Sachin Natesh

1 Temporal Discretization of KdV

Suppose that we want to solve the Korteweg de Vries equation

$$\phi_t = \mathcal{K}(\phi) = -\phi_{xxx} - 3(\phi^2)_x, \tag{1.1}$$

where $\mathcal{K}(\phi)$ is the functional on the right hand side. On an unbounded line, (1.1) admits soliton solutions in the form of

$$\phi_{\text{sol}}(x,t) = \frac{c}{2}\text{sech}^2\Big(\frac{\sqrt{c}}{2}(x-ct)\Big). \tag{1.2}$$

A pseudo-spectral method approximates the r.h.s of (1.1) as

$$\hat{\mathcal{K}} = \mathcal{F}(\mathcal{K}) = F(\hat{\phi}) = ik^3 \boxdot \hat{\phi} - 3ik \boxdot \mathcal{F}\Big(\big(\mathcal{F}^{-1}(\hat{\phi})\big)\boxed{2}\Big), \tag{1.3}$$

where $\mathcal{F}$ is the Fourier transform. Under this transformation, the KdV equation becomes

$$\hat{\phi}_t = F(\hat{\phi}) = A\hat{\phi} + B(\hat{\phi}). \tag{1.4}$$

We consider two schemes to solve (1.4). The first is the AB2BDF2 semi-implicit order 2 method, which after isolating the explicit term is given by

$$(I - \frac{2\Delta t}{3}A)\hat{\phi}^{n+1} = \frac{4}{3}\hat{\phi}^n - \frac{1}{3}\hat{\phi}^{n-1} + \frac{2\Delta t}{3}(2B(\hat{\phi}^n)) - B(\hat{\phi}^{n-1}). \tag{1.5}$$

The second scheme is ETDRK2, an order 2 exponential time differencing 2 stage Runge-Kutta scheme given by

$$\begin{aligned}
\hat{\phi}^{n+1,\star} &= e^{A\Delta t}\hat{\phi}^n + A^{-1}(e^{A\Delta t} - I)B(\hat{\phi}^n) \quad \text{(predictor)} \\
\hat{\phi}^{n+1} &= \hat{\phi}^{n+1,\star} + A^{-2}\Big(\frac{(e^{A\Delta t} - I - A\Delta t)}{\Delta t}\Big)(B(\hat{\phi}^{n+1,\star}) - B(\hat{\phi}^n)) \quad \text{(corrector)}
\end{aligned} \tag{1.6}$$

To avoid cancellation errors in precomputing the inversion operators in (1.6) for their specific instance in (1.3), we repeat A (which is treated as a vector) M times, add the i^{th} complex root of unity to matrix resulting from repeating A, for $i = 1, \cdots, M$, and average over the rows when computing matrix exponentials. That is,

1. Define the array r containing $M = N/4$ complex roots of unity
2. Define AR to be the sum of the matrix formed by repeating the single column of A M times added to each row of r repeated N times. That is,

$$AR = \Delta t \begin{bmatrix} | & \cdots & | \\ A & \cdots & A \\ | & \cdots & | \end{bmatrix} + \begin{bmatrix} | & \cdots & | \\ e^{\frac{2i\pi}{2M}} & \cdots & e^{\frac{2i\pi(2M-1)}{2M}} \\ | & \cdots & | \end{bmatrix} \tag{1.7}$$

$$A^{-1}(e^{A\Delta t} - I) = \Delta t \begin{bmatrix} \frac{1}{M}\sum_{j=1}^{M}(\texttt{expm1}(AR_{1j})/AR_{1j}) \\ \vdots \\ \frac{1}{M}\sum_{j=1}^{M}(\texttt{expm1}(AR_{Nj})/AR_{Nj}) \end{bmatrix} \tag{1.8}$$

$$A^{-2}(e^{A\Delta t} - I - A\Delta t) = \Delta t \begin{bmatrix} \frac{1}{M}\sum_{j=1}^{M}((\texttt{expm1}(AR_{1j}) - AR_{1j})/AR_{1j}^2) \\ \vdots \\ \frac{1}{M}\sum_{j=1}^{M}((\texttt{expm1}(AR_{Nj}) - AR_{Nj})/AR_{Nj}^2) \end{bmatrix} \tag{1.9}$$

1.1 Absolute Stability of the two methods

Here, we analyze the regions of absolute stability for both AB2BDF2 and ETDRK2 applied to the linear model ODE

$$\frac{du}{dt} = Au + B(u) = Au + Bu = aiu - b(1+i)u, \tag{1.10}$$

where $a, b \in \mathbb{R}$ and $b > 0$. First, we consider AB2BDF2. Following the common approach for absolute stability analysis, we substitute the forms of A and B in (1.10) into (1.5) and compute the growth factor $r = \frac{u_{n+1}}{u_n}$. We have

$$u_{n+1} = (4u_n - u_{n-1} - 4\Delta t b(1+i)u_n + 2\Delta t b(1+i)u_{n-1})/(3 - 2\Delta t a i).$$

Let $\alpha = a\Delta t$ and $\beta = b\Delta t$. Then, we have the following quadratic for r:

$$r^2(3 - 2\alpha i) + r(4\beta(1+i) - 4) + 1 - 2\beta(1+i) = 0.$$

We define the region of absolute stability to be

$$\begin{aligned} S &= \Big\{(\alpha, \beta) : |r(\alpha, \beta)| \leq 1\Big\}, \\ r(\alpha, \beta) &= \frac{-4(\beta(1+i) - 1) \pm \sqrt{(4\beta(1+i) - 4)^2 - 4(3 - 2\alpha i)(1 - 2\beta(1+i))}}{2(3 - 2\alpha i)}. \end{aligned} \tag{1.11}$$

We can compute the region of absolute stability by sampling $r(\alpha, \beta)$ on a fine grid in the $\alpha\beta$ half plane and putting down a point wherever r evaluates to less than 1. The stability region found with this approach is provided below:

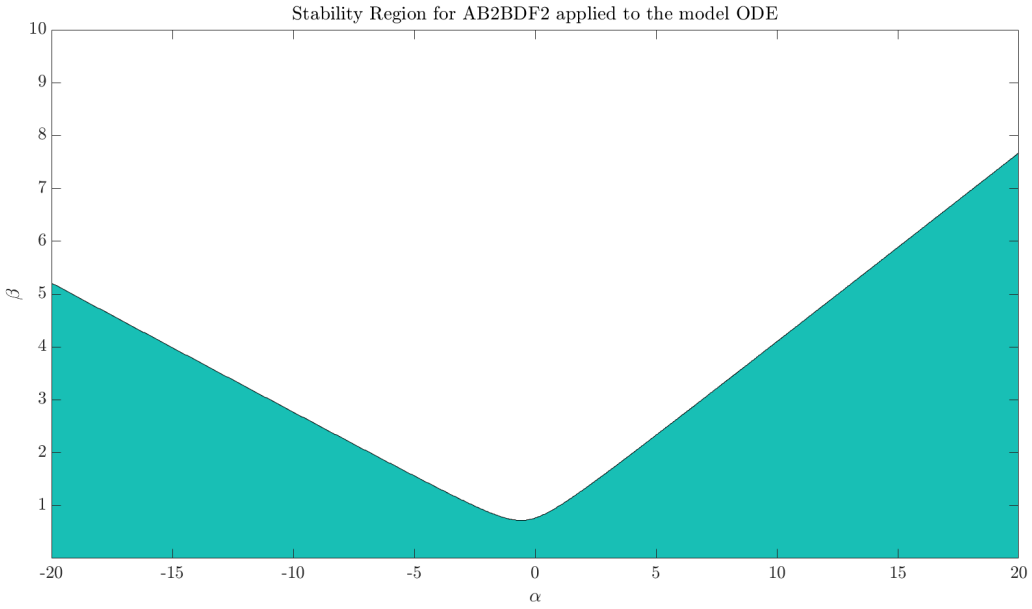

Figure 1: Stability region of AB2BDF2 applied to (1.10). The blue shading corresponds to the region of stability.

Moving on to ETDRK2, we substitute the predictor in (1.6) into the corrector, define α, β and r as before, and arrive at the following definition of the region of stability:

$$\begin{aligned} S &= \Big\{(\alpha,\beta) : |r(\alpha,\beta)| \leq 1\Big\}, \\ r(\alpha,\beta) &= e^{\alpha i \Delta t} - \frac{\beta}{\alpha i}(e^{\alpha i} - 1) - \frac{\beta}{(\alpha i)^3}(\alpha i - \beta)(e^{\alpha i} - 1)(e^{\alpha i} - 1 - \alpha i). \end{aligned} \tag{1.12}$$

Using the same approach as with AB2BDF2, we plot the stability region of ETDRK2 below:

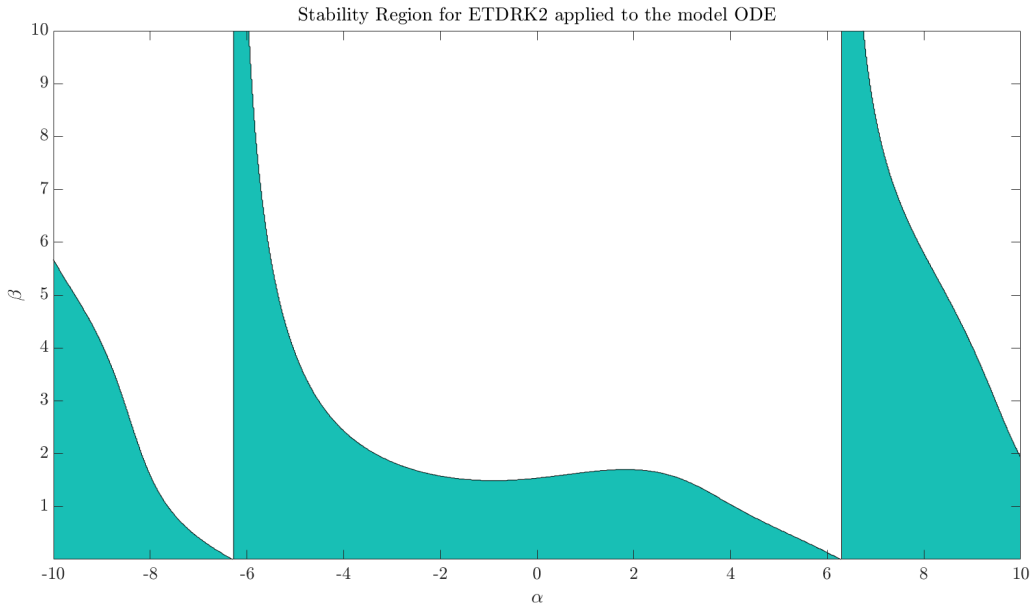

Figure 2: Stability region of ETDRK2 applied to (1.10). The blue shading corresponds to the region of stability.

When there are many Fourier modes in A, the large modes (high wave number) evolve on short time scales. This can be seen by considering (1.3) without the non-linearity:

$$\frac{d\hat{\phi}}{dt} = ik^3 t \Rightarrow \hat{\phi}(t) \propto \cos(k^3 t) + i\sin(k^3 t). \tag{1.13}$$

We see that for larger k, the time required to complete one period decreases as $O(k^{-3})$. Thus, if the modes associated to high wave numbers are significant, we require small time steps to resolve their evolution. This also provides some intuition as to why ODEs with higher order derivatives are considered stiff. In the Fourier basis, higher order derivative operators result in faster time scales of evolution for large modes, as the operators are simply powers of k. The stability regions for AB2BDF2 and ETDRK3 applied to the linear model ODE include a large range of values for $\alpha\Delta t$. Thus, the methods can handle the stiffness caused by large Fourier modes without requiring exceedingly small Δt, and also handle small modes.

1.2 Accuracy, Stability and Robustness for a Single Soliton

1.2.1 Empirical Stability

We find that the stability limit on the stepsize is around $\Delta t = 0.182$ for grids of size 64,128 and 256. The limit was computed based on the growth in norm of the solution compared to the true solution and qualitative comparisons. There doesn't seem to be a reasonable stable time step for $n = 32$ as the solution becomes unstable even for $\Delta t = 0.03$. The stability limit doesn't strongly depend on the resolution, though unstable behavior develops (qualitatively) a bit earlier for lower resolutions. This makes sense as a consequence of under-resolution. Higher modes which evolve on shorter time scales are neglected, contributing to unstable behavior. Moreover, since our stability analysis showed relatively large support for values of $\alpha\Delta t$, it is likely the eigenvalues of the non-linear term in (1.3) that realize the common stability limit for the larger 3 grid sizes.

1.2.2 Accuracy of ODE Solver

In Fig. 3, we plot $\tilde{e}_{\Delta t} = ||\hat{\phi}_{\Delta t}(T) - \hat{\phi}_{\Delta T/2}(T)||_2$. To achieve clear second order accuracy in either method, we required $\Delta t \leq 0.03$, corresponding to 2000 steps for one period of a $c = 1$ solition on a domain of length $L = 60$.

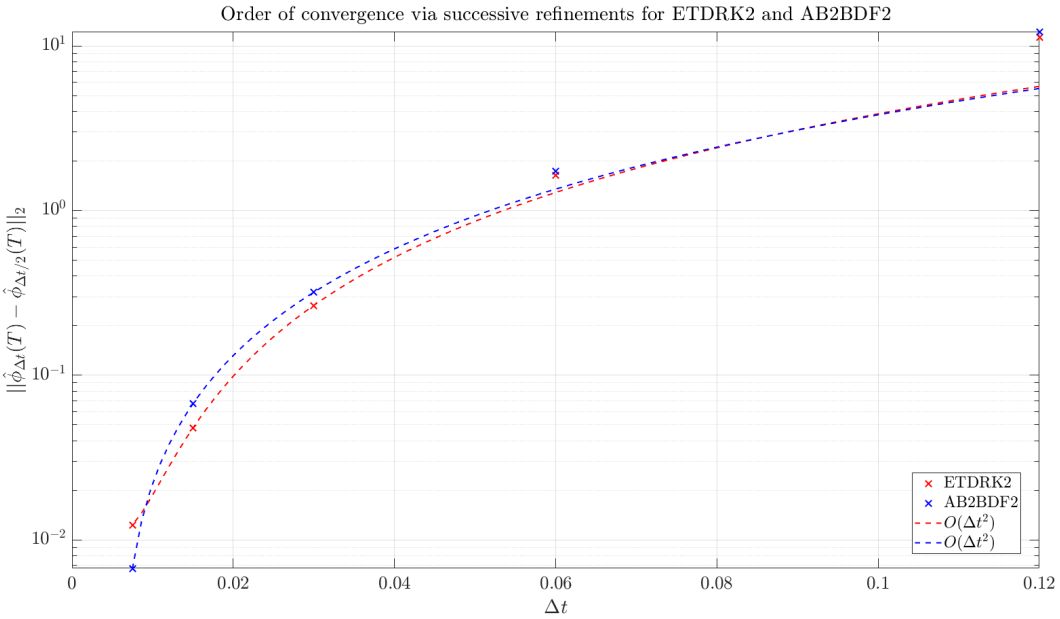

Figure 3: 2-norm difference between solutions to the ODE in (1.3) after one period obtained using successively refined time steps and 256 modes. As $\Delta t \to 0$, the curve should approach the expected $O(\Delta t^2)$ accuracy. We see that for $\Delta t \leq 0.03$, the observed order of accuracy matches the expected order, which is overlayed in black dashed lines. Solutions were interpolated onto grids of 512 points to calculate $\tilde{e}_{\Delta t}$

We can explicitly calculate the observed order of accuracy as follows. Let $\hat{\phi}_1, \hat{\phi}_2$ and $\hat{\phi}_3$ be solutions on a fine time grid, a coarser grid and a coarsest grid respectively. Assume the refinement ratio between subsequent grids is $r = 2$. Then, expanding the discretization error in terms of $\hat{\phi}_{true}$ the exact solution, discarding higher order terms and carrying out the algebra, we solve directly for p via

$$\begin{aligned}
\hat{\phi}_1 &= \hat{\phi}_{true} + g_p \Delta t^p + O(\Delta t^3) \\
\hat{\phi}_2 &= \hat{\phi}_{true} + g_p (r\Delta t)^p + O(\Delta t^3) \\
\hat{\phi}_3 &= \hat{\phi}_{true} + g_p (r^2\Delta t)^p + O(\Delta t^3) \\
&\Rightarrow p = \log_r \left(\frac{\hat{\phi}_3 - \hat{\phi}_2}{\hat{\phi}_2 - \hat{\phi}_1} \right).
\end{aligned} \tag{1.14}$$

Using (1.14) on the solutions from the 3 smallest time steps, we compute the observed order of accuracy to be $p = 1.96$. This matches fairly well to the expected formal order of 2.

1.2.3 Accuracy of the PDE solver

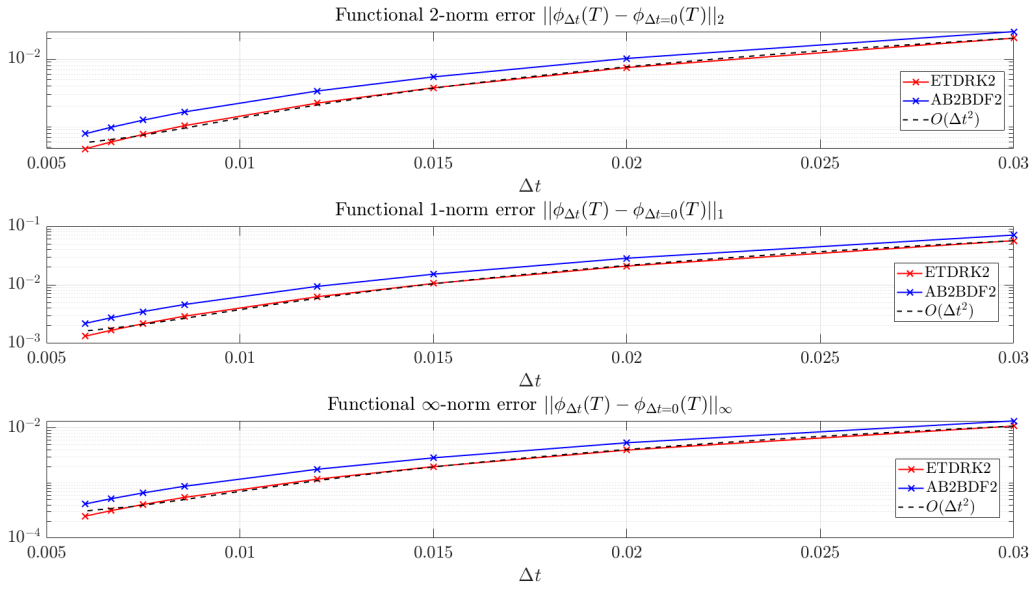

Figure 4: Difference under 3 norms between solutions to the PDE in (1.1) obtained after one period for several time steps and 256 modes. As $\Delta t \to 0$, the curve should approach the expected $O(\Delta t^2)$ accuracy. Since the largest time step used was smaller than 0.03, which we found was the boundary of the asymptotic regime, all errors match the asymptotic rate. Solutions were interpolated onto grids of 512 points to calculate norms. For both solvers and under all norms, we see that the functional error is $O(\Delta t^2)$, indicated by the dashed black line. This test confirms that convergent solves of the ODE lead to convergent solutions to the PDE.

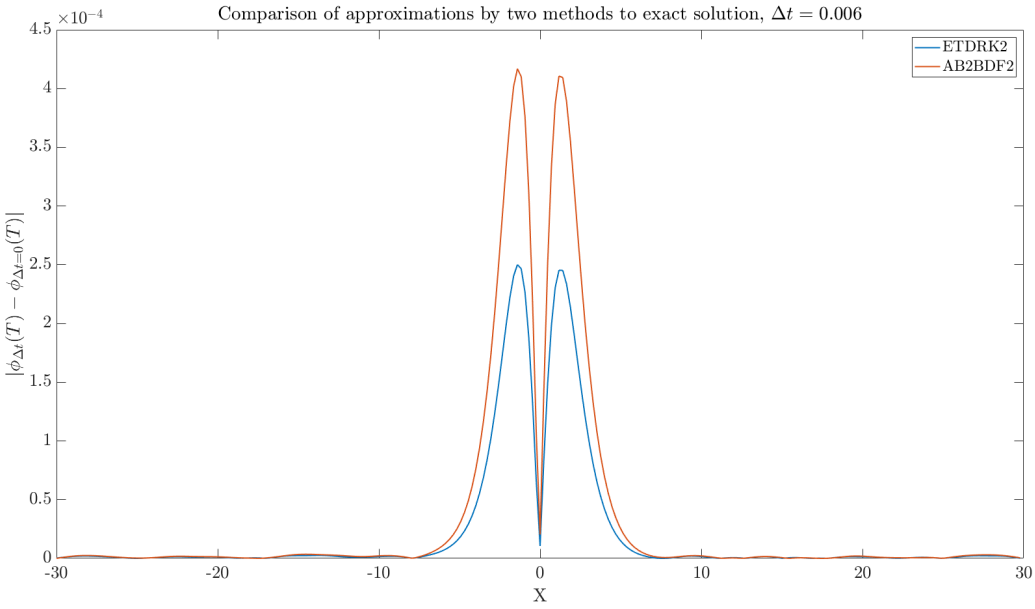

Figure 5: Absolute difference between solutions to the PDE obtained via ETDRK2 and AB2BDF2 for $\Delta t = 0.006$ and 256 modes after one period, compared to the true solution. Similar differences were found for smaller step sizes. From this, we can conclude that for the same step size, ETDRK2 yields lower error.

1.2.4 Robustness of PDE solver

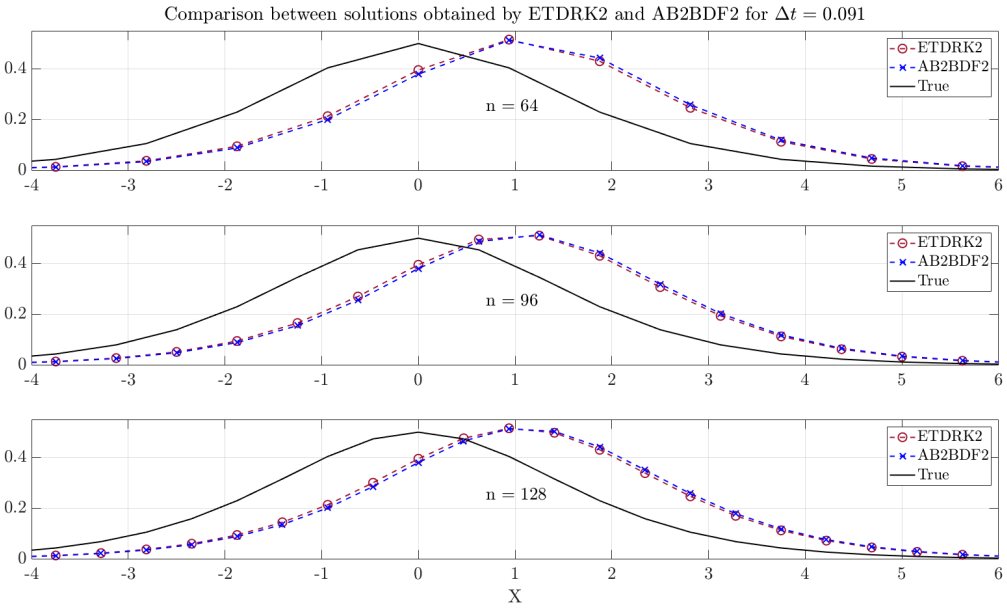

Figure 6: Comparison of solutions obtained by the two solvers for $\Delta t = 0.091$ and n = 64,96,128. It seems that the solvers mainly fail in that they increase the speed of the wave so that it exceeds the correct final position after one period when the problem is under-resolved. It's not completely clear whether one of the methods is more robust than the other, but we can definitely say that ETDRK2 increases the speed slightly less than does AB2BDF2 - which may have a more dramatic affect for longer time integrations. Both methods result in very similar increases in the height of the soliton when compared to the true solution.

1.3 Collisions between two solitons

As an initial condition, I used a superposition of two solitions moving to the right- one with speed 3 and the other with speed 1.

$$\phi_0(x) = \phi_{sol}(x + 15, c = 3) + \phi_{sol}(x - 5, c = 1), \tag{1.15}$$

It is depicted in the following figure:

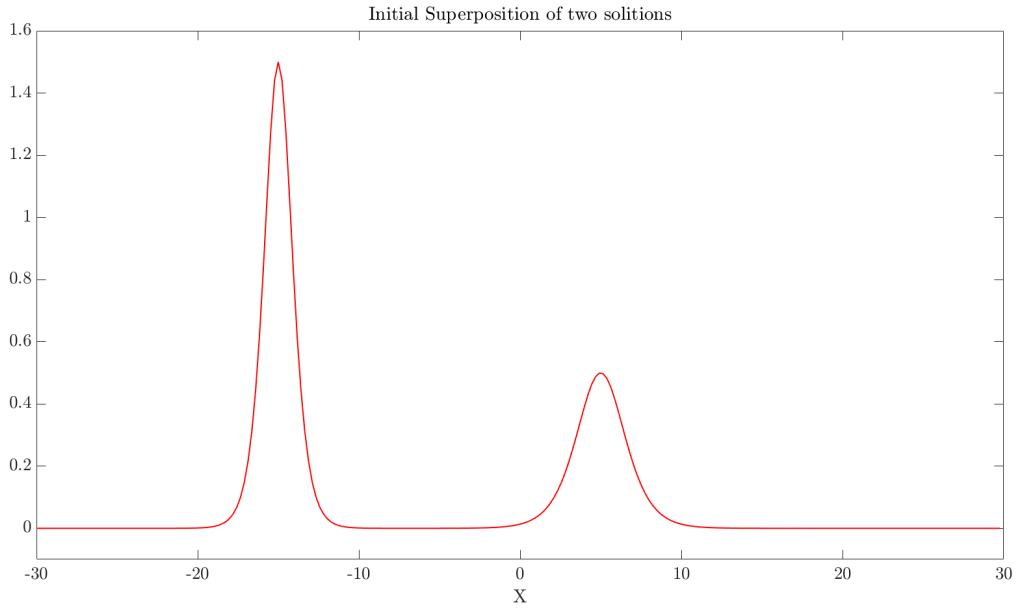

Figure 7: Initial superposition of two solitons. The left one moves to the right with speed 3, and the right one moves to the right with speed 1.

The domain was of length $L = 60.$ and $n = 256$ points were used to sample the initial condition. This number of points was determined based on the Fourier spectrum of the initial condition, which showed that around 400 points were required so that higher wave numbers were near 0. Based on our investigation in this assignment, I knew that using fewer points, say 256, would still produce qualitatively reasonable results since the stability limit isn't strongly tied with the number of modes. To determine the cheapest timestep that could be used to produce animations, I qualitatively checked for unstable behavior and increased the number of steps until the behavior was stable over a time interval of length L (the animation time). This turned out to be 4000 steps, corresponding to $\Delta t = 0.015$. The movie can be viewed in `soliton_passthru.avi`.