c}{2}\left(\xi^n e^{l\theta(i+1)} - \xi^n e^{l\theta(i-1)}\right) + \frac{c^2}{2}\left(\xi^n e^{l\theta(i-1)} - 2\xi^n e^{l\theta i} + \xi^n e^{l\theta(i+1)}\right) \\
\Leftrightarrow \frac{\xi^{n+1}}{\xi^n} &= 1 - \frac{c}{2}\left(e^{l\theta} - e^{-l\theta}\right) + \frac{c^2}{2}\left(e^{-l\theta} - 2 + e^{l\theta}\right) \\
\Leftrightarrow \frac{\xi^{n+1}}{\xi^n} &= 1 - c^2 - \frac{c}{2}\left(e^{l\theta} - e^{-l\theta}\right) + \frac{c^2}{2}\left(e^{-l\theta} + e^{l\theta}\right) \\
\Leftrightarrow \frac{\xi^{n+1}}{\xi^n} &= 1 - c^2 - lc\sin(\theta) + c^2\cos(\theta) \\
\Leftrightarrow \frac{\xi^{n+1}}{\xi^n} &= 1 - lc\sin(\theta) + c^2(\cos(\theta) - 1) \\
&= \underbrace{1 - c^2\left(1 - \cos(\theta)\right)}_{\text{Real part}} - \underbrace{(c\sin(\theta))l}_{\text{Imaginary part}} \qquad (\star)
\end{aligned}$$

Note that $G = \frac{\xi^{n+1}}{\xi^n}$ is the Amplification factor - for convergence, we require $|G| \leq 1$, i.e.:

$$\begin{aligned}
\left|\frac{\xi^{n+1}}{\xi^n}\right| &= \sqrt{(1 - c^2(1 - \cos(\theta)))^2 + c^2\sin^2(\theta)} \leq 1 \\
&\Leftrightarrow (1 - c^2(1 - \cos(\theta)))^2 + c^2\sin^2(\theta) \leq 1 \\
&\Leftrightarrow 1 - 2c^2(1 - \cos(\theta)) + c^4(1 - \cos(\theta))^2 + c^2\sin^2(\theta) \leq 1 \\
&\Leftrightarrow -2(1 - \cos(\theta)) + c^2(1 - \cos(\theta))^2 + \sin^2(\theta) \leq 0 \\
&\Leftrightarrow 2\cos(\theta) + c^2(1 - \cos(\theta))^2 + \sin^2(\theta) - 2 \leq 0 \\
&\Leftrightarrow 2\cos(\theta) + c^2(1 - \cos(\theta))^2 - \cos^2(\theta) - 1 \leq 0 \\
&\Leftrightarrow (\cos(\theta) - 1)^2(c^2 - 1) \leq 0
\end{aligned}$$

Note that if $\cos(\theta) = 1$, then $\cos(\theta) - 1 = 0$, and so the stability condition trivially becomes $0 \leq 0$. For all other values of $\cos(\theta)$ however, $(\cos(\theta) - 1)^2 > 0$, and so for the expression to hold we must have $c^2 - 1 \leq 0 \Rightarrow c \leq 1$

Thus, the Von Neumann criterion for stability of the Lax-Wendroff scheme is:

$$\boxed{c \leq 1}$$

2.2 MacCormack method

Recall the MacCormack formulation:

$$\text{Predictor step: } U_i^* = U_i^n - c(U_{i+1}^n - U_i^n),$$
$$\text{Corrector step: } U_i^{n+1} = \frac{1}{2}[(U_i^n + U_i^*) - c(U_i^* - U_{i-1}^*)].$$

Combining these, we get:

$$\begin{aligned} 2U_i^{n+1} &= U_i^n + \left(U_i^n - c(U_{i+1}^n - U_i^n)\right) - c\left(U_i^n - c(U_{i+1}^n - U_i^n) - (U_i^n - c(U_i^n - U_{i-1}^n))\right) \\ &= 2U_i^n - c(U_{i+1}^n - U_{i-1}^n) + c^2(U_{i+1}^n - 2U_i^n + U_{i-1}^n) \end{aligned}$$

Now substituting $U_i^n \to e_i^n := \xi^n e^{l\theta i}$, where l is the imaginary unit as before, we get:

$$\begin{aligned} \Leftrightarrow 2\xi^{n+1}e^{l\theta i} &= 2\xi^n e^{l\theta i} - c\left(\xi^n e^{l\theta(i+1)} - \xi^n e^{l\theta(i-1)}\right) + c^2\left(\xi^n e^{l\theta(i+1)} - 2\xi^n e^{l\theta i} + \xi^n e^{l\theta(i-1)}\right) \\ \Leftrightarrow \frac{\xi^{n+1}}{\xi^n} &= 1 - \frac{c}{2}\left(e^{l\theta} - e^{-l\theta}\right) + \frac{c^2}{2}\left(e^{l\theta} - 2 + e^{-l\theta}\right) \\ \Leftrightarrow \frac{\xi^{n+1}}{\xi^n} &= 1 - lc\sin(\theta) + c^2(\cos(\theta) - 1) \\ &= 1 - c^2\left(1 - \cos(\theta)\right) - (c\sin(\theta))l \end{aligned}$$

Note the this expression is exactly what we were left with in the Von Neumann stability analysis for the Lax-Wendroff method above (see ($\star$)). Thus, the Von Neumann criterion for stability of the MacCormack is also:

$$\boxed{c \leq 1}$$

3 Results

This section presents the outcomes of numerical experiments conducted on linear and non-linear advection problems, employing various computational methods under different Courant number conditions. We begin by analysing the linear advection equation with an exploration of the effects of the Courant number on numerical stability and accuracy of the First Upwind, Lax-Wenroff, and Euler BTCS methods, and conclude with an examination of the non-linear Inviscid Burgers' equation using the Lax and MacCormack methods.

3.1 Advection equation (Linear problem)

3.1.1 Investigating the effect of varying the Courant number

We now investigate the effect of varying the Courant number on the advection equation for the First Upwinding Differencing (7), Lax-Wendroff (6), and Euler BTCS (8) methods. Keeping the spatial step size fixed at $\Delta x = 5$m, we vary the time step Δt with values in the set $\{0.018, 0.01666, 0.0075\}$. Before we discuss these results, recall that we showed earlier that the requirement for stability for both the first upwind differencing and Lax-Wendroff methods is $c \leq 1$, whereas the Euler BTCS method is theoretically unconditionally stable. Given that a = 300m/s and $\Delta x = 5$m, the Courant numbers for the time step-sizes $\{0.018, 0.01666, 0.0075\}$ are $c = \{1.08, 0.9996, 0.45\}$ respectively. Thus, we expect the first upwind differencing and Lax-Wendroff methods to be unstable for the first time step value, whilst being stable for the other two.

Below are individual figures for each of the three methods compared with the exact solution 2 at the various specified time step values, taken at time t = 0.45.

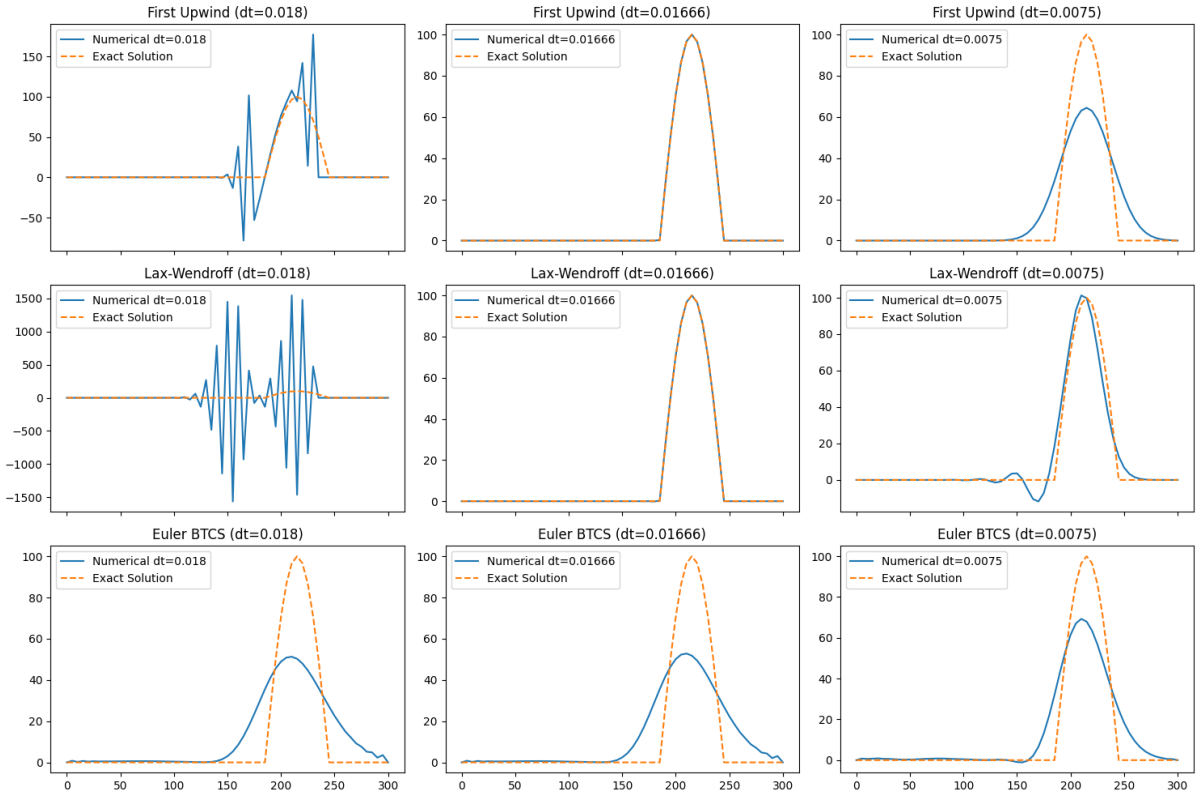

Figure 1: Plot of the Exact vs Numerical solution for each method and various step sizes (separately).

From Figure 1, several observations can be made. For the first time step value of $\Delta t = 0.018$, both the first upwind and Lax-Wendroff methods exhibit significant instability, deviating markedly from the exact solution. This corroborates the theoretical predictions from Von Neumann stability analysis, which posits that these methods become unstable when the Courant number exceeds unity, as is the case for this time step. Conversely, the Euler BTCS method offers a closer approximation to the exact solution, despite the peak's position not aligning precisely with the expected result.

For a reduced time step of $\Delta t = 0.01666$, the first upwind and Lax-Wendroff methods now align closely with the exact solution, displaying an absence of oscillatory errors or numerical instabilities. This aligns with the calculated Courant number for this time step being 0.09996, which is less than or equal to one, satisfying the stability criterion. The Euler BTCS method shows slight improvement with this decreased time step, though its peak still significantly diverges from that of the exact solution and, as such, yields a relatively inferior approximation in comparison to the other methods.

At the smallest time step of $\Delta t = 0.0075$, corresponding to a Courant number of 0.45, an unexpected phenomenon occurs: the first upwind and Lax-Wendroff methods no longer provide a close match to the exact solution. Whilst the peak of the Lax-Wendroff method matches the exact solution pretty closely, oscillatory errors emerge to the left of the peak. As for the first upwind method, the peak is now significantly further away from that of the exact solution. Despite the theoretical assurance of stability for these methods at this Courant number, it's evident that practical discrepancies can still arise. For the first upwind method, this may be due to the intrinsic numerical dissipation of the upwind scheme, which artificially diffuses the wave front, leading to a dampened peak. For the Lax-Wendroff method, the discrepancy may result from the dispersion errors inherent in the method, where the numerical solution can develop non-physical oscillations adjacent to steep gradients, such as those present in the initial condition for this advection equation problem with a significant change in $u(x, 0)$ from 0 to $100 \sin\left(\frac{\pi}{60}(x - 50)\right)$ over a small spatial interval. This underscores the challenges of accurately resolving steep gradients and the sensitivity of numerical outcomes to the chosen discretisation parameters. Finally, despite the smaller step-size leading to a better approximation for Euler BTCS, it still does not fully capture the exact solution.

An interesting additional observation is that the first upwind and Lax-Wendroff methods are the most accurate when the Courant number is the closest to 1 in value - this is something we will discuss in more detail in section 3.1.2, using this observation to improve the accuracy of the Euler BTCS method.

Figure 2 below shows a *combined* plot of each of the methods and the exact solution for the different time step values:

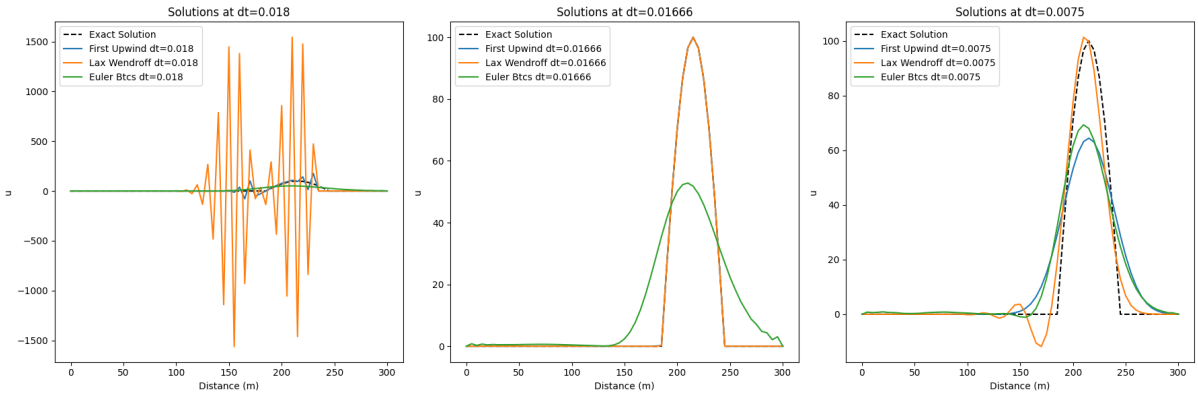

Figure 2: Plot of the Exact vs Numerical solution for each method and various step sizes (combined).

Upon examining the combined plot of the exact and numerical solutions for each method in Figure 2,

we notice that for the initial time-step of $\Delta t = 0.018$, the Lax-Wendroff method displays pronounced oscillatory errors that are substantially larger than those in the other two methods. These considerable oscillations could be attributed to the method's inherent numerical dispersion, which amplifies errors when the Courant-Friedrichs-Lewy (CFL) condition[1] is not satisfied. Specifically, the Lax-Wendroff method is known to produce dispersive waves that travel at incorrect speeds when the time-step is relatively large, as is the case here. Additionally, the steep gradient present in the initial conditions can exacerbate these errors, as the method is sensitive to rapid changes in the solution, leading to the observed overestimation of the wave speed and resulting in the significantly amplified oscillations depicted in the plot. The rightmost plot illustrates that the first upwind method's peak is in fact further away from that of the exact solution when compared to the Euler BTCS method, further accentuating the diminished accuracy of the first upwind method for this reduced time-step value. This discrepancy underscores a notable decrease in performance, as the first upwind method's numerical solution becomes less precise in capturing the peak's location, which is crucial for accurately reflecting the dynamics of the advection equation.

Collectively, these factors underscore the complexity of numerical methods and the importance of empirical testing alongside theoretical analysis. Finally, we quantify the discrepancies between the numerical and exact solutions by calculating the Mean Absolute Error (MAE) and the Infinity Norm, which serve as metrics for the average magnitude of errors and the maximum error across the spatial domain, respectively.

Method	**MAE**	**Infinity Norm Error**
First Upwind with $dt = 0.018$	10.7195	106.5410
First Upwind with $dt = 0.01666$	0.0354	0.2794
First Upwind with $dt = 0.0075$	6.6543	35.5831
Lax-Wendroff with $dt = 0.018$	284.5710	1561.9515
Lax-Wendroff with $dt = 0.01666$	0.0406	0.4160
Lax-Wendroff with $dt = 0.0075$	3.0799	18.5692
Euler BTCS with $dt = 0.018$	9.6738	49.7543
Euler BTCS with $dt = 0.01666$	9.3439	48.2184
Euler BTCS with $dt = 0.0075$	6.2718	33.0185

Table 1: Summary of errors for each method and the different time steps.

The data presented in Table 1 provides a comprehensive quantification of the numerical errors associated with each method at varying time steps. The First Upwind and Lax-Wendroff methods show a dramatic decrease in MAE and the Infinity Norm as the time step is reduced from 0.018 to 0.01666, indicating a significant improvement in accuracy when the Courant number is lower than unity and the methods are theoretically stable. The exceedingly high errors at $dt = 0.018$ exhibited by the Lax-Wendroff method are quantified by the large MAE and the error under the Infinity Norm, reflecting the method's instability at this time-step. However, when the time step size is further reduced to 0.0075, the errors get worse. Overall, the explicit methods seem to be more appropriate for the advection equation, in particular, the best of which was found to be the First Upwind method. We conclude that the First Upwind method is the most suitable because it has the lowest error out of all the methods when using the most suitable Courant number, and it is also "less unstable" than the Lax-Wendroff method when the stability constraint is violated. However, if different initial conditions were to be used, it's possible that the BTCS method would be preferred due to its unconditional stability.

[1]The CFL condition stipulates that the numerical domain of dependence must include the physical domain of dependence for stability. In practical terms, this often means the time step must be small enough relative to the spatial step size and wave speeds.

Finally, we can see that the Euler BTCS method followed a different error reduction trend compared to the other methods, with a consistent decrease in MAE and error under the Infinity Norm as the time step is decreased. This is because the error "benefit" of decreasing the step-size, and consequently bringing the 'peak' of the numerical approximation closer to that of the exact solution, is outweighing the error introduced by the dispersion errors. This motivates us to try to improve the Euler BTCS method by further decreasing the time step-size - we take an in-depth look at this in the subsequent section.

3.1.2 Improving the BTCS method

From the Euler BTCS plots in Figure 1, we can see that the numerical solution is quite far off the exact solution compared to the best numerical solutions of the other two methods. As dt is decreased, the numerical solution does seem to improve - thus, a natural way to improve the BTCS method could be to further decrease the step-size (spoiler: it isn't as simple as this). Below is a plot of the exact solution vs the Euler BTCS method approximation for various values of dt (progressively decreasing it by a factor of a half):

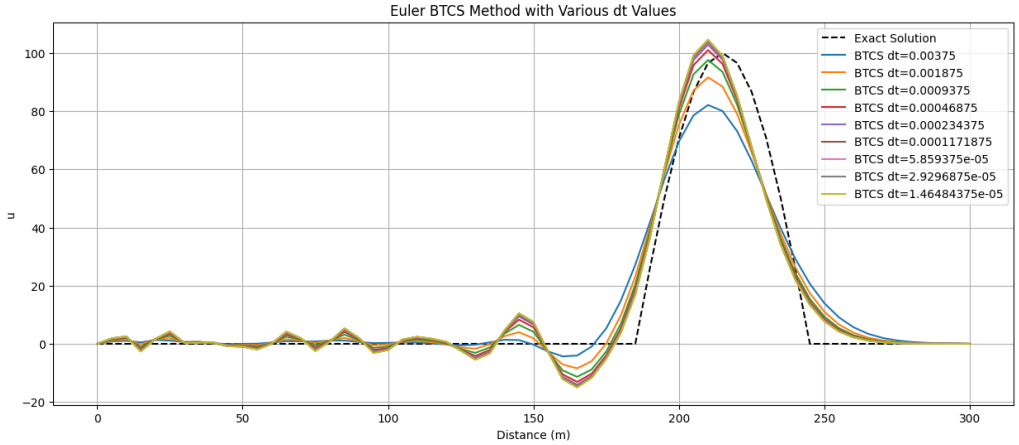

Figure 3: Plot of the exact solution vs the Euler BTCS method approximation for decreasing values of the time step-size.

dt	**MAE**	**Infinity Norm Error**
0.00375	4.6078	27.2017
0.001875	4.0329	23.1359
0.0009375	4.2583	20.3903
0.00046875	4.5686	20.3614
0.000234375	4.7579	20.6694
0.0001171875	4.8625	20.8406
5.859375e-05	4.9176	20.9306
2.9296875e-05	4.9458	20.9768
1.46484375e-05	4.9601	21.0001

Table 2: Numerical Errors for the Euler BTCS Method at various time steps.

The graphical representation of the Euler BTCS method with various dt values, shown in Figure 3, reveals that while the peak of the BTCS method's solution approaches that of the exact solution as the time step dt decreases, dispersion errors on the left-hand side (LHS) of the plot become more pronounced.

An analysis of the tabulated errors confirms this observation: the Mean Absolute Error (MAE) and Infinity Norm initially decrease as dt is reduced, indicative of an improved approximation of the peak. However, past a certain threshold, these errors begin to increase, due to the error introduced by the dispersion errors outweighing the benefit of the higher peak. This behavior is indicative of a trade-off when selecting dt. While smaller time steps generally lead to higher resolution and potentially more accurate solutions, they can also introduce greater numerical errors due to increased computational iterations and the consequent accumulation of round-off errors.

Recall that the first unwinding and Lax-Wendroff method had the best accuracy when the Courant number was close to 1 - as it turns out, doing this for the Euler BTCS method whilst decreasing the step-size provides a better way to improve the method. Recall that Courant number is given by $c = \frac{a\Delta t}{\Delta x}$. Thus, for the Courant number to be close to 1, we require that $\Delta x \approx a\Delta t$. To achieve this, in addition to decreasing the time step-size, we also modify the Δx to be 300* Δt. The results of this are presented in the Figure 4 and Table 3 below:

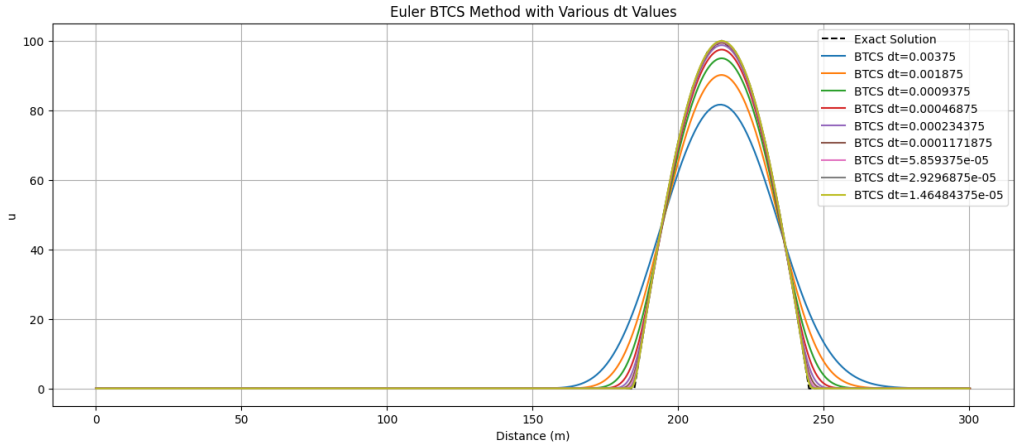

Figure 4: Plot of the Exact vs Numerical solution for the Euler BTCS method and various step sizes, maintaining the Courant number to be unity.

dt	**MAE**	**Infinity Norm Error**
0.00375	3.5888	21.8270
0.001875	2.0632	16.9454
0.0009375	1.1327	12.3073
0.00046875	0.6013	8.8505
0.000234375	0.3124	6.3618
0.0001171875	0.1600	4.5202
5.859375e-05	0.0812	3.2005
2.9296875e-05	0.0409	2.2704
1.46484375e-05	0.0206	1.6067

Table 3: Numerical Errors for the Euler BTCS Method at Decreasing Time Steps and a Courant number of 1.

The progression depicted in Figure 4 and summarised in Table 3, illustrates the enhanced accuracy of the Euler BTCS method as the time step dt is reduced whilst maintaining the Courant number at unity. The alignment of the numerical solution's peak with the exact solution's peak becomes notably more precise with the decreasing dt, a clear indication of improved fidelity in capturing the dynamics of the advection

equation. The corresponding numerical errors, presented in the table, exhibit a consistent decline with the reduction of dt. This trend is evidenced by the Mean Absolute Error, which decreases from 3.5888 to 0.0206, and the Infinity Norm Error, which diminishes from 21.8270 to 1.6067. These results signify not only an average reduction in error across the spatial domain but also a minimisation of the maximum error encountered.

Compared to the optimal Mean Absolute Error of the first upwind and Lax-Wendroff methods listed in Table 1, the MAE for the Euler BTCS method with the smallest dt is notably lower, suggesting enhanced performance. Despite this, the Infinity Norm error for the Euler BTCS method remains relatively high. A further reduction in the time step-size may increase its precision, potentially enabling it to exceed the other methods in terms of Infinity Norm accuracy. However, such increases in precision are accompanied by higher computational demands, leading to a trade-off between accuracy and computational efficiency.

The competitive performance of the first upwind and Lax-Wendroff methods, achieved with considerably larger step-sizes, underscores their potential suitability for efficiently solving this advection equation. The promise shown by these methods suggests that applying the technique of reducing Δt while proportionally modifying Δx to ensure the Courant number equals 1, as successfully implemented for the Euler BTCS method, might also elevate their accuracy. Such an adjustment, aimed at maintaining stability and precision, could render these methods even more effective for addressing the nuances of the problem.

Maintaining the Courant-Friedrichs-Lewy (CFL) number close to 1, by simultaneously adapting the time step Δt and the spatial step Δx, enhances the accuracy of numerical methods for PDEs in several ways. Firstly, it allows the numerical solution to advance the wave in a manner that closely emulates its physical propagation, thus improving accuracy. It also ensures that grid points being used for numerical approximation line up exactly with those used in the exact solution - when they're misaligned, numerical dispersion errors are introduced, which is why decreasing dt is making the dispersion error worse in the original attempt at improving the Euler BTCS method. Finally, this alignment between the numerical and physical domains of dependence results in a more faithful representation of the actual wave behavior. Such an approach mitigates numerical dispersion, reducing the artificial spread of the wave and yielding more precise wave-forms.

3.2 Inviscid Burgers' equation (Non-linear problem)

We now discuss results relating to our second problem — the Inviscid Burgers' equation. Recall that we tackle this problem using two numerical methods, namely the Lax method and the MacCormack method. We investigate their efficacy by producing three plots for each method, one for each $\Delta t \in \{0.14, 0.1, 0.05\}$, and plotted at times 1.2, 1.8, and 2.4 (or as close to these values as possible), keeping $\Delta x = 0.1$ in all cases. Note that each time step in solving the Inviscid Burgers' equation can be viewed as addressing an advection problem where the initial value U^0 and the progression of the numerical solution help anticipate behavior at subsequent times T. When the Courant number $C^0 > 1$, instabilities may arise due to nonlinear growth, potentially leading to an increase in C^n as $\max_i(U_i^n)$ grows. Conversely, stability is expected when $C^0 \leq 1$, since the solution behaves conservatively with $\max_i(U_i^n) \leq \max_i(U_i^{n-1})$. This condition mimics the stability considerations typical of advection equations, effectively linearizing the Inviscid Burgers' equation at each time step.

Thus, the Courant numbers for the time step-sizes $\{0.14, 0.1, 0.05\}$ are $c = \{1.4, 1, 0.5\}$ respectively. Based on the constraints from Von Neumann analysis ($c \leq 1$ for both methods for stability), we expect the two methods to be unstable for the first time-step size, and stable for the other two. Finally, recall that there is no exact solution provided for us to compare our numerical methods with - Figure 5 shows a plot of shock waves that we should expect to see to aid our interpretation of the results.

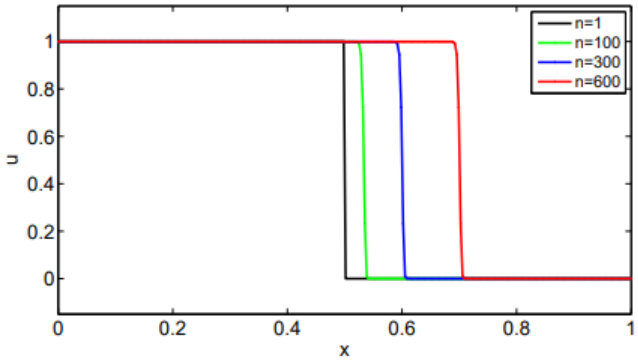

Figure 5: Rough plot of the shock waves we that expect to see for the Burger's equation (note that the x-axis/n values are slightly different to our set-up though). Taken from [2].

3.2.1 Lax method simulations

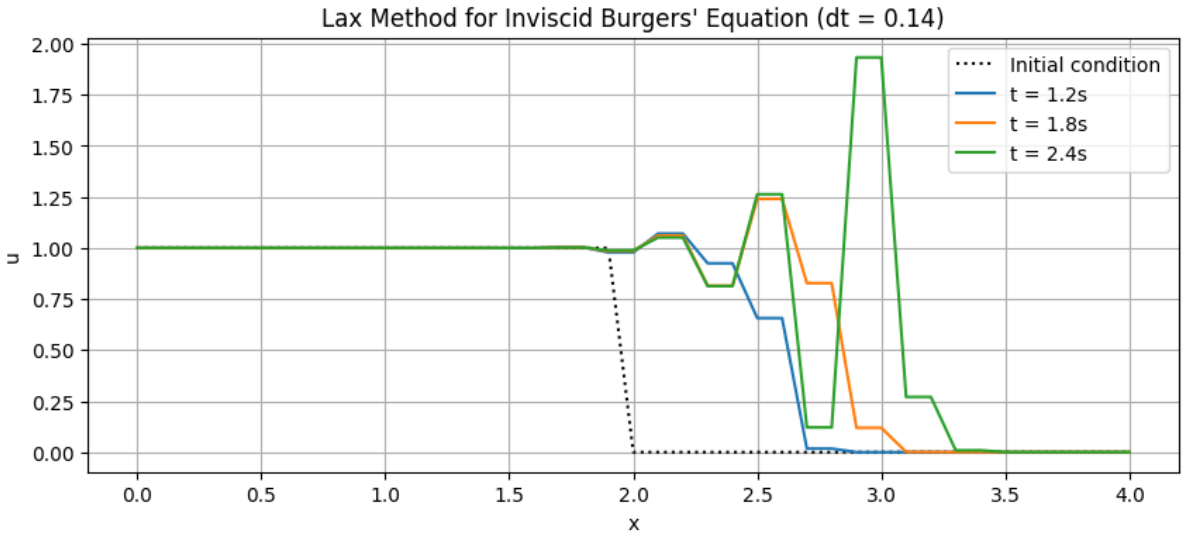

Figure 6: Plot of the Lax method solution at times 1.2, 1.8, 2.4, and the initial condition, using $\Delta t = 0.14$.

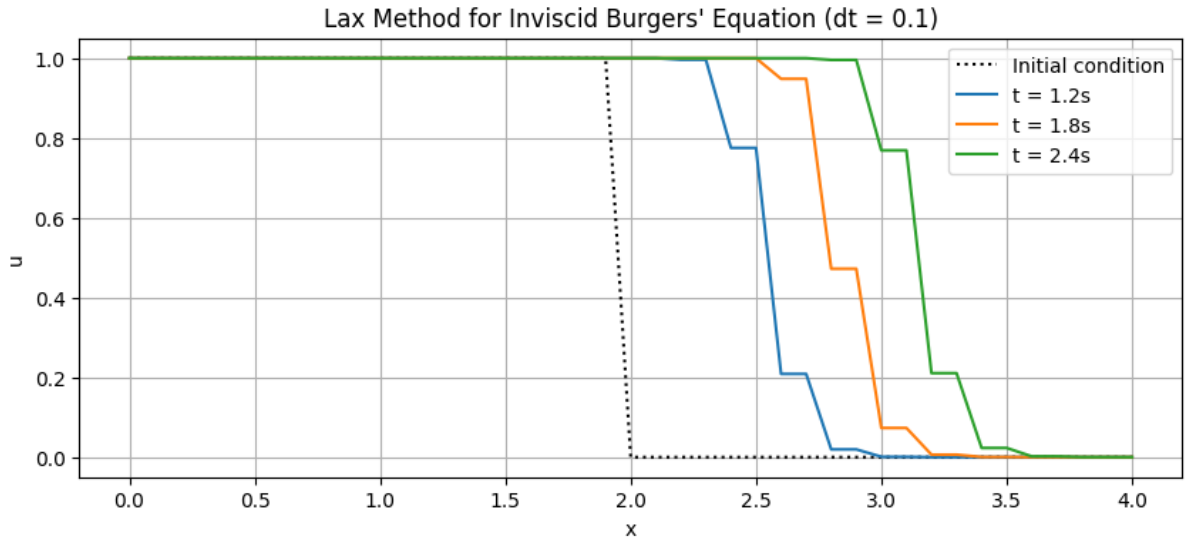

Figure 7: Plot of the Lax method solution at times 1.2, 1.8, 2.4, and the initial condition, using $\Delta t = 0.1$.

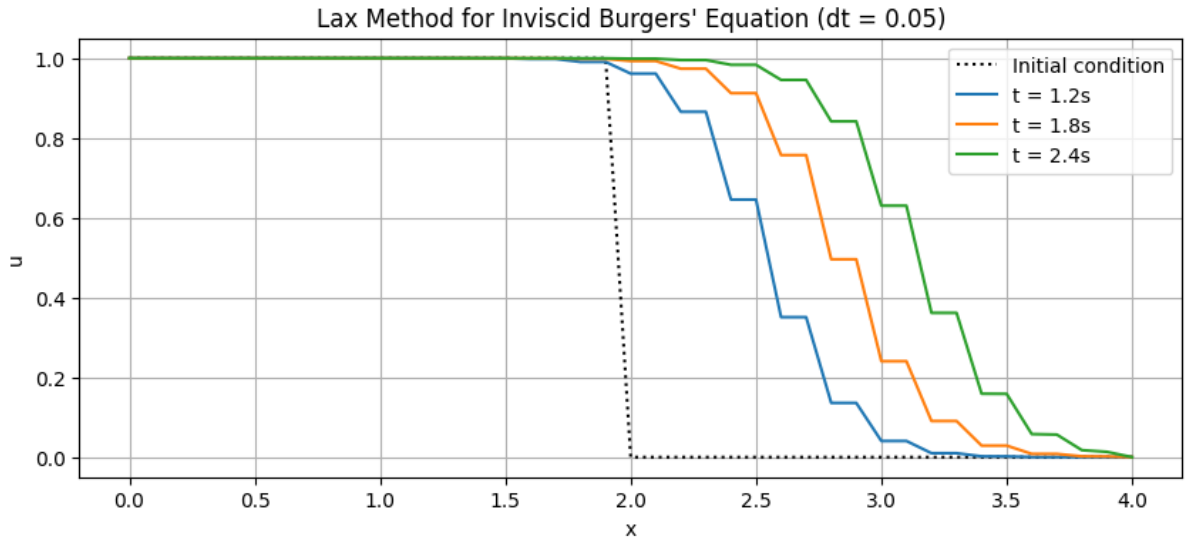

Figure 8: Plot of the Lax method solution at times 1.2, 1.8, 2.4, and the initial condition, using $\Delta t = 0.05$.

The initial observation from the plots is that the numerical solution exhibits a step-like behavior. This characteristic arises due to the averaging of adjacent values to compute u at each spatial point, coupled with the prescribed initial condition. Specifically, the initial condition comprises two distinct horizontal levels. When the numerical solutions are computed at $t = \Delta t$, the initial timestep, the values at points adjacent to the discontinuity are influenced by their neighboring points, one of which is 1 and the other is 0. Consequently, these points obtain a value that reflects the average of their neighbors. This initial step-like pattern then perpetuates in subsequent timesteps as each intermediate point between the original levels continues to reflect an average of its neighbors. This process results in the non-smooth, step-like curve observed across all plots.

For $\Delta t = 0.14$, the Lax method applied to the inviscid Burgers' equation demonstrates a proclivity

for a steep gradient development suggestive of shock formation, a feature common in non-linear wave equations. As the solution evolves to $t = 2.4s$, an overshoot is evident, manifesting a divergence from the anticipated physical discontinuity of an actual shock. This is characteristic of numerical solutions that exceed the critical Courant number, leading to instabilities that can distort the representation of the wave physics. When the time step is diminished to $\Delta t = 0.1$, aligning with a Courant number of unity, the numerical solution more aptly resolves the shock wave. The propensity for overshooting is mitigated, although the solution is not without fault - a discernible discrepancy persists. Proceeding to the smallest time step investigated, $\Delta t = 0.05$, the Lax method's representation of the shock degrades, as evidenced by the discontinuities along the horizontal axis within each velocity profile. This deterioration suggests that excessive refinement in time stepping may introduce non-physical artefacts, undermining the fidelity of the numerical solution. In summary, a time step yielding a Courant number commensurate with unity provides the most satisfactory accuracy. Despite this, the solutions obtained are still imperfect and present avenues for refinement, particularly in the nuanced capture of shock phenomena inherent to the inviscid Burgers' equation.

3.2.2 MacCormack method simulations

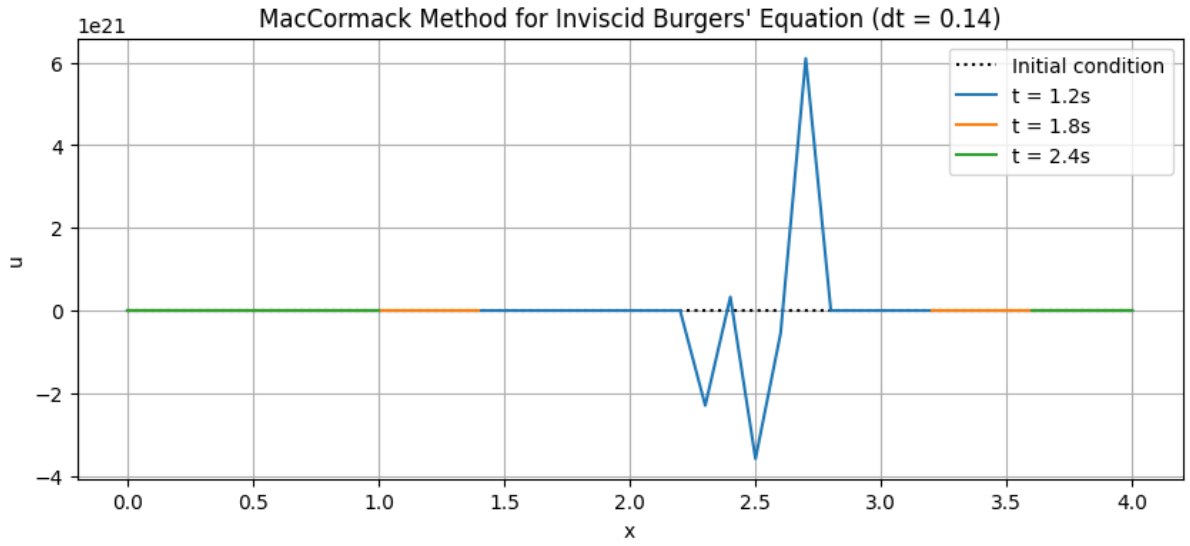

Figure 9: Plot of the MacCormack method solution at times 1.2, 1.8, 2.4, and the initial condition, using $\Delta t = 0.14$.

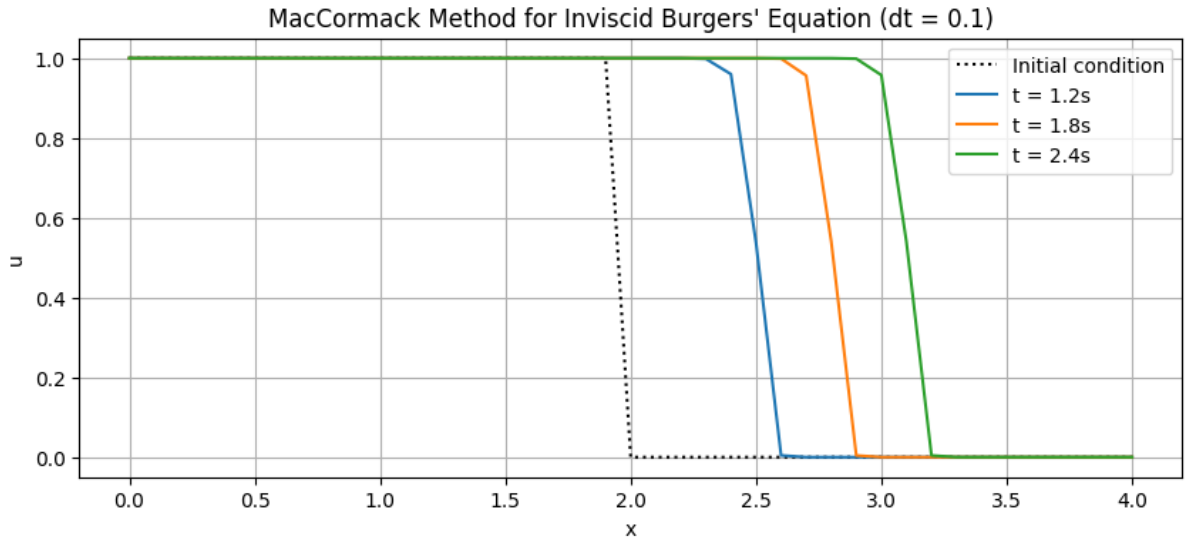

Figure 10: Plot of the MacCormack method solution at times 1.2, 1.8, 2.4, and the initial condition, using $\Delta t = 0.1$.

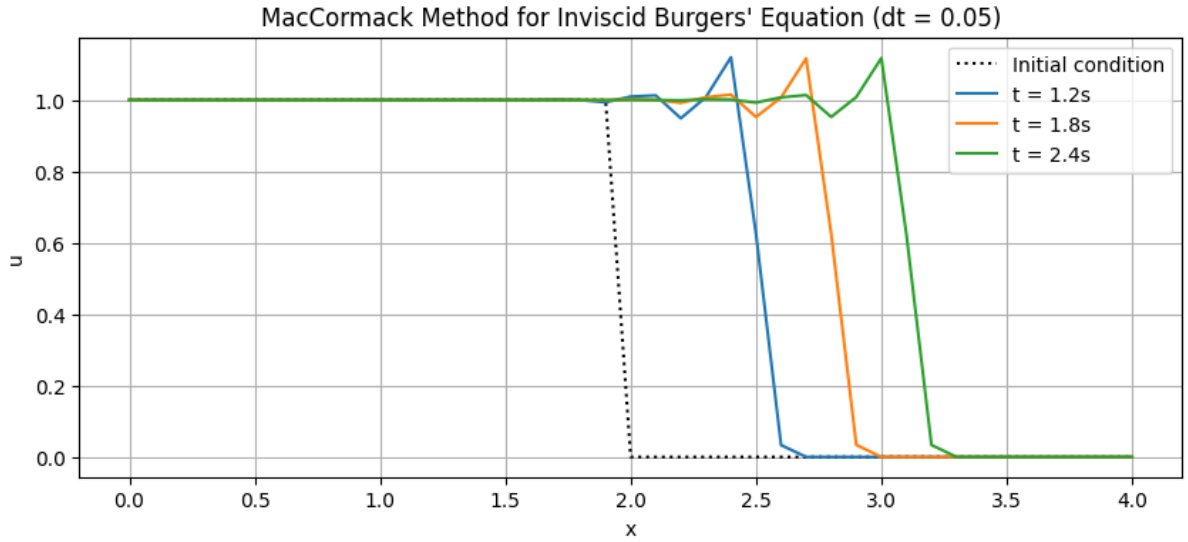

Figure 11: Plot of the MacCormack method solution at times 1.2, 1.8, 2.4, and the initial condition, using $\Delta t = 0.05$.

The largest time step, depicted in the attached image with $\Delta t = 0.14$, unveils considerable numerical instability in the MacCormack method's solution to the inviscid Burgers' equation. Notably, the simulation emitted runtime warnings indicating an overflow in the computation of the squared terms and invalid values during array operations, which are symptomatic of numerical issues arising from the chosen discretisation parameters. These warnings signal the potential for exponentially growing errors or non-physical results. Correspondingly, the solution's unrealistic and unbounded spike at $t = 1.2s$ typifies the method's inadequacy in accurately resolving the evolving shockwaves, integral to the nonlinear nature of the equation. This graphical manifestation and the accompanying computational alerts affirm the MacCormack method's conditional stability limits, which hinge on a Courant number constrained to unity or lower. Surpassing this boundary due to an overly expansive time step precipitates instability

and an inability to suitably capture the abrupt transitions and pronounced gradients emblematic of shock-waves. Consequently, the solution can devolve into a non-physical state with significant overshooting and spurious numerical artefacts.

With $\Delta t = 0.1$, corresponding to a Courant number of unity, the numerical solution demonstrates a more accurate and stable wave propagation, without any overshooting. The solution exhibits a relatively smooth profile, apart from a minor departure from smoothness at the crest of the velocity curve. This minor aberration does not significantly detract from the overall fidelity of the solution. In contrast, reducing the time step further to $\Delta t = 0.05$ leads to an unexpected overshooting at each plotted time step, indicative of an over-refinement which induces numerical inaccuracies. These tendencies mirror the behavior observed when solving the linear advection equation as discussed in Section 3.1, reinforcing the insight that the optimal performance of the MacCormack method aligns with a Courant number approaching unity. Despite the theoretical stability ensured by the Von Neumann analysis for Courant numbers less than unity, the practical accuracy of the solution may be compromised, as clearly illustrated by the smallest Δt evaluated.

Moreover, the phenomena associated with shock waves, manifest in the numerical solutions. The inviscid Burgers' equation is emblematic of equations that describe non-linear wave propagation where shocks — abrupt and substantial alterations in the solution gradient — naturally emerge. The MacCormack method, while adept at capturing a range of wave dynamics, can falter in accurately resolving the steep gradients that characterize shock waves, particularly when the temporal discretisation does not align with the spatial variations as prescribed by the Courant-Friedrichs-Lewy (CFL) condition. The numerical anomalies at $\Delta t = 0.05$ and $\Delta t = 0.14$ underscore the intricacies of simulating shock formation and propagation, necessitating a careful balance between resolution and numerical stability.

3.2.3 Lax vs MacCormack

In summary, the MacCormack simulations demonstrate superior performance when compared to those utilising the Lax method, particularly at a time step of $\Delta t = 0.1$. The success of the MacCormack method can be attributed to its classification as a 'predictor-corrector' method. Such methods are renowned for their efficacy in handling nonlinear hyperbolic equations, like the Inviscid Burgers' equation, by predicting an intermediate solution and then correcting it, which typically results in a higher accuracy for capturing shocks and complex wave structures. Despite its general robustness, the MacCormack method's performance is heavily influenced by the choice of the Courant number. When the Courant number exceeds 1 (e.g. when $\Delta t = 0.14$ was used), the method exhibited extreme instability with error magnitudes on the order of 10^{21}, in contrast to the Lax method which showed errors of a much smaller scale, on the order of 10^{1}. This vast difference in the scale of errors underlines the importance of adhering to the stability condition — specifically, the Courant-Friedrichs-Lewy (CFL) condition — for the MacCormack method to ensure stable and accurate results.

4 Conclusion

In this project, we investigated the numerical solutions of the advection equation and the inviscid Burgers' equation using various computational methods. For the advection equation, our analysis focused on three methods: the First Upwinding Differencing Method, the Lax-Wendroff Method, and the Euler BTCS Method. Each method was evaluated based on its performance at different Courant numbers, maintaining a fixed spatial step size while varying the time step. Our results demonstrated that the stability and accuracy of the solutions are strongly dependent on adherence to the Courant-Friedrichs-Lewy (CFL) condition, which is dictated by the chosen Courant number. The Lax-Wendroff and the First Upwinding Differencing Methods exhibited pronounced instability when the Courant number exceeded unity. In contrast, the Euler BTCS Method, while unconditionally stable, manifested significant numerical dispersion. This dispersion, as elucidated in Section 3.1.2, could be mitigated by adjusting both the time step and spatial resolution to maintain the Courant number at unity. The strategy of proportionally decreasing Δt while tuning Δx enhanced the fidelity of the Euler BTCS method, as evidenced by the consistent reduction in both the Mean Absolute Error and Infinity Norm Error. Ultimately, this balanced approach to temporal and spatial discretisation provided a more accurate and stable numerical solution that more faithfully captured the dynamics of the advection equation. However, despite the improvements made to the Euler BTCS method, the First Upwind and Lax-Wendroff methods still demonstrated superior performance overall. The First Upwind method, in particular, was found to be the best for the advection problem addressed here, offering the lowest errors and demonstrating less instability under violation of the stability constraints. Nonetheless, it's conceivable that under different initial conditions, the BTCS method could be favored due to its unconditional stability.

For the non-linear inviscid Burgers' equation, the Lax and MacCormack methods were employed to address the challenge of accurately capturing shock formations that are characteristic of non-linear wave propagation. The MacCormack method, particularly with the Courant number maintained close to one, showed a pronounced superiority in handling the steep gradients and discontinuities indicative of shock waves. This was in contrast to the Lax method, which contended with overshoots and numerical artifacts, as clearly evidenced by the disparities in their respective results at $\Delta t = 0.14$ and $\Delta t = 0.1$. While both methods theoretically adhere to the Von Neumann stability criteria for $c \leq 1$, the MacCormack method's predictor-corrector mechanism conferred a distinct advantage. It allowed for a more nuanced approach to the approximation of discontinuities, diminishing the non-physical overshooting that plagued the Lax method at larger time steps, as seen in Figures 6 and 7. However, as the time step was further reduced to $\Delta t = 0.05$, both methods began to exhibit limitations in precision. This suggests a complex interplay between time step size, numerical method, and the inherent characteristics of the Burgers' equation in the accurate resolution of shocks.

Overall, the project not only reinforced the critical importance of choosing appropriate numerical methods and parameters for solving partial differential equations but also highlighted the nuanced behavior of linear versus non-linear equations under numerical simulation. Future work could explore further optimisations in numerical schemes and parameter selections to enhance the fidelity and efficiency of simulations, especially in complex real-world applications where these equations find extensive use.

References

[1] Irene Kyaza *MT5846 Lecture Notes*. University of St. Andrew's School of Mathematics and Statistics.

[2] Nazm'I Oyar *Invscid Burgers Equations and its Numerical Solutions*. The graduate school of natural and applied sciences of Middle East Technical University.

[3] A. Salih *Inviscid Burgers' Equation* Indian Institute of Space Science and Technology, 2015.