Nonconservative discontinuous Galerkin methods for shallow water moment models

by

Caleb Logemann

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Applied Mathematics

Program of Study Committee:
James Rossmanith, Major Professor
Hailiang Liu
Songting Luo
Alric Rothmayer
Jue Yan

Iowa State University

Ames, Iowa

2021

DEDICATION

I would like to dedicate this thesis to my wife Glenda and to my daughter Alice without whose support I would not have been able to complete this work.

Contents

List of Tables

List of Figures

ACKNOWLEDGMENTS

I would like to take this opportunity to express my thanks to those who helped me with various aspects of conducting research and the writing of this thesis. First and foremost, Dr. Susan D. Ross for her guidance, patience and support throughout this research and the writing of this thesis. Her insights and words of encouragement have often inspired me and renewed my hopes for completing my graduate education. I would also like to thank my committee members for their efforts and contributions to this work: Dr. August Tanner and Dr. Lewis Hargrave. I would additionally like to thank Dr. Tanner for his guidance throughout the initial stages of my graduate career and Dr. Hargrave for his inspirational teaching style.

ABSTRACT

I present a discontinuous Galerkin method for the generalized shallow water equations first introduced by Kowalski and Torrilhon. These generalized shallow water equations introduce vertical moments into the shallow flow's velocity profile. As a result of these additional moments a nonconservative term appears in these hyperbolic equations. We use the Dal Maso, Le Floch, and Murat theory of nonlinear hyperbolic systems in nonconservative form to correctly discretize this nonconservative term. Using this theory a high order discontinuous Galerkin method is presented for the generalized equations in two dimensions.

Chapter1. INTRODUCTION

In this thesis, I present discontinuous Galerkin methods for shallow water moment models. The shallow water moment equations (SWME) were first introduced by Kowalski and Torrilhon[9] in 2017. One of the main drawbacks of the standard shallow water equations is the lack of vertical resolution in the velocity. In other words the velocity of the fluid is identical at the top and bottom of the fluid. This may not always be a reasonable physical assumption. In particular when the model includes some drag along the bottom topography, the constant velocity restriction allows the drag to have a disproportionate effect on the height of the fluid. However users would like to avoid these issues without needing to go to the computational expense of a fully vertically resolved model. The shallow water moment equations propose a solution to this problem, by adding moments to the velocity profile. This allows for a more complicated description of the velocity, while still retaining relatively few equations in a lower dimensional problem.

One shortcoming of SWME model is that when there is more than one moment, the system is no longer globally hyperbolic. A modified model introduced by Koellermeier and Pimentel-Garcia[8] in 2020, resolves this issue. This model is known as the shallow water linearized moment equations (SWLME), and is provable hyperbolic. This model is derived by removing some terms from the SWME, while still retaining many of it's desirable features. In this work I expand on this model slightly by proving that it is hyperbolic in two dimensions as well as in one dimension.

I chose to numerically solve these systems using the discontinuous Galerkin method. The discontinuous Galerkin method was first introduced by Reed and Hill[10] in their 1973 report on neutron transport. The method was later formalized and popularized by a series of papers by Cockburn and Shu[3, 4, 2, 1, 5] in the 1990s. The discontinuous Galerkin method is a finite element method where the solution space and test space are discontinuous over the mesh element faces.

The main numerical challenge in discretizing the SWME or the SWLME, is that they both contain nonconservative products. In other words they can't be written conservative or divergence form. The traditional theory of weak solutions using distributions does not apply to nonconservative equations. Instead this thesis relies on the theory laid out by Dal Maso, LeFloch, and Murat[6] in their 1995 paper. The main idea of this theory is to include a regularization path at any possible discontinuities. This regularization path smooths out any discontinuities and then value of the solution is examined as the limit approaches the discontinuity. Through this limiting process Dal Maso et. al. are able to define a path integral as a measure which represents the nonconservative product at discontinuities.

Rhebergen et al.[11] were the first to translate the nonconservative product theory of Dal Maso, LeFloch, and Murat, to the discontinuous Galerkin method.

Chapter2. Discontinuous Galerkin Method

In this chapter, I describe the standard discontinuous Galerkin method for hyperbolic balance laws. I introduce the notation that I will be using throughout the thesis, and I describe some of the keys ideas needed for the implementation of these methods.

2.1 Generic Formulation

Consider a partial differential equation of the form

$$\underline{q}_t + \nabla \cdot_{\underline{x}} \underline{\underline{f}}\Big(\underline{q}, \underline{x}, t\Big) = \underline{s}\Big(\underline{q}, \underline{x}, t\Big) \quad \text{for } \underline{x} \in \Omega \subset \mathbb{R}^d \tag{2.1}$$

where $\underline{q}$ is a vector of N_e equations, $\underline{\underline{f}}$ is the flux function, and $\underline{s}$ is the source function. The flux function maps values in $\mathbb{R}^{N_e} \times \mathbb{R}^d \times \mathbb{R}^+$ into matrices in $\mathbb{R}^{N_e \times d}$. Sometimes the flux function is considered as a set of vector functions, where there is one vector for each spatial dimension. I will however use the matrix notation. The divergence of the flux function is the sum of the spatial derivatives of the columns of $\underline{\underline{f}}$, or in other words the divergence is over the last index of the matrix. The source function is a vector function from $\mathbb{R}^{N_e} \times \mathbb{R}^d \times \mathbb{R}^+$ into $\mathbb{R}^{N_e}$. These type of equations are known as balance laws and if the source function is zero, then they are called conservation laws. These equations need initial conditions and boundary conditions at all inflow points on the boundary $\partial\Omega$ to be well-defined. In other words we also have

$$\underline{q}(\underline{x}, 0) = \underline{q}_0(\underline{x}) \tag{2.2}$$

$$\underline{q}(\underline{x}, t) = \underline{q}_b(\underline{x}, t), \quad \underline{x} \in \partial\Omega \tag{2.3}$$

A boundary point is an inflow point if the eigenvalues of the jacobian of the flux function dotted into the outward point normal vector, $\underline{n} \cdot \underline{\underline{f}}'$, are negative. Specifically I am interested in when these type of equations are hyperbolic. Equations of this form are hyperbolic when the flux jacobian along any normal vector has real eigenvalues and is diagonalizable, that is when $\underline{n} \cdot \underline{\underline{f}}'$ is diagonalizable.

One interesting feature of hyperbolic equations is that they may form discontinuities even when the initial condition and boundary conditions are smooth. In contrast this is not true for elliptic and parabolic partial differential equations, which have much stricter regularity theory. Because the solutions of these equations may contain discontinuities, the theory focuses on what are known as weak solutions instead of pointwise solutions, which are also known as strong solutions. The discontinuous Galerkin method is based on the idea of weak solutions to these PDEs. Finding weak solutions to the original PDE require searching an infinite dimensional space of functions. The discontinuous Galerkin method instead approximates the solution using a finite dimensional space. The way the DG method does this is by partitioning the domain Ω as the set of elements K_i which I will label as $\Omega_h = \{K_i\}_{i=1}^N$. The DG method then tries to find a solution that is polynomial on each element. Mathematically we denote the set of possible solutions as

$$V_h^k = \left\{ \underline{q} \in L^1(\Omega \times \mathbb{R}^+) |\ \underline{q}\Big|_{K_i} \in \mathbb{P}^k(K_i) \right\}. \tag{2.4}$$

Another way of writing this is with a basis expansion on each element,

$$\underline{q}(\underline{x}, t)\Big|_{K_i} = \sum_{j=1}^{k} \left(\underline{Q}_i^j \phi_i^j(\underline{x}) \right) = \underline{\underline{Q}}_i \underline{\phi}_i(\underline{x}). \tag{2.5}$$

To specify the DG method we need a set of linearly independent polynomials to form a basis on each element. To make things simpler I will use a single basis on a canonical element, $\mathcal{K}$, and linear transformations from each mesh element and the canonical element. Let the spatial dimensions on the canonical element be denoted as $\underline{\xi}$, and I will denote the linear transformation from the mesh elements to the canonical elements and back as $\underline{c}_i(\underline{x}) : K_i \to \mathcal{K}$ and $\underline{b}_i(\underline{\xi}) : \mathcal{K} \to K_i$. Then if $\{\phi\}$ is a basis of $\mathbb{P}^k(\mathcal{K})$, we can describe a basis on each element with the linear transformations as follows,

$$\phi_i^k(\underline{x}) = \phi^k(\underline{c}_i(\underline{x})) \text{ and } \phi^k(\underline{\xi}) = \phi_i^k(b_i(\underline{\xi})). \tag{2.6}$$

The local statements of the discontinuous galerkin method

$$\int_{K_i} \underline{q}_t \underline{\phi}_i^T(\underline{x}) \,\mathrm{d}\underline{x} = \int_{K_i} \underline{\underline{f}}\Big(\underline{q}, \underline{x}, t\Big) \Big(\underline{\phi}_i'(\underline{x})\Big)^T \mathrm{d}\underline{x} - \int_{\partial K_i} \underline{\underline{f}}^* \underline{n} \underline{\phi}_i^T(\underline{x}) \,\mathrm{d}s + \int_{K_i} \underline{s}\Big(\underline{q}, \underline{x}, t\Big) \underline{\phi}_i^T(\underline{x}) \,\mathrm{d}\underline{x} \tag{2.7}$$

On each element, K_i the discontinuous Galerkin solution can be written as an expansion of the basis, that is $\underline{q}\Big|_{K_i} = \underline{\underline{Q}}_i \underline{\phi}_i(\underline{x})$. Substituting this expression into the statement of the method gives,

$$\int_{K_i} \underline{\underline{Q}}_{i,t} \underline{\phi}_i(\underline{x}) \underline{\phi}_i^T(\underline{x})\,\mathrm{d}\underline{x} = \int_{K_i} \underline{\underline{f}}\Big(\underline{\underline{Q}}_i \underline{\phi}_i(\underline{x}), \underline{x}, t\Big)\Big(\underline{\phi}_i'(\underline{x})\Big)^T \mathrm{d}\underline{x} - \int_{\partial K_i} \underline{\underline{f}}^* \underline{n} \underline{\phi}_i^T(\underline{x})\,\mathrm{d}s + \int_{K_i} \underline{s}\Big(\underline{\underline{Q}}_i \underline{\phi}_i(\underline{x}), \underline{x}, t\Big) \underline{\phi}_i^T(\underline{x})\,\mathrm{d}\underline{x} \tag{2.8}$$

Ideally we would like to only work with the basis functions on the canonical element, therefore using the function $\underline{b}_i(\underline{\xi})$, the integrals can be transformed onto the canonical element with a change of variables. The integral of the numerical flux on the boundary of the element, will be left on the mesh element as in each dimension this integral looks very different. More details are given in future sections. The DG formulation is now

$$\int_{\mathcal{K}} \underline{\underline{Q}}_{i,t} \underline{\phi}(\underline{\xi}) \underline{\phi}^T(\underline{\xi}) m_i\,\mathrm{d}\underline{\xi} = \int_{\mathcal{K}} \underline{\underline{f}}\Big(\underline{\underline{Q}}_i \underline{\phi}(\underline{\xi}), \underline{b}_i(\underline{\xi}), t\Big)\Big(\underline{\phi}'(\underline{\xi}) \underline{c}_i'(\underline{b}_i(\underline{\xi}))\Big)^T m_i\,\mathrm{d}\underline{\xi} \tag{2.9}$$

$$- \int_{\partial K_i} \underline{\underline{f}}^* \underline{n} \underline{\phi}_i^T(\underline{x})\,\mathrm{d}s + \int_{\mathcal{K}} \underline{s}\Big(\underline{\underline{Q}}_i \underline{\phi}(\underline{\xi}), \underline{b}_i(\underline{\xi}), t\Big) \underline{\phi}^T\Big(\underline{\xi}\Big) m_i\,\mathrm{d}\underline{\xi}, \tag{2.10}$$

where $m_i = \frac{|K_i|}{|\mathcal{K}|} = |b_i'(\underline{xi})|$ is the element metric and satisfies

$$\int_{K_i} \mathrm{d}\underline{x} = \int_{\mathcal{K}} m_i\,\mathrm{d}\underline{\xi}. \tag{2.11}$$

Simplifying and solving for $\underline{\underline{Q}}_{i,t}$ gives

$$\underline{\underline{Q}}_{i,t} = \int_{\mathcal{K}} \underline{\underline{f}}\Big(\underline{\underline{Q}}_i \underline{\phi}(\underline{\xi}), \underline{b}_i(\underline{\xi}), t\Big)\Big(\underline{\phi}'(\underline{\xi}) \underline{c}_i'(\underline{b}_i(\underline{\xi}))\Big)^T \mathrm{d}\underline{\xi} \underline{\underline{M}}^{-1} \tag{2.12}$$

$$- \int_{\partial K_i} \underline{\underline{f}}^* \underline{n} \underline{\phi}_i^T(\underline{x})\,\mathrm{d}s \underline{\underline{M}}^{-1} \frac{1}{m_i} + \int_{\mathcal{K}} \underline{s}\Big(\underline{\underline{Q}}_i \underline{\phi}(\underline{\xi}), \underline{b}_i(\underline{\xi}), t\Big) \underline{\phi}^T\Big(\underline{\xi}\Big)\,\mathrm{d}\underline{\xi} \underline{\underline{M}}^{-1}, \tag{2.13}$$

where M is the mass matrix on the canonical element. The mass matrix of a given basis on the canonical element is given by

$$M_{ij} = \int_{\mathcal{K}} \phi^i(\underline{\xi}) \phi^k(\underline{\xi})\,\mathrm{d}\underline{\xi} \tag{2.14}$$

or

$$M = \int_{\mathcal{K}} \underline{\phi}(\underline{\xi}) \underline{\phi}^T(\underline{\xi})\,\mathrm{d}\underline{\xi}. \tag{2.15}$$

In order to specify the discontinuous Galerkin method for a specify dimension and type of mesh element, a canonical element, $\mathcal{K}$, the linear transformations, $\underline{c}_i$ and $\underline{b}_i$, the basis $\underline{\phi}$, and boundary integral all need to be described.

2.2 One Dimension

Consider the one dimensional balance law given below.

$$\underline{q}_t + \underline{f}\left(\underline{q}, x, t\right)_x = \underline{s}(\underline{q}, x, t) \tag{2.16}$$

In one dimension the elements are $K_i = \left[x_{i-1/2}, x_{i+1/2}\right]$, where the center of the element is given by x_i and $\Delta x_i = |K_i| = x_{i+1/2} - x_{i-1/2}$. The canonical element is $\mathcal{K} = [-1, 1]$, and the linear transformations are $c_i(x) = (x - x_i)\frac{2}{\Delta x_i}$ and $b_i(\xi) = \frac{\Delta x_i}{2}\xi + x_i$. Then the element metric will be $m_i = \frac{\Delta x_i}{2}$. The boundary integral of the numerical flux is just the point value at the two boundary points.

The the DG method in one dimension can be expressed as

$$\underline{\underline{Q}}_{i,t} = \frac{2}{\Delta x_i}\int_{-1}^{1} \underline{\underline{f}}\left(\underline{\underline{Q}}_i\underline{\phi}(\xi), \underline{b}_i(\xi), t\right)\underline{\phi}_\xi^T(\xi)\,\mathrm{d}\xi \underline{\underline{M}}^{-1} \tag{2.17}$$

$$-\frac{2}{\Delta x_i}\left(\underline{f}^*_{i+1/2}\underline{\phi}^T(1) - \underline{f}^*_{i-1/2}\underline{\phi}^T(-1)\right)\underline{\underline{M}}^{-1} + \int_{-1}^{1} \underline{s}\left(\underline{\underline{Q}}_i\underline{\phi}(\xi), b_i(\xi), t\right)\underline{\phi}^T(\xi)\,\mathrm{d}\xi \underline{\underline{M}}^{-1}. \tag{2.18}$$

If the basis on the canonical element is orthonormal with orthogonality condition,

$$\frac{1}{2}\int_{-1}^{1} \phi^j(\xi)\phi^k(\xi)\,\mathrm{d}\xi = \delta_{jk}, \tag{2.19}$$

then the mass matrix and its inverse are given by $M = 2I$ and $M^{-1} = \frac{1}{2}I$. The DG method can then be simplified even further as

$$\underline{\underline{Q}}_{i,t} = \frac{1}{\Delta x_i}\int_{-1}^{1} \underline{\underline{f}}\left(\underline{\underline{Q}}_i\underline{\phi}(\xi), \underline{b}_i(\xi), t\right)\underline{\phi}_\xi^T(\xi)\,\mathrm{d}\xi \tag{2.20}$$

$$-\frac{1}{\Delta x_i}\left(\underline{f}^*_{i+1/2}\underline{\phi}^T(1) - \underline{f}^*_{i-1/2}\underline{\phi}^T(-1)\right) + \frac{1}{2}\int_{-1}^{1} \underline{s}\left(\underline{\underline{Q}}_i\underline{\phi}(\xi), b_i(\xi), t\right)\underline{\phi}^T(\xi)\,\mathrm{d}\xi. \tag{2.21}$$

The integrals can be evaluated easily using gaussian quadrature.

2.3 Two Dimensions

In two dimensions the flux function is a matrix function of size $N_e \times 2$. Often it is denoted as two vector functions $\underline{f}_1$ and $\underline{f}_2$ or $\underline{f}$ and $\underline{g}$, however I will denote it as the matrix function $\underline{\underline{f}} = \left[\underline{f}_1, \underline{f}_2\right] = \left[\underline{f}, \underline{g}\right]$. Also in two dimensions the boundary integral of the numerical flux is a line integral. A line integral can be expressed as a one dimensional integral through a parameterization of that line. Suppose we have a line $L(\underline{x}) = 0$, that can be parameterized by $\underline{l}(t) = \underline{x}$ for $t \in [t_1, t_2]$. Then the line integral can be written as

$$\int_L h(\underline{x}) \,\mathrm{d}s = \int_{t_1}^{t_2} h(\underline{l}(t)) \|\underline{l}'(t)\| \,\mathrm{d}t. \tag{2.22}$$

In two dimensions the canonical element will have a set of faces, $\mathcal{F} = \{f_j\}$. I will have a parameterization of each face of the canonical element, $r_j(s)$, with $s \in [-1, 1]$. Having $s \in [-1, 1]$ is convenient as 1D quadrature rules won't need to be transformed from their canonical intervals. The actual integral is over the faces of the mesh element, so the actual parameterization for the faces of the mesh element will be $\underline{b}_i(\underline{r}_j(t))$. In this way I will handle the transformation to the canonical element and the parameterization of the line in one step. Therefore the boundary integral of the numerical flux can be written as

$$\int_{\partial K_i} \underline{\underline{f}}^* \underline{n} \underline{\phi}_i^T(\underline{x}) \,\mathrm{d}s = \sum_{f_j \in \mathcal{F}} \left(\int_{-1}^{1} \underline{\underline{f}}^* \underline{n} \underline{\phi}^T(\underline{r}_j(s)) \left\| \underline{b}_i'(\underline{r}_j(s)) \underline{r}_j'(s) \right\| \,\mathrm{d}s \right) \tag{2.23}$$

In two dimensions the discontinuous galerkin formulation is therefore

$$\underline{\underline{Q}}_{i,t} = \int_{\mathcal{K}} \underline{\underline{f}}\left(\underline{\underline{Q}}_i \underline{\phi}(\underline{\xi}), \underline{b}_i(\underline{\xi}), t\right) \left(\underline{\phi}'(\underline{\xi}) \underline{c}_i'(\underline{b}_i(\underline{\xi}))\right)^T \mathrm{d}\underline{\xi} \underline{\underline{M}}^{-1} \tag{2.24}$$

$$- \sum_{f_j \in \mathcal{F}} \left(\int_{-1}^{1} \underline{\underline{f}}^* \underline{n} \underline{\phi}^T(\underline{r}_j(s)) \left\| \underline{b}_i'(\underline{r}_j(s)) \underline{r}_j'(s) \right\| \,\mathrm{d}s \right) \underline{\underline{M}}^{-1} \frac{1}{m_i} + \int_{\mathcal{K}} \underline{s}\left(\underline{\underline{Q}}_i \underline{\phi}(\underline{\xi}), \underline{b}_i(\underline{\xi}), t\right) \underline{\phi}^T\left(\underline{\xi}\right) \mathrm{d}\underline{\xi} \underline{\underline{M}}^{-1} \tag{2.25}$$

2.3.1 Rectangular Elements

Consider if the mesh contain rectangular elements, then $K_i = \left[x_{i-1/2}, x_{i+1/2}\right] \times \left[y_{i-1/2}, y_{i+1/2}\right]$. The center of the element is (x_i, y_i) with $\Delta x_i = x_{i+1/2} - x_{i-1/2}$ and $\Delta y_i = y_{i+1/2} - y_{i-1/2}$. The canonical element is $\mathcal{K} = [-1, 1] \times [-1, 1]$ with coordinates $\underline{\xi} = [\xi, \eta]$. The linear transformations are given by

$$\underline{b}_i(\underline{\xi}) = \left[\frac{\Delta x_i}{2}\xi + x_i, \frac{\Delta y_i}{2}\eta + y_i\right]^T \tag{2.26}$$

$$\underline{c}_i(\underline{x}) = \left[\frac{2}{\Delta x_i}(x - x_i), \frac{2}{\Delta y_i}(y - y_i)\right]^T \tag{2.27}$$

with Jacobians

$$\underline{\underline{b}}_i' = \begin{pmatrix} \frac{\Delta x_i}{2} & 0 \\ 0 & \frac{\Delta y_i}{2} \end{pmatrix} \tag{2.28}$$

$$\underline{\underline{c}}_i' = \begin{pmatrix} \frac{2}{\Delta x_i} & 0 \\ 0 & \frac{2}{\Delta y_i} \end{pmatrix} \tag{2.29}$$

The metric of element i is $m_i = \frac{\Delta x_i \Delta y_i}{4}$. Also the parameterizations of the left, right, bottom, and top faces, r_l, r_r, r_b, r_t respectively, are given by

$$r_l(s) = [-1, s] \tag{2.30}$$

$$r_r(s) = [1, s] \tag{2.31}$$

$$r_b(s) = [s, -1] \tag{2.32}$$

$$r_t(s) = [s, 1] \tag{2.33}$$

for $s \in [-1, 1]$. We can easily compute $\left\|\underline{\underline{b}}_i'(\underline{r}_f(s))\underline{r}_f'(s)\right\|$ for each face as well

$$\left\|\underline{\underline{b}}_i'(\underline{r}_l(s))\underline{r}_l'(s)\right\| = \frac{\Delta y_i}{2} \tag{2.34}$$

$$\left\|\underline{\underline{b}}_i'(\underline{r}_r(s))\underline{r}_r'(s)\right\| = \frac{\Delta y_i}{2} \tag{2.35}$$

$$\left\|\underline{\underline{b}}_i'(\underline{r}_b(s))\underline{r}_b'(s)\right\| = \frac{\Delta x_i}{2} \tag{2.36}$$

$$\left\|\underline{\underline{b}}_i'(\underline{r}_t(s))\underline{r}_t'(s)\right\| = \frac{\Delta x_i}{2} \tag{2.37}$$

Substituting all these into the formulation gives,

$$\underline{\underline{Q}}_{i,t} = \int_{\mathcal{K}} \frac{2}{\Delta x_i} \underline{f}_1\left(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i(\underline{\xi}), t\right) \underline{\phi}_\xi^T + \frac{2}{\Delta y_i} \underline{f}_2\left(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i(\underline{\xi}), t\right) \underline{\phi}_\eta^T \, \mathrm{d}\underline{\xi} \underline{\underline{M}}^{-1} \tag{2.38}$$

$$+ \frac{2}{\Delta x_i} \int_{-1}^{1} \underline{f}_1^*(b_i(\xi = -1, \eta)) \underline{\phi}^T(\xi = -1, \eta) \, \mathrm{d}s \underline{\underline{M}}^{-1} \tag{2.39}$$

$$- \frac{2}{\Delta x_i} \int_{-1}^{1} \underline{f}_1^*(b_i(\xi = 1, \eta)) \underline{\phi}^T(\xi = 1, \eta) \, \mathrm{d}s \underline{\underline{M}}^{-1} \tag{2.40}$$

$$+ \frac{2}{\Delta y_i} \int_{-1}^{1} \underline{f}_2^*(b_i(\xi, \eta = -1)) \underline{\phi}^T(\xi, \eta = -1) \, \mathrm{d}s \underline{\underline{M}}^{-1} \tag{2.41}$$

$$- \frac{2}{\Delta y_i} \int_{-1}^{1} \underline{f}_2^*(b_i(\xi, \eta = 1)) \underline{\phi}^T(\xi, \eta = 1) \, \mathrm{d}s \underline{\underline{M}}^{-1} \tag{2.42}$$

For the case of a legendre orthogonal basis with orthogonality condition

$$\frac{1}{4}\int_{\mathcal{K}} \phi^i(\underline{\xi})\phi^j(\underline{\xi})\,\mathrm{d}\xi = \delta_{ij},$$

then the mass matrix and it's inverse become $M = 4I$ and $M^{-1} = \frac{1}{4}I$. So the full method becomes,

$$\underline{\underline{Q}}_{i,t} = \int_{\mathcal{K}} \frac{1}{2\Delta x_i}\underline{f}_1\left(\underline{\underline{Q}}_i\underline{\phi}, \underline{b}_i(\underline{\xi}), t\right)\underline{\phi}_{\xi}^T + \frac{1}{2\Delta y_i}\underline{f}_2\left(\underline{\underline{Q}}_i\underline{\phi}, \underline{b}_i(\underline{\xi}), t\right)\underline{\phi}_{\eta}^T\,\mathrm{d}\underline{\xi} \tag{2.43}$$

$$+\frac{1}{2\Delta x_i}\int_{-1}^{1} \underline{f}_1^*(b_i(\xi = -1, \eta))\underline{\phi}^T(\xi = -1, \eta)\,\mathrm{d}s \tag{2.44}$$

$$-\frac{1}{2\Delta x_i}\int_{-1}^{1} \underline{f}_1^*(b_i(\xi = 1, \eta))\underline{\phi}^T(\xi = 1, \eta)\,\mathrm{d}s \tag{2.45}$$

$$+\frac{1}{2\Delta y_i}\int_{-1}^{1} \underline{f}_2^*(b_i(\xi, \eta = -1))\underline{\phi}^T(\xi, \eta = -1)\,\mathrm{d}s \tag{2.46}$$

$$-\frac{1}{2\Delta y_i}\int_{-1}^{1} \underline{f}_2^*(b_i(\xi, \eta = 1))\underline{\phi}^T(\xi, \eta = 1)\,\mathrm{d}s \tag{2.47}$$

2.3.2 Triangular Elements

Consider a mesh with triangular elements. That is each mesh element is given by three vertices in $\mathbb{R}^2$, $\{\underline{v}_1, \underline{v}_2, \underline{v}_3\}$. The coordinates of each vertex are given by $\underline{v}_i = [x_i, y_i]$. The canonical element that I will use is a right triangle with vertices, $[-1, 1]$, $[-1, -1]$, $[1, -1]$. The linear transformations between mesh elements and the canonical element are given by

$$\underline{b}_i(\underline{\xi}) = [b_{00}\xi + b_{01}\eta + b_{02}, b_{10}\xi + b_{11}\eta + b_{12}] \tag{2.48}$$

$$\underline{c}_i(\underline{x}) = [c_{00}x + c_{01}y + c_{02}, c_{10}x + c_{11}y + c_{12}] \tag{2.49}$$

where the coefficients are

$$b_{00} = \frac{1}{2}(x_3 - x_2) \tag{2.50}$$

$$b_{01} = \frac{1}{2}(x_1 - x_2) \tag{2.51}$$

$$b_{02} = \frac{1}{2}(x_1 + x_3) \tag{2.52}$$

$$b_{10} = \frac{1}{2}(y_3 - y_2) \tag{2.53}$$

$$b_{11} = \frac{1}{2}(y_1 - y_2) \tag{2.54}$$

$$b_{12} = \frac{1}{2}(y_1 + y_3) \tag{2.55}$$

$$c_{00} = \frac{-2(y_1 - y_2)}{y_1(x_2 - x_3) - x_1(y_2 - y_3) + y_2x_3 - x_2y_3} \tag{2.56}$$

$$c_{01} = \frac{2(x_1 - x_2)}{y_1(x_2 - x_3) - x_1(y_2 - y_3) + y_2x_3 - x_2y_3} \tag{2.57}$$

$$c_{02} = \frac{y_1(x_2 + x_3) - x_1(y_2 + y_3) - y_2x_3 + x_2y_3}{y_1(x_2 - x_3) - x_1(y_2 - y_3) + y_2x_3 - x_2y_3} \tag{2.58}$$

$$c_{10} = \frac{-2(y_2 - y_3)}{y_1(x_2 - x_3) - x_1(y_2 - y_3) + y_2x_3 - x_2y_3} \tag{2.59}$$

$$c_{11} = \frac{2(x_2 - x_3)}{y_1(x_2 - x_3) - x_1(y_2 - y_3) + y_2x_3 - x_2y_3} \tag{2.60}$$

$$c_{12} = \frac{x_1(y_2 - y_3) - y_1(x_2 - x_3) + y_2x_3 - x_2y_3}{y_1(x_2 - x_3) - x_1(y_2 - y_3) + y_2x_3 - x_2y_3} \tag{2.61}$$

These coefficients were found by doing a linear solve such that the vertices of the mesh element would be transformed to the vertices of the canonical element.

The jacobians of the linear transformations are

$$\underline{b}_i'(\underline{\xi}) = \begin{pmatrix} b_{00} & b_{01} \\ b_{10} & b_{11} \end{pmatrix} \tag{2.62}$$

$$\underline{c}_i'(\underline{x}) = \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix} \tag{2.63}$$

The metric of the element will be $m_i = \det(\underline{b}_i') = b_{00}b_{11} - b_{10}b_{01}$.

Also we can parameterize the left, bottom and hypotenuse faces of the canonical element as

$$r_l(s) = [-1, s] \tag{2.64}$$

$$r_b(s) = [s, -1] \tag{2.65}$$

$$r_h(s) = [s, -s] \tag{2.66}$$

for $s \in [-1, 1]$. We can easily compute $\left\|\underline{b}_i'(\underline{r}_f(s))\underline{r}_f'(s)\right\|$ for each face as well

$$\|\underline{b}_i'(\underline{r}_l(s))\underline{r}_l'(s)\| = \sqrt{b_{01}^2 + b_{11}^2} = \frac{1}{2}\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{2.67}$$

$$\|\underline{b}_i'(\underline{r}_b(s))\underline{r}_b'(s)\| = \sqrt{b_{00}^2 + b_{10}^2} = \frac{1}{2}\sqrt{(x_3 - x_2)^2 + (y_3 - y^2)^2} \tag{2.68}$$

$$\|\underline{b}_i'(\underline{r}_h(s))\underline{r}_h'(s)\| = \sqrt{(b_{00} - b_{01})^2 + (b_{10} - b_{11})^2} = \frac{1}{2}\sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2} \tag{2.69}$$

For the case of an orthonormal modal basis with orthogonality condition,

$$\frac{1}{2}\int_{\mathcal{K}} \phi^i(\underline{\xi})\phi^j(\underline{\xi})\,\mathrm{d}\underline{\xi} = \delta_{ij} \tag{2.70}$$

then the mass matrix and it's inverse will be $\underline{\underline{M}} = 2I$ and $\underline{\underline{M}}^{-1} = \frac{1}{2}I$.

Chapter3. The Models

3.1 Shallow Water Moment Models

The shallow water moment equations (SWME) were first introduced by Kowalski and Torrilhon. The goal of this new model is to add vertical resolution to the velocity of the shallow water equations. The standard shallow water equations make several key assumptions. The shallow water equations assume hydrostatic pressure and that the horizontal velocity is constant in the vertical direction. The assumption that the horizontal velocity is constant in the vertical direction is particularly restricting. One common approach to add vertical resolution the the shallow water models is the so-called multilayer shallow water model. The multilayer shallow water model assumes that the horizontal velocity consists of multiple layers of constant velocity. This approach can reflect nature, where the oceans and atmosphere do have multiple layers. However the multilayer model has a significant numerical downside. The multilayer model is not globally hyperbolic, which means that the problem can become ill-posed. When the velocities of the different layers become too different the system is no longer hyperbolic. In this case the fluid should create vortices at the interface between the layers. However the multilayer shallow water model does not allow for these roll-ups and so becomes ill-posed.

Kowalski and Torrilhon have introduced a new approach to adding vertical resolution to the shallow water equations, which has better hyperbolicity properties. The main idea of their approach is to approximate the horizontal velocity as an Ansatz expansion in the vertical direction, that is the velocities can be represented as

$$u(x, y, z, t) = u_m(x, y, t) + \sum_{j=1}^{N} (\alpha_j(x, y, t)\phi_j(z)) \tag{3.1}$$

$$v(x, y, z, t) = v_m(x, y, t) + \sum_{j=1}^{N} (\beta_j(x, y, t)\phi_j(z)), \tag{3.2}$$

where $u_m(x, y, t)$ and $v_m(x, y, t)$ are the mean velocities in the x and y directions respectively. In general the functions ϕ_j can be arbitrary. In fact if ϕ_j are characteristic functions, then the multilayer shallow water model can be derived. However in this work we will assume that ϕ_j are polynomials. This approach maintains computational efficiency compared with fully vertically resolved models.

3.1.1 Derivation

We begin by considering the Navier-Stokes equations,

$$\nabla \cdot \underline{u} = 0 \tag{3.3}$$

$$\underline{u}_t + \nabla \cdot (\underline{u}\underline{u}) = -\frac{1}{\rho}\nabla p + \frac{1}{\rho}\nabla \cdot \sigma + \underline{g}, \tag{3.4}$$

where $\underline{u} = [u, v, w]^T$ is the vector of velocities, p is the pressure, ρ is the constant density, σ is the deviatoric stress tensor, and $\underline{g}$ is the gravitational force vector. We also have two boundaries, the bottom topography $h_b(t, x, y)$, and the free surface $h_s(t, x, y)$. At both of these boundaries the kinematic boundary conditions are in effect and can be expressed as

$$(h_s)_t + [u(t, x, y, h_s), v(t, x, y, h_s)]^T \cdot \nabla h_s = w(t, x, y, h_s) \tag{3.5}$$

$$(h_b)_t + [u(t, x, y, h_b), v(t, x, y, h_b)]^T \cdot \nabla h_b = w(t, x, y, h_b). \tag{3.6}$$

In practice the bottom topography is unchanging in time, but we express h_b with time dependence to allow for a symmetric representation of the boundary conditions.

3.1.1.1 Dimensional Analysis

Now we consider the characteristic scales of the problem. Let L be the characteristic horizontal length scale, and let H be the characteristic vertical length scale. For this problem we assume that $H << L$ and we denote the ratio of these lengths as $\varepsilon = H/L$. With these characteristic lengths we can scale the length variables to a nondimensional form

$$x = L\hat{x}, \quad y = L\hat{y}, \quad z = H\hat{z}. \tag{3.7}$$

Now let U be the characteristic horizontal velocity, then because of the shallowness the characteristic vertical velocity will be εU. Therefore the velocity variables can be scaled as follows,

$$u = U\hat{u}, \quad v = U\hat{v}, \quad w = \varepsilon U\hat{w}. \tag{3.8}$$

Now with the characteristic length and velocity, the time scaling can be described as

$$t = \frac{L}{U}\hat{t} \tag{3.9}$$

The pressure will be scaled by the characteristic height, H, and the stresses will be scaled by a characteristic stress, S. It is assumed that the basal shear stresses, σ_{xz} and σ_{yz} are of larger order than the lateral shear stress, σ_{xy}, and the normal stresses, σ_{xx}, σ_{yy}, and σ_{zz}, so that

$$p = \rho g H\hat{p}, \quad \sigma_{xz/yz} = S\hat{\sigma}_{xz/yz}, \quad \sigma_{xx/xy/yy/zz} = \varepsilon S\hat{\sigma}_{xx/xy/yy/zz}. \tag{3.10}$$

Substituting all of these scaled variables into the Navier-Stokes system gives,

$$\hat{u}_{\hat{x}} + \hat{v}_{\hat{y}} + \hat{w}_{\hat{z}} = 0 \tag{3.11}$$

$$\varepsilon F^2\Big(\hat{u}_{\hat{t}} + \left(\hat{u}^2\right)_{\hat{x}} + (\hat{u}\hat{v})_{\hat{y}} + (\hat{u}\hat{w})_{\hat{z}}\Big) = -\varepsilon\hat{p}_{\hat{x}} + G\Big(\varepsilon^2(\hat{\sigma}_{xx})_{\hat{x}} + \varepsilon^2(\hat{\sigma}_{xy})_{\hat{y}} + (\hat{\sigma}_{xz})_{\hat{z}}\Big) + e_x \tag{3.12}$$

$$\varepsilon F^2\left(\hat{v}_{\hat{t}} + (\hat{u}\hat{v})_{\hat{x}} + \left(\hat{v}^2\right)_{\hat{y}} + (\hat{v}\hat{w})_{\hat{z}}\right) = -\varepsilon\hat{p}_{\hat{y}} + G\Big(\varepsilon^2(\hat{\sigma}_{xy})_{\hat{x}} + \varepsilon^2(\hat{\sigma}_{yy})_{\hat{y}} + (\hat{\sigma}_{yz})_{\hat{z}}\Big) + e_y \tag{3.13}$$

$$\varepsilon^2 F^2\Big(\hat{w}_{\hat{t}} + (\hat{u}\hat{w})_{\hat{x}} + (\hat{v}\hat{w})_{\hat{x}} + \left(\hat{w}^2\right)_{\hat{z}}\Big) = -\hat{p}_{\hat{z}} + \varepsilon G\Big((\hat{\sigma}_{xz})_{\hat{x}} + (\hat{\sigma}_{yz})_{\hat{y}} + (\hat{\sigma}_{zz})_{\hat{z}}\Big) + e_z \tag{3.14}$$

$$F = \frac{U}{\sqrt{gH}} \approx 1, \quad G = \frac{S}{\rho g H} < 1 \tag{3.15}$$

Drop terms with ε^2 and εG, giving

$$\hat{u}_{\hat{x}} + \hat{v}_{\hat{y}} + \hat{w}_{\hat{z}} = 0 \tag{3.16}$$

$$\varepsilon F^2\Big(\hat{u}_{\hat{t}} + \left(\hat{u}^2\right)_{\hat{x}} + (\hat{u}\hat{v})_{\hat{y}} + (\hat{u}\hat{w})_{\hat{z}}\Big) = -\varepsilon\hat{p}_{\hat{x}} + G(\hat{\sigma}_{xz})_{\hat{z}} + e_x \tag{3.17}$$

$$\varepsilon F^2\left(\hat{v}_{\hat{t}} + (\hat{u}\hat{v})_{\hat{x}} + \left(\hat{v}^2\right)_{\hat{y}} + (\hat{v}\hat{w})_{\hat{z}}\right) = -\varepsilon\hat{p}_{\hat{y}} + G(\hat{\sigma}_{yz})_{\hat{z}} + e_y \tag{3.18}$$

$$\hat{p}_{\hat{z}} = e_z \tag{3.19}$$

where we can solve for the hydrostatic pressure

$$\hat{p}(\hat{t}, \hat{x}, \hat{y}) = \Big(\hat{h}_s(\hat{t}, \hat{x}, \hat{y}) - \hat{z}\Big)e_z \tag{3.20}$$

For the rest of the derivation we will transform back into dimensional variables for readability purposes.

$$u_x + v_y + w_z = 0 \tag{3.21}$$

$$u_t + \left(u^2\right)_x + (uv)_y + (uw)_z = -\frac{1}{\rho}p_x + \frac{1}{\rho}(\sigma_{xz})_z + ge_x \tag{3.22}$$

$$v_t + (uv)_x + \left(v^2\right)_y + (vw)_z = -\frac{1}{\rho}p_y + \frac{1}{\rho}(\sigma_{yz})_z + ge_y \tag{3.23}$$

$$p(t,x,y,z) = (h_s(t,x,y) - z)\rho g e_z \tag{3.24}$$

3.1.1.2 Mapping

In order to make this system more accessible we will map the vertical variable z to the normalized variable ζ, through the transformation

$$\zeta(t,x,y,z) = \frac{z - h_b(t,x,y)}{h(t,x,y)}, \tag{3.25}$$

or equivalently

$$z(t,x,y,\zeta) = h(t,x,y)\zeta + h_b(t,x,y) \tag{3.26}$$

where $h(t,x,y) = h_s(t,x,y) - h_b(t,x,y)$. This transformation maps the vertical variable, z onto $\zeta \in [0,1]$. In order to transform the partial differential equations we consider a function $\Psi(t,x,y,z)$, then it's mapped counterpart $\tilde{\Psi}(t,x,y,\zeta)$ can be described as

$$\tilde{\Psi}(t,x,y,\zeta) = \Psi(t,x,y,z(t,x,y,\zeta)) = \Psi(t,x,y,h(t,x,y)\zeta + h_b(t,x,y)), \tag{3.27}$$

or equivalently

$$\Psi(t,x,y,z) = \tilde{\Psi}(t,x,y,\zeta(t,x,y,z)) = \tilde{\Psi}\left(t,x,y,\frac{z - h_b(t,x,y)}{h(t,x,y)}\right). \tag{3.28}$$

We also need to be able to map derivatives of functions in order to be able to map the differential equations. This can be described

$$\Psi_z(t,x,y,z) = \left(\tilde{\Psi}(t,x,y,\zeta(t,x,y,z))\right)_z \tag{3.29}$$

$$\Psi_z(t,x,y,z) = \tilde{\Psi}_\zeta(t,z,y,\zeta(t,x,y,z))\zeta_z(t,x,y,z) \tag{3.30}$$

$$\Psi_z(t,x,y,z) = \tilde{\Psi}_\zeta(t,z,y,\zeta(t,x,y,z))\frac{1}{h(t,x,y)} \tag{3.31}$$

$$h(t,x,y)\Psi_z(t,x,y,z) = \tilde{\Psi}_\zeta(t,z,y,\zeta(t,x,y,z)) \tag{3.32}$$

$$h\Psi_z = \tilde{\Psi}_\zeta \tag{3.33}$$

For the other variables, $\{t,x,y\}$, the partial derivatives are identical. Let $s \in \{t,x,y\}$, then

$$\zeta_s(t,x,y,z) = \left(\frac{z-h_b(t,x,y)}{h(t,x,y)}\right)_s \tag{3.34}$$

$$= -\frac{(z-h_b(t,x,y))h_s(t,x,y)}{h(t,x,y)^2} - \frac{(h_b)_s(t,x,y)}{h(t,x,y)} \tag{3.35}$$

$$= -\zeta(t,x,y,z)\frac{h_s(t,x,y)}{h(t,x,y)} - \frac{(h_b)_s(t,x,y)}{h(t,x,y)} \tag{3.36}$$

$$= -\frac{\zeta(t,x,y,z)h_s(t,x,y) + (h_b)_s(t,x,y)}{h(t,x,y)} \tag{3.37}$$

and

$$\Psi_s(t,x,y,z) = \left(\tilde{\Psi}(t,x,y,\zeta(t,x,y,z))\right)_s \tag{3.38}$$

$$\Psi_s(t,x,y,z) = \tilde{\Psi}_s(t,x,y,\zeta(t,x,y,z)) + \tilde{\Psi}_\zeta(t,x,y,\zeta(t,x,y,z))\zeta_s(t,x,y,z) \tag{3.39}$$

$$\Psi_s(t,x,y,z) = \tilde{\Psi}_s(t,x,y,\zeta) - \tilde{\Psi}_\zeta(t,x,y,\zeta)\left(\frac{\zeta h_s(t,x,y) + (h_b)_s(t,x,y)}{h(t,x,y)}\right) \tag{3.40}$$

$$h(t,x,y)\Psi_s(t,x,y,z) = h(t,x,y)\tilde{\Psi}_s(t,x,y,\zeta) - \tilde{\Psi}_\zeta(t,x,y,\zeta)(\zeta h_s(t,x,y) + (h_b)_s(t,x,y)) \tag{3.41}$$

$$h(t,x,y)\Psi_s(t,x,y,z) = h(t,x,y)\tilde{\Psi}_s(t,x,y,\zeta) - \tilde{\Psi}_\zeta(t,x,y,\zeta)(\zeta h_s(t,x,y) + (h_b)_s(t,x,y)) \tag{3.42}$$

$$h\Psi_s = h\tilde{\Psi}_s - \tilde{\Psi}_\zeta(\zeta h_s + (h_b)_s) \tag{3.43}$$

$$h\Psi_s = h\tilde{\Psi}_s + h_s\tilde{\Psi} - h_s\tilde{\Psi} - \tilde{\Psi}_\zeta(\zeta h + h_b)_s \tag{3.44}$$

$$h\Psi_s = \left(h\tilde{\Psi}\right)_s - \left(h_s\tilde{\Psi} + \tilde{\Psi}_\zeta(\zeta h + h_b)_s\right) \tag{3.45}$$

$$h\Psi_s = \left(h\tilde{\Psi}\right)_s - \left(\left((\zeta h + h_b)_\zeta\right)_s\tilde{\Psi} + \tilde{\Psi}_\zeta(\zeta h + h_b)_s\right) \tag{3.46}$$

$$h\Psi_s = \left(h\tilde{\Psi}\right)_s - \left(((\zeta h + h_b)_s)_\zeta\tilde{\Psi} + \tilde{\Psi}_\zeta(\zeta h + h_b)_s\right) \tag{3.47}$$

$$h\Psi_s = \left(h\tilde{\Psi}\right)_s - \left((\zeta h + h_b)_s\tilde{\Psi}\right)_\zeta \tag{3.48}$$

Mapping of the Mass Balance Equation Now we can use these differential transformations to map the continuity equation or mass balance equation onto the normalized space. We begin by multiplying the continuity equation by h

$$h(u_x + v_y + w_z) = 0, \tag{3.49}$$

and then transforming from z to ζ

$$(h\tilde{u})_x + (h\tilde{v})_y + \left(\tilde{w} - (\zeta h + h_b)_x\tilde{u} - (\zeta h + h_b)_y\tilde{v}\right)_\zeta = 0. \tag{3.50}$$

We can then integrate over ζ to find an explicit expression for w the vertical velocity.

$$\begin{gathered}\tilde{w}(t,x,y,\zeta) - \tilde{w}(t,x,y,0) = \\ -\int_0^\zeta (h\tilde{u})_x \,\mathrm{d}\zeta' - \int_0^\zeta (h\tilde{v})_y \,\mathrm{d}\zeta' + (\zeta h + h_b)_x\tilde{u} + (\zeta h + h_b)_y\tilde{v} - (h_b)_x\tilde{u} - (h_b)_y\tilde{v}.\end{gathered} \tag{3.51}$$

This can be simplified using the kinematic boundary condition at the bottom surface, to show that the vertical velocity can be expressed as

$$\tilde{w}(t,x,y,\zeta) = -\int_0^\zeta (h\tilde{u})_x \,\mathrm{d}\zeta' - \int_0^\zeta (h\tilde{v})_y \,\mathrm{d}\zeta' + (\zeta h + h_b)_x \tilde{u} + (\zeta h + h_b)_y \tilde{v}. \tag{3.52}$$

Lastly by consider the vertical velocity at the free surface and using the kinematic boundary condition at that surface we arrive at the mass conservation equation,

$$h_t + (hu_m)_x + (hv_m)_y = 0, \tag{3.53}$$

where $u_m = \int_0^1 \tilde{u}\,\mathrm{d}\zeta$ and $v_m = \int_0^1 \tilde{v}\,\mathrm{d}\zeta$ are the mean velocities in the x and y directions respectively. This mass conservation equation is identical to the corresponding equation in the standard shallow water equations.

Mapping of the Momentum Equations Next we map the conservation of momentum equations. Again we multiply by h,

$$hu_t + h\left(u^2\right)_x + h(uv)_y + h(uw)_z + \frac{1}{\rho}hp_x = \frac{1}{\rho}h(\sigma_{xz})_z + ghe_x \tag{3.54}$$

and transform from z to ζ,

$$(h\tilde{u})_t + \left(h\tilde{u}^2\right)_x + (h\tilde{u}\tilde{v})_y + \left(\tilde{u}\left(\tilde{w} - (\zeta h + h_b)_t - (\zeta h + h_b)_x\tilde{u} - (\zeta h + h_b)_y\tilde{v}\right)\right)_\zeta \tag{3.55}$$

$$+\frac{1}{\rho}(h\tilde{p})_x - \frac{1}{\rho}((\zeta h + h_b)_x\tilde{p})_\zeta = \frac{1}{\rho}(\tilde{\sigma}_{xz})_\zeta + ghe_x. \tag{3.56}$$

The hydrostatic pressure can be mapped onto ζ as

$$\tilde{p}(t,x,y,\zeta) = h(t,x,y)(1-\zeta)\rho g e_z, \tag{3.57}$$

and then the pressure terms in the momentum equation can be simplified in the following way,

$$\frac{1}{\rho}(h\tilde{p})_x - \frac{1}{\rho}((\zeta h + h_b)_x\tilde{p})_\zeta = \left(\frac{1}{2}h^2 g e_z\right)_x + (h_b)_x h g e_z. \tag{3.58}$$

The resulting momentum balance equation is

$$(h\tilde{u})_t + \left(h\tilde{u}^2 + \frac{1}{2}h^2 g e_z\right)_x + (h\tilde{u}\tilde{v})_y + \left(\tilde{u}\left(\tilde{w} - (\zeta h + h_b)_t - (\zeta h + h_b)_x\tilde{u} - (\zeta h + h_b)_y\tilde{v}\right)\right)_\zeta \tag{3.59}$$

$$= \frac{1}{\rho}(\tilde{\sigma}_{xz})_\zeta + gh(e_x - (h_b)_x e_z) \tag{3.60}$$

Next we consider the vertical coupling term, ω

$$\omega = \tilde{w} - (\zeta h + h_b)_t - (\zeta h + h_b)_x \tilde{u} - (\zeta h + h_b)_y \tilde{v}, \tag{3.61}$$

Using the expression for vertical velocity in (3.52), we find that,

$$\omega = -\left(h \int_0^\zeta \tilde{u}\, \mathrm{d}\zeta'\right)_x - \left(h \int_0^\zeta \tilde{v}\, \mathrm{d}\zeta'\right)_y - \zeta h_t, \tag{3.62}$$

and then using (3.53), the vertical coupling becomes

$$\omega = -\left(h \int_0^\zeta \tilde{u}_d\, \mathrm{d}\zeta'\right)_x - \left(h \int_0^\zeta \tilde{v}_d\, \mathrm{d}\zeta'\right)_y, \tag{3.63}$$

where

$$\tilde{u}_d = \tilde{u} - u_m \quad \tilde{v}_d = \tilde{v} - v_m. \tag{3.64}$$

Thus the x momentum equation can be concisely written as

$$(h\tilde{u})_t + \left(h\tilde{u}^2 + \frac{1}{2}h^2 g e_z\right)_x + (h\tilde{u}\tilde{v})_y + (\tilde{u}\omega)_\zeta = \frac{1}{\rho}(\tilde{\sigma}_{xz})_\zeta + gh(e_x - (h_b)_x e_z) \tag{3.65}$$

Similarly the y-momentum equation can be written

$$(h\tilde{v})_t + (h\tilde{u}\tilde{v})_x + \left(h\tilde{v}^2 + \frac{1}{2}h^2 g e_z\right)_y + (\tilde{u}\omega)_\zeta = \frac{1}{\rho}(\tilde{\sigma}_{yz})_\zeta + gh\Big(e_y - (h_b)_y e_z\Big) \tag{3.66}$$

Newtonian Closure For the remainder of the derivation, a Newtonian flow is considered. In this case the deviatoric stress tensor terms are related to the velocities according to

$$\sigma_{xz} = \mu u_z, \qquad \sigma_{yz} = \mu v_z \tag{3.67}$$

where μ is the dynamic viscosity of the fluid. These terms can be mapped from the z domain to the zeta domain as follows,

$$\frac{1}{\rho}\tilde{\sigma}_{xz} = \frac{\nu}{h}\tilde{u}_\zeta, \qquad \frac{1}{\rho}\tilde{\sigma}_{yz} = \frac{\nu}{h}\tilde{v}_\zeta, \tag{3.68}$$

where $\nu = \frac{\mu}{\rho}$ is the kinematic viscosity. Boundary conditions also need to be specified at the bottom topography and at the free surface. Stress free conditions will be assumed at the free surface,

$$u_z|_{z=h_s} = 0 \qquad v_z|_{z=h_s} = 0. \tag{3.69}$$

At the bottom topography I will use a slip boundary condition of the form

$$(u - \lambda u_z)|_{z=h_b} = 0 \qquad (v - \lambda v_z)|_{z=h_b} = 0. \tag{3.70}$$

Mapping these boundary conditions from z to zeta gives

$$\tilde{u}_\zeta|_{\zeta=1} = 0 \qquad \tilde{v}_\zeta|_{\zeta=1} = 0 \tag{3.71}$$

$$\tilde{u}_\zeta|_{\zeta=0} = \frac{h}{\lambda}\,\tilde{u}|_{\zeta=0} \qquad \tilde{v}_\zeta|_{\zeta=0} = \frac{h}{\lambda}\,\tilde{v}|_{\zeta=0}\,. \tag{3.72}$$

Moment Closure Now that we have mapped the mass, momentum, and boundary equations onto the new reference frame of $\zeta \in [0,1]$, I will drop the tilde notation for readability. We can see the vertically resolved reference system, has the form,

$$h_t + (hu_m)_x + (hv_m)_y = 0 \tag{3.73}$$

$$(hu)_t + \left(hu^2 + \frac{1}{2}h^2 g e_z\right)_x + (huv)_y + (u\omega)_\zeta = \frac{\nu}{h}(u_\zeta)_\zeta + gh(e_x - (h_b)_x e_z) \tag{3.74}$$

$$(hv)_t + (huv)_x + \left(hv^2 + \frac{1}{2}h^2 g e_z\right)_y + (u\omega)_\zeta = \frac{\nu}{h}(v_\zeta)_\zeta + gh\Big(e_y - (h_b)_y e_z\Big) \tag{3.75}$$

The final step in deriving this system of equations is depth-averaging the equations. This removes the explicit dependence on ζ, and creates evolution equations for all of the individual moments. This is done by using the moment expansion

$$u(x,y,\zeta,t) = u_m(x,y,t) + \sum_{j=1}^{N} (\alpha_j(x,y,t)\phi_j(\zeta)) \tag{3.76}$$

$$v(x,y,\zeta,t) = v_m(x,y,t) + \sum_{j=1}^{N} (\beta_j(x,y,t)\phi_j(\zeta)), \tag{3.77}$$

for the vertically resolved velocities. The velocity is approximated by this polynomial expansion, and the conservation of momentum equation can be used to derive time evolution equations for the mean velocities, u_m and v_m, and the moment coefficients, α_j and β_j. The polynomials that are used in this work are Legendre polynomials orthogonal on $[0,1]$ and scaled so that $\phi_j(0) = 1$. The first few polynomials are given by

$$\phi_0(\zeta) = 1, \qquad \phi_1(\zeta) = 1 - 2\zeta, \qquad \phi_2(\zeta) = 1 - 6\zeta + 6\zeta^2. \tag{3.78}$$

Note that the mean velocities can be interpreted as α_0 and β_0 if desired. By multiplying the momentum equations by the legendre polynomials and depth averaging, orthogonality will give equations for the mean velocities and moments. Note that the boundary conditions are weakly enforced by the conditions

$$-\frac{\nu}{h}\, u_\zeta\big|_{\zeta=0}\,\zeta = 1 = \frac{\nu}{\lambda}\, u\big|_{\zeta=0}\,. \tag{3.79}$$

Also the vertical coupling ω disappears because it is zero at the bottom topography and at the free surface. The resulting system after depth averaging is

$$h_t + (hu)_x + (hv)_x = 0 \tag{3.80}$$

$$\begin{aligned}(hu)_t + \left(hu^2 + h\sum_{j=1}^{N}\left(\frac{1}{2j+1}\alpha_j^2\right) + \frac{1}{2}ge_z h^2\right)_x + \left(huv + h\sum_{j=1}^{N}\left(\frac{1}{2j+1}\alpha_j\beta_j\right)\right)_y \\ = -\frac{\nu}{\lambda}\left(u + \sum_{j=1}^{N}(\alpha_j)\right) + hge_x - hge_z(h_b)_x\end{aligned} \tag{3.81}$$

$$\begin{aligned}(hv)_t + \left(huv + h\sum_{j=1}^{N}\left(\frac{1}{2j+1}\alpha_j\beta_j\right)\right)_x + \left(hv^2 + h\sum_{j=1}^{N}\left(\frac{1}{2j+1}\beta_j^2\right) + \frac{1}{2}ge_z h^2\right)_y \\ = -\frac{\nu}{\lambda}\left(v + \sum_{j=1}^{N}(\beta_j)\right) + hge_y - hge_z(h_b)_y\end{aligned} \tag{3.82}$$

$$\begin{aligned}(h\alpha_i)_t + \left(2hu\alpha_i + h\sum_{j=1}^{N}\left(\sum_{k=1}^{N}(A_{ijk}\alpha_j\alpha_k)\right)\right)_x + \left(hu\beta_i + hv\alpha_i + h\sum_{j=1}^{N}\left(\sum_{k=1}^{N}(A_{ijk}\alpha_j\beta_k)\right)\right)_y \\ = u_m D_i - \sum_{j=1}^{N}\left(D_j\sum_{k=1}^{N}(B_{ijk}\alpha_k)\right) - (2i+1)\frac{\nu}{\lambda}\left(u + \sum_{j=1}^{N}\left(\left(1 + \frac{\lambda}{h}C_{ij}\right)\alpha_j\right)\right)\end{aligned} \tag{3.83}$$

$$\begin{aligned}(h\beta_i)_t + \left(hu\beta_i + hv\alpha_i + h\sum_{j=1}^{N}\left(\sum_{k=1}^{N}(A_{ijk}\alpha_j\beta_k)\right)\right)_x + \left(2hv\beta_i + h\sum_{j=1}^{N}\left(\sum_{k=1}^{N}(A_{ijk}\beta_j\beta_k)\right)\right)_y \\ = v_m D_i - \sum_{j=1}^{N}\left(D_j\sum_{k=1}^{N}(B_{ijk}\beta_k)\right) - (2i+1)\frac{\nu}{\lambda}\left(v + \sum_{j=1}^{N}\left(\left(1 + \frac{\lambda}{h}C_{ij}\right)\beta_j\right)\right)\end{aligned} \tag{3.84}$$

where

$$A_{ijk} = (2i+1)\int_0^1 \phi_i\phi_j\phi_k \, \mathrm{d}\zeta \tag{3.85}$$

$$B_{ijk} = (2i+1)\int_0^1 \phi_i'\left(\int_0^\zeta \phi_j \, \mathrm{d}\hat{\zeta}\right)\phi_k \, \mathrm{d}\zeta \tag{3.86}$$

$$C_{ij} = \int_0^1 \phi_i'\phi_j' \, \mathrm{d}\zeta \tag{3.87}$$

$$D_i = (h\alpha_i)_x + (h\beta_i)_y. \tag{3.88}$$

3.1.2 Example Systems

3.1.2.1 One Dimensional Equations

In one dimension the generalized shallow water equations will have the following form,

$$\underline{q}_t + \underline{f}\left(\underline{q}\right)_x = g(\underline{q})\underline{q}_x + \underline{p}. \tag{3.89}$$

In this case the unknown $\underline{q}$ will have the form

$$\underline{q} = [h, hu, h\alpha_1, h\alpha_2, \ldots]^T, \tag{3.90}$$

where the number of components depends on the number of moments in the velocity profiles.

The wavespeed of this system is given by the the eigenvalues of the matrix when the system is in quasilinear form. For this system we need to look at the eigenvalues of the matrix $f'(\underline{q}) - g(\underline{q})$. If all of the eigenvalues are real, then this system is considered hyperbolic. Also if the gradient of the eigenvalue with respect to the conserved variables dotted with the corresponding eigenvector is always positive or always negative, then we say that the system is convex. That is if

$$\nabla\lambda_i.\underline{v}_i < 0 \quad \text{or} \quad \nabla\lambda_i.\underline{v}_i > 0,$$

then the system is convex. If the dot product is zero, then we have a degenerate wave.

Zeroth Order The Zeroth order system is equivalent to standard shallow water equations. The flux function and source function are given by,

$$\underline{f}\left(\underline{q}\right) = \begin{pmatrix} hu \\ \frac{1}{2}e_z gh^2 + hu^2 \end{pmatrix} \tag{3.91}$$

$$\underline{p}\left(\underline{q}\right) = \begin{pmatrix} 0 \\ -(e_z \frac{\partial}{\partial x} h_b - e_x)gh - \frac{\nu u}{\lambda} \end{pmatrix}. \tag{3.92}$$

Thy hyperbolicity of the system can be checked with the flux jacobian, which is equivalent to the quasilinear matrix, as this system is conservative. The flux jacobian and its eigenvalues, λ, are shown below as

$$\underline{f}'(\underline{q}) = A = \begin{pmatrix} 0 & 1 \\ e_z gh - u^2 & 2u \end{pmatrix} \tag{3.93}$$

Quasilinear Matrix Eigenvalues

$$\lambda = u \pm \sqrt{gh} \tag{3.94}$$

Also this system is convex, which can be verified as follows,

$$\nabla \lambda_i \cdot \underline{v}_i = \pm \frac{3}{2} \frac{g}{\sqrt{gh}}. \tag{3.95}$$

First Order For the first order system, one moment is introduced into the velocity. In this case the velocity can be described as a line in the vertical direction. With the addition of this moment the nonconservative matrix is introduced, and is shown below with the flux function and source function.

$$\underline{f}\Big(\underline{q}\Big) = \begin{pmatrix} hu \\ \frac{1}{2}e_z gh^2 + \frac{1}{3}\alpha_1^2 h + hu^2 \\ 2\alpha_1 hu \end{pmatrix} \tag{3.96}$$

$$\underline{\underline{g}}\Big(\underline{q}\Big) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & u \end{pmatrix} \tag{3.97}$$

$$\underline{p}\Big(\underline{q}\Big) = \begin{pmatrix} 0 \\ -(e_z \frac{\partial}{\partial x} h_b - e_x)gh - \frac{\nu(\alpha_1+u)}{\lambda} \\ -\frac{3((\frac{4\lambda}{h}+1)\alpha_1+u)\nu}{\lambda} \end{pmatrix} \tag{3.98}$$

The hyperbolicity of this system can be checked by computing the flux jacobian and the quasilinear matrix, $A = \underline{f}'(\underline{q}) - g(q)$, of this system.

$$\underline{f}'(\underline{q}) = \begin{pmatrix} 0 & 1 & 0 \\ e_z gh - \frac{1}{3}\alpha_1^2 - u^2 & 2u & \frac{2}{3}\alpha_1 \\ -2\alpha_1 u & 2\alpha_1 & 2u \end{pmatrix} \tag{3.99}$$

$$A = \begin{pmatrix} 0 & 1 & 0 \\ e_z gh - \frac{1}{3}\alpha_1^2 - u^2 & 2u & \frac{2}{3}\alpha_1 \\ -2\alpha_1 u & 2\alpha_1 & u \end{pmatrix} \tag{3.100}$$

The eigenvalues of the quasilinear matrix can be computed as

$$\lambda = u \pm \sqrt{gh + \alpha_1^2}, u. \tag{3.101}$$

Also the convexity of the system can be checked with the following dot products.

$$\nabla\lambda_1 \cdot \underline{v}_1 = -\frac{1}{2}\Big(3gh + 4\alpha_1^2\Big)\frac{\sqrt{gh + \alpha_1^2}}{gh^2 + h\alpha_1^2} \tag{3.102}$$

$$\nabla\lambda_2 \cdot \underline{v}_2 = \frac{\sqrt{gh + \alpha_1^2}}{h} \tag{3.103}$$

$$\nabla\lambda_3 \cdot \underline{v}_3 = -\frac{1}{2}\Big(2gh + \alpha_1^2\Big)\frac{\sqrt{gh + \alpha_1^2}}{gh^2 + h\alpha_1^2} \tag{3.104}$$

Second Order For the second order system there are two moments, and the velocity in the vertical direction is parabolic. The flux function, nonconservative matrix, and source function are given as

$$\underline{f}\left(\underline{q}\right) = \begin{pmatrix} hu \\ \frac{1}{2}e_z g h^2 + hu^2 + \frac{1}{15}(5\alpha_1^2 + 3\alpha_2^2)h \\ \frac{4}{5}\alpha_1\alpha_2 h + 2\alpha_1 hu \\ 2\alpha_2 hu + \frac{2}{21}(7\alpha_1^2 + 3\alpha_2^2)h \end{pmatrix} \tag{3.105}$$

$$g\left(\underline{q}\right) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{5}\alpha_2 + u & \frac{1}{5}\alpha_1 \\ 0 & 0 & \alpha_1 & \frac{1}{7}\alpha_2 + u \end{pmatrix} \tag{3.106}$$

$$\underline{p}\left(\underline{q}\right) = \begin{pmatrix} 0 \\ -(e_z \frac{\partial}{\partial x} h_b - e_x) g h - \frac{\nu(\alpha_1 + \alpha_2 + u)}{\lambda} \\ -\frac{3((\frac{4\lambda}{h}+1)\alpha_1 + \alpha_2 + u)\nu}{\lambda} \\ -\frac{5((\frac{12\lambda}{h}+1)\alpha_2 + \alpha_1 + u)\nu}{\lambda} \end{pmatrix} \tag{3.107}$$

The flux jacobian and quasilinear matrix can be computed to be

$$\underline{f}'(\underline{q}) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ e_z g h - \frac{1}{3}\alpha_1^2 - \frac{1}{5}\alpha_2^2 - u^2 & 2u & \frac{2}{3}\alpha_1 & \frac{2}{5}\alpha_2 \\ -\frac{4}{5}\alpha_1\alpha_2 - 2\alpha_1 u & 2\alpha_1 & \frac{4}{5}\alpha_2 + 2u & \frac{4}{5}\alpha_1 \\ -\frac{2}{3}\alpha_1^2 - \frac{2}{7}\alpha_2^2 - 2\alpha_2 u & 2\alpha_2 & \frac{4}{3}\alpha_1 & \frac{4}{7}\alpha_2 + 2u \end{pmatrix} \tag{3.108}$$

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ e_z g h - \frac{1}{3}\alpha_1^2 - \frac{1}{5}\alpha_2^2 - u^2 & 2u & \frac{2}{3}\alpha_1 & \frac{2}{5}\alpha_2 \\ -\frac{4}{5}\alpha_1\alpha_2 - 2\alpha_1 u & 2\alpha_1 & \alpha_2 + u & \frac{3}{5}\alpha_1 \\ -\frac{2}{3}\alpha_1^2 - \frac{2}{7}\alpha_2^2 - 2\alpha_2 u & 2\alpha_2 & \frac{1}{3}\alpha_1 & \frac{3}{7}\alpha_2 + u \end{pmatrix} \tag{3.109}$$

This system is no longer globally hyperbolic. The eigenvalues of the system can be computed numerically, but there are values of α_1 and α_2, which result in a non hyperbolic system. Intuitively this occurs when the values of α_1 and α_2 would physically result in a vortex, turbulent behavior, or some other physical behavior, that can't be captured or described by the system. In this case the problem becomes ill-posed.

Third Order For three moments, the velocity in the vertical direction is cubic, and the flux function, nonconservative matrix, and source function are shown below,

$$\underline{f}\left(\underline{q}\right) = \begin{pmatrix} hu \\ \frac{1}{2}e_z g h^2 + hu^2 + \frac{1}{105}(35\alpha_1^2 + 21\alpha_2^2 + 15\alpha_3^2)h \\ 2\alpha_1 hu + \frac{2}{35}(14\alpha_1\alpha_2 + 9\alpha_2\alpha_3)h \\ 2\alpha_2 hu + \frac{2}{21}(7\alpha_1^2 + 3\alpha_2^2 + 9\alpha_1\alpha_3 + 2\alpha_3^2)h \\ 2\alpha_3 hu + \frac{2}{15}(9\alpha_1\alpha_2 + 4\alpha_2\alpha_3)h \end{pmatrix} \tag{3.110}$$

$$g\left(\underline{q}\right) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{5}\alpha_2 + u & \frac{1}{5}\alpha_1 - \frac{3}{35}\alpha_3 & \frac{3}{35}\alpha_2 \\ 0 & 0 & \alpha_1 - \frac{3}{7}\alpha_3 & \frac{1}{7}\alpha_2 + u & \frac{2}{7}\alpha_1 + \frac{1}{21}\alpha_3 \\ 0 & 0 & \frac{6}{5}\alpha_2 & \frac{4}{5}\alpha_1 + \frac{2}{15}\alpha_3 & \frac{1}{5}\alpha_2 + u \end{pmatrix} \tag{3.111}$$

$$\underline{p}\left(\underline{q}\right) = \begin{pmatrix} 0 \\ -(e_z\frac{\partial}{\partial x}h_b - e_x)gh - \frac{\nu(\alpha_1+\alpha_2+\alpha_3+u)}{\lambda} \\ -\frac{3((\frac{4\lambda}{h}+1)\alpha_1+(\frac{4\lambda}{h}+1)\alpha_3+\alpha_2+u)\nu}{\lambda} \\ -\frac{5((\frac{12\lambda}{h}+1)\alpha_2+\alpha_1+\alpha_3+u)\nu}{\lambda} \\ -\frac{7((\frac{4\lambda}{h}+1)\alpha_1+(\frac{24\lambda}{h}+1)\alpha_3+\alpha_2+u)\nu}{\lambda} \end{pmatrix} \tag{3.112}$$

The flux jacobian and quasilinear matrix can be computed as

$$\underline{f}'(\underline{q}) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ e_z gh - \frac{1}{3}\alpha_1^2 - \frac{1}{5}\alpha_2^2 - \frac{1}{7}\alpha_3^2 - u^2 & 2u & \frac{2}{3}\alpha_1 & \frac{2}{5}\alpha_2 & \frac{2}{7}\alpha_3 \\ -\frac{4}{5}\alpha_1\alpha_2 - \frac{18}{35}\alpha_2\alpha_3 - 2\alpha_1 u & 2\alpha_1 & \frac{4}{5}\alpha_2 + 2u & \frac{4}{5}\alpha_1 + \frac{18}{35}\alpha_3 & \frac{18}{35}\alpha_2 \\ -\frac{2}{3}\alpha_1^2 - \frac{2}{7}\alpha_2^2 - \frac{6}{7}\alpha_1\alpha_3 - \frac{4}{21}\alpha_3^2 - 2\alpha_2 u & 2\alpha_2 & \frac{4}{3}\alpha_1 + \frac{6}{7}\alpha_3 & \frac{4}{7}\alpha_2 + 2u & \frac{6}{7}\alpha_1 + \frac{8}{21}\alpha_3 \\ -\frac{6}{5}\alpha_1\alpha_2 - \frac{8}{15}\alpha_2\alpha_3 - 2\alpha_3 u & 2\alpha_3 & \frac{6}{5}\alpha_2 & \frac{6}{5}\alpha_1 + \frac{8}{15}\alpha_3 & \frac{8}{15}\alpha_2 + 2u \end{pmatrix} \tag{3.113}$$

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ e_z gh - \frac{1}{3}\alpha_1^2 - \frac{1}{5}\alpha_2^2 - \frac{1}{7}\alpha_3^2 - u^2 & 2u & \frac{2}{3}\alpha_1 & \frac{2}{5}\alpha_2 & \frac{2}{7}\alpha_3 \\ -\frac{4}{5}\alpha_1\alpha_2 - \frac{18}{35}\alpha_2\alpha_3 - 2\alpha_1 u & 2\alpha_1 & \alpha_2 + u & \frac{3}{5}\alpha_1 + \frac{3}{5}\alpha_3 & \frac{3}{7}\alpha_2 \\ -\frac{2}{3}\alpha_1^2 - \frac{2}{7}\alpha_2^2 - \frac{6}{7}\alpha_1\alpha_3 - \frac{4}{21}\alpha_3^2 - 2\alpha_2 u & 2\alpha_2 & \frac{1}{3}\alpha_1 + \frac{9}{7}\alpha_3 & \frac{3}{7}\alpha_2 + u & \frac{4}{7}\alpha_1 + \frac{1}{3}\alpha_3 \\ -\frac{6}{5}\alpha_1\alpha_2 - \frac{8}{15}\alpha_2\alpha_3 - 2\alpha_3 u & 2\alpha_3 & 0 & \frac{2}{5}\alpha_1 + \frac{2}{5}\alpha_3 & \frac{1}{3}\alpha_2 + u \end{pmatrix} \tag{3.114}$$

The eigenvalues of the quasilinear matrix need to be computed numerically, and as in the second order case they are no longer always real. There are cases where the system is no longer hyperbolic.

3.1.2.2 2D Equations

In two dimensions the generalized shallow water equations will have the following form,

$$\underline{q}_t + \underline{f}_1\left(\underline{q}\right)_x + \underline{f}_2\left(\underline{q}\right)_y = g_1(\underline{q})\underline{q}_x + g_2(\underline{q})\underline{q}_y + \underline{p}. \tag{3.115}$$

In this case the unknown $\underline{q}$ will have the form

$$\underline{q} = [h, hu, hv, h\alpha_1, h\beta_1, h\alpha_2, h\beta_2, \ldots]^T, \tag{3.116}$$

where the number of components depends on the number of moments in the velocity profiles.

The wavespeeds of the two dimensional system in the direction $\underline{n} = [n_1, n_2]$, are given by the eigenvalues of the matrix

$$n_1\left(\underline{f}_1'(\underline{q}) - g_1(\underline{q})\right) + n_2\left(\underline{f}_2'(\underline{q}) - g_2(\underline{q})\right). \tag{3.117}$$

If this matrix is diagonalizable with real eigenvalues for all directions $\underline{n}$, then this system is considered hyperbolic.

Zeroth Order The zeroth order system is exactly the standard shallow water equations, where only the average velocity is considered. This velocity profiles in this system only consider the constant moment. In this case the nonconservative product disappears and the equation has the following form.

$$\underline{q}_t + \underline{f}_1\left(\underline{q}\right)_x + \underline{f}_2\left(\underline{q}\right)_y = \underline{p}. \tag{3.118}$$

where

$$\underline{f}_1(\underline{q}) = \begin{pmatrix} hu \\ \frac{1}{2}e_z g h^2 + hu^2 \\ huv \end{pmatrix}, \qquad \underline{f}_2(\underline{q}) = \begin{pmatrix} hv \\ huv \\ \frac{1}{2}e_z g h^2 + hv^2 \end{pmatrix} \tag{3.119}$$

$$\underline{p} = \begin{pmatrix} 0 \\ -\left(e_z \dfrac{\partial}{\partial x}(h_b) - e_x\right) gh - \frac{\nu}{\lambda} u \\ -\left(e_z \dfrac{\partial}{\partial y}(h_b) - e_y\right) gh - \frac{\nu}{\lambda} v \end{pmatrix} \tag{3.120}$$

The flux jacobians and quasilinear matrices, $A = \underline{f}_1'(\underline{q}) - g_1(\underline{q}), B = \underline{f}_2'(\underline{q}) - g_2(\underline{q})$, are given below,

$$\underline{f}_1'(\underline{q}) = \begin{pmatrix} 0 & 1 & 0 \\ e_z gh - u^2 & 2u & 0 \\ -uv & v & u \end{pmatrix}, \qquad \underline{f}_2'(\underline{q}) = \begin{pmatrix} 0 & 0 & 1 \\ -uv & v & u \\ e_z gh - v^2 & 0 & 2v \end{pmatrix} \tag{3.121}$$

$$A = \begin{pmatrix} 0 & 1 & 0 \\ e_z gh - u^2 & 2u & 0 \\ -uv & v & u \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 & 1 \\ -uv & v & u \\ e_z gh - v^2 & 0 & 2v \end{pmatrix} \tag{3.122}$$

The wavespeed of this system in the direction of $\underline{n} = [n_1, n_2]$ is given by the following eigenvalues,

$$\lambda_{1,2} = n_1 u + n_2 v \pm \sqrt{e_z gh(n_0^2 + n_1^2)} \tag{3.123}$$

$$\lambda_3 = n_1 u + n_2 v \tag{3.124}$$

First Order The first order system describes the velocity in the vertical direction as a line and adds two additional moments, one in the x-direction and one in the y-direction. The flux function, nonconservative matrices, and source term are given below.

$$\underline{f}_1(\underline{q}) = \begin{pmatrix} hu \\ \frac{1}{2}e_z g h^2 + \frac{1}{3}\alpha_1^2 h + hu^2 \\ \frac{1}{3}\alpha_1\beta_1 h + huv \\ 2\alpha_1 hu \\ \beta_1 hu + \alpha_1 hv \end{pmatrix}, \quad \underline{f}_2(\underline{q}) = \begin{pmatrix} hv \\ \frac{1}{3}\alpha_1\beta_1 h + huv \\ \frac{1}{2}e_z g h^2 + \frac{1}{3}\beta_1^2 h + hv^2 \\ \beta_1 hu + \alpha_1 hv \\ 2\beta_1 hv \end{pmatrix} \tag{3.125}$$

$$g_1(\underline{q}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & u & 0 \\ 0 & 0 & 0 & v & 0 \end{pmatrix}, \quad g_2(\underline{q}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & u \\ 0 & 0 & 0 & 0 & v \end{pmatrix} \tag{3.126}$$

$$\underline{p} = \begin{pmatrix} 0 \\ -(e_z \frac{\partial}{\partial x} h_b - e_x) g h - \frac{\nu(\alpha_1 + u)}{\lambda} \\ -(e_z \frac{\partial}{\partial y} h_b - e_y) g h - \frac{\nu(\beta_1 + v)}{\lambda} \\ -\frac{3((\frac{4\lambda}{h}+1)\alpha_1 + u)\nu}{\lambda} \\ -\frac{3((\frac{4\lambda}{h}+1)\beta_1 + v)\nu}{\lambda} \end{pmatrix} \tag{3.127}$$

The flux jacobians and quasilinear matrices for this system are

$$\underline{f}_1'(\underline{q}) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ e_z gh - \frac{1}{3}\alpha_1^2 - u^2 & 2u & 0 & \frac{2}{3}\alpha_1 & 0 \\ -\frac{1}{3}\alpha_1\beta_1 - uv & v & u & \frac{1}{3}\beta_1 & \frac{1}{3}\alpha_1 \\ -2\alpha_1 u & 2\alpha_1 & 0 & 2u & 0 \\ -\beta_1 u - \alpha_1 v & \beta_1 & \alpha_1 & v & u \end{pmatrix} \tag{3.128}$$

$$\underline{f}_2'(\underline{q}) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ -\frac{1}{3}\alpha_1\beta_1 - uv & v & u & \frac{1}{3}\beta_1 & \frac{1}{3}\alpha_1 \\ e_z gh - \frac{1}{3}\beta_1^2 - v^2 & 0 & 2v & 0 & \frac{2}{3}\beta_1 \\ -\beta_1 u - \alpha_1 v & \beta_1 & \alpha_1 & v & u \\ -2\beta_1 v & 0 & 2\beta_1 & 0 & 2v \end{pmatrix} \tag{3.129}$$

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ e_z gh - \frac{1}{3}\alpha_1^2 - u^2 & 2u & 0 & \frac{2}{3}\alpha_1 & 0 \\ -\frac{1}{3}\alpha_1\beta_1 - uv & v & u & \frac{1}{3}\beta_1 & \frac{1}{3}\alpha_1 \\ -2\alpha_1 u & 2\alpha_1 & 0 & u & 0 \\ -\beta_1 u - \alpha_1 v & \beta_1 & \alpha_1 & 0 & u \end{pmatrix} \tag{3.130}$$

$$B = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ -\frac{1}{3}\alpha_1\beta_1 - uv & v & u & \frac{1}{3}\beta_1 & \frac{1}{3}\alpha_1 \\ e_z gh - \frac{1}{3}\beta_1^2 - v^2 & 0 & 2v & 0 & \frac{2}{3}\beta_1 \\ -\beta_1 u - \alpha_1 v & \beta_1 & \alpha_1 & v & 0 \\ -2\beta_1 v & 0 & 2\beta_1 & 0 & v \end{pmatrix}. \tag{3.131}$$

The wavespeed of the system in the direction of $\underline{n} = [n_1, n_2]$ can be computed as the following eigenvalues of $n_1 A + n_2 B$,

$$\lambda_{1,2} = n_0 u + n_1 v \pm \sqrt{e_z gh(n_0^2 + n_1^2) + (\alpha_1 n_0 + \beta_1 n_1)^2} \tag{3.132}$$

$$\lambda_3 = n_0 u + n_1 v \tag{3.133}$$

$$\lambda_{4,5} = n_0 u + n_1 v \pm \frac{\sqrt{3}}{3}(\alpha_1 n_0 + \beta_1 n_1) \tag{3.134}$$

Second Order The second order system has seven equations and four additional moments from the standard shallow water model. The velocity in the vertical direction can be described as parabolas, The flux functions, nonconservative matrices, and source function are shown below,

$$\underline{f}_1(\underline{q}) = \begin{pmatrix} hu \\ \frac{1}{2}e_z gh^2 + hu^2 + \frac{1}{15}(5\alpha_1^2 + 3\alpha_2^2)h \\ huv + \frac{1}{15}(5\alpha_1\beta_1 + 3\alpha_2\beta_2)h \\ \frac{4}{5}\alpha_1\alpha_2 h + 2\alpha_1 hu \\ \beta_1 hu + \alpha_1 hv + \frac{2}{5}(\alpha_2\beta_1 + \alpha_1\beta_2)h \\ 2\alpha_2 hu + \frac{2}{21}(7\alpha_1^2 + 3\alpha_2^2)h \\ \beta_2 hu + \alpha_2 hv + \frac{2}{21}(7\alpha_1\beta_1 + 3\alpha_2\beta_2)h \end{pmatrix}, \quad \underline{f}_2(\underline{q}) = \begin{pmatrix} hv \\ huv + \frac{1}{15}(5\alpha_1\beta_1 + 3\alpha_2\beta_2)h \\ \frac{1}{2}e_z gh^2 + hv^2 + \frac{1}{15}(5\beta_1^2 + 3\beta_2^2)h \\ \beta_1 hu + \alpha_1 hv + \frac{2}{5}(\alpha_1\beta_1 + \alpha_2\beta_2)h \\ \frac{4}{5}\beta_1\beta_2 h + 2\beta_1 hv \\ \beta_2 hu + \alpha_2 hv + \frac{2}{21}(7\alpha_1\beta_1 + 3\alpha_2\beta_2)h \\ 2\beta_2 hv + \frac{2}{21}(7\beta_1^2 + 3\beta_2^2)h \end{pmatrix} \tag{3.135}$$

$$g_1(\underline{q}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{5}\alpha_2 + u & 0 & \frac{1}{5}\alpha_1 & 0 \\ 0 & 0 & 0 & -\frac{1}{5}\beta_2 + v & 0 & \frac{1}{5}\beta_1 & 0 \\ 0 & 0 & 0 & \alpha_1 & 0 & \frac{1}{7}\alpha_2 + u & 0 \\ 0 & 0 & 0 & \beta_1 & 0 & \frac{1}{7}\beta_2 + v & 0 \end{pmatrix}, \quad g_2(\underline{q}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{5}\alpha_2 + u & 0 & \frac{1}{5}\alpha_1 \\ 0 & 0 & 0 & 0 & -\frac{1}{5}\beta_2 + v & 0 & \frac{1}{5}\beta_1 \\ 0 & 0 & 0 & 0 & \alpha_1 & 0 & \frac{1}{7}\alpha_2 + u \\ 0 & 0 & 0 & 0 & \beta_1 & 0 & \frac{1}{7}\beta_2 + v \end{pmatrix} \tag{3.136}$$

$$\underline{p} = \begin{pmatrix} 0 \\ -(e_z \frac{\partial}{\partial x} h_b - e_x)gh - \frac{\nu(\alpha_1+\alpha_2+u)}{\lambda} \\ -(e_z \frac{\partial}{\partial y} h_b - e_y)gh - \frac{\nu(\beta_1+\beta_2+v)}{\lambda} \\ -\frac{3((\frac{4\lambda}{h}+1)\alpha_1+\alpha_2+u)\nu}{\lambda} \\ -\frac{3((\frac{4\lambda}{h}+1)\beta_1+\beta_2+v)\nu}{\lambda} \\ -\frac{5((\frac{12\lambda}{h}+1)\alpha_2+\alpha_1+u)\nu}{\lambda} \\ -\frac{5((\frac{12\lambda}{h}+1)\beta_2+\beta_1+v)\nu}{\lambda} \end{pmatrix} \tag{3.137}$$

The flux jacobians are computed below as,

$$\underline{f}_1'(\underline{q}) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ e_z gh - \frac{1}{3}\alpha_1^2 - \frac{1}{5}\alpha_2^2 - u^2 & 2u & 0 & \frac{2}{3}\alpha_1 & 0 & \frac{2}{5}\alpha_2 & 0 \\ -\frac{1}{3}\alpha_1\beta_1 - \frac{1}{5}\alpha_2\beta_2 - uv & v & u & \frac{1}{3}\beta_1 & \frac{1}{3}\alpha_1 & \frac{1}{5}\beta_2 & \frac{1}{5}\alpha_2 \\ -\frac{4}{5}\alpha_1\alpha_2 - 2\alpha_1 u & 2\alpha_1 & 0 & \frac{4}{5}\alpha_2 + 2u & 0 & \frac{4}{5}\alpha_1 & 0 \\ -\frac{2}{5}\alpha_2\beta_1 - \frac{2}{5}\alpha_1\beta_2 - \beta_1 u - \alpha_1 v & \beta_1 & \alpha_1 & \frac{2}{5}\beta_2 + v & \frac{2}{5}\alpha_2 + u & \frac{2}{5}\beta_1 & \frac{2}{5}\alpha_1 \\ -\frac{2}{3}\alpha_1^2 - \frac{2}{7}\alpha_2^2 - 2\alpha_2 u & 2\alpha_2 & 0 & \frac{4}{3}\alpha_1 & 0 & \frac{4}{7}\alpha_2 + 2u & 0 \\ -\frac{2}{3}\alpha_1\beta_1 - \frac{2}{7}\alpha_2\beta_2 - \beta_2 u - \alpha_2 v & \beta_2 & \alpha_2 & \frac{2}{3}\beta_1 & \frac{2}{3}\alpha_1 & \frac{2}{7}\beta_2 + v & \frac{2}{7}\alpha_2 + u \end{pmatrix} \tag{3.138}$$

$$\underline{f}_2'(\underline{q}) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{1}{3}\alpha_1\beta_1 - \frac{1}{5}\alpha_2\beta_2 - uv & v & u & \frac{1}{3}\beta_1 & \frac{1}{3}\alpha_1 & \frac{1}{5}\beta_2 & \frac{1}{5}\alpha_2 \\ e_z gh - \frac{1}{3}\beta_1^2 - \frac{1}{5}\beta_2^2 - v^2 & 0 & 2v & 0 & \frac{2}{3}\beta_1 & 0 & \frac{2}{5}\beta_2 \\ -\frac{2}{5}\alpha_1\beta_1 - \frac{2}{5}\alpha_2\beta_2 - \beta_1 u - \alpha_1 v & \beta_1 & \alpha_1 & \frac{2}{5}\beta_1 + v & \frac{2}{5}\alpha_1 + u & \frac{2}{5}\beta_2 & \frac{2}{5}\alpha_2 \\ -\frac{4}{5}\beta_1\beta_2 - 2\beta_1 v & 0 & 2\beta_1 & 0 & \frac{4}{5}\beta_2 + 2v & 0 & \frac{4}{5}\beta_1 \\ -\frac{2}{3}\alpha_1\beta_1 - \frac{2}{7}\alpha_2\beta_2 - \beta_2 u - \alpha_2 v & \beta_2 & \alpha_2 & \frac{2}{3}\beta_1 & \frac{2}{3}\alpha_1 & \frac{2}{7}\beta_2 + v & \frac{2}{7}\alpha_2 + u \\ -\frac{2}{3}\beta_1^2 - \frac{2}{7}\beta_2^2 - 2\beta_2 v & 0 & 2\beta_2 & 0 & \frac{4}{3}\beta_1 & 0 & \frac{4}{7}\beta_2 + 2v \end{pmatrix}. \tag{3.139}$$

The quasilinear matrices are

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ e_z gh - \frac{1}{3}\alpha_1^2 - \frac{1}{5}\alpha_2^2 - u^2 & 2u & 0 & \frac{2}{3}\alpha_1 & 0 & \frac{2}{5}\alpha_2 & 0 \\ -\frac{1}{3}\alpha_1\beta_1 - \frac{1}{5}\alpha_2\beta_2 - uv & v & u & \frac{1}{3}\beta_1 & \frac{1}{3}\alpha_1 & \frac{1}{5}\beta_2 & \frac{1}{5}\alpha_2 \\ -\frac{4}{5}\alpha_1\alpha_2 - 2\alpha_1 u & 2\alpha_1 & 0 & \alpha_2 + u & 0 & \frac{3}{5}\alpha_1 & 0 \\ -\frac{2}{5}\alpha_2\beta_1 - \frac{2}{5}\alpha_1\beta_2 - \beta_1 u - \alpha_1 v & \beta_1 & \alpha_1 & \frac{3}{5}\beta_2 & \frac{2}{5}\alpha_2 + u & \frac{1}{5}\beta_1 & \frac{2}{5}\alpha_1 \\ -\frac{2}{3}\alpha_1^2 - \frac{2}{7}\alpha_2^2 - 2\alpha_2 u & 2\alpha_2 & 0 & \frac{1}{3}\alpha_1 & 0 & \frac{3}{7}\alpha_2 + u & 0 \\ -\frac{2}{3}\alpha_1\beta_1 - \frac{2}{7}\alpha_2\beta_2 - \beta_2 u - \alpha_2 v & \beta_2 & \alpha_2 & -\frac{1}{3}\beta_1 & \frac{2}{3}\alpha_1 & \frac{1}{7}\beta_2 & \frac{2}{7}\alpha_2 + u \end{pmatrix} \tag{3.140}$$

$$B = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{1}{3}\alpha_1\beta_1 - \frac{1}{5}\alpha_2\beta_2 - uv & v & u & \frac{1}{3}\beta_1 & \frac{1}{3}\alpha_1 & \frac{1}{5}\beta_2 & \frac{1}{5}\alpha_2 \\ e_z gh - \frac{1}{3}\beta_1^2 - \frac{1}{5}\beta_2^2 - v^2 & 0 & 2v & 0 & \frac{2}{3}\beta_1 & 0 & \frac{2}{5}\beta_2 \\ -\frac{2}{5}\alpha_1\beta_1 - \frac{2}{5}\alpha_2\beta_2 - \beta_1 u - \alpha_1 v & \beta_1 & \alpha_1 & \frac{2}{5}\beta_1 + v & \frac{2}{5}\alpha_1 + \frac{1}{5}\alpha_2 & \frac{2}{5}\beta_2 & -\frac{1}{5}\alpha_1 + \frac{2}{5}\alpha_2 \\ -\frac{4}{5}\beta_1\beta_2 - 2\beta_1 v & 0 & 2\beta_1 & 0 & \beta_2 + v & 0 & \frac{3}{5}\beta_1 \\ -\frac{2}{3}\alpha_1\beta_1 - \frac{2}{7}\alpha_2\beta_2 - \beta_2 u - \alpha_2 v & \beta_2 & \alpha_2 & \frac{2}{3}\beta_1 & -\frac{1}{3}\alpha_1 & \frac{2}{7}\beta_2 + v & \frac{1}{7}\alpha_2 \\ -\frac{2}{3}\beta_1^2 - \frac{2}{7}\beta_2^2 - 2\beta_2 v & 0 & 2\beta_2 & 0 & \frac{1}{3}\beta_1 & 0 & \frac{3}{7}\beta_2 + v \end{pmatrix}. \tag{3.141}$$

As in the one dimensional case the second order wavespeeds need to be computed numerically and the system is no longer globally hyperbolic.

3.2 Shallow Water Linearized Moment Equations

As was seen in the previous section, the shallow water moment equations are not guaranteed hyperbolic when there is two or more moments in velocity profile. Further exploration of these equations by Koellermeier[8, 7] and others indicate that the model can lose hyperbolicity in regimes arbitrarily close to equilibrium for certain numbers of moments. To prevent this lose of hyperbolicity Koellermeier introduced a few different models, which didn't suffer from this problem. One of these models is known as the shallow water linearized moment equations, and is a small change or linearization of the original shallow water moment equations. Recall the shallow water moment equations, (3.80),(3.81),(3.82),(3.83), and (3.84). The shallow water linearized moment equations are created by assuming that $\alpha_i = O(\varepsilon)$ and $\beta_i = O(\varepsilon)$. Then all terms of $O(\varepsilon^2)$ in the moment equations are dropped. The momentum equations are unchanged even though they contain terms of $O(\varepsilon^2)$. This is important to maintain conservation. The resulting system is

$$
\begin{gathered}
h_t + (hu)_x + (hv)_y = 0 \\
(hu_m)_t + \left(h\left(u^2 + \sum_{j=1}^{N} \frac{\alpha_j^2}{2j+1}\right) + \frac{1}{2} g e_z h^2\right)_x + \left(h\left(uv + \sum_{j=1}^{N} \frac{\alpha_j \beta_j}{2j+1}\right)\right)_y \\
= -\frac{\nu}{\lambda}\left(u + \sum_{j=1}^{N} \alpha_j\right) + hg(e_x - e_z (h_b)_x) \\
(hv)_t + \left(h\left(v^2 + \sum_{j=1}^{N} \frac{\alpha_j \beta_j}{2j+1}\right) + \frac{1}{2} g e_z h^2\right)_y + \left(h\left(uv + \sum_{j=1}^{N} \frac{\alpha_j \beta_j}{2j+1}\right)\right)_x \\
= -\frac{\nu}{\lambda}\left(v + \sum_{j=1}^{N} \beta_j\right) + hg\Big(e_y - e_z (h_b)_y\Big) \\
(h\alpha_i)_t + (2hu\alpha_i)_x + (hu\beta_i + hv\alpha_i)_y = uD_i - (2i+1)\frac{\nu}{\lambda}\left(u + \sum_{j=1}^{N}\left(1 + \frac{\lambda}{h} C_{ij}\right)\alpha_j\right) \\
(h\beta_i)_t + (hu\beta_i + hv\alpha_i)_x + (2hv\beta_i)_y = vD_i - (2i+1)\frac{\nu}{\lambda}\left(v + \sum_{j=1}^{N}\left(1 + \frac{\lambda}{h} C_{ij}\right)\beta_j\right) \\
C_{ij} = \int_0^1 \phi_i' \phi_j' \, \mathrm{d}\zeta \\
D_i = (h\alpha_i)_x + (h\beta_i)_y
\end{gathered}
$$

Koellermeier examined this system in one dimension. In one dimension this system is of the form

$$
\underline{q}_t + \underline{f}(\underline{q})_x = \underline{\underline{g}}\,\underline{q}_x - \frac{\nu}{\lambda}\underline{p}. \tag{3.142}
$$

The quasilinear matrix is $A = \underline{f}'(\underline{q}) - \underline{\underline{g}}$, and it was shown to by diagonalizable with eigenvalues

$$
\lambda_{1,2} = u \pm \sqrt{gh + \sum_{i=1}^{N}\left(\frac{3\alpha_i^2}{2i+1}\right)} \quad \text{and} \quad \lambda_i = u \quad \text{for } i = 3, \ldots, N+2. \tag{3.143}
$$

Therefore the system is hyperbolic for any number of moments.

I will now show that the two dimensional system is hyperbolic for any number of moments as well. In two dimensions the SWLME will be of the form,

$$\underline{q}_t + \underline{f}_1\left(\underline{q}\right)_x + \underline{f}_2\left(\underline{q}\right)_y = g_1(\underline{q})\underline{q}_x + g_2(\underline{q})\underline{q}_y + \underline{p}. \tag{3.144}$$

The flux functions are given by,

$$\underline{f}_1(\underline{q}) = \begin{pmatrix} hu \\ \frac{1}{2}e_z gh^2 + hu^2 + \sum_{j=1}^{N}\left(\frac{1}{2j+1}h\alpha_j^2\right) \\ huv + \sum_{j=1}^{N}\left(\frac{1}{2j+1}h\alpha_j\beta_j\right) \\ 2hu\alpha_1 \\ \beta_1 hu + \alpha_1 hv \\ \vdots \\ 2hu\alpha_N \\ hu\beta_N + hv\alpha_N \end{pmatrix}, \quad \text{and} \quad \underline{f}_2(\underline{q}) = \begin{pmatrix} hv \\ huv + \sum_{j=1}^{N}\left(\frac{1}{2j+1}h\alpha_j\beta_j\right) \\ \frac{1}{2}e_z gh^2 + hv^2 + \sum_{j=1}^{N}\left(\frac{1}{2j+1}h\beta_j^2\right) \\ hu\beta_1 + hv\alpha_1 \\ 2hv\beta_1 \\ \vdots \\ hu\beta_N + hv\alpha_N \\ 2hv\beta_N \end{pmatrix}. \tag{3.145}$$

Also the nonconservative matrices are of the form,

Nonconservative Matrices

$$g_1(\underline{q}) = \begin{pmatrix} 0 & & & & & & & \\ 0 & 0 & & & & & & \\ & 0 & 0 & & & & & \\ & & 0 & u & & & & \\ & & & v & 0 & & & \\ & & & & 0 & u & & \\ & & & & & v & 0 & \\ & & & & & & 0 & \ddots \\ & & & & & & & \ddots \end{pmatrix}, \quad \text{and} \quad g_2(\underline{q}) = \begin{pmatrix} 0 & 0 & & & & & & & \\ & 0 & 0 & & & & & & \\ & & 0 & 0 & & & & & \\ & & & 0 & u & & & & \\ & & & & v & 0 & & & \\ & & & & & 0 & u & & \\ & & & & & & v & 0 & \\ & & & & & & & 0 & \ddots \\ & & & & & & & & \ddots \end{pmatrix}. \tag{3.146}$$

The wavespeeds of the two dimensional system in the direction $\underline{n} = [n_1, n_2]$, are given by the eigenvalues of the matrix

$$n_1\Big(\underline{f}_1'(\underline{q}) - g_1(\underline{q})\Big) + n_2\Big(\underline{f}_2'(\underline{q}) - g_2(\underline{q})\Big).$$

If this matrix is diagonalizable with real eigenvalues for all directions $\underline{n}$, then this system is considered hyperbolic.

In order to compute the eigenvalues of this matrix, I first compute the flux jacobians, which are

$$\underline{f}_1'(\underline{q}) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ e_z g h - u^2 - \sum_{i=1}^{N}\left(\frac{1}{2*i+1}\alpha_i^2\right) & 2u & 0 & \frac{2}{3}\alpha_1 & 0 & \cdots & \frac{2}{2N+1}\alpha_N & 0 \\ -uv - \sum_{i=1}^{N}\left(\frac{1}{2N+1}\alpha_i\beta_i\right) & v & u & \frac{1}{3}\beta_1 & \frac{1}{3}\alpha_1 & \cdots & \frac{1}{2N+1}\beta_N & \frac{1}{2N+1}\alpha_N \\ -2u\alpha_1 & 2\alpha_1 & 0 & 2u & & & & \\ -u\beta_1 - v\alpha_1 & \beta_1 & \alpha_1 & v & u & & & \\ \vdots & \vdots & \vdots & & \ddots & \ddots & & \\ -2u\alpha_N & 2\alpha_N & 0 & & & 0 & 2u & \\ -u\beta_N - v\alpha_N & \beta_N & \alpha_N & & & & v & u \end{pmatrix} \tag{3.147}$$

and

$$\underline{f}_2'(\underline{q}) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ -uv - \sum_{i=1}^{N}\left(\frac{1}{2i+1}\alpha_i\beta_i\right) & v & u & \frac{1}{3}\beta_1 & \frac{1}{3}\alpha_1 & \cdots & \frac{1}{2N+1}\beta_N & \frac{1}{2N+1}\alpha_N \\ e_z g h - v^2 - \sum_{i=1}^{N}\left(\frac{1}{2i+1}\beta_i^2\right) & 0 & 2v & 0 & \frac{2}{3}\beta_1 & \cdots & 0 & \frac{2}{2N+1}\beta_N \\ -u\beta_1 - \alpha_1 v & \beta_1 & \alpha_1 & v & u & & & \\ -2v\beta_1 & 0 & 2\beta_1 & & 2v & 0 & & \\ \vdots & \vdots & \vdots & & & \ddots & \ddots & \\ -u\beta_N - v\alpha_N & \beta_N & \alpha_N & & & & v & u \\ -2u\beta_N & 0 & 2\beta_N & & & & & 2v \end{pmatrix}. \tag{3.148}$$

The quasilinear matrices $Q_x = \underline{f}_1'(\underline{q}) - g_1(\underline{q})$ and $Q_y = \underline{f}_2'(\underline{q}) - g_2(\underline{q})$, are given by

$$Q_x = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ e_z gh - u^2 - \sum_{i=1}^{N}\left(\frac{1}{2i+1}\alpha_i^2\right) & 2u & 0 & \frac{2}{3}\alpha_1 & 0 & \cdots & \frac{2}{2N+1}\alpha_N & 0 \\ -uv - \sum_{i=1}^{N}\left(\frac{1}{2i+1}\alpha_i\beta_i\right) & v & u & \frac{1}{3}\beta_1 & \frac{1}{3}\alpha_1 & \cdots & \frac{1}{2N+1}\beta_N & \frac{1}{2N+1}\alpha_N \\ -2u\alpha_1 & 2\alpha_1 & 0 & u & & & & \\ -u\beta_1 - v\alpha_1 & \beta_1 & \alpha_1 & & u & & & \\ \vdots & \vdots & \vdots & & & \ddots & & \\ -2u\alpha_N & 2\alpha_N & 0 & & & & u & \\ -u\beta_N - v\alpha_N & \beta_N & \alpha_N & & & & & u \end{pmatrix} \tag{3.149}$$

and

$$Q_y = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ -uv - \sum_{i=1}^{N}\left(\frac{1}{2i+1}\alpha_i\beta_i\right) & v & u & \frac{1}{3}\beta_1 & \frac{1}{3}\alpha_1 & \cdots & \frac{1}{2N+1}\beta_N & \frac{1}{2N+1}\alpha_N \\ e_z gh - v^2 - \sum_{i=1}^{N}\left(\frac{1}{2i+1}\beta_i^2\right) & 0 & 2v & 0 & \frac{2}{3}\beta_1 & \cdots & 0 & \frac{2}{2N+1}\beta_N \\ -u\beta_1 - \alpha_1 v & \beta_1 & \alpha_1 & v & & & & \\ -2v\beta_1 & 0 & 2\beta_1 & & v & & & \\ \vdots & \vdots & \vdots & & & \ddots & & \\ -u\beta_N - v\alpha_N & \beta_N & \alpha_N & & & & v & \\ -2u\beta_N & 0 & 2\beta_N & & & & & v \end{pmatrix} \tag{3.150}$$

I now want to find the eigenvalues for the matrix $n_1 Q_x + n_2 Q_y$ for any values of n_1 and n_2. In order to make this computation a little more tractable I will introduce the following convenient constants,

$$\tilde{\lambda} = \lambda - n_1 u - n_2 v$$

$$d_0^1 = n_1\left(e_z gh - u^2 - \sum_{i=1}^{N}\left(\frac{1}{2i+1}\alpha_i^2\right)\right) + n_2\left(-uv - \sum_{i=1}^{N}\left(\frac{1}{2i+1}\alpha_i\beta_i\right)\right)$$

$$d_0^2 = n_1\left(-uv - \sum_{i=1}^{N}\left(\frac{1}{2i+1}\alpha_i\beta_i\right)\right) + n_2\left(e_z gh - v^2 - \sum_{i=1}^{N}\left(\frac{1}{2i+1}\beta_i^2\right)\right)$$

$$d_i^1 = n_1(-2u\alpha_i) + n_2(-u\beta_i - \alpha_i v)$$

$$d_i^2 = n_1(-u\beta_i - v\alpha_i) + n_2(-2v\beta_i)$$

$$b_i^1 = n_1\frac{2}{2i+1}\alpha_i + n_2\frac{1}{2i+1}\beta_i$$

$$b_i^2 = n_2\frac{1}{2i+1}\alpha_i$$

$$b_i^3 = n_1\frac{1}{2i+1}\beta_i$$

$$b_i^4 = n_1\frac{1}{2i+1}\alpha_i + n_2\frac{2}{2i+1}\beta_i$$

$$c_i^1 = n_1 2\alpha_i + n_2\beta_i$$

$$c_i^2 = n_1\beta_i$$

$$c_i^3 = n_2\alpha_i$$

$$c_i^4 = n_1\alpha_i + n_2 2\beta_i.$$

We can now compute $\det(n_1 Q_x + n_2 Q_y - \lambda I)$,

$$\det(n_1 Q_x + n_2 Q_y - \lambda I) = \begin{vmatrix} -\tilde{\lambda} - n_1 u - n_2 v & n_1 & n_2 & 0 & 0 & \cdots & 0 & 0 \\ d_0^1 & n_1 u - \tilde{\lambda} & n_2 u & b_1^1 & b_1^2 & \cdots & b_N^1 & b_N^2 \\ d_0^2 & n_1 v & n_2 v - \tilde{\lambda} & b_1^3 & b_1^4 & \cdots & b_N^3 & b_N^4 \\ d_1^1 & c_1^1 & c_1^3 & -\tilde{\lambda} & & & & \\ d_1^2 & c_1^2 & c_1^4 & & -\tilde{\lambda} & & & \\ \vdots & \vdots & \vdots & & & \ddots & & \\ d_N^1 & c_N^1 & c_N^3 & & & & -\tilde{\lambda} & \\ d_N^2 & c_N^2 & c_N^4 & & & & & -\tilde{\lambda} \end{vmatrix}. \tag{3.151}$$

This can be interpreted as a block matrix and the determinant of a block matrix

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D| \Big| A - BD^{-1}C \Big|. \tag{3.152}$$

Therefore

$$\det(n_1 Q_x + n_2 Q_y - \lambda I) = |D| \Big| A - BD^{-1}C \Big|, \tag{3.153}$$

where

$$A = \begin{pmatrix} -\tilde{\lambda} - n_1 u - n_2 v & n_1 & n_2 \\ d_0^1 & n_1 u - \tilde{\lambda} & n_2 u \\ d_0^2 & n_1 v & n_2 v - \tilde{\lambda} \end{pmatrix} \tag{3.154}$$

$$B = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ b_1^1 & b_1^2 & \cdots & b_N^1 & b_N^2 \\ b_1^3 & b_1^4 & \cdots & b_N^3 & b_N^4 \end{pmatrix} \tag{3.155}$$

$$C = \begin{pmatrix} d_1^1 & c_1^1 & c_1^3 \\ d_1^2 & c_1^2 & c_1^4 \\ \vdots & \vdots & \vdots \\ d_N^1 & c_N^1 & c_N^3 \\ d_N^2 & c_N^2 & c_N^4 \end{pmatrix} \tag{3.156}$$

$$D = \begin{pmatrix} -\tilde{\lambda} & & & & \\ & -\tilde{\lambda} & & & \\ & & \ddots & & \\ & & & -\tilde{\lambda} & \\ & & & & -\tilde{\lambda} \end{pmatrix} \tag{3.157}$$

It is easy to see that

$$\det(D) = \tilde{\lambda}^{2N}. \tag{3.158}$$

Also we can compute

$$BD^{-1}C = \frac{1}{-\tilde{\lambda}} \begin{pmatrix} 0 & 0 & 0 \\ s_1 & s_2 & s_3 \\ s_4 & s_5 & s_6 \end{pmatrix} \tag{3.159}$$

$$\tag{3.160}$$

where

$$s_1 = \sum_{i=1}^{N} \left(b_i^1 d_i^1 + b_i^2 d_i^2 \right) = \sum_{i=1}^{N} \left(-\frac{1}{2i+1} (4\alpha_i u n_1 + \beta_i u n_2 + 3\alpha_i v n_2)(\alpha_i n_1 + \beta_i n_2) \right) \tag{3.161}$$

$$s_2 = \sum_{i=1}^{N} \left(b_i^1 c_i^1 + b_i^2 c_i^2 \right) = \sum_{i=1}^{N} \left(\frac{1}{2i+1} (4\alpha_i n_1 + \beta_i n_2)(\alpha_i n_1 + \beta_i n_2) \right) \tag{3.162}$$

$$s_3 = \sum_{i=1}^{N} \left(b_i^1 c_i^3 + b_i^2 c_i^4 \right) = \sum_{i=1}^{N} \left(\frac{3}{2i+1} \alpha_i n_2 (\alpha_i n_1 + \beta_i n_2) \right) \tag{3.163}$$

$$s_4 = \sum_{i=1}^{N} \left(b_i^3 d_i^1 + b_i^4 d_i^2 \right) = \sum_{i=1}^{N} \left(-\frac{1}{2i+1} (3\beta_i u n_1 + \alpha_i v n_1 + 4\beta_i v n_2)(\alpha_i n_1 + \beta_i n_2) \right) \tag{3.164}$$

$$s_5 = \sum_{i=1}^{N} \left(b_i^3 c_i^1 + b_i^4 c_i^2 \right) = \sum_{i=1}^{N} \left(\frac{3}{2i+1} \beta_i n_1 (\alpha_i n_1 + \beta_i n_2) \right) \tag{3.165}$$

$$s_6 = \sum_{i=1}^{N} \left(b_i^3 c_i^3 + b_i^4 c_i^4 \right) = \sum_{i=1}^{N} \left(\frac{1}{2i+1} (\alpha_i n_1 + 4\beta_i n_2)(\alpha_i n_1 + \beta_i n_2) \right). \tag{3.166}$$

Then the determinant of $A - BD^{-1}C$ can be

$$\det\left(A - BD^{-1}C\right) = -\tilde{\lambda}^3 + \left(s_2 + s_6 + (un_1 + vn_2)^2 + d_0^1 n_1 + d_0^2 n_2\right)\tilde{\lambda} \tag{3.167}$$

$$+(vn_1 s_3 + un_2 s_5 + un_1 s_2 + vn_2 s_6 + s_1 n_1 + n_2 s_4) \tag{3.168}$$

$$+\left(s_3 s_5 - s_2 s_6 + \left(u^2 n_1 n_2 + uvn_2^2 + d_0^1 n_2\right) s_5 + \left(uvn_1^2 + v^2 n_1 n_2 + d_0^2 n_1\right) s_3\right)\tilde{\lambda}^{-1} \tag{3.169}$$

$$-\left(\left(u^2 n_1^2 + uvn_1 n_2 + d_0^1 n_1\right) s_6 + \left(uvn_1 n_2 + v^2 n_2^2 + d_0^2 n_2\right) s_2\right)\tilde{\lambda}^{-1} \tag{3.170}$$

$$+((s_3 s_5 - s_2 s_6)(un_1 + vn_2) + (s_3 s_4 - s_1 s_6) n_1 + (s_1 s_5 - s_2 s_4) n_2)\tilde{\lambda}^{-2} \tag{3.171}$$

This can be simplified using the definitions of all of the constants to

$$\det\left(A - BD^{-1}C\right) = -\frac{1}{\tilde{\lambda}}\left(\tilde{\lambda}^2 - \sum_{i=1}^{N}\left(\frac{1}{2i+1}(\alpha_i n_1 + \beta_i n_2)^2\right)\right) \tag{3.172}$$

$$\left(\tilde{\lambda}^2 - \left(3\left(\sum_{i=1}^{N}\left(\frac{1}{2i+1}(\alpha_i n_1 + \beta_i n_2)^2\right)\right) + e_z gh\left(n_1^2 + n_2^2\right)\right)\right) \tag{3.173}$$

Therefore the determinant of the whole quasilinear matrix is

$$\det(n_1 Q_x + n_2 Q_y + \lambda I) = -\tilde{\lambda}^{2N-1}\left(\tilde{\lambda}^2 - \sum_{i=1}^{N}\left(\frac{1}{2i+1}(\alpha_i n_1 + \beta_i n_2)^2\right)\right) \tag{3.174}$$

$$\left(\tilde{\lambda}^2 - \left(3\left(\sum_{i=1}^{N}\left(\frac{1}{2i+1}(\alpha_i n_1 + \beta_i n_2)^2\right)\right) + e_z gh\left(n_1^2 + n_2^2\right)\right)\right) \tag{3.175}$$

This shows that the eigenvalues are

$$\lambda_{1,2} = un_1 + vn_2 \pm \sqrt{e_z gh(n_1^2 + n_2^2) + 3\left(\sum_{i=1}^{N}\left(\frac{1}{2i+1}(\alpha_i n_1 + \beta_i n_2)^2\right)\right)}$$

$$\lambda_{3,4} = un_1 + vn_2 \pm \sqrt{\sum_{i=1}^{N}\left(\frac{1}{2i+1}(\alpha_i n_1 + \beta_i n_2)^2\right)}$$

$$\lambda_i = un_1 + vn_2 \quad \text{for} i = 5, \ldots 2N + 3$$

These eigenvalues will always be real for any values of u, v, α_i, and β_i. Also it can easily seen that these eigenvalues will not be defective, that is they will have a full set of linearly independent eigenvectors. The distinct eigenvalues are guaranteed to have an eigenvector, and the repeated eigenvalue $\lambda = un_1 + vn_2$ will have a full set of eigenvectors, because in this case for $n_1 Q_x + n_2 Q_y + \lambda I$, the submatrix $D = 0$.

Since this system is guaranteed hyperbolic and behaves very similarly to the SWME as demonstrated by Koellermeier, this is the system that I will use in all of my numerical tests. The shallow water linearized moment equations are identical to the shallow water moment equations for zero and one moments. There are only slight differences to the system for two or more moments.

Chapter4. Nonconservative Discontinuous Galerkin Methods

4.1 Definition

Consider the nonconservative product

$$g(\underline{q})\frac{\mathrm{d}\underline{q}}{\mathrm{d}x},$$

where $g(\underline{q}) : \mathbb{R}^p \to \mathbb{R}^p \times \mathbb{R}^p$ is continuous, but $\underline{q}$ is possibly discontinuous. In this case, the product is traditionally not well-defined at the discontinuities of $\underline{q}$. In order to define this product for discontinuous functions, $\underline{q}$, it is possible to regularize $\underline{q}$ with a path ϕ at discontinuities according to the theory laid out by Dal Maso, Le Floch, and Murat. To this end consider Lipschitz continuous paths, $\underline{\psi} : [0,1] \times \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^p$, that satisfy the following properties.

1. $\forall \underline{q}_L, \underline{q}_R \in \mathbb{R}^p$, $\underline{\psi}(0, \underline{q}_L, \underline{q}_R) = \underline{q}_L$ and $\underline{\psi}(1, \underline{q}_L, \underline{q}_R) = \underline{q}_R$
2. $\exists k > 0$, $\forall \underline{q}_L, \underline{q}_R \in \mathbb{R}^p$, $\forall s \in [0,1]$, $\left|\frac{\partial \underline{\psi}}{\partial s}(s, \underline{q}_L, \underline{q}_R)\right| \leq k\left|\underline{q}_L - \underline{q}_R\right|$ elementwise
3. $\exists k > 0$, $\forall \underline{q}_L, \underline{q}_R, \underline{u}_L, \underline{u}_R \in \mathbb{R}^p$, $\forall s \in [0,1]$, elementwise

$$\left|\frac{\partial \underline{\psi}}{\partial s}(s, \underline{q}_L, \underline{q}_R) - \frac{\partial \underline{\psi}}{\partial s}(s, \underline{u}_L, \underline{u}_R)\right| \leq k\left(\left|\underline{q}_L - \underline{u}_L\right| + \left|\underline{q}_R - \underline{u}_R\right|\right)$$

Once we have these paths, $\underline{\psi}$, we can define the nonconservative product.

Let $\underline{q} : [a,b] \to \mathbb{R}^p$ be a function of bounded variation, let $g : \mathbb{R}^p \to \mathbb{R}^p \times \mathbb{R}^p$ be a continuous function, and let $\underline{\psi}$ satisfy the properties given above. Then there exists a unique real-valued bounded Borel measure $\underline{\mu}$ on $[a,b]$ characterized by the two following properties.

1. If q is continuous on a Borel set $B \subset [a,b]$, then

$$\underline{\mu}(B) = \int_B g(\underline{q})\frac{\mathrm{d}\underline{q}}{\mathrm{d}x}\,\mathrm{d}x$$

2. If q is discontinuous at a point $x_0 \in [a, b]$, then

$$\underline{\mu}(x_0) = \int_0^1 g(\underline{\psi}(s; \underline{q}(x_0^-), \underline{q}(x_0^+))) \frac{\partial \underline{\psi}}{\partial s}(s; \underline{q}(x_0^-), \underline{q}(x_0^+))\, \mathrm{d}s$$

By definition, this measure $\underline{\mu}$ is the nonconservative product $g(\underline{q})\frac{\mathrm{d}\underline{q}}{\mathrm{d}x}$ and will be denoted by

$$\underline{\mu} = \left[g(\underline{q})\frac{\mathrm{d}\underline{q}}{\mathrm{d}x}\right]_{\underline{\psi}}$$

Note that if there exists a function $\underline{f}\left(\underline{q}\right)$ such that $\underline{f}'\left(\underline{q}\right) = g(\underline{q})$, then

$$\int_0^1 g(\underline{\psi}(s; \underline{q}(x_0^-), \underline{q}(x_0^+))) \frac{\partial \underline{\psi}}{\partial s}(s; \underline{q}(x_0^-), \underline{q}(x_0^+))\, \mathrm{d}s = \underline{f}\left(\underline{q}(x_0^+)\right) - \underline{f}\left(\underline{q}(x_0^-)\right) \tag{4.1}$$

for any path $\underline{\psi}$ that satisfies the conditions 1 — 3.

4.1.1 Higher Dimensions

In higher dimensions the paths, $\underline{\psi}$ must also have the property that

4. $\underline{\psi}(s, \underline{q}_L, \underline{q}_R) = \underline{\psi}(1 - s, \underline{q}_L, \underline{q}_R)$

Then the following Theorem can be given in spacetime Let $\underline{q} : \Omega \to \mathbb{R}^m$ be a bounded function of bounded variation defined on an open subset Ω of $\mathbb{R}^{n+1}$ and $\underline{t} : \mathbb{R}^m \to \mathbb{R}^m$ be a locally bounded Borel function. Then there exists a unique family of real-valued bounded Borel measures μ_i on Ω, $i = 1, 2, \ldots, m$ such that

1. if B is a continuous Borel subset of Ω, then

   $$\mu_i(B) = \int_B t_{ik}(q)\underline{q}_{x_k}\, \mathrm{d}\lambda$$

   where λ is the Borel measure;

2. if B is a discontinuous subset of Ω of approximate jump, then

   $$\mu_i(B) = \int_B \int_0^1 t_{ik}(\underline{\psi}(s, \underline{q}^L, \underline{q}^R)) \frac{\partial \underline{\psi}}{\partial s}(s, \underline{q}^L, \underline{q}^R)\, \mathrm{d}s \underline{n}_k^L\, \mathrm{d}H^n$$

   with $\underline{q}^L$ and $\underline{q}^R$ the left and right traces at the discontinuity, where H^n is the n-dimensional Hausdorf measure and where we choose $\underline{n}^L$ the outward normal with respect to the left state,

3. if B is an irregular Borel subset of Ω, then $\mu_i(B) = 0$

4.2 Weak Solutions

A function $\underline{q}$ of bounded variation is a weak solution to

$$\underline{q}_t + g(\underline{q})\underline{q}_x = 0 \tag{4.2}$$

if

$$\underline{q}_t + \left[g(\underline{q})\underline{q}_x\right]_\phi = 0 \tag{4.3}$$

as a bounded Borel measure on $\mathbb{R} \times \mathbb{R}_+$. This is equivalent to finding $\underline{q}$ that satisfies,

$$\int_{\mathbb{R}_+}\int_{\mathbb{R}} v_t(t,x)\underline{q}(t,x)\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{R}_+}\int_{\mathbb{R}} v(t,\cdot)\left[g(\underline{q}(t,\cdot))\underline{q}_x(t,\cdot)\right]_\psi \mathrm{d}t = \underline{0} \tag{4.4}$$

for all functions $v \in C_0^\infty(\mathbb{R}_t \times \mathbb{R})$.

4.3 DG Weak Formulation

4.3.1 Rhebergen Weak Formulation

Find $\underline{q} \in V_h$ such that for all $\underline{v} \in V_h$,

$$\sum_j \left(\int_{K_j} \underline{v}^T \underline{q}_t - \underline{v}_x^T \underline{f}(\underline{q}) + \underline{v}^T g(\underline{q})\underline{q}_x \,\mathrm{d}x\right) + \sum_S \left(\int_S \left(\underline{v}^L - \underline{v}^R\right)^T \underline{\hat{P}}^{nc}\,\mathrm{d}S\right) \tag{4.5}$$

$$+ \sum_S \left(\int_S \frac{1}{2}\left(\underline{v}^R + \underline{v}^L\right)^T \int_0^1 g\left(\underline{\psi}\left(\tau, \underline{q}^L, \underline{q}^R\right)\right) \frac{\partial \underline{\psi}}{\partial \tau}(\tau, \underline{q}^L, \underline{q}^R)\,\mathrm{d}\tau\,\mathrm{d}S\right) \tag{4.6}$$

where $\underline{\hat{P}}^{nc}$ is the nonconservative numerical flux, if symmetrical wave speeds are assumed, then the Rusanov or Local Lax Friedrichs flux can be used, otherwise the nonconservative product will affect the numerical flux.

4.3.2 Standard Hyperbolic Conservation Law DG Formulation

Let $\{K_j\}$ be a mesh of the domain $[a,b]$. Also denote the DG space as

$$V_h = \left\{v \in L^1([a,b]) \middle|\, v|_{K_j} \in \mathbb{P}^M(K_j)\right\}$$

Consider the hyperbolic conservation law given below with the corresponding classical and semi discrete weak solutions.

$$\underline{q}_t + \underline{f}\left(\underline{q}\right)_x = 0 \tag{4.7}$$

$$\int_a^b v\underline{q}_t - v_x\underline{f}\left(\underline{q}\right) \mathrm{d}x = 0 \tag{4.8}$$

The DG formulation requires finding $\underline{q}_h \in V_h$ for all $v_h \in V_h$ such that

$$\int_a^b v_h\underline{q}_{h,t} + v_h\underline{f}\left(\underline{q}_h\right)_x \mathrm{d}x = 0 \tag{4.9}$$

$$\sum_j \left(\int_{K_j} v_h\underline{q}_{h,t} + v_h\underline{f}\left(\underline{q}_h\right)_x \mathrm{d}x\right) = 0 \tag{4.10}$$

$$\sum_j \left(\int_{K_j} v_h\underline{q}_{h,t}\,\mathrm{d}x\right) + \sum_j \left(\int_{K_j} v_h\underline{f}\left(\underline{q}_h\right)_x \mathrm{d}x\right) = 0 \tag{4.11}$$

$$\sum_j \left(\int_{K_j} v_h\underline{q}_{h,t}\,\mathrm{d}x\right) + \sum_j \left(\hat{v}_h\hat{f}\left(\underline{q}_h\right)\Big|_{x_{j-1/2}}^{x_{j+1/2}} - \int_{K_j} v_{h,x}\underline{f}\left(\underline{q}_h\right) \mathrm{d}x\right) = 0 \tag{4.12}$$

Usually the value of $\hat{v}_h$ is the interior value of the test function on the element integral that is being integrated by parts. That is

$$\sum_j \left(\int_{K_j} v_h\underline{q}_{h,t}\,\mathrm{d}x\right) + \sum_j \left(v_h\left(x^-_{j+1/2}\right)\hat{f}_{j+1/2} - v_h\left(x^+_{j-1/2}\right)\hat{f}_{j-1/2} - \int_{K_j} v_{h,x}\underline{f}\left(\underline{q}_h\right) \mathrm{d}x\right) = 0 \tag{4.13}$$

$$\sum_j \left(\int_{K_j} v_h\underline{q}_{h,t}\,\mathrm{d}x\right) + \sum_{I_j} \left(\left(v_h(x^-) - v_h\left(x^+\right)\right)\hat{f}\right) - \sum_j \left(\int_{K_j} v_{h,x}\underline{f}\left(\underline{q}_h\right) \mathrm{d}x\right) = 0 \tag{4.14}$$

Using these values for the test functions at the interfaces and then grouping the interfaces together reveals jump terms in the test functions at the interfaces.

4.3.3 Pure Nonconservative DG Formulation

Consider the 1D nonconservative equation shown below,

$$\underline{q}_t + g(\underline{q})\underline{q}_x = \underline{0} \qquad x \in [a,b], 0 < t < T$$

Now the semi discrete DG formulation for this problem becomes finding $\underline{q}_h \in V_h$ for all $v_h \in V_h$ that satisfies

$$\int_a^b v_h \underline{q}_{h,t}\,\mathrm{d}x + \int_a^b v_h \left[g(\underline{q}_h)\underline{q}_{h,x}\right]_\psi \mathrm{d}x = \underline{0} \tag{4.15}$$

$$\sum_j \left(\int_{K_j} v_h \underline{q}_{h,t}\,\mathrm{d}x\right) + \sum_j \left(\int_{K_j} v_h g(\underline{q}_h)\underline{q}_{h,x}\,\mathrm{d}x\right) + \sum_I \left(\hat{v}_h \int_0^1 g(\psi(s, \underline{q}_h^L, \underline{q}_h^R))\frac{\partial \psi}{\partial s}(s, \underline{q}_h^L, \underline{q}_h^R)\,\mathrm{d}s\right) = 0 \tag{4.16}$$

Consider the case where there exists a function $\underline{f}\left(\underline{q}_h\right)$ such that $\underline{f}'\left(\underline{q}_h\right) = g(\underline{q})$.

$$\sum_j \left(\int_{K_j} v_h \underline{q}_{h,t}\,\mathrm{d}x\right) + \sum_j \left(\int_{K_j} v_h \underline{f}\left(\underline{q}_h\right)_x \mathrm{d}x\right) + \sum_I \left(\hat{v}_h \int_0^1 f'(\psi(s, \underline{q}_h^L, \underline{q}_h^R))\frac{\partial \psi}{\partial s}(s, \underline{q}_h^L, \underline{q}_h^R)\,\mathrm{d}s\right) = 0 \tag{4.17}$$

$$\sum_j \left(\int_{K_j} v_h \underline{q}_{h,t}\,\mathrm{d}x\right) + \sum_j \left(\int_{K_j} v_h \underline{f}\left(\underline{q}_h\right)_x \mathrm{d}x\right) + \sum_I \left(\hat{v}_h\left(\underline{f}\left(\underline{q}_h^R\right) - \underline{f}\left(\underline{q}_h^L\right)\right)\right) = 0 \tag{4.18}$$

$$\sum_j \left(\int_{K_j} v_h \underline{q}_{h,t}\,\mathrm{d}x\right) - \sum_j \left(\int_{K_j} v_{h,x} \underline{f}\left(\underline{q}_h\right) \mathrm{d}x\right) + \sum_I \left(v_h^L \underline{f}\left(\underline{q}_h^L\right) - v_h^R \underline{f}\left(\underline{q}_h^R\right)\right) + \sum_I \left(\hat{v}_h\left(\underline{f}\left(\underline{q}_h^R\right) - \underline{f}\left(\underline{q}_h^L\right)\right)\right) = 0 \tag{4.19}$$

$$\sum_j \left(\int_{K_j} v_h \underline{q}_{h,t}\,\mathrm{d}x\right) - \sum_j \left(\int_{K_j} v_{h,x} \underline{f}\left(\underline{q}_h\right) \mathrm{d}x\right) + \sum_I \left(\left(\hat{v} - v_h^R\right)\underline{f}\left(\underline{q}_h^R\right) + \left(v_h^L - \hat{v}\right)\underline{f}\left(\underline{q}_h^L\right)\right) = 0 \tag{4.20}$$

Now we want to choose $\hat{v}$ such that this is equivalent to the traditional DG formulation. However we don't have any numerical flux terms so instead we want the interface terms to look like $\left(v_h^L - v_h^R\right)\frac{1}{2}\left(\underline{f}\left(\underline{q}_h^R\right) + \underline{f}\left(\underline{q}_h^L\right)\right)$. At least this is what Rhebergen does and then replaces the flux average with the numerical flux.

$$\left(\hat{v}-v_h^R\right)\underline{f}\left(\underline{q}_h^R\right)+\left(v_h^L-\hat{v}\right)\underline{f}\left(\underline{q}_h^L\right)=\left(v_h^L-v_h^R\right)\frac{1}{2}\left(\underline{f}\left(\underline{q}_h^R\right)+\underline{f}\left(\underline{q}_h^L\right)\right) \tag{4.21}$$

$$\left(\hat{v}-v_h^R\right)\underline{f}\left(\underline{q}_h^R\right)+\left(v_h^L-\hat{v}\right)\underline{f}\left(\underline{q}_h^L\right)=\frac{1}{2}\left(v_h^L-v_h^R\right)\underline{f}\left(\underline{q}_h^R\right)+\frac{1}{2}\left(v_h^L-v_h^R\right)\underline{f}\left(\underline{q}_h^L\right) \tag{4.22}$$

$$\left(\hat{v}-v_h^R\right)=\frac{1}{2}\left(v_h^L-v_h^R\right) \tag{4.23}$$

$$\hat{v}=\frac{1}{2}\left(v_h^L+v_h^R\right) \tag{4.24}$$

$$\left(v_h^L-\hat{v}\right)=\frac{1}{2}\left(v_h^L-v_h^R\right) \tag{4.25}$$

$$-\hat{v}=\frac{1}{2}\left(-v_h^L-v_h^R\right) \tag{4.26}$$

$$\hat{v}=\frac{1}{2}\left(v_h^L+v_h^R\right) \tag{4.27}$$

$$ \tag{4.28}$$

We see that the appropriate numerical flux for the test function when multiplying the nonconservative product at the interface should be the average value. This agrees with the results given in Rhebergen. My one question about this is the swap from the average value of f to the numerical flux of f. I am tempted to use the numerical flux of f when integrating by parts, but then in order to agree with the traditional method $\hat{v}$ should be zero.

4.3.4 DG Formulation

Consider the 1D PDE below with a conservative and nonconservative term,

$$\underline{q}_t+\underline{f}\left(\underline{q}\right)_x+g\left(\underline{q}\right)\underline{q}_x=\underline{s}\left(\underline{q}\right) \qquad x\in[a,b],0<t<T \tag{4.29}$$

$$ \tag{4.30}$$

The semi discrete DG formulation is finding $\underline{q}_h \in V_h$ for all $v_h \in V_h$ such that

$$\int_a^b v_h \underline{q}_t \,\mathrm{d}x + \int_a^b v_h \underline{f}\left(\underline{q}\right)_x \mathrm{d}x + \int_a^b v_h \left[g\left(\underline{q}\right)\underline{q}_x\right]_{\underline{\psi}} = \int_a^b v_h \underline{s}\left(\underline{q}\right)\mathrm{d}x \tag{4.31}$$

$$\sum_j \left(\int_{K_j} v_h \underline{q}_t \,\mathrm{d}x\right) + \sum_j \left(\int_{K_j} v_h \underline{f}\left(\underline{q}\right)\mathrm{d}x\right) + \sum_j \left(\int_{K_j} v_h g\left(\underline{q}\right)\underline{q}_x \,\mathrm{d}x\right) \tag{4.32}$$

$$+\sum_I \left(\hat{v}_h \int_0^1 g\left(\underline{\psi}\left(s, \underline{q}_h^L, \underline{q}_h^R\right)\right)\frac{\partial \underline{\psi}}{\partial s}\left(s, \underline{q}_h^L, \underline{q}_h^R\right)\mathrm{d}s\right) = \int_a^b v_h \underline{s}\left(\underline{q}\right)\mathrm{d}x \tag{4.33}$$

$$\sum_j \left(\int_{K_j} v_h \underline{q}_t \,\mathrm{d}x\right) - \sum_j \left(\int_{K_j} v_{h,x} \underline{f}\left(\underline{q}\right)\mathrm{d}x\right) + \sum_I \left(\left(v_h^L - v_h^R\right)\hat{\underline{f}}\right) + \sum_j \left(\int_{K_j} v_h g\left(\underline{q}\right)\underline{q}_x \,\mathrm{d}x\right) \tag{4.34}$$

$$+\sum_I \left(\hat{v}_h \int_0^1 g\left(\underline{\psi}\left(s, \underline{q}_h^L, \underline{q}_h^R\right)\right)\underline{\psi}_s\left(s, \underline{q}_h^L, \underline{q}_h^R\right)\mathrm{d}s\right) = \int_a^b v_h \underline{s}\left(\underline{q}\right)\mathrm{d}x \tag{4.35}$$

As shown earlier if we choose $\hat{v}_h = \frac{1}{2}\left(v_h^R + v_h^L\right)$, then in the case where there exists $\underline{h}$ such that $\underline{h}' = g$, then this formulation will reduce to the standard DG formulation of the conservative PDE $\underline{q}_t + \left(\underline{f}\left(\underline{q}\right) + \underline{h}\left(\underline{q}\right)\right)_x = \underline{s}\left(\underline{q}\right)$.

Consider the case where $v_h = \phi_i^k$, that is the kth order basis function on the element K_i. In order to consider all of the basis functions on a cell K_i, I will use the test function $\underline{\phi}_i^T$. Let $\underline{\phi}$ be the vector of basis functions on the canonical element $[-1, 1]$ and let c_i be the linear transformation from $K_i \to [-1, 1]$, then $\underline{\phi}_i(x) = \underline{\phi}(c_i(x))$

$$\int_a^b \underline{q}_t \underline{\phi}_i^T \,\mathrm{d}x + \int_a^b \underline{f}\left(\underline{q}\right)_x \underline{\phi}_i^T \,\mathrm{d}x + \int_a^b \left[g\left(\underline{q}\right)\underline{q}_x\right]_{\underline{\psi}} \underline{\phi}_i^T = \int_a^b \underline{s}\left(\underline{q}\right)\underline{\phi}_i^T \,\mathrm{d}x \tag{4.36}$$

$$\int_{K_i} \underline{q}_t \underline{\phi}_i^T \,\mathrm{d}x + \int_{K_i} \underline{f}\left(\underline{q}\right)_x \underline{\phi}_i^T \,\mathrm{d}x + \int_{K_i} g\left(\underline{q}\right)\underline{q}_x \underline{\phi}_i^T \,\mathrm{d}x + \int_0^1 g(\underline{\psi}(s, \underline{q}_{i-1/2}^-, \underline{q}_{i-1/2}^+))\underline{\psi}_s(s, \underline{q}_{i-1/2}^-, \underline{q}_{i-1/2}^+)\,\mathrm{d}s\, \hat{\underline{\phi}}_{i-1/2}^T \tag{4.37}$$

$$+\int_0^1 g(\underline{\psi}(s, \underline{q}_{i+1/2}^-, \underline{q}_{i+1/2}^+))\underline{\psi}_s(s, \underline{q}_{i+1/2}^-, \underline{q}_{i+1/2}^+)\,\mathrm{d}s\, \hat{\underline{\phi}}_{i+1/2}^T = \int_{K_i} \underline{s}\left(\underline{q}\right)\underline{\phi}_i^T \,\mathrm{d}x \tag{4.38}$$

The value of $\underline{q}$ restricted to cell K_i can be expressed as an expansion of coefficients over the basis functions, that is $\underline{q}(x,t)\Big|_{K_i} = Q_i(t)\underline{\phi}_i(x)$

$$\int_{K_i} Q_i' \underline{\phi}_i \underline{\phi}_i^T \,\mathrm{d}x + \int_{K_i} \underline{f}\left(Q_i \underline{\phi}_i\right)_x \underline{\phi}_i^T \,\mathrm{d}x + \int_{K_i} g\left(Q_i \underline{\phi}_i\right) Q_i \underline{\phi}_{i,x} \underline{\phi}_i^T \,\mathrm{d}x \tag{4.39}$$

$$+\int_0^1 g(\underline{\psi}(s, Q_{i-1}\underline{\phi}(1), Q_i\underline{\phi}(-1)))\underline{\psi}_s(s, Q_{i-1}\underline{\phi}(1), Q_i\underline{\phi}(-1))\,\mathrm{d}s\, \hat{\underline{\phi}}_{i-1/2}^T \tag{4.40}$$

$$+\int_0^1 g(\underline{\psi}(s, Q_i\underline{\phi}(1), Q_{i+1}\underline{\phi}(-1)))\underline{\psi}_s(s, Q_i\underline{\phi}(1), Q_{i+1}\underline{\phi}(-1))\,\mathrm{d}s\, \hat{\underline{\phi}}_{i+1/2}^T = \int_{K_i} \underline{s}\left(Q_i \underline{\phi}_1\right)\underline{\phi}_i^T \,\mathrm{d}x \tag{4.41}$$

Integrate by Parts

$$\int_{K_i} Q_i' \underline{\phi}_i \underline{\phi}_i^T \,\mathrm{d}x + \underline{\hat{f}}_{i+1/2} \underline{\phi}^T(1) - \underline{\hat{f}}_{i-1/2} \underline{\phi}^T(-1) + \int_{K_i} g\Big(Q_i \underline{\phi}_i\Big) Q_i \underline{\phi}_{i,x} \underline{\phi}_i^T \,\mathrm{d}x \tag{4.42}$$

$$+ \int_0^1 g(\underline{\psi}(s, Q_{i-1}\underline{\phi}(1), Q_i \underline{\phi}(-1))) \underline{\psi}_s(s, Q_{i-1}\underline{\phi}(1), Q_i\underline{\phi}(-1)) \,\mathrm{d}s\, \underline{\hat{\phi}}_{i-1/2}^T \tag{4.43}$$

$$+ \int_0^1 g(\underline{\psi}(s, Q_i\underline{\phi}(1), Q_{i+1}\underline{\phi}(-1))) \underline{\psi}_s(s, Q_i\underline{\phi}(1), Q_{i+1}\underline{\phi}(-1)) \,\mathrm{d}s\, \underline{\hat{\phi}}_{i+1/2}^T = \int_{K_i} \underline{s}\Big(Q_i \underline{\phi}_1\Big) \underline{\phi}_i^T \,\mathrm{d}x \tag{4.44}$$

Change to canonical basis with linear transformation, and rearrange equation

$$\int_{K_i} Q_i' \underline{\phi}(c_i(x)) \underline{\phi}^T(c_i(x)) \,\mathrm{d}x = \int_{K_i} \underline{f}\Big(Q_i \underline{\phi}(c_i(x))\Big) \partial_x \underline{\phi}^T(c_i(x)) \,\mathrm{d}x - \Big(\underline{\hat{f}}_{i+1/2} \underline{\phi}^T(1) - \underline{\hat{f}}_{i-1/2} \underline{\phi}^T(-1)\Big) \tag{4.45}$$

$$- \int_{K_i} g\Big(Q_i \underline{\phi}(c_i(x))\Big) Q_i \partial_x \underline{\phi}(c_i(x)) \underline{\phi}^T(c_i(x)) \,\mathrm{d}x - \int_0^1 g(\underline{\psi}(s, Q_{i-1}\underline{\phi}(1), Q_i\underline{\phi}(-1))) \underline{\psi}_s(s, Q_{i-1}\underline{\phi}(1), Q_i\underline{\phi}(-1)) \,\mathrm{d}s\, \underline{\hat{\phi}}_{i-1/2}^T \tag{4.46}$$

$$- \int_0^1 g(\underline{\psi}(s, Q_i\underline{\phi}(1), Q_{i+1}\underline{\phi}(-1))) \underline{\psi}_s(s, Q_i\underline{\phi}(1), Q_{i+1}\underline{\phi}(-1)) \,\mathrm{d}s\, \underline{\hat{\phi}}_{i+1/2}^T + \int_{K_i} \underline{s}\Big(Q_i \underline{\phi}(c_i(x))\Big) \underline{\phi}^T(c_i(x)) \,\mathrm{d}x \tag{4.47}$$

Convert integrals to canonical element, denote $c_i^{-1}(\xi) = b_i(\xi)$ and $m_i = \dfrac{\partial b_i}{\partial \xi}$. Performing two operations, $\partial_x \underline{\phi}(c_i(x)) = \underline{\phi}_\xi(c_i(x)) c_i'(x) = \frac{1}{m_i} \underline{\phi}(c_i(x))$ and transforming integrals from K_i to $[-1, 1]$ results in $x = b_i(\xi)$ or $c_i(x) = \xi$ and multiply by measure m_i. Drop explicit dependence on ξ, $\phi = \phi(\xi)$

$$\int_{-1}^1 Q_i' \underline{\phi \phi}^T m_i \,\mathrm{d}\xi = \int_{-1}^1 \underline{f}\Big(Q_i \underline{\phi}\Big) \frac{1}{m_i} \underline{\phi}_\xi^T m_i \,\mathrm{d}\xi - \Big(\underline{\hat{f}}_{i+1/2} \underline{\phi}^T(1) - \underline{\hat{f}}_{i-1/2} \underline{\phi}^T(-1)\Big) \tag{4.48}$$

$$- \int_{-1}^1 g\Big(Q_i \underline{\phi}\Big) Q_i \frac{1}{m_i} \underline{\phi}_\xi \underline{\phi}^T m_i \,\mathrm{d}\xi - \int_0^1 g(\underline{\psi}(s, Q_{i-1}\underline{\phi}(1), Q_i\underline{\phi}(-1))) \underline{\psi}_s(s, Q_{i-1}\underline{\phi}(1), Q_i\underline{\phi}(-1)) \,\mathrm{d}s\, \underline{\hat{\phi}}_{i-1/2}^T \tag{4.49}$$

$$- \int_0^1 g(\underline{\psi}(s, Q_i\underline{\phi}(1), Q_{i+1}\underline{\phi}(-1))) \underline{\psi}_s(s, Q_i\underline{\phi}(1), Q_{i+1}\underline{\phi}(-1)) \,\mathrm{d}s\, \underline{\hat{\phi}}_{i+1/2}^T + \int_{-1}^1 \underline{s}\Big(Q_i \underline{\phi}\Big) \underline{\phi}^T m_i \,\mathrm{d}\xi \tag{4.50}$$

Simplify, note that as before $\underline{\hat{\phi}}$ is the interface average, so in this case it results in half the interior value.

$$m_i Q_i' \int_{-1}^{1} \underline{\phi}\underline{\phi}^T \, \mathrm{d}\xi = \int_{-1}^{1} \underline{f}\Big(Q_i \underline{\phi}\Big) \underline{\phi}_\xi^T \, \mathrm{d}\xi - \Big(\underline{\hat{f}}_{i+1/2} \underline{\phi}^T(1) - \underline{\hat{f}}_{i-1/2} \underline{\phi}^T(-1)\Big) \tag{4.51}$$

$$- \int_{-1}^{1} g\Big(Q_i \underline{\phi}\Big) Q_i \underline{\phi}_\xi \underline{\phi}^T \, \mathrm{d}\xi - \frac{1}{2} \int_0^1 g(\underline{\psi}(s, Q_{i-1}\underline{\phi}(1), Q_i \underline{\phi}(-1))) \underline{\psi}_s(s, Q_{i-1}\underline{\phi}(1), Q_i \underline{\phi}(-1)) \, \mathrm{d}s \, \underline{\phi}^T(-1) \tag{4.52}$$

$$-\frac{1}{2} \int_0^1 g(\underline{\psi}(s, Q_i \underline{\phi}(1), Q_{i+1}\underline{\phi}(-1))) \underline{\psi}_s(s, Q_i \underline{\phi}(1), Q_{i+1}\underline{\phi}(-1)) \, \mathrm{d}s \, \underline{\phi}^T(1) + m_i \int_{-1}^{1} \underline{s}\Big(Q_i \underline{\phi}\Big) \underline{\phi}^T \, \mathrm{d}\xi \tag{4.53}$$

The integral on the left hand side gives the mass matrix, right multiplying by M^{-1} and dividing m_i across gives

$$Q_i' = \frac{1}{m_i} \int_{-1}^{1} \underline{f}\Big(Q_i \underline{\phi}\Big) \underline{\phi}_\xi^T \, \mathrm{d}\xi \, M^{-1} - \frac{1}{m_i} \Big(\underline{\hat{f}}_{i+1/2} \underline{\phi}^T(1) - \underline{\hat{f}}_{i-1/2} \underline{\phi}^T(-1)\Big) M^{-1} \tag{4.54}$$

$$- \frac{1}{m_i} \int_{-1}^{1} g\Big(Q_i \underline{\phi}\Big) Q_i \underline{\phi}_\xi \underline{\phi}^T \, \mathrm{d}\xi M^{-1} \tag{4.55}$$

$$- \frac{1}{2m_i} \int_0^1 g(\underline{\psi}(s, Q_{i-1}\underline{\phi}(1), Q_i \underline{\phi}(-1))) \underline{\psi}_s(s, Q_{i-1}\underline{\phi}(1), Q_i \underline{\phi}(-1)) \, \mathrm{d}s \, \underline{\phi}^T(-1) M^{-1} \tag{4.56}$$

$$- \frac{1}{2m_i} \int_0^1 g(\underline{\psi}(s, Q_i \underline{\phi}(1), Q_{i+1}\underline{\phi}(-1))) \underline{\psi}_s(s, Q_i \underline{\phi}(1), Q_{i+1}\underline{\phi}(-1)) \, \mathrm{d}s \, \underline{\phi}^T(1) M^{-1} + \int_{-1}^{1} \underline{s}\Big(Q_i \underline{\phi}\Big) \underline{\phi}^T \, \mathrm{d}\xi \, M^{-1} \tag{4.57}$$

4.3.5 Two Dimensions

Consider the two dimensional balance law given by

$$\underline{q}_t + \nabla \cdot \underline{f}_j\Big(\underline{q}, \underline{x}, t\Big) + \underline{\underline{G}}_j\Big(\underline{q}, \underline{x}, t\Big) \underline{q}_{x_j} = \underline{s}\Big(\underline{q}, \underline{x}, t\Big) \tag{4.58}$$

Note that the flux function is a matrix or two index tensor, so the divergence is a vector quantity, and the nonconservative term is a sum over the dimensions. It could also be written as

$$\underline{q}_t + \underline{f}_1\Big(\underline{q}, \underline{x}, t\Big)_x + \underline{f}_2\Big(\underline{q}, \underline{x}, t\Big)_y + \underline{\underline{G}}_1\Big(\underline{q}, \underline{x}, t\Big) \underline{q}_x + \underline{\underline{G}}_2\Big(\underline{q}, \underline{x}, t\Big) \underline{q}_y = \underline{s}\Big(\underline{q}, \underline{x}, t\Big) \tag{4.59}$$

The local statements of the weak discontinuous Galerkin form are given by

$$\int_{K_i} \underline{q}_t \phi_i^k(\underline{x}) - \underline{f}_j\left(\underline{q}, \underline{x}, t\right)\phi_{i,x_j}^k(\underline{x}) + \underline{\underline{G}}_j\left(\underline{q}, \underline{x}, t\right)\underline{q}_{x_j}\phi_i^k(\underline{x})\,\mathrm{d}\underline{x} \tag{4.60}$$

$$= -\int_{\partial K_i} \underline{\underline{f}}^* \underline{n}\phi_i^k(\underline{x})\,\mathrm{d}s - \frac{1}{2}\int_{\partial K_i}\int_0^1 \underline{\underline{G}}(\underline{\psi}\left(\tau, \underline{q}_L, \underline{q}_R\right), \underline{x}, t)\underline{\psi}_\tau\left(\tau, \underline{q}_L, \underline{q}_R\right)\mathrm{d}\tau \underline{n}_L \phi_i^k\,\mathrm{d}s + \int_{K_i} \underline{s}\left(\underline{q}, \underline{x}, t\right)\phi_i^k(\underline{x})\,\mathrm{d}\underline{x} \tag{4.61}$$

The vector $\underline{n}_L$ is the outward normal facing vector with respect to the left state, i.e. the state of $\underline{q}_L$. We can also consider all of the basis components at once, by using the test function $\underline{\phi}_i^T$ instead of ϕ_i^k.

$$\int_{K_i} \underline{q}_t \underline{\phi}_i^T(\underline{x}) - \underline{\underline{f}}\left(\underline{q}, \underline{x}, t\right)\underline{\underline{D\phi}}_i^T(\underline{x}) + \underline{\underline{G}}_j\left(\underline{q}, \underline{x}, t\right)\underline{q}_{x_j}\underline{\phi}_i^T(\underline{x})\,\mathrm{d}\underline{x} \tag{4.62}$$

$$= -\int_{\partial K_i} \underline{\underline{f}}^* \underline{n}\underline{\phi}_i^T(\underline{x})\,\mathrm{d}s - \frac{1}{2}\int_{\partial K_i}\int_0^1 \underline{\underline{G}}\left(\underline{\psi}\left(\tau, \underline{q}_L, \underline{q}_R\right), \underline{x}, t\right)\underline{\psi}_\tau\left(\tau, \underline{q}_L, \underline{q}_R\right)\mathrm{d}\tau \underline{n}_L \underline{\phi}_i^T(\underline{x})\,\mathrm{d}s + \int_{K_i} \underline{s}\left(\underline{q}, \underline{x}, t\right)\underline{\phi}_i^T(\underline{x})\,\mathrm{d}\underline{x} \tag{4.63}$$

Using the fact that $\underline{q}\Big|_{K_i} = \underline{\underline{Q}}_i\underline{\phi}_i$, and dropping the explicit dependence on $\underline{x}$ for $\underline{\phi}_i$.

$$\int_{K_i} \underline{\underline{Q}}_{i,t}\underline{\phi}_i\underline{\phi}_i^T - \underline{\underline{f}}\left(\underline{\underline{Q}}_i\underline{\phi}_i, \underline{x}, t\right)\underline{\underline{D\phi}}_i^T + \underline{\underline{G}}_j\left(\underline{\underline{Q}}_i\underline{\phi}_i, \underline{x}, t\right)\underline{\underline{Q}}_i\underline{\phi}_{i,x_j}\underline{\phi}_i^T\,\mathrm{d}\underline{x} \tag{4.64}$$

$$= -\int_{\partial K_i} \underline{\underline{f}}^* \underline{n}\underline{\phi}_i^T\,\mathrm{d}s - \frac{1}{2}\int_{\partial K_i}\int_0^1 G\left(\underline{\psi}\left(\tau, \underline{q}_L, \underline{q}_R\right), \underline{x}, t\right)\underline{\psi}_\tau\left(\tau, \underline{q}_L, \underline{q}_R\right)\mathrm{d}\tau \underline{n}_L \underline{\phi}_i^T\,\mathrm{d}s + \int_{K_i} \underline{s}\left(\underline{\underline{Q}}_i\underline{\phi}_i, \underline{x}, t\right)\underline{\phi}_i^T\,\mathrm{d}\underline{x} \tag{4.65}$$

Rearranging to solve for $\underline{\underline{Q}}_{i,t}$.

$$\underline{\underline{Q}}_{i,t}\int_{K_i} \underline{\phi}_i\underline{\phi}_i^T\,\mathrm{d}\underline{x} = \int_{K_i} \underline{\underline{f}}\left(\underline{\underline{Q}}_i\underline{\phi}_i, \underline{x}, t\right)\underline{\underline{D\phi}}_i^T - \underline{\underline{G}}_j\left(\underline{\underline{Q}}_i\underline{\phi}_i, \underline{x}, t\right)\underline{\underline{Q}}_i\underline{\phi}_{i,x_j}\underline{\phi}_i^T\,\mathrm{d}\underline{x} \tag{4.66}$$

$$-\int_{\partial K_i} \underline{\underline{f}}^* \underline{n}\underline{\phi}_i^T\,\mathrm{d}s - \frac{1}{2}\int_{\partial K_i}\int_0^1 G\left(\underline{\psi}\left(\tau, \underline{q}_L, \underline{q}_R\right), \underline{x}, t\right)\underline{\psi}_\tau\left(\tau, \underline{q}_L, \underline{q}_R\right)\mathrm{d}\tau \underline{n}_L \underline{\phi}_i^T\,\mathrm{d}s + \int_{K_i} \underline{s}\left(\underline{\underline{Q}}_i\underline{\phi}_i, \underline{x}, t\right)\underline{\phi}_i^T\,\mathrm{d}\underline{x} \tag{4.67}$$

Transforming integrals to canonical element, where $\underline{c}'_{ij}$ is the jth column of the jacobian of the function $\underline{c}_i(\underline{x})$, which transforms the element K_i to the canonical element $\mathcal{K}$. Also let f be the faces of $\mathcal{K}$, with parameterizations $\underline{r}_f(s)$.

$$\underline{\underline{Q}}_{i,t} m_i \underline{\underline{M}} = \int_{\mathcal{K}} \underline{\underline{f}}\Big(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i(\underline{\xi}), t\Big) \underline{\underline{D\phi^T}} \underline{c}'_i m_i \, \mathrm{d}\underline{\xi} - \int_{\mathcal{K}} \sum_{j=1}^{d} \Big(\underline{\underline{G}}_j\Big(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i(\underline{\xi}), t\Big) \underline{\underline{Q}}_i \underline{\phi}' \underline{c}'_{ij}\Big) \underline{\phi}^T \, \mathrm{d}\underline{\xi} \tag{4.68}$$

$$- \sum_{f \in mcK} \left(\int \underline{\underline{f}}^* \Big(\underline{b}_i\Big(\underline{r}_f(s)\Big)\Big) \underline{n} \underline{\phi}^T \Big(\underline{r}_f(s)\Big) \Big\| \underline{b}'_i\Big(\underline{r}_f(s)\Big) \underline{r}'_f(s) \Big\| \, \mathrm{d}s \right) \tag{4.69}$$

$$-\frac{1}{2} \sum_{f \in \mathcal{K}} \left(\int \sum_{j=1}^{d} \left(\int_0^1 \underline{\underline{G}}_j \Big(\underline{\psi}\Big(\tau, \underline{q}_L, \underline{q}_R\Big), \underline{b}_i(\underline{r}(s)), t\Big) \underline{\psi}_\tau \Big(\tau, \underline{q}_L, \underline{q}_R\Big) \, \mathrm{d}\tau \, n_{L,j} \right) \underline{\phi}^T \Big(\underline{r}_f(s)\Big) \Big\| \underline{b}'_i\Big(\underline{r}_f(s)\Big) \underline{r}'_f(s) \Big\| \, \mathrm{d}s \right) \tag{4.70}$$

$$+ \int_{\mathcal{K}} \underline{s}\Big(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i(\underline{\xi}), t\Big) \underline{\phi}^T m_i \, \mathrm{d}\underline{\xi} \tag{4.71}$$

Solving for $\underline{\underline{Q}}_{i,t}$ gives

$$\underline{\underline{Q}}_{i,t} = \int_{\mathcal{K}} \underline{\underline{f}}\Big(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i\Big(\underline{\xi}\Big), t\Big) \underline{\underline{D\phi^T}} \underline{c}'_i \, \mathrm{d}\underline{\xi} \underline{\underline{M}}^{-1} - \int_{\mathcal{K}} \sum_{j=1}^{d} \Big(\underline{\underline{G}}_j\Big(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i(\underline{\xi}), t\Big) \underline{\underline{Q}}_i \underline{\phi}' \underline{c}'_{ij}\Big) \underline{\phi}^T \, \mathrm{d}\underline{\xi} \underline{\underline{M}}^{-1} \tag{4.72}$$

$$-\frac{1}{m_i} \sum_{f \in mcK} \left(\int \underline{\underline{f}}^* \Big(\underline{b}_i\Big(\underline{r}_f(s)\Big)\Big) \underline{n} \underline{\phi}^T \Big(\underline{r}_f(s)\Big) \Big\| \underline{b}'_i\Big(\underline{r}_f(s)\Big) \underline{r}'_f(s) \Big\| \, \mathrm{d}s \right) \underline{\underline{M}}^{-1} \tag{4.73}$$

$$-\frac{1}{2m_i} \sum_{f \in \mathcal{K}} \left(\int \sum_{j=1}^{d} \left(\int_0^1 \underline{\underline{G}}_j \Big(\underline{\psi}\Big(\tau, \underline{q}_L, \underline{q}_R\Big), \underline{b}_i(\underline{r}(s)), t\Big) \underline{\psi}_\tau \Big(\tau, \underline{q}_L, \underline{q}_R\Big) \, \mathrm{d}\tau \, n_{L,j} \right) \underline{\phi}^T \Big(\underline{r}_f(s)\Big) \Big\| \underline{b}'_i\Big(\underline{r}_f(s)\Big) \underline{r}'_f(s) \Big\| \, \mathrm{d}s \right) \underline{\underline{M}}^{-1} \tag{4.74}$$

$$+ \int_{\mathcal{K}} \underline{s}\Big(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i(\underline{\xi}), t\Big) \underline{\phi}^T \, \mathrm{d}\underline{\xi} \underline{\underline{M}}^{-1} \tag{4.75}$$

4.3.5.1 Rectangular Elements

Consider if the mesh contain rectangular elements, then $K_i = \Big[x_{i-1/2}, x_{i+1/2}\Big] \times \Big[y_{i-1/2}, y_{i+1/2}\Big]$. The center of the element is (x_i, y_i) with $\Delta x_i = x_{i+1/2} - x_{i-1/2}$ and $\Delta y_i = y_{i+1/2} - y_{i-1/2}$. The canonical element is $\mathcal{K} = [-1, 1] \times [-1, 1]$ with coordinates $\underline{\xi} = [\xi, \eta]$. The linear transformations are given by

$$\underline{b}_i(\underline{\xi}) = \left[\frac{\Delta x_i}{2}\xi + x_i, \frac{\Delta y_i}{2}\eta + y_i\right]^T \tag{4.76}$$

$$\underline{c}_i(\underline{x}) = \left[\frac{2}{\Delta x_i}(x - x_i), \frac{2}{\Delta y_i}(y - y_i)\right]^T \tag{4.77}$$

with Jacobians

$$\underline{b}_i' = \begin{pmatrix} \frac{\Delta x_i}{2} & 0 \\ 0 & \frac{\Delta y_i}{2} \end{pmatrix} \tag{4.78}$$

$$\underline{c}_i' = \begin{pmatrix} \frac{2}{\Delta x_i} & 0 \\ 0 & \frac{2}{\Delta y_i} \end{pmatrix} \tag{4.79}$$

The metric of element i is $m_i = \frac{\Delta x_i \Delta y_i}{4}$. Also the parameterizations of the left, right, bottom, and top faces, r_l, r_r, r_b, r_t respectively, are given by

$$r_l(t) = [-1, t] \tag{4.80}$$

$$r_r(t) = [1, t] \tag{4.81}$$

$$r_b(t) = [t, -1] \tag{4.82}$$

$$r_t(t) = [t, 1] \tag{4.83}$$

for $t \in [-1, 1]$. We can easily compute $\left\| \underline{b}_i'(\underline{r}_f(t))\underline{r}_f'(t) \right\|$ for each face as well

$$\|\underline{b}_i'(\underline{r}_l(t))\underline{r}_l'(t)\| = \frac{\Delta y_i}{2} \tag{4.84}$$

$$\|\underline{b}_i'(\underline{r}_r(t))\underline{r}_r'(t)\| = \frac{\Delta y_i}{2} \tag{4.85}$$

$$\|\underline{b}_i'(\underline{r}_b(t))\underline{r}_b'(t)\| = \frac{\Delta x_i}{2} \tag{4.86}$$

$$\|\underline{b}_i'(\underline{r}_t(t))\underline{r}_t'(t)\| = \frac{\Delta x_i}{2} \tag{4.87}$$

Substituting all of these into the formulation gives,

$$\underline{\underline{Q}}_{i,t} = \int_{\mathcal{K}} \frac{2}{\Delta x_i} \underline{f}_1\Big(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i\big(\underline{\xi}\big), t\Big) \underline{\phi}_{\xi}^T + \frac{2}{\Delta y_i} \underline{f}_2\Big(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i\big(\underline{\xi}\big), t\Big) \underline{\phi}_{\eta}^T \, \mathrm{d}\underline{\xi} \underline{\underline{M}}^{-1} \tag{4.88}$$

$$- \int_{\mathcal{K}} \frac{2}{\Delta x_i} \underline{\underline{G}}_1\Big(\underline{\underline{Q}}_i \underline{\phi}_i, \underline{b}_i\big(\underline{\xi}\big), t\Big) \underline{\underline{Q}}_i \underline{\phi}_{\xi} \underline{\phi}^T + \underline{\underline{G}}_2\Big(\underline{\underline{Q}}_i \underline{\phi}_i, \underline{b}_i\big(\underline{\xi}\big), t\Big) \underline{\underline{Q}}_i \underline{\phi}_{\eta} \underline{\phi}^T \, \mathrm{d}\underline{\xi} \underline{\underline{M}}^{-1} \tag{4.89}$$

$$+ \frac{2}{\Delta x_i} \int_{-1}^{1} \underline{f}_1^*(\underline{b}_i(\xi = -1, \eta)) \underline{\phi}^T(\xi = -1, \eta) \, \mathrm{d}\eta \underline{\underline{M}}^{-1} \tag{4.90}$$

$$- \frac{2}{\Delta x_i} \int_{-1}^{1} \underline{f}_1^*(\underline{b}_i(\xi = 1, \eta)) \underline{\phi}^T(\xi = 1, \eta) \, \mathrm{d}\eta \underline{\underline{M}}^{-1} \tag{4.91}$$

$$+ \frac{2}{\Delta y_i} \int_{-1}^{1} \underline{f}_2^*(\underline{b}_i(\xi, \eta = -1)) \underline{\phi}^T(\xi, \eta = -1) \, \mathrm{d}\xi \underline{\underline{M}}^{-1} \tag{4.92}$$

$$- \frac{2}{\Delta y_i} \int_{-1}^{1} \underline{f}_2^*(\underline{b}_i(\xi, \eta = 1)) \underline{\phi}^T(\xi, \eta = 1) \, \mathrm{d}\xi \underline{\underline{M}}^{-1} \tag{4.93}$$

$$- \frac{1}{\Delta x_i} \int_{-1}^{1} \int_{0}^{1} \underline{\underline{G}}_1\Big(\underline{\psi}\Big(\tau, \underline{q}_L, \underline{q}_R\Big), \underline{b}_i(\xi = -1, \eta), t\Big) \underline{\psi}_{\tau}\Big(\tau, \underline{q}_L, \underline{q}_R\Big) \, \mathrm{d}\tau \underline{\phi}^T(\xi = -1, \eta) \, \mathrm{d}\eta \underline{\underline{M}}^{-1} \tag{4.94}$$

$$- \frac{1}{\Delta x_i} \int_{-1}^{1} \int_{0}^{1} \underline{\underline{G}}_1\Big(\underline{\psi}\Big(\tau, \underline{q}_L, \underline{q}_R\Big), \underline{b}_i(\xi = 1, \eta), t\Big) \underline{\psi}_{\tau}\Big(\tau, \underline{q}_L, \underline{q}_R\Big) \, \mathrm{d}\tau \underline{\phi}^T(\xi = 1, \eta) \, \mathrm{d}\eta \underline{\underline{M}}^{-1} \tag{4.95}$$

$$- \frac{1}{\Delta y_i} \int_{-1}^{1} \int_{0}^{1} \underline{\underline{G}}_2\Big(\underline{\psi}\Big(\tau, \underline{q}_L, \underline{q}_R\Big), \underline{b}_i(\xi, \eta = -1), t\Big) \underline{\psi}_{\tau}\Big(\tau, \underline{q}_L, \underline{q}_R\Big) \, \mathrm{d}\tau \underline{\phi}^T(\xi, \eta = -1) \, \mathrm{d}\xi \underline{\underline{M}}^{-1} \tag{4.96}$$

$$- \frac{1}{\Delta y_i} \int_{-1}^{1} \int_{0}^{1} \underline{\underline{G}}_2\Big(\underline{\psi}\Big(\tau, \underline{q}_L, \underline{q}_R\Big), \underline{b}_i(\xi, \eta = 1), t\Big) \underline{\psi}_{\tau}\Big(\tau, \underline{q}_L, \underline{q}_R\Big) \, \mathrm{d}\tau \underline{\phi}^T(\xi, \eta = 1) \, \mathrm{d}\xi \underline{\underline{M}}^{-1} \tag{4.97}$$

$$+ \int_{\mathcal{K}} \underline{s}\Big(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i(\underline{\xi}), t\Big) \underline{\phi}^T \, \mathrm{d}\underline{\xi} \underline{\underline{M}}^{-1} \tag{4.98}$$

For the case of a legendre orthogonal basis with orthogonality condition

$$\frac{1}{4} \int_{\mathcal{K}} \phi^i(\underline{\xi}) \phi^j(\underline{\xi}) \, \mathrm{d}\xi = \delta_{ij},$$

then the mass matrix and it's inverse become $\underline{\underline{M}} = 4I$ and $\underline{\underline{M}}^{-1} = \frac{1}{4}I$.

$$\underline{\underline{Q}}_{i,t} = \int_{\mathcal{K}} \frac{1}{2\Delta x_i} \underline{f}_1\Big(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i\Big(\underline{\xi}\Big), t\Big) \underline{\phi}_{\xi}^T + \frac{1}{2\Delta y_i} \underline{f}_2\Big(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i\Big(\underline{\xi}\Big), t\Big) \underline{\phi}_{\eta}^T \,\mathrm{d}\underline{\xi} \tag{4.99}$$

$$- \int_{\mathcal{K}} \frac{1}{2\Delta x_i} \underline{\underline{G}}_1\Big(\underline{\underline{Q}}_i \underline{\phi}_i, \underline{b}_i\Big(\underline{\xi}\Big), t\Big) \underline{\underline{Q}}_i \underline{\phi}_{\xi} \underline{\phi}^T + \frac{1}{2\Delta y_i} \underline{\underline{G}}_2\Big(\underline{\underline{Q}}_i \underline{\phi}_i, \underline{b}_i\Big(\underline{\xi}\Big), t\Big) \underline{\underline{Q}}_i \underline{\phi}_{\eta} \underline{\phi}^T \,\mathrm{d}\underline{\xi} \tag{4.100}$$

$$+ \frac{1}{2\Delta x_i} \int_{-1}^{1} \underline{f}_1^*(\underline{b}_i(\xi = -1, \eta)) \underline{\phi}^T(\xi = -1, \eta) \,\mathrm{d}\eta \tag{4.101}$$

$$- \frac{1}{2\Delta x_i} \int_{-1}^{1} \underline{f}_1^*(\underline{b}_i(\xi = 1, \eta)) \underline{\phi}^T(\xi = 1, \eta) \,\mathrm{d}\eta \tag{4.102}$$

$$+ \frac{1}{2\Delta y_i} \int_{-1}^{1} \underline{f}_2^*(\underline{b}_i(\xi, \eta = -1)) \underline{\phi}^T(\xi, \eta = -1) \,\mathrm{d}\xi \tag{4.103}$$

$$- \frac{1}{2\Delta y_i} \int_{-1}^{1} \underline{f}_2^*(\underline{b}_i(\xi, \eta = 1)) \underline{\phi}^T(\xi, \eta = 1) \,\mathrm{d}\xi \tag{4.104}$$

$$- \frac{1}{4\Delta x_i} \int_{-1}^{1} \int_{0}^{1} \underline{\underline{G}}_1\Big(\underline{\psi}\Big(\tau, \underline{q}_L, \underline{q}_R\Big), \underline{b}_i(\xi = -1, \eta), t\Big) \underline{\psi}_{\tau}\Big(\tau, \underline{q}_L, \underline{q}_R\Big) \,\mathrm{d}\tau \underline{\phi}^T(\xi = -1, \eta) \,\mathrm{d}\eta \tag{4.105}$$

$$- \frac{1}{4\Delta x_i} \int_{-1}^{1} \int_{0}^{1} \underline{\underline{G}}_1\Big(\underline{\psi}\Big(\tau, \underline{q}_L, \underline{q}_R\Big), \underline{b}_i(\xi = 1, \eta), t\Big) \underline{\psi}_{\tau}\Big(\tau, \underline{q}_L, \underline{q}_R\Big) \,\mathrm{d}\tau \underline{\phi}^T(\xi = 1, \eta) \,\mathrm{d}\eta \tag{4.106}$$

$$- \frac{1}{4\Delta y_i} \int_{-1}^{1} \int_{0}^{1} \underline{\underline{G}}_2\Big(\underline{\psi}\Big(\tau, \underline{q}_L, \underline{q}_R\Big), \underline{b}_i(\xi, \eta = -1), t\Big) \underline{\psi}_{\tau}\Big(\tau, \underline{q}_L, \underline{q}_R\Big) \,\mathrm{d}\tau \underline{\phi}^T(\xi, \eta = -1) \,\mathrm{d}\xi \tag{4.107}$$

$$- \frac{1}{4\Delta y_i} \int_{-1}^{1} \int_{0}^{1} \underline{\underline{G}}_2\Big(\underline{\psi}\Big(\tau, \underline{q}_L, \underline{q}_R\Big), \underline{b}_i(\xi, \eta = 1), t\Big) \underline{\psi}_{\tau}\Big(\tau, \underline{q}_L, \underline{q}_R\Big) \,\mathrm{d}\tau \underline{\phi}^T(\xi, \eta = 1) \,\mathrm{d}\xi \tag{4.108}$$

$$+ \frac{1}{4} \int_{\mathcal{K}} \underline{s}\Big(\underline{\underline{Q}}_i \underline{\phi}, \underline{b}_i(\underline{\xi}), t\Big) \underline{\phi}^T \,\mathrm{d}\underline{\xi} \tag{4.109}$$

4.3.5.2 Triangular Elements

Still need to type this up

Chapter5. Results

5.1 One Dimensional Results

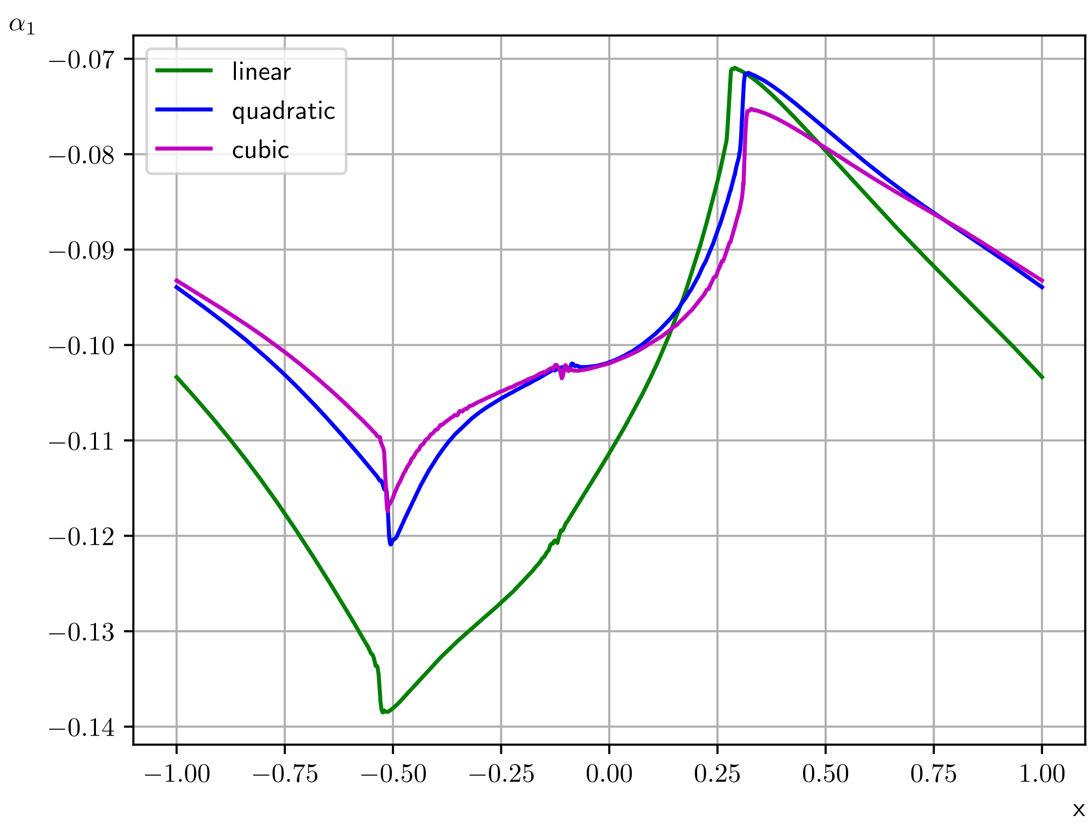
5.1.1

5.1.2 Effect of Drag

One interesting test to run

$$h_0 = 1 + e^{3\cos(\pi(x-0.5))-4} \tag{5.1}$$

$$u_0 = 0.25 \tag{5.2}$$



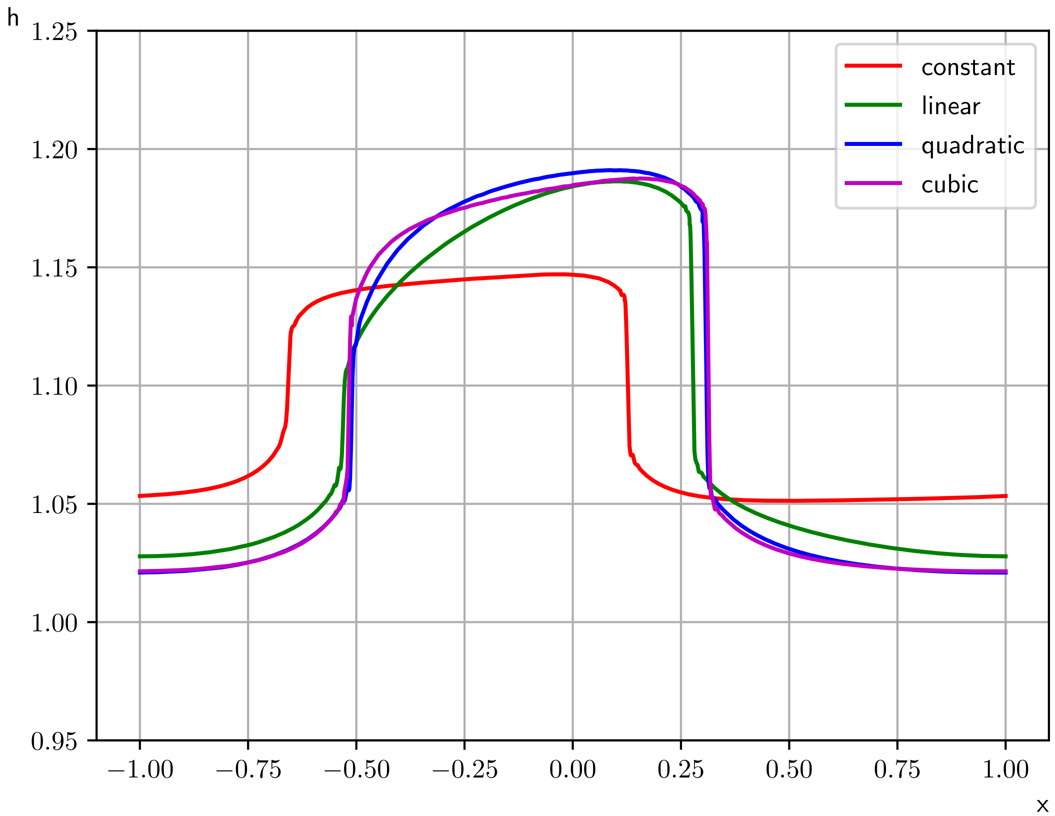
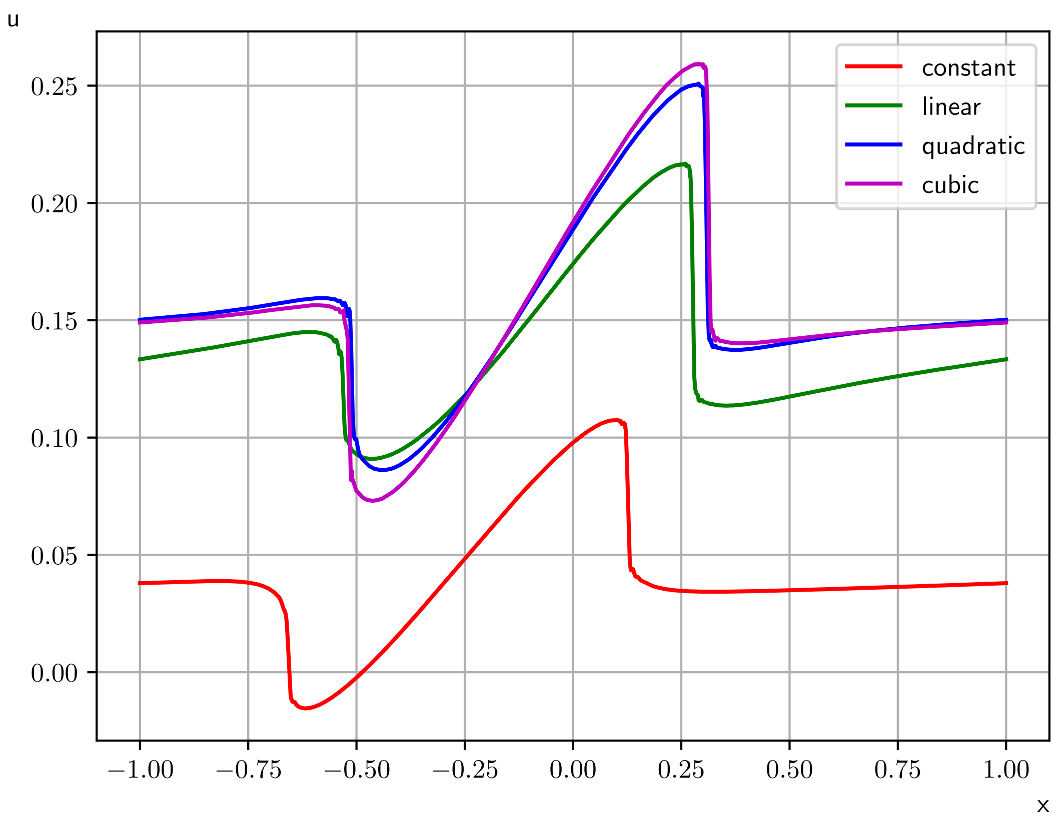
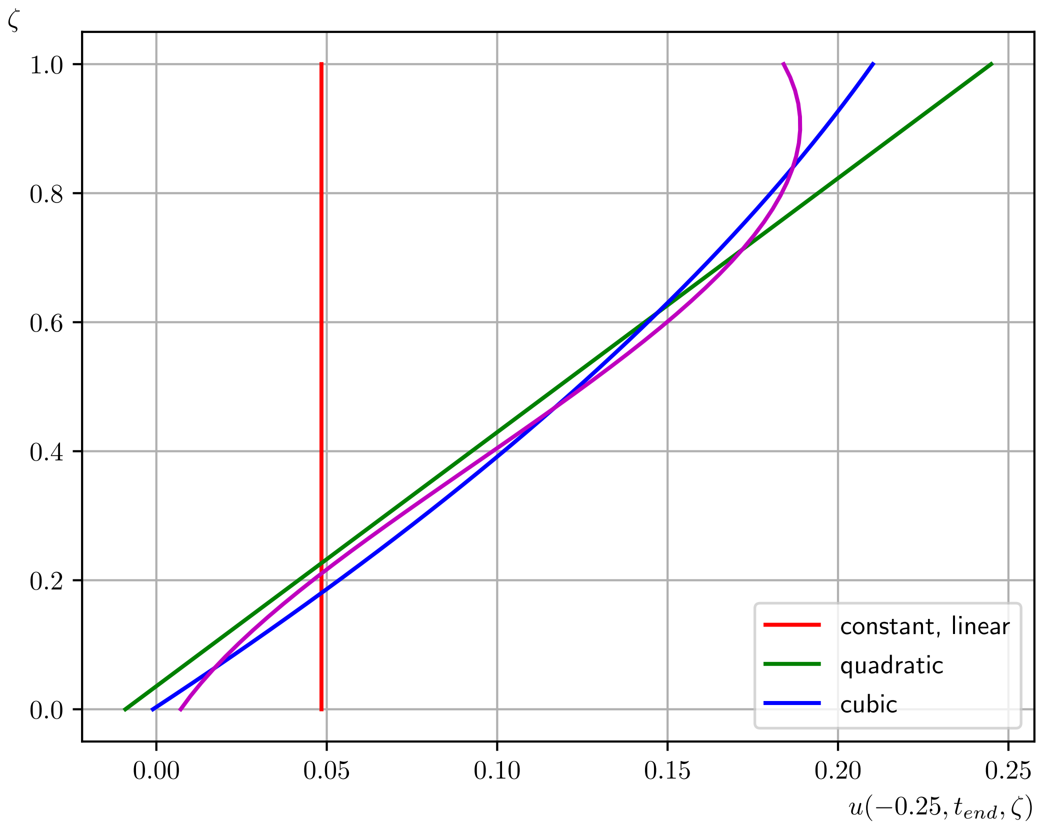
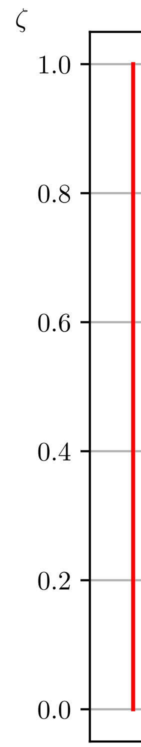

h
1.25
1.20
1.15
1.10
1.05
1.00
0.95
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
x
constant
linear
quadratic
cubic

Figure 5.1


u
0.25
0.20
0.15
0.10
0.05
0.00
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
x
constant
linear
quadratic
cubic





5.1.3 Manufactured Solution

I also ran a manufactured solution example to verify the order of convergence of the method.

$$q_i(x,t) = 0.1\sin(2\pi(x - t - 0.1i)) \tag{5.3}$$

5.2 Two Dimensional Cartesian Results

5.3 Two Dimensional Unstructured Results

This is still in progress and should be done before thesis is due for submission.

	1st Order		2nd Order		3rd Order		4th Order		5th Order	
n	error	order	error	order	error	order	error	order	error	order
20	0.226	—	8.57×10^{-3}	—	1.67×10^{-4}	—	3.17×10^{-6}	—	7.61×10^{-8}	0.00
40	0.117	0.96	2.17×10^{-3}	1.98	2.07×10^{-5}	3.02	1.98×10^{-7}	4.00	2.38×10^{-9}	5.00
80	0.058	1.00	5.40×10^{-4}	2.01	2.57×10^{-6}	3.01	1.24×10^{-8}	4.00	7.71×10^{-11}	4.95
160	0.028	1.06	1.35×10^{-4}	2.00	3.21×10^{-7}	3.00	7.76×10^{-10}	4.00	4.04×10^{-11}	0.93
320	0.014	0.99	3.37×10^{-5}	2.00	4.01×10^{-8}	3.00	4.85×10^{-11}	4.00	8.09×10^{-11}	-1.00

Table 5.1 Convergence Table for standard shallow water equations or shallow water moment equations with zero moments.

	1st Order		2nd Order		3rd Order		4th Order		5th Order	
n	error	order	error	order	error	order	error	order	error	order
20	2.53×10^{-1}	—	9.97×10^{-3}	—	1.71×10^{-3}	—	1.14×10^{-4}	—	2.68×10^{-5}	—
40	1.30×10^{-1}	0.96	2.52×10^{-3}	1.98	3.85×10^{-4}	2.15	1.74×10^{-5}	2.72	8.01×10^{-7}	5.06
80	6.47×10^{-2}	1.00	6.28×10^{-4}	2.00	6.13×10^{-5}	2.65	7.50×10^{-7}	4.53	1.53×10^{-8}	5.71
160	3.13×10^{-2}	1.05	1.57×10^{-4}	2.00	9.09×10^{-6}	2.75	1.25×10^{-7}	2.59	4.04×10^{-10}	5.25
320	1.58×10^{-2}	0.99	3.92×10^{-5}	2.00	1.73×10^{-6}	2.39	8.79×10^{-9}	3.83	8.40×10^{-11}	2.27

Table 5.2 Convergence table for shallow water moment equations with one moment.

	1st Order		2nd Order		3rd Order		4th Order		5th Order	
n	error	order	error	order	error	order	error	order	error	order
20	2.78×10^{-1}	—	1.14×10^{-2}	—	5.35×10^{-3}	—	3.69×10^{-4}	—	5.19×10^{-5}	—
40	1.42×10^{-1}	0.96	2.88×10^{-3}	1.98	6.47×10^{-4}	3.05	2.46×10^{-5}	3.91	1.12×10^{-6}	5.53
80	7.12×10^{-2}	1.00	7.19×10^{-4}	2.00	7.83×10^{-5}	3.04	1.40×10^{-6}	4.13	1.93×10^{-8}	5.86
160	3.45×10^{-2}	1.04	1.80×10^{-4}	2.00	1.27×10^{-5}	2.63	1.14×10^{-7}	3.62	5.86×10^{-10}	5.04
320	1.74×10^{-2}	0.99	4.49×10^{-5}	2.00	2.55×10^{-6}	2.32	1.09×10^{-8}	3.39	8.79×10^{-11}	2.74

Table 5.3 Convergence table for shallow water moments equations with two moments

	1st Order		2nd Order		3rd Order		4th Order		5th Order	
n	error	order	error	order	error	order	error	order	error	order
20	3.02×10^{-1}	—	1.30×10^{-2}	—	7.02×10^{-3}	—	3.17×10^{-4}	—	5.57×10^{-5}	—
40	1.56×10^{-1}	0.96	3.28×10^{-3}	1.99	6.99×10^{-4}	3.33	2.38×10^{-5}	3.73	1.10×10^{-6}	5.66
80	7.81×10^{-2}	0.99	8.19×10^{-4}	2.00	1.18×10^{-4}	2.56	2.51×10^{-6}	3.25	2.64×10^{-8}	5.38
160	3.80×10^{-2}	1.04	2.05×10^{-4}	2.00	2.55×10^{-5}	2.22	3.17×10^{-7}	2.99	1.37×10^{-9}	4.27
320	1.92×10^{-2}	0.99	5.12×10^{-5}	2.00	5.11×10^{-6}	2.32	4.68×10^{-8}	2.76	1.17×10^{-10}	3.55

Table 5.4 Convergence table for shallow water moment equations with three moments or a cubic velocity profile

	1st Order		2nd Order		3rd Order		4th Order		5th Order	
n	error	order	error	order	error	order	error	order	error	order
5	9.26×10^{-1}	—	9.42×10^{-2}	—	1.24×10^{-2}	—	2.25×10^{-3}	—	4.62×10^{-4}	—
10	5.02×10^{-1}	0.88	2.12×10^{-2}	2.16	2.77×10^{-3}	2.17	2.79×10^{-4}	3.02	4.12×10^{-5}	3.49
20	2.79×10^{-1}	0.85	5.09×10^{-3}	2.06	7.16×10^{-4}	1.95	3.43×10^{-5}	3.02	2.40×10^{-6}	4.10
40	1.43×10^{-1}	0.96	1.26×10^{-3}	2.02	1.60×10^{-4}	2.16	3.08×10^{-6}	3.48	1.23×10^{-7}	4.28
80	7.36×10^{-2}	0.96	3.13×10^{-4}	2.00	3.37×10^{-5}	2.25	2.76×10^{-7}	3.48	6.09×10^{-9}	4.34

Table 5.5 Convergence table for standard shallow water equations or shallow water moment equations with zero moments

	1st Order		2nd Order		3rd Order		4th Order		5th Order	
n	error	order	error	order	error	order	error	order	error	order
5	1.20	—	1.16×10^{-1}	—	2.07×10^{-2}	—	8.11×10^{-3}	—	1.52×10^{-3}	—
10	6.69×10^{-1}	0.84	2.46×10^{-2}	2.24	7.20×10^{-3}	1.53	8.54×10^{-4}	3.25	1.44×10^{-4}	3.40
20	3.81×10^{-1}	0.81	5.86×10^{-3}	2.07	2.10×10^{-3}	1.78	1.14×10^{-4}	2.90	9.01×10^{-6}	4.00
40	2.00×10^{-1}	0.93	1.45×10^{-3}	2.02	4.78×10^{-4}	2.13	1.18×10^{-5}	3.28	4.62×10^{-7}	4.28
80	1.03×10^{-1}	0.95	3.61×10^{-4}	2.00	1.01×10^{-4}	2.25	1.07×10^{-6}	3.47	2.14×10^{-8}	4.43

Table 5.6 Convergence table for shallow water moment equations with one moment.

	1st Order		2nd Order		3rd Order		4th Order		5th Order	
n	error	order	error	order	error	order	error	order	error	order
5	1.42	—	1.40×10^{-1}	—	2.77×10^{-2}	—	1.08×10^{-2}	—	2.24×10^{-3}	—
10	8.01×10^{-1}	0.83	2.95×10^{-2}	2.24	1.04×10^{-2}	1.41	1.22×10^{-3}	3.15	2.08×10^{-4}	3.43
20	4.62×10^{-1}	0.80	7.01×10^{-3}	2.07	2.98×10^{-3}	1.81	1.62×10^{-4}	2.91	1.22×10^{-5}	4.09
40	2.44×10^{-1}	0.92	1.73×10^{-3}	2.02	6.60×10^{-4}	2.18	1.57×10^{-5}	3.37	6.11×10^{-7}	4.32
80	1.26×10^{-1}	0.95	4.33×10^{-4}	2.00	1.37×10^{-4}	2.27	1.42×10^{-6}	3.47	2.86×10^{-8}	4.42

Table 5.7 Convergence table for shallow water moment equations with two moments

Chapter6. Conclusion

6.1 Summary

6.2 Future Work

BIBLIOGRAPHY

[1] Cockburn, B., Hou, S., and Shu, C.-W. The runge-kutta local projection discontinuous galerkin finite element method for conservation laws. iv. the multidimensional case. *Mathematics of Computation 54*, 190 (1990), 545–581.

[2] Cockburn, B., Lin, S.-Y., and Shu, C.-W. Tvb runge-kutta local projection discontinuous galerkin finite element method for conservation laws iii: one-dimensional systems. *Journal of Computational Physics 84*, 1 (1989), 90–113.

[3] Cockburn, B., and Shu, C.-W. Tvb runge-kutta local projection discontinuous galerkin finite element method for conservation laws. ii. general framework. *Mathematics of computation 52*, 186 (1989), 411–435.

[4] Cockburn, B., and Shu, C.-W. The runge-kutta local projection-discontinuous-galerkin finite element method for scalar conservation laws. *ESAIM: Mathematical Modelling and Numerical Analysis 25*, 3 (1991), 337–361.

[5] Cockburn, B., and Shu, C.-W. The runge–kutta discontinuous galerkin method for conservation laws v: multidimensional systems. *Journal of Computational Physics 141*, 2 (1998), 199–224.

[6] Dal Maso, G., Lefloch, P. G., and Murat, F. Definition and weak stability of nonconservative products. *Journal de mathématiques pures et appliquées 74*, 6 (1995), 483–548.

[7] Huang, Q., Koellermeier, J., and Yong, W.-A. Equilibrium stability analysis of hyperbolic shallow water moment equations. *arXiv preprint arXiv:2011.08571* (2020).

[8] Koellermeier, J., and Pimentel-Garcia, E. Steady states and well-balanced schemes for shallow water moment equations with topography. *arXiv preprint arXiv:2011.07667* (2020).

[9] Kowalski, J., and Torrilhon, M. Moment approximations and model cascades for shallow flow. *arXiv preprint arXiv:1801.00046* (2017).

[10] Reed, W. H., and Hill, T. Triangular mesh methods for the neutron transport equation. Tech. rep., Los Alamos Scientific Lab., N. Mex.(USA), 1973.

[11] Rhebergen, S., Bokhove, O., and van der Vegt, J. J. Discontinuous galerkin finite element methods for hyperbolic nonconservative partial differential equations. *Journal of Computational Physics 227*, 3 (2008), 1887–1922.