ERDC Internship Report

S. Gerberding

July 27, 2022

1 Introduction:

A major drawback of the Serre equations is the presence of a third order space derivative. This third order term makes for a nasty CFL condition if the method is treated purely explicitly. Informally, for example, if the mesh sized is halved, then the time stepped must be reduced by a factor of 8 (2^{-3}). This restriction makes for a highly inefficient method. One solutions is to split the method into an explicit part and an implicit part. In particular, we treat the hyperbolic part of the problem explicitly, and then solve the dispersive part implicitly. Indeed, our long term goal is to develop and implement such a method. Once the theory is developed, we aim to integrate the method into AdH.

Now, to better under the Serre-Greene equations, we first analyze a simpler equation: the KdV equation. This equation contains a third order space derivative (and so comes with the same CFL condition issue), and thus solving the KdV will give insights on how to address the Serre-Greene equations. This report thus largely reflects work done on the KdV equation.

This report is composed as follows. Section two address what we shall call the “linear KdV equation” (the transport term is linear). It includes the weak form of the problem and analysis thereof. It also includes a finite element method, which is composed of an explicit step and an implicit step, for which we prove stability. Section three gives details for implementing the method, including the algebraic form of the scheme and the matrix required for solving the implicit problem. In particular, we give the details for both Dirichlet and periodic boundary conditions.

The last section concerns several numerical illustrations. First, to highlight the explicit step (which corresponds to a hyperbolic problem), we present two numerical illustrations. One experiment is when the method includes a graph viscosity term and one experiment when it isn’t.

2 The linear KdV equation

We consider the 1D linear KdV equation:

$$\partial_t u + a\partial_x u + K\partial_{xxx} u = 0; \quad u(x,0) = u_0(x); \quad x \in [b,c]; \quad t \in [0,T] \tag{1}$$

We assume that we have Dirichlet or periodic boundary conditions.
We want to approximate the solution using finite elements. Let us first address the weak formulation. Let $V := H^1(D)$ and let $S := C^1([0,T];V)$. Then, we seek $u : [0,T] \to V$ such that, for a.e. t and all $v \in V$:

$$\int_D \partial_t uv + a\partial_x uv \, \mathrm{d}x + \int_D K\partial_{xxx} uv \, \mathrm{d}x = 0 \text{ for all test functions } v\text{, for all } t \in [0,T] \tag{2}$$

This problem is difficult–the third order term causes problems in the analysis. Instead we consider a different problem by introducing a new variable: $z := \partial_x u$.
The new problem is: Find $(u,z) \in S \times S$ such that, for a.e. $t \in [0,T]$ and all $(v,q) \in V \times V$:

$$\int_D \partial_t uv + a\partial_x uv \, \mathrm{d}x - K\int_D \partial_x z \partial_x v \, \mathrm{d}x = 0; \tag{3}$$

$$\int_D z\partial_x q \ \mathrm{d}x = \int_D \partial_x u \partial_x q \ \mathrm{d}x \text{ for all } (v,q) \in V \times V. \tag{4}$$

Let us quickly gives several conservation results.

Lemma 1. *Let (u,z) solve* (3) *and* (4)*. Then for all $t \in [0,T]$, $\|u(\cdot,t)\|_{L^2(D)} = \|u_0\|_{L^2(D)}$.*

Proof. Test (3) with $v = u$ and (4) with $q = z$. We arrive at:

$$\int_D \partial_t uu + a\partial_x uu - K\partial_x z \partial_x u \ \mathrm{d}x = 0.$$
$$\int_D z\partial_x z \ \mathrm{d}x = \int_D \partial_x u \partial_x z \ \mathrm{d}x.$$

Using the fundamental theorem of calculus and the boundary conditions, we deduce:

$$\partial_t \frac{\|u\|^2_{L^2(D)}}{2} - K\int_D \partial_x z \partial_x u \ \mathrm{d}x = 0.$$
$$0 = \int_D z\partial_x z \ \mathrm{d}x = \int_D \partial_x u \partial_x z \ \mathrm{d}x.$$

Substituting the second equation into the first, we achieve

$$\partial_t \frac{\|u\|^2_{L^2(D)}}{2} = 0,$$

which proves the result.

□

Lemma 2. *Let u solve (2). Then for all $t \in [0,T]$ and for all $p \in [1,\infty]$, we have $\|u(\cdot,t)\|_{L^p(D)} = \|u_0\|_{L^p(D)}$.*

Proof. □

2.1 FE discretizations

2.1.1 Space discretization

Let us now set up the discretization in space. Let $D := [a,b]$ be the interval, and let $\{\mathcal{K}_h\}_{h>0}$ be a sequence of shape-regular meshes. Now let $h > 0$ be given. For each $K \in \mathcal{K}_h$, let $T_K : [0,1] \to K$ be an affine geometric mapping. For $n \in \mathbb{N}$, let $\widehat{\mathcal{P}}_n$ be the set of all polynomials of degree n on $[0,1]$. Furthermore, let $\mathcal{V} := \{1 : I\}$ be an enumeration of the degrees of freedom. Now we consider the space:

$$V_h := \{v : v \circ T_k \in \widehat{P} \text{ for all } K \in K_h\}.$$

We assume that a basis exists such that it has the unity of partition property. That is, $\{\phi_i\}_{i\in\mathcal{V}}$ is a basis for V_h and that $\sum_{i\in\mathcal{V}} \phi_i(x) = 1$ for all $x \in D$. Using the basis functions, we define several quantities:

$$m_i := \int_D \phi_i(x) \ \mathrm{d}x; \quad c_{ij} := \int_D \phi_i(x)\partial_x\phi_j(x) \ \mathrm{d}x; \quad a_{ij} := \int_D \partial_x\phi_i(x)\partial_x\phi_j(x) \ \mathrm{d}x; \quad d_{ij} = a|c_{ij}|. \tag{5}$$

Assume throughout that $m_i > 0$. We introduce a norm on V_h. For $v_h \in V_h$ with a basis expansion $v_h = \sum_{i\in\mathcal{V}} V_i\phi_i$, define

$$\|v_h\|^2_{l^2,m} := \sum_{i\in\mathcal{I}} V_i^2 m_i.$$

Let $I(i) := \{j : m(\mathrm{supp}(\phi_j) \cap \mathrm{supp}(\phi_i)) \neq 0\}$; that is, when ϕ_i and ϕ_j share support. This set is also known as the stencil. Let us give a few key properties.

$$c_{ij} = -c_{ji}; \qquad d_{ij} = d_{ji}; \qquad \sum_{j\in I(i)} c_{ij} = 0.$$

For $i \in \{2 : N-1\}$, we have, $c_{ii} = 0$. The coefficients c_{11} and $c_{N,N}$ will depend on the boundary conditions.

2.1.2 Time discretization

Let t^n be the current time, $\tau_n > 0$, and $t^{n+1} := t^n + \tau_n$. For simplicity, we now assume $\tau_n = \tau$ for all n. We break the time discretization into two steps: (1) a hyperbolic prediction based only on the first two terms of the PDE and (2) an implicit step that includes the constraint condition. Throughout, we assume that $u_h^0(x) := \sum_{j\in\mathcal{I}} U_j^0 \phi_j(x)$ is a reasonable approximation of the initial data. Let $U^0 \in \mathbb{R}^I$ be column vector where the entries are the coefficients for u_h^0. Similarly, let $U^n \in \mathbb{R}^I$ be a column vector whose values are the coefficients for the approximation at time $t = t_n$, ie, $u_h^n(x) = \sum_{j\in\mathcal{I}} U_j^n \phi_j(x)$.

2.2 The scheme

Let $U^0 \in \mathbb{R}^I$. Suppose we have computed $U^1, ..., U^n$. We find U^{n+1} as follows.

2.2.1 Step 1: Hyperbolic prediction

We compute the hyperbolic prediction, which we shall call W^{n+1}, explicitly via:

$$m_i \frac{W_i^{n+1} - U_i^n}{\tau} = -\sum_{j\in I(i)} a U_j^n c_{ij} + \sum_{j\in I(i)} d_{ij}(U_j^n - U_i^n). \tag{6}$$

2.2.2 Step 2: Implicit update

Suppose we have computed W^{n+1} as outlined above. We now compute U^{n+1} and Z^{n+1} by solving the following system of equations:

$$m_i \frac{U_i^{n+1} - W_i}{\tau} - K \sum_{j\in I(i)} Z_j^{n+1} a_{ij} = 0 \tag{7}$$

$$\sum_{j\in I(i)} Z_j^{n+1} c_{ji} = \sum_{j\in I(i)} U_j^{n+1} a_{ij} \tag{8}$$

$$Z_0^{n+1} = 0. \tag{9}$$

2.3 l^2 estimates

Lemma 3. *Let $U^n \in \mathbb{R}^I$. Let W^{n+1} be defined as in (6). Define $\tau^* > 0$ by:*

$$\tau^* := \min_{i\in\mathcal{V}} \frac{m_i}{4\sum_{j\in I(i)} d_{ij}} = \min_{i\in\mathcal{V}} \frac{m_i}{4\sum_{j\in I(i)} a|c_{ij}|}.$$

Then for all $0 < t \le \tau^*$, we have $\|W^{n+1}\|_{l^2,m}^2 \le \|U^n\|_{l^2,m}^2$.

Proof. Multiplying (6) by $2\tau W_i^{n+1}$, we derive:

$$2m_i W_i^{n+1}(W_i^{n+1} - U_i^n) = -2\tau \sum_{j\in I(i)} W_i^{n+1} a U_j^n c_{ij} + 2\tau \sum_{j\in I(i)} W_i^{n+1} d_{ij}(U_j^n - U_i^n).$$

Using the identity $2a(a-b) = a^2 - b^2 + (a-b)^2$ and summing over i, we achieve:

$$\begin{aligned}\sum_{i\in\mathcal{V}}(W_i^{n+1})^2 m_i - \sum_{i\in\mathcal{V}}(U_i^n)^2 m_i + \sum_{i\in\mathcal{V}}(W_i^{n+1} - U_i^n)^2 m_i = \underbrace{-2\tau \sum_{i\in\mathcal{V}}\sum_{j\in I(i)} W_i^{n+1} a U_j^n c_{ij}}_{R_1} + \\ \underbrace{2\tau \sum_{i\in\mathcal{V}}\sum_{j\in I(i)} d_{ij}(U_j^n - U_i^n) W_i^{n+1}}_{R_2}.\end{aligned}$$

Let us analyze R_1 and R_2.

$$
\begin{aligned}
R_1 = -2\tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} W_i a U_j c_{ij} &= 2\tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} W_i^{n+1}(aU_j^n - aU_i^n)c_{ij} \\
&= 2\tau \int_D a\partial_x(u_h^n)(-w_h^{n+1}) \text{ d}x \\
&= 2\tau \int_D a\partial_x(u_h^n)(u_h^n - w_h^{n+1}) \text{ d}x \quad \left(\text{follows from } \int_D a\partial_x u_h^n u_h^n \text{ d}x = 0\right) \\
&= 2\tau \sum_{i\in\mathcal{V}} \sum_{jin\mathcal{V}} (U_i^n - W_i^{n+1})(aU_j^n - aU_i^n)c_{ij} \\
&= 2\tau \sum_{i\in\mathcal{V}} \sum_{j\in\mathcal{V}} (U_i^n - W_i^{n+1})(U_j^n - U_i^n)ac_{ij} \\
&\leq 2\tau \sum_{i\in\mathcal{V}} \sum_{j\in\mathcal{V}} |U_i^n - W_i^{n+1}|\ |U_j^n - U_i^n| d_{ij} \\
&\leq \tau \sum_{i\in\mathcal{V}} \sum_{j\in\mathcal{V}} d_{ij}(\epsilon(U_i - W_i)^2) + \tau \sum_{i\in\mathcal{V}} \sum_{j\in\mathcal{V}} d_{ij} \frac{(U_j - U_i)^2}{\epsilon}.
\end{aligned}
$$

In the last line, we used the inequality $2ab \leq \epsilon a^2 + \frac{b^2}{\epsilon}$ for all $\epsilon > 0$. (We will choose a suitable $\epsilon > 0$ later.)

$$
R_2 = 2\tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij}(U_j^n - U_i^n)W_i^{n+1} = \underbrace{2\tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij}(U_j^n - U_i^n)(W_i^{n+1} - U_i^n)}_{R_2.i} + \underbrace{2\tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij}(U_j^n - U_i^n)U_i^n}_{R_2.ii}.
$$

Further bounding these pieces:

$$
R_2.i = 2\tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij}(U_j^n - U_i^n)(W_i^{n+1} - U_i^n) \leq \tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij} \frac{(U_j - U_i)^2}{\epsilon} + \tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij}\epsilon(W_i^{n+1} - U_i^n)^2.
$$

$$
\begin{aligned}
R_2.ii = 2\tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij}(U_j^n - U_i^n)U_i &= \tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij}((U_j^n)^2 - (U_i^n)^2 - (U_j^n - U_i^n)^2) = \\
&\underbrace{\tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij}(U_j^n)^2 - \tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij}(U_i^n)^2}_{0} - \tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij}(U_i^n - U_j^n)^2.
\end{aligned}
$$

Let's put these computations together. First, we bring the remaining term of $R_2.ii$ to the left hand side. We then combine like terms between R_1 and $R_2.ii$. Finally, we subtract the graph viscosity term (ie, the $(d_{ij}(U_j^n - U_n^i)^2$ term) from both sides. We arrive at the following inequality:

$$
\|w_h^{n+1}\|_{l^2,m}^2 - \|u_h^n\|_{l^2,m}^2 + \|w_h^{n+1} - u_h^n\|_{l^2,m}^2 + \tau(1 - \frac{2}{\epsilon}) \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij}(U_j^n - U_i^n)^2 \leq 2\tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij}\epsilon(W_i^{n+1} - U_i^n)^2.
$$

Let us take $\epsilon = 2$, so the inequality simplifies to:

$$
\|w_h^{n+1}\|_{l^2,m}^2 - \|u_h^n\|_{l^2,m}^2 + \|w_h^{n+1} - u_h^n\|_{l^2,m}^2 \leq 4\tau \sum_{i\in\mathcal{V}} \sum_{j\in I(i)} d_{ij}(W_i^{n+1} - U_i^n)^2.
$$

Since $\tau < \tau^*$ by assumption, we have:

$$\|w_h^{n+1}\|_{l^2,m}^2 - \|u_h^n\|_{l^2,m}^2 + \|w_h^{n+1} - u_h^n\|_{l^2,m}^2 \leq \sum_{i\in\mathcal{V}} m_i (W_i^{n+1} - U_i^n)^2 = \|w_h^{n+1} - u_h^n\|_{l^2,m}^2.$$

Thus, we conclude:

$$\|w_h^{n+1}\|_{l^2,m}^2 \leq \|u_h^n\|_{l^2,m}^2. \tag{10}$$

□

Lemma 4. *Let $W^{n+1} \in \mathbb{R}^I$. Let U^{n+1} and Z^{n+1} be computed as in (7) and (8). Then the following estimates holds:*

$$\|U^{n+1}\|_{l^2,m} \leq \|W^{n+1}\|_{l^2,m} \tag{11}$$

Furthermore, there exists a $C_H > 0$, dependent upon only the space dimension, such that:

$$|z_h^{n+1}|_{H^1(D)}^2 \leq C_H \left\| \frac{u_h^{n+1} - w_h^{n+1}}{\tau} \right\|_{l^2,m}^2. \tag{12}$$

Proof. Using the definition of a_{ij}, we can write equation (7) as

$$m_i \frac{U_i^{n+1} - W_i^{n+1}}{\tau} - K \int_D \partial_x z_h^{n+1} \partial_x \phi_i \ \mathrm{d}x = 0. \tag{13}$$

Multiplying the above equation by $2\tau U_i^{n+1}$, we obtain:

$$m_i 2 U_i^{n+1}(U_i^{n+1} - W_i^{n+1}) - 2\tau K \int_D \partial_x z_h^{n+1} \partial_x U_i^{n+1} \partial_x \phi_i \ \mathrm{d}x = 0.$$

Using the identity $2a(a-b) = a^2 - b^2 + (a-b)^2$, we achieve:

$$m_i \left((U_i^{n+1})^2 - (W_i^{n+1})^2 + (U_i^{n+1} - W_i^{n+1})^2 \right) - 2\tau K U_i^{n+1} \int_D \partial_x z_h^{n+1} \partial_x \phi_i \ \mathrm{d}x = 0.$$

Now we sum over i:

$$\sum_{i\in\mathcal{I}} m_i (U_i^{n+1})^2 + \sum_{i\in\mathcal{I}} (U_i^{n+1} - W_i^{n+1})^2 - 2\tau K \int_D \partial_x z_h^{n+1} u_h^{n+1} \ \mathrm{d}x = \sum_{i\in\mathcal{I}} m_i (W_i^{n+1})^2.$$

Testing the second equation with $q_h = z_h^{n+1}$ we deduce the identity:

$$\int_D z_h^{n+1} \partial_x z_h^{n+1} \ \mathrm{d}x = \int_D \partial_x u_h^{n+1} \partial_x z_h^{n+1} \ \mathrm{d}x.$$

However, $\int_D z_h^{n+1} \partial_x z_h^{n+1} \ \mathrm{d}x = \int_D \partial_x \left(\frac{z_h^{n+1}}{2} \right)^2 \ \mathrm{d}x = 0$. Therefore, combining these equations, we arrive at:

$$\sum_{i\in\mathcal{I}} m_i (U_i^{n+1})^2 + \sum_{i\in\mathcal{I}} m_i (U_i^{n+1} - W_i^{n+1})^2 = \sum_{i\in\mathcal{I}} m_i (W_i^{n+1})^2.$$

After dropping the (positive) middle term, we obtain (11):

$$\|U^{n+1}\|_{l^2,m}^2 \leq \|W^{n+1}\|_{l^2,m}^2. \tag{14}$$

For equation (12), we multiply (13) this time by Z_i^{n+1} and sum over i. We arrive at:

$$K \int_D \partial_x z_h^{n+1} \partial_x z_h^{n+1} \ \mathrm{d}x = \sum_{i\in\mathcal{V}} Z_i^{n+1} \left(\frac{U_i^{n+1} - W_i^{n+1}}{\tau} \right) m_i.$$

We bound the right hand side as follows (we will choose $\epsilon > 0$ later):

$$\sum_{i\in\mathcal{V}} Z_i^{n+1}\left(\frac{U_i^{n+1}-W_i^{n+1}}{\tau}\right)m_i \le \sum_{i\in\mathcal{V}}\left(\frac{\epsilon}{2}\left(Z_i^{n+1}\right)^2 + \frac{1}{2\epsilon}\left(\frac{U_i^{n+1}-W_i^{n+1}}{\tau}\right)^2\right)m_i$$
$$= \frac{\epsilon}{2}\|z_h^{n+1}\|_{l^2,m}^2 + \frac{1}{2\epsilon}\left\|\frac{u_h^{n+1}-w_h^{n+1}}{\tau}\right\|_{l^2,m}^2$$

Since all norms are equivalent on finite dimensional spaces, there exists $C_d > 0$ such that:

$$\|z_h^{n+1}\|_{l^2,m}^2 \le C_d\|z_h^{n+1}\|_{L^2(D)}^2.$$

Now, since z_h^{n+1} is zero on one endpoint, we may invoke a Poincare inequality, namely, that there exists a $C_p > 0$ such that

$$\left\|z_h^{n+1}\right\|_{L^2(D)}^2 \le C_p|z_h^{n+1}|_{H^1(D)}^2.$$

We thus arrive at:

$$K|z_h^{n+1}|_{H^1(D)}^2 \le \frac{\epsilon}{2}C_dC_p|z_h^{n+1}|_{H^1(D)}^2 + \frac{1}{2\epsilon}\left\|\frac{u_h^{n+1}-w_h^{n+1}}{\tau}\right\|_{l^2,m}^2.$$

Choose $\epsilon > 0$ such that $\frac{\epsilon}{2}C_dC_p < K$. Thus:

$$(K - \frac{\epsilon}{2}C_dC_p)|z_h^{n+1}|_{H^1(D)}^2 \le \frac{1}{2\epsilon}\left\|\frac{u_h^{n+1}-w_h^{n+1}}{\tau}\right\|_{l^2,m}^2.$$

Divide the inequality by $(K - \frac{\epsilon}{2}C_dC_p)$, and we arrive at (12) with

$$C_H := \frac{1}{2\epsilon(K - \frac{\epsilon}{2}C_dC_p)}.$$

□

Remark 5. The estimates (11) and (12) imply that the implicit step gives rise to a unique U_i^{n+1} for each i, whereas Z_i^{n+1} is unique up to a constant. Indeed, this is part of reason why we insist upon Dirichlet boundary conditions for z_h^{n+1}: it makes the implicit problem well-posed.

Remark 6. We would like to emphasize that the above result does not depend up on the time step. Indeed, this is the advantage of the implicit step: stability holds regardless of τ.

Theorem 7. *Let $U^0 \in \mathbb{R}^I$. Let W^{n+1} and U^{n+1} be computed via (6) and (7) with (8) respectively. Define $\tau^* > 0$ by:*

$$\tau^* := \min_{i\in\mathcal{V}} \frac{m_i}{4\sum_{j\in I(i)} d_{ij}} = \min_{i\in\mathcal{V}} \frac{m_i}{4\sum_{j\in I(i)} a|c_{ij}|}.$$

Then for all $0 < \tau \le \tau^*$ and all $n \in \mathbb{N}$, the following estimate holds:

$$\|U^N\|_{l^2,m} \le \|U^0\|_{l^2,m}.$$

Proof. Let $U^0 \in \mathbb{R}^I$ and $0 < \tau \le \tau^*$. We proceed by induction. The $n = 0$ case is trivial. Now suppose $\|U^n\|_{l^2,m} \le \|U^0\|_{l^2,m}$. We wish to show that $\|U^{n+1}\|_{l^2,m} \le \|U^0\|_{l^2,m}$. By lemma 3, we have $\|W^{n+1}\|_{l^2,m} \le \|U^n\|_{l^2,m}$. Then, by lemma 4, we have (independent of τ) $\|U^{n+1}\|_{l^2,m} \le \|W^{n+1}\|_{l^2,m}$. Combining these two bounds gives the result. □

3 Implementation

3.1 Preliminaries

Assume $D = [0,1]$. Let us use piecewise linear elements. Let $h_i := x_{i+1} - x_i$. Let $i \in \{1 : N\} := \mathcal{V}$. Let x_i be a node. Then, for $i \in \{2 : N-1\}$, the global basis function corresponding to x_i is:

$$\phi_i(x) := \begin{cases} 0 & x \le x_{i-1} \\ \dfrac{x - x_{i-1}}{x_i - x_{i-1}} & x_{i-1} < x \le x_i \\ \dfrac{x - x_{i+1}}{x_i - x_{i+1}} & x_i < x < x_{i+1} \\ 0 & x \ge x_{i+1} \end{cases} = \begin{cases} 0 & x \le x_{i-1} \\ \dfrac{x - x_{i-1}}{h_{i-1}} & x_{i-1} < x \le x_i \\ \dfrac{x - x_{i+1}}{-h_i} & x_i < x < x_{i+1} \\ 0 & x \ge x_{i+1} \end{cases}.$$

Furthermore, the (weak) derivative of $\phi_i(x)$ is given by:

$$\partial_x \phi_i(x) = \begin{cases} 0 & x < x_{i-1} \\ 1/h_{i-1} & x_{i-1} < x < x_i \\ -1/h_i & x_i < x < x_{i+1} \\ 0 & x > x_{i+1}. \end{cases}$$

For $i \in \{1, N\}$, we have:

$$\phi_1(x) := \begin{cases} \dfrac{x - x_2}{-h_1} & x_1 < x \le x_2 \\ 0 & \text{else} \end{cases}; \quad \phi_N(x) := \begin{cases} \dfrac{x - x_{N-1}}{h_{N-1}} & x_{N-1} < x \le x_N \\ 0 & \text{else} \end{cases};$$

$$\partial_x \phi_1(x) = \begin{cases} -1/h_1 & x_1 < x < x_2 \\ 0 & \text{else} \end{cases}; \quad \partial_x \phi_N(x) = \begin{cases} 1/h_{N-1} & x_{N-1} < x < x_N \\ 0 & \text{else} \end{cases}.$$

We can now explicitly compute the coefficients found in equations (6), (7), and (8).
For $i \in \{2 : N-1\}$, we have:

$$m_i = \frac{h_i + h_{i-1}}{2}; \quad c_{i,i-1} = -\frac{1}{2}; \quad c_{i,i+1} = \frac{1}{2}; \quad a_{i,i-1} = -\frac{1}{h_i}; \quad a_{ii} = \frac{1}{h_{i-1}} + \frac{1}{h_{i-1}}; \quad a_{i,i+1} = -\frac{1}{h_i}.$$

3.2 (Linear Algebra section...?)

As before, assume $U^0 \in \mathbb{R}^I$ is a reasonable approximation of the initial data. Suppose we have computed $U^0, ..., U^N$. As outlined in section [...], we first, we compute W^{N+1} explicitly using equation (6) then implicitly solve for U^{N+1} and Z^{N+1}. In this section, we give the algebraic representation for both of these steps. To alleviate the notation, we simply to refer to W, U, and Z in this section.

3.2.1 Hyperbolic prediction

<u>Generic case</u>

Let $\gamma_i := \tau / m_i$. Expanding equation (6) and collecting like terms gives the explicit updates by:

$$W_1 = U_1(1 - \gamma_1(ac_{1,1} + d_{1,2})) + U_2\gamma_1(d_{1,2} - ac_{1,2}) \tag{15}$$

$$W_i = U_{i-1}\gamma_i(d_{i,i-1} - ac_{i,i-1}) + U_i(1 - \gamma_i(d_{i,i+1} + d_{i,i+1})) + U_{i+1}\gamma_i(d_{i,i+1} - ac_{i,i+1}) \quad i \in \{2 : N-1\} \tag{16}$$

$$W_N = U_N(1 - \gamma_N(ac_{N,N} + d_{N,N-1})) + U_{N-1}\gamma_N(d_{N,N-1} - ac_{N,N-1}). \tag{17}$$

<u>**Dirichlet Boundary Conditions:**</u>

In the case of Dirichlet boundary conditions, our system instead takes the form:

$$W_1 = 0 \tag{18}$$

$$W_i = U_{i-1}\gamma_i(d_{i,i-1} - ac_{i,i-1}) + U_i(1 - \gamma_i(d_{i,i+1} + d_{i,i+1})_)) + U_{i+1}\gamma_i(d_{i,i+1} - ac_{i,i+1}) \quad i \in \{2 : N-1\} \tag{19}$$

$$W_N = 0. \tag{20}$$

<u>**Periodic Boundary Conditions:**</u> In the case of periodic boundary conditions, (that is, we insist $W_N = W_1$ and identify U_0 with U_{N-1}, $U_{N+1} = U_2$, etc.), we need to further define the coefficients:

$$c_{1,0} := -1/2; \quad c_{N,N+1} = 1/2; \quad c_{11} = 0$$

Also: $d_{1,0}$ is defined similarly as d_{ij} with $c_{0,1}$. Our system of equation is thus:

$$W_1 = U_1(1 - \gamma_1(ac_{1,1} + d_{1,0} + d_{1,2})) + U_2\gamma_1(d_{1,2} - c_{1,2}) + U_{N-1}\gamma_1(d_{1,0} - c_{1,0}). \tag{21}$$

$$W_i = U_{i-1}\gamma_i(d_{i,i-1} - ac_{i,i-1}) + U_i(1 - \gamma_i(d_{i,i+1} + d_{i,i+1})_)) + U_{i+1}\gamma_i(d_{i,i+1} - ac_{i,i+1}) \quad i \in \{2 : N-1\} \tag{22}$$

$$W_N = W_1. \tag{23}$$

3.2.2 Dispersive Update

Now we show how to implicitly solve for U^{N+1} and Z^{N+1}. To do so, we want to pose the problem in form of: find $\mathbf{x} \in \mathbb{R}^{\mathbf{2I}}$ such that $A\mathbf{x} = \mathbf{b}$. Naturally, we need to assemble the matrix A in a sensible way that solves both U and Z from (7) and (8). To do so, we set several conventions. First, we insist that the *odd* rows of A solve the equations corresponding to (7) and that the *even* rows of A solve the equations corresponding to (8). Second, this implies that, for $X \in \mathbb{R}^{2I}$, odd components correspond to the U_i and even components correspond to the Z_i. Formally, this is: $U_i = X_{2i-1}$ and $Z_i = X_{2i}$, ie:

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ \vdots \\ X_{2N-1} \\ X_{2N} \end{bmatrix} \to \begin{bmatrix} U_1 \\ Z_1 \\ U_2 \\ Z_2 \\ \vdots \\ U_N \\ Z_N \end{bmatrix}.$$

This convention also implies that the odd components of b correspond to W and are 0 otherwise. That is, $b_{2i+1} = W_i$ and $b_{2i} = 0$ for $i \in \mathcal{I}$.

With this convention, we recast equations (7) and (8) in terms of X_i (for simplicity, let $\xi := -\tau K$):

$$m_i X_{2i-1} + \xi\left(X_{2(i-1)}a_{i,i-1} + X_{2i}a_i i + X_{2(i+1)}a_{i,i+1}\right) = m_i W_i \tag{24}$$

$$X_{2i-2}c_{i-1,i} + X_{2i+2)}c_{i+1,i} - X_{2i-3}a_{i,i-1} - X_{2i-1}a_{i,i} - X_{2i+1}a_{i,i+1} = 0. \tag{25}$$

$$X_2 = 0 \tag{26}$$

<u>**Generic case**</u>

First, let us assemble the generic matrix A, that is, A without taking boundary conditions into considerations. With this system of equations, we can write a formula for the components of A:

$$\text{if } i \text{ is odd;} \quad A_{ij} = \begin{cases} \xi a_{\frac{i+1}{2},\frac{i+1}{2}-1} & j = i-1 \\ m_i & j = i \\ \xi a_{\frac{i+1}{2},\frac{i+1}{2}} & j = i+1 \\ \xi a_{\frac{i+1}{2},\frac{i+1}{2}+1} & j = i+3 \\ 0 & \text{else} \end{cases} \tag{27}$$

$$\text{if } i > 2 \text{ is even;}\quad A_{ij} = \begin{cases} -a_{\frac{i}{2},\frac{i}{2}-1} & j = i-3 \\ c_{\frac{i}{2}-1,\frac{i}{2}} & j = i-2 \\ a_{\frac{i}{2},\frac{i}{2}} & j = i-1 \\ -a_{\frac{i}{2},\frac{i}{2}-1} & j = i+1 \\ c_{\frac{i}{2}+1,\frac{i}{2}} & j = i+2 \\ 0 & \text{else} \end{cases}. \tag{28}$$

$$A_{ij} = 1 \quad i = j = 2.$$

Similarly, the components of b are given by:

$$b_i = \begin{cases} W_{\frac{i+1}{2}} m_{\frac{i+1}{2}} & i \text{ is even} \\ 0 & \text{else} \end{cases}. \tag{29}$$

Therefore, the generic matrix A is of the form:

$$A = \begin{bmatrix} m_1 & \xi a_{1,1} & 0 & \xi a_{1,2} & 0 & 0 & 0 & \dots & 0 \\ -a_{1,1} & 0 & -a_{1,2} & c_{2,1} & 0 & 0 & 0 & \dots & 0 \\ 0 & \xi a_{2,1} & m_2 & \xi a_{2,2} & 0 & \xi a_{2,3} & 0 & \dots & 0 \\ -a_{2,1} & c_{1,2} & -a_{2,2} & 0 & -a_{2,3} & c_{3,2} & 0 & \dots & 0 \\ \vdots & & & & \ddots & & & & \\ 0 & \dots & & & 0 & 0 & \xi a_{N,N_1} & m_N & \xi a_{N,N} \\ 0 & \dots & & & 0 & -a_{N,N-1} & c_{N-1,N} & -a_{N,N} & 0 \end{bmatrix};$$

and the vector b takes the form:

$$b = \begin{bmatrix} W_1 m_1 \\ 0 \\ W_2 m_2 \\ 0 \\ \vdots \\ W_N m_N \\ 0 \end{bmatrix}.$$

Dirichlet Boundary Conditions

In the case with Dirichlet boundary conditions, we enforce $U^1 = Z^1 = U^N = Z^n = 0$. In terms of X, that is to say that $X^1 = X^2 = X^{2N-1} = X^{2N} = 0$. Thus, our system of equations becomes:

$$X_1 = 0 \tag{30}$$

$$X_2 = 0 \tag{31}$$

$$m_i X_{2i-1} + \xi\left(X_{2(i-1)} a_{i,i-1} + X_{2i} a_i i + X_{2(i+1)} a_{i,i+1}\right) = m_i W_i \quad i \in \{2 : N-1\} \tag{32}$$

$$X_{2i-2} c_{i-1,i} + X_{2i+2)} c_{i+1,i} - X_{2i-3} a_{i,i-1} - X_{2i-1} a_{i,i} - X_{2i+1} a_{i,i+1} = 0 \quad i \in \{2 : N-1\} \tag{33}$$

$$X_{2N-1} = 0 \tag{34}$$

$$X_{2N} = 0. \tag{35}$$

With these considerations, the matrix A takes the form:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & \xi a_{2,1} & m_2 & \xi a_{2,2} & 0 & \xi a_{2,3} & 0 & \dots & 0 \\ -a_{2,1} & c_{1,2} & -a_{2,2} & 0 & -a_{2,3} & c_{3,2} & 0 & \dots & 0 \\ \vdots & & & & \ddots & & & & \\ 0 & \dots & & & 0 & 0 & 0 & 1 & 0 \\ 0 & \dots & & & 0 & 0 & 0 & 0 & 1 \end{bmatrix};$$

and the vector b takes the form:

$$b = \begin{bmatrix} 0 \\ 0 \\ W_2 m_2 \\ 0 \\ \vdots \\ W_{N-1} m_{N-1} \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Periodic Boundary Conditions

In the case of periodic boundary conditions, we proceed as follows: The idea is to treat U^N and U^1 as the same, as well as Z^N and Z^1. That is, the effects felt by U^N are transferred to U^1 and likewise for Z^1 and Z^N. Formally, this requires that $U_N = U_1$, $Z_N = Z_1$, and that $U_0 = U_{N-1}$ and $Z_0 = Z_{N-1}$. In terms of equations (24) and (25), this is equivalent to saying $X_1 = X_{2N-1}$, $X_2 = X_{2N}$, $X_{-1} = X_{2N-2}$, and $X_{-2} = X_{2N-3}$. Note that these changes only affect the first two and last two rows of the matrix A and the vector b. We also must define more coefficients: $a_{1,0} := -1/h_1$ and $c_{0,1} := -1/2$. With these considerations in mind, our new system of equations is:

$$m_1 X_1 - \xi(X_{2N-2} a_{1,0} + X_2 a_{1,1} + X_4 a_{1,2}) = m_1 W_1 \tag{36}$$

$$X_2 = 0 \tag{37}$$

$$m_i X_{2i-1} + \xi \left(X_{2(i-1)} a_{i,i-1} + X_{2i} a_i i + X_{2(i+1)} a_{i,i+1} \right) = m_i W_i \quad i \in \{2 : N-1\} \tag{38}$$

$$X_{2i-2} c_{i-1,i} + X_{2i+2)} c_{i+1,i} - X_{2i-3} a_{i,i-1} - X_{2i-1} a_{i,i} - X_{2i+1} a_{i,i+1} = 0 \quad i \in \{2 : N-1\} \tag{39}$$

$$X_{2N-1} = X_1 \tag{40}$$

$$X_{2N} = X_2. \tag{41}$$

This system gives rise to the matrix A:

$$A = \begin{bmatrix} m_1 & \xi a_{1,1} & 0 & \xi a_{1,2} & 0 & \dots & 0 & \xi a_{1,0} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & \xi a_{2,1} & m_2 & \xi a_{2,2} & 0 & \xi a_{2,3} & 0 & \dots & 0 & 0 \\ -a_{2,1} & c_{1,2} & -a_{2,2} & 0 & -a_{2,3} & c_{3,2} & 0 & \dots & 0 & 0 \\ \vdots & & & & \ddots & & & & & \\ 1 & 0 & 0 & \dots & & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & \dots & & & 0 & 0 & 0 & 0 & -1 \end{bmatrix};$$

and the vector b takes the form:

$$
b = \begin{bmatrix} W_1 m_1 \\ 0 \\ W_2 m_2 \\ 0 \\ \vdots \\ W_{N-1} m_{N-1} \\ 0 \\ 0 \\ 0 \end{bmatrix}.
$$

4 KdV Equation

e Let $D = [b, c]$, $b \neq c$, with time domain $[0, T]$. We now proceed to the KdV equation.

$$
\partial_t u + a u \partial_x u + K \partial_{xxx} u = 0; \quad u(0, x) = u_0(x). \tag{42}
$$

As in the linear case, we introduce a variable, $z = \partial_x u$. Equation (42) then becomes the system:

$$
\partial_t u + a u \partial_x u + K \partial_{xx} z = 0; \quad u(x, 0) = u_0(x) \tag{43}
$$

$$
z = \partial_x u. \tag{44}
$$

Let us multiply (43) by a test function v and (44) by $\partial_x q$, where q is a test function. Now integrate. Applying integration by parts to the third term of (43), we arrive at the weak form: Find $(u, z) \in S \times S$ such that, for all $(v, q) \in V \times V$ and a.e. t:

$$
\int_D \partial_t u v + a u \partial_x u v - K \partial_x z \partial_x v \, \mathrm{d}x = 0 \tag{45}
$$

$$
\int_D z \partial_x q \, \mathrm{d}x = \int_D \partial_x u \partial_x q \, \mathrm{d}x.. \tag{46}
$$

Lemma 8. *Let (u, z) solve (45) and (46). Then, for all $t \in [0, T]$, we have:*

$$
\|u(\cdot, t)\|_{L^2[a,b]} = \|u_0\|_{L^2[a,b]}.
$$

Proof. Test (45) with $v = u$ and (46) with $q = z$. The system thus becomes:

$$
\int_D \partial_t u u \, \mathrm{d}x + \int_D u^2 \partial_x u \, \mathrm{d}x - K \int_D \partial_x z \partial_x u \, \mathrm{d}x = 0.
$$

$$
\int_D z \partial_x z \, \mathrm{d}x = \int_D \partial_x u \partial_x z \, \mathrm{d}x.
$$

The first equation is equivalent to:

$$
\partial_t \frac{\|u(\cdot, t)\|^2_{L^2(D)}}{2} + \int_D \partial_x \left(\frac{u^3}{3} \right) \, \mathrm{d}x - K \int_D \partial_x z \partial_x u \, \mathrm{d}x = 0.
$$

As in the linear case, the middle terms vanishes. Furthermore, in the second equation, the left hand side term is also zero. Therefore, the first equation becomes;

$$
\partial_t \frac{\|u(\cdot, t)\|^2_{L^2(D)}}{2} = 0.
$$

This completes the proof. □

4.1 KdV: FE discretization

4.1.1 Space discretization

Let us now set up the discretization in space. Let $D := [a,b]$ be the interval, and let $\{\mathcal{K}\}_{h>0}$ be a sequence of shape-regular meshes. Now let $h > 0$ be given. For each $K \in \mathcal{K}_h$, define $T_K : [0,1] \to K$ be an affine geometric mapping. For $n \in \mathbb{N}$, let $\widehat{\mathcal{P}}_n$ be the set of all polynomials of degree n on $[0,1]$. Furthermore, let $\mathcal{V} := \{1 : I\}$ be an enumeration of the degrees of freedom. Now we consider the space:

$$V_h := \{v : v \circ T_k \in \widehat{P} \text{ for all } K \in K_h\}.$$

We assume that a global basis exists such that it has the unity of partition property. That is, $\{\phi_j\}_{j\in\mathcal{V}}$ is a basis for V_h and that $\sum_{j\in\mathcal{V}} \phi_j(x) = 1$ for all $x \in D$. Using the basis functions, we define several quantities:

$$m_i := \int_D \phi_i(x)\,\mathrm{d}x; \quad c_{ij} := \int_D \phi_i(x)\partial_x\phi_j(x)\,\mathrm{d}x; \quad a_{ij} := \int_D \partial_x\phi_i(x)\partial_x\phi_j(x)\,\mathrm{d}x. \tag{47}$$

Assume throughout that $m_i > 0$. We introduce a norm on V_h :
For $v_h \in V_h$ with a basis expansion $v_h = \sum_{i\in\mathcal{V}} V_i\phi_i$, define

$$\|v_h\|^2_{l^2,m} := \sum_{i\in\mathcal{I}} V_i^2 m_i.$$

Let $I(i) := \{j : m(\mathrm{supp}(\phi_j) \cap \mathrm{supp}(\phi_i)) \neq 0\}$; that is, when ϕ_i and ϕ_j share support. This set is also known as the stencil. Let us give a few key properties.
We further define:

$$d^n_{ij} := \max(|U^n_i|, |U^n_j|)|c_{ij}|.$$

Let us give a few key properties:

$$c_{ij} = -c_{ji}; \qquad d_{ij} = d_{ji}; \qquad \sum_{j\in I(i)} c_{ij} = 0.$$

For $i \in \{2 : N-1\}$, we have, $c_{ii} = 0$. The coefficients c_{11} and $C_{N,N}$ will be altered depending on the boundary conditions.

4.1.2 Time discretization

Let t^n be the current time, $\tau^n > 0$, and $t^{n+1} := t^n + \tau_n$. We break the time discretization into two steps: (1) a hyperbolic prediction based only on the first two terms of the PDE, and (2) an implicit step that includes the constraint condition. Throughout, we assume that $u^0_h(x) := \sum_{j\in\mathcal{I}} U^0_j\phi_j(x)$ is a reasonable approximation of the initial data. Let $U^0 \in \mathbb{R}^I$ be column vector where the entries are the coefficients for u^0_h. Similarly, let $U^n \in \mathbb{R}^I$ be a column vector whose values are the coefficients for the approximation at time $t = t_n$, ie, $u^n_h(x) = \sum_{j\in\mathcal{I}} U^n_j\phi_j(x)$.

4.2 The scheme

Let $U^0 \in \mathbb{R}^I$, and suppose we have computed $U_1, ..., U_n$. We find U^{n+1} as follows.

4.2.1 Step 1: Hyperbolic prediction

We first compute the hyperbolic prediction, which we shall call W^{n+1}, explicitly via:

$$m_i \frac{W^{n+1} - U^n_i}{\tau_n} = -\sum_{j\in I(i)} \frac{(U^n_j)^2}{2} c_{ij} + \sum_{j\in I(i)} d^n_{ij}(U^n_j - U^n_i). \tag{48}$$

Remark 9. We draw the reader's attention to the difference between this discretization and that for the linear equation. In the linear equation, we were able to assume that $\tau_n = \tau > 0$ for all n. We cannot make such an assumption here. Indeed, at each step, we much compute a new CFL number before stepping forward in time. Indeed, this is a key difference between the linear case and the nonlinear case.

4.2.2 Step 2: Implicit update

Suppose we have computed W^{n+1} as outlined above. Now we compute U^{n+1} and Z^{n+1} by solving the following system of equations.

$$m_i \frac{U_i^{n+1} - W_i}{\tau_n} - K \sum_{j \in I(i)} Z_j^{n+1} a_{ij} = 0 \tag{49}$$

$$\sum_{j \in I(i)} Z_j^{n+1} c_{ji} = \sum_{j \in I(i)} U_j^{n+1} a_{ij}. \tag{50}$$

$$Z_0^{n+1} = 0. \tag{51}$$

4.3 l^2 Estimates

Because of the nonlinearity, we must first prove an additional lemma, namely that $\int_D \partial_x \left(\sum_j \frac{U_j^2}{2}\phi_j\right) \sum_i U_i \phi_i \, \mathrm{d}x = 0$ (compare with lemma 3). For simplicity of notation (and indeed, for more generality), we present the lemma in the case of a generic flux $f : \mathbb{R} \to \mathbb{R}$. The KdV equation is the case when $f(u) = a\frac{u^2}{2}$.

Lemma 10. *Let $\{\phi_j\}_{j\in\mathcal{V}}$ be a basis for piecewise linear finite element space V_h with the partition of unity property. Let $u_h \in V_h$. Then:*

$$\int_D \partial_x \left(\sum_{j \in \mathcal{V}} f(U_j)\phi_j \right) u_h \, \mathrm{d}x = 0.$$

Proof. Let $u_h = \sum_{j\in\mathcal{V}} U_i \phi_i$. Then:

$$\int_D \partial_x \left(\sum_j f(U_j)\phi_j \right) \sum_i U_i \phi_i \, \mathrm{d}x = \sum_j \sum_i f(U_j) U_i \int_D \partial_x \phi_j \phi_i \, \mathrm{d}x.$$

Now, the partition of unity property implies the following:

$$\sum_j f(U_i) \int_D \partial_x \phi_j \phi_i \, \mathrm{d}x = f(U_i) \int_D \partial_x (\sum_j \phi_j)\phi_i \, \mathrm{d}x = f(U_i) \int_D \partial_x (1) \phi_i \, \mathrm{d}x = 0.$$

Let $f_{ij} := \dfrac{f(U_j) - f(U_i)}{U_j - U_i}$. Observe that $f_{ij} = f_{ji}$. Now, we compute:

$$\begin{aligned}
\sum_j \sum_i f(U_j) U_i \int_D \partial_x \phi_j \phi_i \, \mathrm{d}x &= \sum_j \sum_i U_i (f(U_j) - f(U_i)) \int_D \partial_x \phi_j \phi_i \, \mathrm{d}x \\
&= \sum_j \sum_i U_i (U_j - U_i) \frac{f(U_j) - f(U_i)}{U_j - U_i} \int_D \partial_x \phi_j \phi_i \, \mathrm{d}x \\
&= \sum_j \sum_i (U_i^2 - U_j^2 - (U_j - U_i)^2) f_{ij} \int_D \partial_x \phi_j \phi_i \, \mathrm{d}x \\
&= \sum_i \sum_j \frac{U_i^2}{2} f_{ij} \int_D \partial_x \phi_j \phi_i \, \mathrm{d}x - \sum_j \sum_i \frac{U_j^2}{2} f_{ij} \int_D \partial_x \phi_j \phi_i \, \mathrm{d}x - \\
&\sum_j \sum_i (U_j - U_i)^2 f_{ij} \int_D \partial_x \phi_j \phi_i \, \mathrm{d}x.
\end{aligned}$$

We observe that the summand of the last term is skew-symmetric, and thus the sum is 0. Using a change of index with the second term and the fact that $\int_D \partial_x \phi_i \phi_j \, \mathrm{d}x = -\int_D \partial_x \phi_j \phi_i \, \mathrm{d}x$, the equation becomes:

$$\sum_j \sum_i U_i^2 f_{ij} \int_D \partial_x \phi_j \phi_i \, \mathrm{d}x.$$

Let $\{K_h\}$ denote the cells. Then, since ϕ_j is piecewise linear, $\partial_x \phi_j$ is constant over each cell. Furthermore, $\int_{K_i} \phi_i \, \mathrm{d}x \frac{1}{|K_i|} = \beta > 0$ for all i. Thus:

$$\begin{aligned}
\sum_j \sum_i U_i^2 f_{ij} \sum_{K \in K_h} \partial_x(\phi_j|_K) \int_K \phi_i \, \mathrm{d}x &= \sum U_i^2 f_{ij} \sum_{K \in K_h} \partial_x(\phi_j|_K) \frac{|K|}{|K|} \int_D \phi_i \, \mathrm{d}x \\
&= \sum_j \sum_i U_i^2 f_{ij} \sum_{K \in K_h} \int_K \partial_x \phi_j \, \mathrm{d}x \beta \\
&= \sum_j \sum_i U_i^2 f_{ij} \beta \int_D \partial_x \phi_j \, \mathrm{d}x \\
&= 0.
\end{aligned}$$

$\square$

With this result, we may now prove an l^2 estimate for the hyperbolic prediction.

Lemma 11. *Let $U^n \in \mathbb{R}^I$. Let W^{n+1} be computed via* (48). *Let $\tau_n^* > 0$ be defined by:*

$$\tau_n^* = \frac{1}{4} \min_{i \in \mathcal{V}} \frac{m_i}{\sum_{j \in I(i)} d_{ij}} = \frac{1}{4} \min_{i \in \mathcal{V}} \frac{m_i}{\max(|U_{i-1}^n|, |U_i^n|) + \max(|U_{i+1}^n|, |U_i^n|)}.$$

Then, for all $\tau_n \leq \tau_n^*$, the following estimate holds:

$$\|W^{n+1}\| l^2, m \leq \|U^n\|_{l^2,m}.$$

Proof. For simplicity, we remove the n from W and U, but keep the dependence on τ to emphasize the importance that τ_n is time-dependent. Multiply (48) $2\tau_n W_i^{n+1}$. This gives rise to:

$$2m_i W_i (W_i - U_i) = -2\tau_n W_i \sum_j \frac{U_j^2}{2} c_{ij} + 2\tau_n W_i \sum_j d_{ij}^n (U_j - U_i).$$

Using the identity $2a(a-b) = a^2 - b^2 + (a-b)^2$ and summing over i, we obtain:

$$\|W\|_{l^2,m}^2 - \|U\|_{l^2,m}^2 + \|W - U\|_{l^2,m}^2 = -2\tau_n \sum_i W_i \sum_j \frac{U_j^2}{2} c_{ij} + 2\tau_n \sum_i W_i \sum_j d_{ij}^n (U_j - U_i).$$

Let us analyze the first term of the right hand side.

$$
\begin{aligned}
-\sum_i 2\tau_n W_i \sum_j \frac{U_j^2}{2} c_{ij} &= -2\tau_n \sum_i \sum_j W_i \frac{U_j^2}{2} \int_D \partial_x \phi_j \phi_i \, \mathrm{d}x \\
&= 2\tau_n \sum_j \frac{U_j^2}{2} \int_D \partial_x \phi_j (-w_h) \, \mathrm{d}x \\
&= 2\tau_n \int_D \partial_x \left(\sum_j \frac{U_j^2}{2} \phi_j \right) (-w_h) \, \mathrm{d}x \\
&= 2\tau_n \int_D \partial_x \left(\sum_j \frac{U_j^2}{2} \phi_j \right) (u_h - w_h) \, \mathrm{d}x \quad \text{by lemma 10} \\
&= 2\tau_n \int_D \partial_x \left(\sum_j \frac{U_j^2}{2} \phi_j \right) \left(\sum_i (U_i - W_i) \phi_i \right) \, \mathrm{d}x \\
&= 2\tau_n \sum_j \sum_i \frac{U_j^2}{2} (U_i - W_i) \int_D \partial_x \phi_j \phi_i \, \mathrm{d}x \\
&= 2\tau_n \sum_j \sum_i (U_i - W_i)(U_j - U_i) \\
&= 2\tau_n \sum_j \sum_i (U_i - W_i)(U_j - U_i) \frac{1}{2} \frac{U_j^2 - U_i^2}{U_j - U_i} c_{ij} \\
&\le 2\tau_n |U_i - W_i||U_j - U_i| \frac{1}{2} \frac{U_j^2 - U_i^2}{U_j - U_i} \frac{c_{ij}}{|c_{ij}|} |c_{ij}| \\
&\le 2\tau_n \sum_i \sum_j |U_i - W_i||U_j - U_i| d_{ij}^n \\
&\le \tau_n \sum_j \sum_i d_{ij}^n \epsilon (U_i - W_i)^2 + \tau_n \sum_j \sum_i d_{ij}^n \frac{(U_j - U_i)^2}{\epsilon}.
\end{aligned}
$$

Therefore, we have:

$$
\|W\|_{l^2,m}^2 - \|U\|_{l^2,m}^2 + \|W - U\|_{l^2,m}^2 \le \tau_n \sum_j \sum_i d_{ij}^n \epsilon (U_i - W_i)^2 + \tau_n \sum_j \sum_i d_{ij}^n \frac{(U_j - U_i)^2}{\epsilon} + 2\tau_n \sum_i W_i \sum_j d_{ij}^n (U_j - U_i).
$$

Now, the last term can be analyzed as follows:

$$
\begin{aligned}
2\tau_n \sum_j \sum_i d_{ij}^n W_i (U_j - U_i) &= 2\tau_n \sum_j \sum_i d_{ij}^n (W_i - U_i + U_i)(U_j - U_i) \\
&= 2\tau_n \sum_j \sum_i d_{ij}^n (W_i - U_i)(U_j - U_i) + 2\tau_n \sum_j \sum_i U_i (U_j - U_i) d_{ij}^n.
\end{aligned}
$$

We observe that these are the $R_2.i$ and $R_2.ii$ terms as in lemma (3). Indeed, the rest of the proof follows as in lemma (3), but with the new τ_n^* condition. $\square$

Lemma 12. *Let $W^{n+1} \in \mathbb{R}^I$. Let U^{n+1} and Z^{n+1} be computed as in (49) and (50). Then the following estimates holds:*

$$
\|U^{n+1}\|_{l^2,m} \le \|W^{n+1}\|_{l^2,m} \tag{52}
$$

Furthermore, there exists a $C_H > 0$, dependent upon only the space dimension, such that:

$$|z_h^{n+1}|^2_{H^1(D)} \leq C_H \left\| \frac{u_h^{n+1} - w_h^{n+1}}{\tau} \right\|^2_{l^2,m}. \tag{53}$$

Proof. The proof is nearly identical to that of lemma (4). □

Theorem 13. *Let $U_0 \in \mathbb{R}^I$. Let W^{n+1}, U^{n+1}, and Z^{n+1} be computed as in (48), (49), and (50). Let $n \in \{1 : N\}$, and let τ_n^* be defined by:*

$$\tau_n^* = \frac{1}{4} \min_{i \in \mathcal{V}} \frac{m_i}{\max(|U_{i-1}^n|, |U_i^n|) + \max(|U_{i+1}^n|, |U_i^n|)}.$$

Then, for $\tau_n \leq \tau_n^*$, we have:

$$\|U^N\|_{l^2,m} \leq \|U^0\|_{l^2,m}.$$

Proof. We proceed by induction. The case $n = 0$ is trivial. Suppose that $\|U^n\|_{l^2,m} \leq \|U_0\|_{l^2,m}$. Then, by lemma (11) , $\|W^{n+1}\|_{l^2,m} \leq \|U^n\|_{l^2.m}$. By lemma (12), $\|U^{n+1}\|_{l^2,m} \leq \|W^{n+1}\|_{l^2,m}$. These two estimates give the result. □

5 Numerical illustrations

We demonstrate the methods described above. These methods were implemented using python in PyCharm.

5.1 Linear Transport

These tests demonstrate the hyperbolic update on its own. That is to say, in equation (6) we replace W^{n+1} with U^{n+1} and disregard the implicit part entirely. Furthermore, we examine two cases: when graph viscosity is included and when it is not (ie, when the $d_{ij} = 0$ and when they don't).

5.1.1 Linear transport with graph viscosity

#Dofs	L1 Error	rate	L2 Error	rate	Linf Error	rate
20	4.09 E-01	–	4.05 E-01	–	4.06 E-01	–
40	2.24 E-01	0.837	2.24 E-01	0.823	2.24 E-01	0.826
80	1.18 E-01	0.908	1.17 E-01	0.920	1.17 E-01	0.920
160	6.02 E-02	0.962	6.02 E-02	0.950	6.02 E-02	0.950
320	3.05 E-02	0.976	3.05 E-02	0.976	3.05 E-02	0.976

Table 1: Error: Scheme with graph viscosity

5.1.2 Linear transport without graph viscosity

#Dofs	L1 Error	rate	L2 Error	rate	Linf Error	rate
20	6.17 E-02	–	5.97 E-02	–	5.70 E-02	–
40	1.41 E-02	2.05	1.39 E-02	2.02	1.35 E-02	2.00
80	3.37 E-03	2.02	3.35 E-03	2.01	3.31 E-03	1.99
160	8.25 E-04	2.01	8.22 E-04	2.00	8.17 E-04	2.00
320	2.04 E-04	2.00	2.03 E-04	2.00	2.03 E-04	1.99

Table 2: Error: Scheme without graph viscosity