2003

Numerical solutions of hyperbolic conservation laws: incorporating multi-resolution viscosity methods into the finite element framework

Marcus Calhoun-Lopez

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Mathematics Commons

Recommended Citation

https://lib.dr.iastate.edu/rtd/1426

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Numerical solutions of hyperbolic conservation laws:
Incorporating multi-resolution viscosity methods into the finite element framework

by

Marcus Calhoun-Lopez

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:
Max Gunzburger, Major Professor
Scott Hansen
L. Steven Hou
Janet Peterson
Sunder Sethuraman

Iowa State University
Ames, Iowa
2003
Copyright © Marcus Calhoun-Lopez, 2003. All rights reserved.
Graduate College
Iowa State University

This is to certify that the Doctoral dissertation of
Marcus Calhoun-Lopez
has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Committee Member
Signature was redacted for privacy.
Committee Member
Signature was redacted for privacy.
Committee Member
Signature was redacted for privacy.
Committee Member
Signature was redacted for privacy.
Major Professor
Signature was redacted for privacy.
For the Major Program
DEDICATION

This work is dedicated to my family.

To my parents, Dr. Thomas R. Lopez and Carolyn C. Lopez, who have been my role models and guides. Their encouragement, love, and support have meant more to me than I can say.

To my grandparents, Dr. Paul B. Calhoun, Ruth S. Calhoun, Lucy C. Lopez, and Thomas Lopez, for their love and support.

To my siblings, Lt. Thomas Calhoun-Lopez and Elizabeth S. Gue, for their friendship and encouragement.

To the greatest nieces ever, Miranda R. Gue and Olivia R. Gue.
TABLE OF CONTENTS

LIST OF TABLES ............................................................ viii

LIST OF FIGURES ........................................................... x

ABSTRACT ................................................................. xiii

1. HYPERBOLIC CONSERVATION LAWS ................................. 1
   1.1 Shocks and characteristic curves .............................. 2
   1.2 Entropy solutions ............................................... 3
      1.2.1 Entropy condition ........................................ 4
      1.2.2 Viscous solutions ......................................... 4
   1.3 Riemann problems ............................................... 5

2. NUMERICAL SCHEMES .................................................. 7
   2.1 Finite difference schemes for one dimension ................. 8
      2.1.1 Lax-Friedrichs ............................................ 9
      2.1.2 Godunov .................................................. 9
      2.1.3 Flux-limiter methods ..................................... 10
      2.1.4 Slope-limiter methods ................................... 10
   2.2 Finite difference schemes for higher dimensions ............ 10
   2.3 Finite volume methods ......................................... 11
   2.4 Finite element methods ....................................... 12
      2.4.1 Discontinuous Galerkin methods ......................... 12
      2.4.2 Streamline diffusion .................................... 13
3. MODEL PROBLEMS .......................................................... 14

3.1 Burgers' equation ......................................................... 14
  3.1.1 Problem with steady state solution for long time ............... 14
  3.1.2 Periodic problem .................................................. 15

3.2 Euler's equations ...................................................... 16
  3.2.1 Shock tube problem ............................................... 17
  3.2.2 Supersonic flow over a wedge .................................. 19

4. BACKGROUND ............................................................. 21

4.1 Spectral viscosity methods ........................................... 21
  4.1.1 Overview of SV method .......................................... 21
  4.1.2 Post-processing of SV solution ................................ 22

4.2 Hierarchical finite element basis functions ......................... 23
  4.2.1 Hierarchical bases on polygonal domains ....................... 23
  4.2.2 Hierarchical bases on general domains ......................... 24
  4.2.3 Properties of hierarchical bases ................................ 25
  4.2.4 Implementation issues of hierarchical bases .................. 32

5. FINITE ELEMENT MULTI-RESOLUTION VISCOSITY METHOD ............ 34

5.1 Implementation of the method ........................................ 35
  5.2 Advantages of hierarchical bases over spectral bases ........... 36
    5.2.1 Adaptive diffusion ............................................ 36
    5.2.2 Edge detection ................................................ 36

5.3 Comparison of hierarchical and nodal bases ........................ 40

6. PROOF OF CONVERGENCE ............................................... 41

6.1 The Div-Curl lemma and Murat's lemma ............................ 41

6.2 Problem ............................................................... 42
  6.2.1 Notation ......................................................... 43
  6.2.2 FE Formulation ................................................ 43

6.3 Sufficient conditions for convergence ............................... 44
6.4 Properties of the FE system of ODEs ........................................ 44
6.5 Estimates of \( u_N \) .......................................................... 45
  6.5.1 Estimates from the FE formulation .................................... 45
  6.5.2 \( H^1(U_T) \) estimates of \( u_N \) for linear polynomials .......... 46
  6.5.3 \( H^1(U_T) \) estimates of \( u_N \) for higher degree polynomials 47
6.6 Strong Convergence of \( \{u_N\} \) ............................................ 48
  6.6.1 \( L^2(U_T) \) bound on \{div \( v_N\)\} .................................. 49
  6.6.2 \( H^{-1}(U_T) \) bound on \{div \( v_N\)\} .......................... 51
  6.6.3 \{div \( v_N\)\} lies in a compact subset of \( H^{-1}(U_T) \) .......... 52
  6.6.4 \{curl \( w_N\)\} lies in a compact subset of \( H^{-1}(U_T) \) .......... 52
  6.6.5 Subsequence of \( u_N \) converges strongly in \( L^2(U_T) \) .......... 53
6.7 Convergence to a solution .................................................. 55
6.8 Convergence to the entropy solution ..................................... 55
6.9 Free Parameters ............................................................ 59
  6.9.1 Satisfying the convergence criteria .................................. 60
  6.9.2 Inverse estimate of the low frequencies ............................ 60
  6.9.3 \( H^1 \) error estimates of the low frequencies ..................... 61
7. NUMERICAL RESULTS ...................................................... 62
  7.1 Free parameters .......................................................... 62
    7.1.1 Choosing \( \epsilon_N \) ................................................ 62
    7.1.2 Choosing \( m \) so that \( Q_{N;m,n} = 0 \) for \( n < m \) ........... 62
    7.1.3 Choosing the form of \( Q_N \) ...................................... 62
    7.1.4 Choosing the quadratic hierarchical structure .................... 63
    7.1.5 Ordinary differential equation solver ............................. 63
    7.1.6 Post-processing strategy .......................................... 64
  7.2 Convergence rates ....................................................... 64
  7.3 Boundary conditions for the wedge problem ......................... 65
  7.4 Numerical simulations .................................................. 65
LIST OF TABLES

Table 4.1  Comparison of stiffness matrix condition numbers in one dimension (polynomial degree=1) .............................. 28
Table 4.2  Comparison of stiffness matrix condition numbers in one dimension (polynomial degree=2) .............................. 28
Table 4.3  Comparison of mass matrix condition numbers in one dimension (polynomial degree=1) .............................. 29
Table 4.4  Comparison of mass matrix condition numbers in one dimension (polynomial degree=2) .............................. 29
Table 7.1  Convergence rate of steady state Burgers' equation with linear polynomials .................................................. 72
Table 7.2  Convergence rate of steady state Burgers' equation with quadratic linear low frequency polynomials .............................. 72
Table 7.3  Convergence rate of steady state Burgers' equation with quadratic polynomials .................................................. 72
Table 7.4  Convergence rate of periodic Burgers' equation with linear polynomials .................................................. 79
Table 7.5  Convergence rate of periodic Burgers' equation with quadratic linear low frequency polynomials .............................. 79
Table 7.6  Convergence rate of periodic Burgers' equation with quadratic polynomials .................................................. 79
Table 7.7  Convergence rate of shock tube problem with linear polynomials .................................................. 90
Table 7.8  Convergence rate of shock tube problem with quadratic linear low frequency polynomials .................................................. 91
Table 7.9 Convergence rate of shock tube problem with quadratic polynomials . 91
Table 7.10 Convergence rate of wedge problem with linear polynomials . . . . . 99
### LIST OF FIGURES

| Figure 3.1 | Solutions of the Periodic Burgers’ Equation | 15 |
| Figure 3.2 | Solution of the Shock Tube Problem | 18 |
| Figure 3.3 | Solution of supersonic flow over a wedge | 20 |
| Figure 4.1 | Hierarchical and nodal linear basis functions | 24 |
| Figure 4.2 | Quadratic hierarchical basis functions | 25 |
| Figure 4.3 | Condition numbers of standard matrices | 27 |
| Figure 5.1 | Hierarchical decomposition of a piecewise smooth function with a discontinuity | 37 |
| Figure 5.2 | Linear hierarchical coefficient from function values | 39 |
| Figure 5.3 | Quadratic hierarchical coefficient from function values | 39 |
| Figure 7.1 | Solution of steady state Burgers’ equation with linear polynomials (without post-processing) | 66 |
| Figure 7.2 | Solution of steady state Burgers’ equation with linear polynomials (with post-processing) | 67 |
| Figure 7.3 | Solution of steady state Burgers’ equation with quadratic linear low frequency polynomials (without post-processing) | 68 |
| Figure 7.4 | Solution of steady state Burgers’ equation with quadratic linear low frequency polynomials (with post-processing) | 69 |
| Figure 7.5 | Solution of steady state Burgers’ equation with quadratic polynomials (without post-processing) | 70 |
Figure 7.6  Solution of steady state Burgers' equation with quadratic polynomials (with post-processing) .............................. 71
Figure 7.7  Solution of periodic Burgers' equation with linear polynomials (without post-processing) ..................................... 73
Figure 7.8  Solution of periodic Burgers' equation with linear polynomials (with post-processing) ................................. 74
Figure 7.9  Solution of periodic Burgers' equation with quadratic linear low frequency polynomials (without post-processing) ................. 75
Figure 7.10 Solution of periodic Burgers' equation with quadratic linear low frequency polynomials (with post-processing) ......................... 76
Figure 7.11 Solution of periodic Burgers' equation with quadratic polynomials (without post-processing) ......................... 77
Figure 7.12 Solution of periodic Burgers' equation with quadratic polynomials (with post-processing) ........................................ 78
Figure 7.13 Solution of shock tube problem with linear polynomials ($\rho$) ......................................................... 80
Figure 7.14 Solution of shock tube problem with linear polynomials ($\rho u$) .......................................................... 81
Figure 7.15 Solution of shock tube problem with linear polynomials ($\rho e$) ....................................................... 82
Figure 7.16 Solution of shock tube problem with linear low frequency polynomials ($\rho$) .................................................. 83
Figure 7.17 Solution of shock tube problem with linear low frequency polynomials ($\rho u$) .................................................. 84
Figure 7.18 Solution of shock tube problem with linear low frequency polynomials ($\rho e$) .................................................. 85
Figure 7.19 Solution of shock tube problem with quadratic polynomials ($\rho$) .......................................................... 86
Figure 7.20 Solution of shock tube problem with quadratic polynomials ($\rho u$) .......................................................... 87
Figure 7.21 Solution of shock tube problem with quadratic polynomials ($\rho e$) .......................................................... 88
Figure 7.22 Solution of shock tube problem with quadratic polynomials and 3rd order ODE solver (unstable) ......................... 89
Figure 7.23  Poor approximation of wedge problem ($\rho$) because of slip boundary conditions ............................................. 92
Figure 7.24  Solution of wedge problem with linear polynomials ($\rho$) ............................... 93
Figure 7.25  Solution of wedge problem with linear polynomials ($|\rho u|$) ............................... 94
Figure 7.26  Solution of wedge problem with linear polynomials ($\frac{n}{|u|}$) ............................... 95
Figure 7.27  Solution of wedge problem with linear polynomials ($\rho e$) ............................... 96
Figure 7.28  Solution of wedge problem with linear polynomials ($\rho e$ along curves) ..................... 97
Figure 7.29  Grids for the wedge problem ................................................................. 98
Figure 8.1  Discontinuous linear hierarchical basis functions ........................................... 101
ABSTRACT

It is well known that the classic Galerkin finite-element method is unstable when applied to hyperbolic conservation laws, such as the Euler equations for compressible flow. Adding a diffusion term to the equations stabilizes the method but sacrifices too much accuracy to be of any practical use. An elegant solution devised by Eitan Tadmor for spectral methods is to add diffusion only to the high frequency modes of the solution, which stabilizes the method without the sacrifice of accuracy. We incorporate this idea into the finite-element framework by using hierarchical functions as a multi-frequency basis. The result is a new finite element method for solving hyperbolic conservation laws. For this method, we are able to prove convergence for a one-dimensional scalar conservation law. Numerical results are presented for one- and two-dimensional hyperbolic conservation laws.
1 HYPERBOLIC CONSERVATION LAWS

A general system of conservation laws has the form
\[
\begin{aligned}
\frac{\partial \mathbf{q}}{\partial t} + \sum_{j=1}^{d} \frac{\partial}{\partial x_j} f_j (\mathbf{q}) &= 0 \text{ in } \mathbb{R}^d \times (0, \infty) \\
\mathbf{q} (\cdot, 0) &= \mathbf{g} \text{ in } \mathbb{R}^d.
\end{aligned}
\] (1.1)

\(\mathbf{q} : \mathbb{R}^d \times [0, \infty) \rightarrow \mathbb{R}^p\) are the \(p\) conserved variables, \(f_j : \mathbb{R}^p \rightarrow \mathbb{R}^p\) are the \(d\) flux functions. Let
\(A_j : \mathbb{R}^p \rightarrow \mathbb{R}^{p \times p}\) be the derivative of \(f\), so for \(\mathbf{q} \in \mathbb{R}^p\), \(A_j (\mathbf{q})\) is the \(p \times p\) Jacobian matrix of \(f_j\), \(A_j (\mathbf{q}) = \frac{\partial f_j}{\partial \mathbf{q}} (\mathbf{q}) := \left(\frac{\partial (f_j)_{kl}}{\partial q_l} (\mathbf{q})\right)_{1 \leq k,l \leq p} \). Equation (1.1) is hyperbolic if for all solutions \(\mathbf{q}\), any linear combination of \(\{A_j (\mathbf{q})\}_{j=1}^{d}\) has real eigenvalues with eigenvectors that span \(\mathbb{R}^p\).

Equation (1.1) is in differential form. We will also have occasion to use the integral form of the conservation law: for any open bounded set \(U \subseteq \mathbb{R}^d\) with smooth boundary and any open time interval \((t_l, t_r)\),
\[
\begin{aligned}
\frac{1}{|U|} \int_U \mathbf{q} (\cdot, t_r) \, dx - \frac{1}{|U|} \int_U \mathbf{q} (\cdot, t_l) \, dx \\
+ \frac{t_r - t_l}{|U|} \sum_{j=1}^{d} \frac{1}{t_r - t_l} \int_{t_l}^{t_r} \int_{\partial U} f_j (\mathbf{q}) \, n_j \, dS \, dt = 0
\end{aligned}
\] (1.2)

\(n\) is the unit outward normal to the boundary, \(\partial U\), of \(U\). \(|U| = \int_U 1 \, dx\) is the measure of \(U\). By integrating Equation (1.1) over \(U \times (t_l, t_r)\), we can see that it is equivalent to Equation (1.2) for smooth solutions.
The preceding definitions are taken from [9, 7].

1.1 Shocks and characteristic curves

Equation (1.1) does not, in general, have a classical solution because of the spontaneous formation of discontinuities. To see why a discontinuity forms, even with smooth initial conditions, let us consider a scalar conservation law in one dimension:

\[
\begin{cases}
\frac{\partial q}{\partial t} + \frac{\partial}{\partial x} (f(q)) = 0 \text{ in } \mathbb{R} \times (0, \infty) \\
q(\cdot, 0) = g \text{ in } \mathbb{R}.
\end{cases}
\] (1.3)

Let \( a(q) := f'(q) \). Equation (1.3) in non-conservative form is

\[
\begin{cases}
\frac{\partial q}{\partial t} + a(q) \frac{\partial q}{\partial x} = 0 \text{ in } \mathbb{R} \times (0, \infty) \\
q(\cdot, 0) = g \text{ in } \mathbb{R}.
\end{cases}
\] (1.4)

The following argument is taken from [9]. Let us assume \( g \) is smooth. We will also assume that the solution \( q \) is smooth. The characteristic curves, \( \{x_\xi(t)\}_{\xi \in \mathbb{R}} \) of Equation (1.4) are determined by the ordinary differential equation:

\[
\begin{cases}
\frac{dx_\xi}{dt} = a(q(x_\xi(t), t)) \\
x_\xi(0) = \xi.
\end{cases}
\] (1.5)

Along \( x_\xi(t) \), \( q \) is constant:

\[
\frac{d}{dt} q(x_\xi(t), t) = \frac{\partial q}{\partial t} (x_\xi(t), t) + \frac{\partial q}{\partial x} (x_\xi(t), t) \frac{dx_\xi}{dt} (t)
\]

\[
= \frac{\partial q}{\partial t} (x_\xi(t), t) + \frac{\partial q}{\partial x} (x_\xi(t), t) a(q(x_\xi(t), t))
\]

\[
= 0.
\]

Thus, along \( x_\xi(t) \), we can calculate \( q \) from \( q(\cdot, 0) = g \):

\[
q(x_\xi(t), t) = q(x_\xi(0), 0)
\]

\[
= q(\xi, 0)
\]

\[
= g(\xi).
\]
Equation (1.5) becomes

\[
\begin{align*}
\frac{dx}{dt} &= a(g(\xi)) \\
\xi(0) &= \xi,
\end{align*}
\]

which has the solution

\[\xi = \xi + ta(g(\xi)).\]

Therefore if \( q \) is a smooth solution of Equation (1.3), then we can calculate its value from \( g \) by following the characteristic curves \( \{x_{\xi}\}_{\xi \in \mathbb{R}} \).

Problems arise if the characteristics cross. Let us assume that \( a \circ g \) is decreasing on some interval. Then there exists \( \xi_l < \xi_r \) such that \( a(g(\xi_r)) < a(g(\xi_l)) \), and so \( x_{\xi_l} \) and \( x_{\xi_r} \) will intersect at some finite time. So unless \( a \circ g \) is highly specialized, there cannot exist a smooth solution to Equation (1.3) for all time. When the characteristic curves first intersect, \( x_{\xi_1} = x_{\xi_2} \) for \( \xi_1 \neq \xi_2 \), we should not even expect \( q \) to be continuous since \( g(\xi_1) \neq g(\xi_2) \).

1.2 Entropy solutions

As we showed in the previous section, Equation (1.1) will not, in general, have a classical solution for all time. Instead, we must look for a solution \( q \in L^\infty(\mathbb{R}^d \times (0, \infty); \mathbb{R}^p) \) which satisfies Equation (1.1) in the distributional sense:

\[
\int_0^\infty \int_{\mathbb{R}^d} \left[ q \cdot \frac{\partial \phi}{\partial \ell} + \sum_{j=1}^d f_j(q) \cdot \frac{\partial \phi}{\partial x_j} \right] dx \, dt + \int_{\mathbb{R}^d} g \cdot \phi(\cdot, 0) \, dx = 0 \quad (1.6)
\]

for all test functions \( \phi \in C^\infty_0(\mathbb{R}^d \times [0, \infty); \mathbb{R}^p) \). It is clear that for a smooth enough solution, Equations (1.1), (1.2), and (1.6) are all equivalent.

Further, we have seen that after a shock forms, there is an ambiguity as to which characteristic curve to follow to get to the correct solution. Extra conditions must be imposed to get the unique physically relevant solution. The laws of thermodynamics tell us that the entropy of the system should not decrease. Satisfying this requirement is enough to get us the unique physically relevant entropy solution. An equivalent way is to seek a solution which is the limit of solutions of a more general class of equations.
1.2.1 Entropy condition

Let $\Phi, \{\Psi_j\}_{j=1}^d : \mathbb{R}^p \to \mathbb{R}$ be smooth functions. $\Phi$ is an entropy with $\{\Psi_j\}_{j=1}^d$ entropy fluxes of Equation (1.1) if

\[
\begin{cases}
\Phi \text{ is convex and} \\
\nabla_q \Phi^T \frac{\partial \Psi_j}{\partial q} = \nabla_q \Psi_j \text{ in } \mathbb{R}^p \text{ for } 1 \leq j \leq d.
\end{cases}
\]

(1.7)

A trivial calculation gives us that if $q$ is a smooth solution to Equation (1.1), then $\Phi(q)$ satisfies a scalar conservation law with flux $\Psi(q)$:

\[
\frac{\partial}{\partial t} \Phi(q) + \sum_{j=1}^d \frac{\partial}{\partial x_j} \Psi_j(q) = 0 \text{ in } \mathbb{R}^d \times (0, \infty).
\]

(1.8)

In some instances, $\Phi$ can be interpreted as the negative of the physical entropy, so Equation (1.8) says for a smooth solution $u$, then $\Phi \circ q$ satisfies a conservation law with flux functions $\{\Psi_j \circ q\}_{j=1}^d$.

For solutions with shocks, we impose the condition on $q$ that the physical entropy is non-decreasing:

\[
\frac{\partial}{\partial t} \Phi(q) + \sum_{j=1}^d \frac{\partial}{\partial x_j} \Psi_j(q) \leq 0
\]

(1.9)

for every entropy function $\Phi$ with entropy fluxes $\{\Psi\}_{j=1}^d$. Equation (1.9) is an inequality in the distributional sense

\[
\int_0^\infty \int_{\mathbb{R}^d} \left[ \Phi(q) \frac{\partial \phi}{\partial t} + \sum_{j=1}^d \Psi_j(q) \frac{\partial \phi}{\partial x_j} \right] dx \, dt \geq 0
\]

(1.10)

for all $\phi \in C_0^\infty(\mathbb{R}^d \times (0, \infty))$ with $\phi \geq 0$.

1.2.2 Viscous solutions

In deriving Equation (1.1) it was assumed that the flux functions had no dependence on the gradient of $q$, thus ignoring viscous effects. For the class of phenomena which are modeled with conservation laws, viscous effects are small, but they are present and play a role when sharp gradients (such as shocks) start forming. An alternate way of characterizing the unique physically relevant solution of Equation (1.1) is to have $q = \lim_{\varepsilon \to 0} q^{\varepsilon}$ a.e., where
\( q^\varepsilon : \mathbb{R}^d \times [0, \infty) \to \mathbb{R}^p \) is the solution of the equation

\[
\begin{cases}
\frac{\partial q^\varepsilon}{\partial t} + \sum_{j=1}^{d} \frac{\partial}{\partial x_j} f_j(q^\varepsilon) = \varepsilon \Delta q^\varepsilon \text{ in } \mathbb{R}^d \times (0, \infty) \\
q^\varepsilon(\cdot, 0) = g \text{ in } \mathbb{R}^d.
\end{cases}
\]  

(1.11)

In other words, \( q \) is the limit of the viscous solutions as the viscosity goes to zero.

### 1.3 Riemann problems

A particularly important class of problems is the Riemann problem:

\[
\begin{cases}
\frac{\partial q}{\partial t} + \frac{\partial}{\partial x} f(q) = 0 \text{ in } \mathbb{R} \times (0, \infty) \\
q(x, 0) = \begin{cases}
q_L, & x < 0 \\
q_R, & x > 0,
\end{cases}
\end{cases}
\]  

(1.12)

where \( q_L \) and \( q_R \) are constant states. Several numerical schemes are based on having a solution of Equation (1.12).

Let us assume that Equation (1.12) is strictly hyperbolic. Let \( \{\lambda_k\}_{k=1}^p \) be the distinct eigenvalues of the Jacobian of \( f \) with associated right eigenvalues \( \{r_k\}_{k=1}^p \). If \( |q_L - q_R| \) is sufficiently small, then there will exist a solution consisting of at most \( p + 1 \) constant regions.

Each boundary is associated with an eigenpair, \( (\lambda_k, r_k) \), which is either genuinely nonlinear, \( \nabla \lambda_k(q) \cdot r_k(q) \neq 0 \forall q \), or linearly degenerate, \( \nabla \lambda_k(q) \cdot r_k(q) = 0 \). Two constant regions, with values \( q_L \) and \( q_R \), will be separated by one of the following:

- Shock Discontinuity: \( k \)-th characteristic is genuinely nonlinear and \( \lambda_k(q_L) > \lambda_k(q_R) \).

\[
q \left( \frac{x}{t} \right) = \begin{cases}
q_L, & \frac{x}{t} < s_k \\
q_R, & \frac{x}{t} > s_k.
\end{cases}
\]

\( s_k \) is the speed of the shock, and \( \lambda_k(q_L) > s_k > \lambda_k(q_R) \).
Rarefaction Wave: $k$-th characteristic is genuinely nonlinear and $\lambda_k(\bar{q}_L) < \lambda_k(\bar{q}_R)$.

$$q\left(\frac{x}{t}\right) = \begin{cases} \bar{q}_L, & \frac{x}{t} \leq \lambda_k(\bar{q}_L) \\ w\left(\frac{x}{t}\right), & \lambda_k(\bar{q}_L) < \frac{x}{t} < \lambda_k(\bar{q}_R) \\ \bar{q}_R, & \frac{x}{t} \geq \lambda_k(\bar{q}_L), \end{cases}$$

where $w$ is smooth, $w(\lambda_k(\bar{q}_L)) = \bar{q}_L$, and $w(\lambda_k(\bar{q}_R)) = \bar{q}_R$.

Contact Discontinuity: $k$-th characteristic is linearly degenerate.

$$q\left(\frac{x}{t}\right) = \begin{cases} \bar{q}_L, & \frac{x}{t} < s_k \\ \bar{q}_R, & \frac{x}{t} > s_k. \end{cases}$$

$s_k$ is the speed of the discontinuity, and $\lambda_k(\bar{q}_L) = s_k = \lambda_k(\bar{q}_R)$.

The above assertions are taken from [9, 16, 26].
2 NUMERICAL SCHEMES

We will briefly consider some of the families of numerical schemes one can use to solve hyperbolic conservation laws. In particular, we consider finite difference (FD) schemes, finite volume (FV) methods, and finite element (FE) schemes. Of particular interest in the FE category are the discontinuous Galerkin (DG) methods. The distinctions between the various methods are not sharp. Some FV and FE schemes can be written into an equivalent FD formulation. As noted in [5], the FV method can be considered to be a special type of DG method. Nevertheless, we shall list some of the general features of the various methods.

The FD method is the oldest of the numerical methods, and so many variations have been developed. Many successful strategies for solving hyperbolic conservation laws were originally developed in the FD framework then adapted to other methods. As noted in [10], however, FD schemes tend to have difficulties with complex geometries, satisfying prescribed boundary conditions, and rigorous analysis. In fluid dynamics, complex geometries are common, and, as shown in [5], poor approximation of boundary conditions can severely affect a numerical method.

The FV method inherently captures many of the important aspects of conservation laws. FV methods are locally conservative. Information is propagated along the characteristic curves, at least approximately. FV methods use unstructured grids and so can handle complex geometries. High order schemes, however, are difficult to attain.

FE methods are well suited to handle complex geometries and prescribed boundary conditions. Formally high order schemes can be attained by increasing the degree of the approximating polynomial. The price paid is a large increase in the number of unknowns.

The spectral viscosity (SV) method will be considered in Section 4.1. Incorporating the
ideas from the SV method into the FE framework is the subject of this thesis.

2.1 Finite difference schemes for one dimension

For ease of notation, we will assume a uniform grid. The space and time dimensions are divided into grids of size $\Delta x$ and $\Delta t$ respectively. A classic interpretation of the FD method would be to consider $q_j^n$ as an approximation of $q$ at the point $(x_j, t_n)$ := $(j \Delta x, n \Delta t)$. Instead, we will consider $\bar{q}_j^n$ as an approximation of the cell average of $q$ at time $t = t_n$:

$$\bar{q}_j^n \approx \frac{1}{\Delta x} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} q(x, t_n) \, dx .$$

This allows us to use the integral form of the conservation law, Equation (1.2), which is for $d = 1, U = (x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}})$, and $(t_l, t_r) = (t_n, t_{n+1})$:

$$
\begin{aligned}
\frac{1}{\Delta x} \int_U q(\cdot, t_{n+1}) \, dx &= \frac{1}{\Delta x} \int_U q(\cdot, t_n) \, dx \\
&- \frac{\Delta t}{\Delta x} \left[ \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} f_j(\bar{q}(x_{j-\frac{1}{2}}, t)) \, dt - \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} f_j(q(x_{j+\frac{1}{2}}, t)) \, dt \right] \\
&\frac{1}{\Delta x} \int_U q(\cdot, 0) \, dx = \frac{1}{\Delta x} \int_U g \, dx.
\end{aligned}
$$

Equation (2.1) is approximated by

$$\frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} f(q(x_{j+\frac{1}{2}}, t)) \, dt \approx \tilde{F}_{j+\frac{1}{2}}^n,$$

where $\tilde{F}_{j+\frac{1}{2}}^n$ are determined by $\tilde{F}$:

$$\tilde{F}_{j-\frac{1}{2}}^n = \tilde{F}(\bar{q}_{j-k}, \ldots, \bar{q}_{j+k-1})$$

and

$$\tilde{F}_{j+\frac{1}{2}}^n = \tilde{F}(\bar{q}_{j-k+1}, \ldots, \bar{q}_{j+k}) .$$

The conservative FD scheme is now

$$
\bar{q}_{j+1}^n = \bar{q}_j^n - \frac{\Delta t}{\Delta x} \left[ \tilde{F}_{j+\frac{1}{2}}^n - \tilde{F}_{j-\frac{1}{2}}^n \right].
$$

To make the FD scheme consistent, for all $q \in \mathbb{R}^p$, $\tilde{F}(q, \ldots, q) = f(q)$. 


2.1.1 Lax-Friedrichs

The Lax-Friedrichs scheme is a 3-point FD method with numerical flux

\[ F(q_1, q_2) = \frac{f(q_1) + f(q_2)}{2} - \frac{1}{2} \frac{\Delta x}{\Delta t} (q_2 - q_1). \]

The first term in \( \mathbf{F} \) is simply an average of the left and right values of the flux. As noted in [17], the second term is a diffusion term with coefficient \( \frac{\Delta x^2}{2 \Delta t} \) since

\[ \frac{q_2 - q_1}{\Delta x} \approx \frac{\partial q}{\partial x}. \]

So if \( \frac{\Delta x}{\Delta t} \) is kept constant, then we are adding a diffusion term with coefficient of the same order as the spatial grid size. In other words, we are approximating the viscous solutions of Equation (1.11) with \( \varepsilon = O(\Delta x) \).

2.1.2 Godunov

The Lax-Friedrichs scheme shows no directional bias, yet we know that information travels in the direction of the characteristic curves. The Godunov scheme takes advantage of this by solving a Riemann problem to advance the solution to the next time step.

Let us assume we have the values of the cell averages \( \{q_j^n\} \) at time \( t_n \). We define a function \( q_\Delta(x, t) := \sum_j q_j^n \chi_j \), where \( \chi_j \) is the characteristic function on \( (x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}) \). We can now consider \( q_\Delta \) as a series of Riemann problems with initial discontinuities at the center of the cells. As long as \( \Delta t \) is not too large, these Riemann problems will not interact, and we know how to advance \( q_\Delta \) from \( t_n \) to \( t_{n+1} \). We then take cell averages of \( q_\Delta \) to get \( \{q_j^{n+1}\} \).

By taking advantage of the knowledge we have about the solutions of Riemann problems, Godunov's method can be greatly simplified (see [16, 9]). Let \( q_{rs}(\frac{t}{\tau}; q_L, q_R) \) be the solution of the Riemann problem (1.12) with left and right initial data \( q_L \) and \( q_R \) respectively. We can define our numerical flux as

\[ \mathbf{\mathbf{\bar{F}}} (q_1, q_2) = f(q_{rs}(0; q_1, q_2)). \]
2.1.3 Flux-limiter methods

For ease of notation, we will assume a scalar conservation law in this section. Multidimensional problems can be handled in the same manner for each component. Following [16], let us consider two flux functions $\bar{F}_L$ and $\bar{F}_H$. $\bar{F}_L$ is low order but does not produce spurious oscillations, while $\bar{F}_H$ is formally high order but does produce spurious oscillations. Let us construct a flux function

$$ F = \bar{F}_H - (1 - \Phi) (\bar{F}_H - \bar{F}_L), $$

where $\Phi$ depends on the local behavior of the solution. $\Phi$ is the flux limiter. It should be near one in the smooth regions and near zero close to a discontinuity. $(1 - \Phi) (\bar{F}_H - \bar{F}_L)$ acts like a variable diffusion term.

2.1.4 Slope-limiter methods

At each time step, Godunov's method projects the solution into a space of piecewise constant functions. This limits the convergence rate to first order accuracy. A second order scheme can be achieved by projecting the solution into the space of piecewise linear functions at each time step. Since this is no longer a series of Riemann problems, we can not expect to solve it exactly, but the solution can be advanced to the next time step by approximate solvers similar to the approximate Riemann solvers. By limiting the slopes of the piecewise linear solutions from time step to time step, it can be ensured that the total variation of the solution does not increase and so no spurious oscillations form.

2.2 Finite difference schemes for higher dimensions

As noted in [26], higher dimensional solutions can be approximated by either piecing together one dimensional solutions or using "genuinely multidimensional" schemes.

Dimension splitting methods use solutions to one dimensional problems to construct solutions to a multi-dimensional problem. Following [16], let $H^{1}_{\Delta t}(\dot{q}^n)$ represent the one dimensional numerical method which advances $q^n$ to $q^{n+1}$ in the spatial variable $x_j$. The simplest
splitting method, which is limited to first order accuracy, is

\[ q^{n+1} = H_{\Delta t}^1 \cdots H_{\Delta t}^1 q^n. \]

Other splittings are possible. For instance, in two dimensions, the accuracy of the splitting can be increased by using the Strang splitting

\[ q^{n+1} = H_{\Delta t}^2 H_{\Delta t}^2 H_{\Delta t}^2 q^n. \]

See [26] and the references therein for a listing of various genuinely multidimensional schemes.

2.3 Finite volume methods

For ease of notation, we will follow [9] and define the matrix \( F : \mathbb{R}^p \rightarrow \mathbb{M}^{d \times p} \), where \( F = (f_1, \ldots, f_k)^T \). Equation (1.2) becomes

\[
\begin{align*}
\frac{1}{|U|} \int_U q(\cdot, t_r) \, dx - \frac{1}{|U|} \int_U q(\cdot, t_l) \, dx + \frac{t_r - t_l}{|U|} \int_U \int_{\partial U} F(q) \cdot n \, dS \, dt &= 0 \\
\frac{1}{|U|} \int_U q(\cdot, 0) \, dx &= \frac{1}{|U|} \int_U g \, dx.
\end{align*}
\]

Like the FD methods, the FV method is based on the integral formulation of the conservation law (2.2) and approximates averages. The computational domain is partitioned into control volumes \( \{U_i\} \) and their centers. The control volumes could be a triangulation of the computation domain or a dual mesh in which the centers are the nodes of the triangulation. Let \( \Gamma_{ij} \) be the boundary between \( U_i \) and \( U_j \), and \( n_{ij} \) be the outward unit normal from \( U_i \) to \( U_j \) (\( n_{ij} = -n_{ji} \)). We approximate the average over the control volume

\[ q_{ji}^n \approx \frac{1}{|U_j|} \int_{U_j} q(x, t_n) \, dx \]

between \( U_j \) and \( U_i \). We also define a function, \( \Phi : \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}^d \rightarrow \mathbb{R}^p \), to approximate the average of the flux

\[ \Phi(q_{ij}, q_{ji}, n_{ij}) \approx \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \frac{1}{|\Gamma_{ij}|} \int_{\Gamma_{ij}} F(q) \cdot n \, dS \, dt. \]
Φ is required to be locally Lipschitz, satisfy the conservation requirement

\[ Φ(q_{ij}, q_{ji}, n_{ij}) = Φ(q_{ji}, q_{ij}, n_{ji}), \]

and satisfy the consistency requirement

\[ Φ(q, q, n) = F(q) \cdot n. \]

The FV approximation of Equation (2.2) is

\[ q_i^{n+1} = q_i^n - \frac{\Delta t}{|U_i|} \sum_{\Gamma_{ij} \subseteq \partial U_i} |\Gamma_{ij}| Φ(q_{ij}, q_{ji}, n_{ij}). \]

For a first order method, \( Φ(q_{ij}, q_{ji}, n_{ij}) \) can be defined in terms of a solution of a one dimensional Riemann problem between the centers of \( U_i \) and \( U_j \). The solution can either be exact (as in Godunov's scheme) or some approximate Riemann solver such as Roe's scheme.

For a second order method, \( Φ(q_{ij}, q_{ji}, n_{ij}) \) can be calculated as a one dimensional flux-limiting problem between the centers of \( U_i \) and \( U_j \).

2.4 Finite element methods

2.4.1 Discontinuous Galerkin methods

Most of this section is taken from [5].

DG methods take many of the best features from the FE framework. Formally high order methods are obtained by simply increasing the degree of the approximating polynomial. Complex geometries are easily handled. Boundary conditions are easily satisfied.

Owing to the fact that no continuity restrictions are placed on the approximating solution, DG methods have further advantages. They are easy to parallelize. Adaptive strategies are relatively easy to implement. The mass matrix is block diagonal so implicit time stepping schemes are possible.

The lack of continuity of solutions, however, is the cause of the biggest drawback of DG methods. The number of unknowns is drastically increased.
Let $\mathcal{T}_h$ be a partition of the computational domain. For all $K \in T^h$, let $P^k(K)$ be the space of polynomial functions of order $k$ on $K$. We can define an approximation of Equation (1.1), $q^h|_{K} \in P^k(K)$, for each $K \in T^h$ as the solution of: for $1 \leq j \leq p$,

$$
\int_{K} \left[ \frac{\partial}{\partial t} (q^h) \right] v \, dx + \sum_{i=1}^{d} \int_{K} \left[ f_i (q^h) \right] \frac{\partial v}{\partial x_i} \, dx + \int_{\partial K} \hat{h}_j v \, dS = 0 \quad \forall \, v \in P^k(K). \quad (2.3)
$$

$\hat{h}_j \approx \sum_{i=1}^{d} \left[ f_i (q^h) \right] (\hat{n})_i$ on $\partial K$ is described as an approximate Riemann solver. To retain the advantages of DG methods, great care must be taken in discretizing Equation (2.3) in time. A Runge-Kutta type ODE solver and slope limiter have been developed to solve Equation (2.3) for multi-dimensional systems.

### 2.4.2 Streamline diffusion

The shock capturing streamline diffusion method adds a diffusion term to the conservation law, but unlike Equation (1.11), diffusion is added in different amounts in the direction of the characteristic curves (streamline diffusion) and its normal direction (crosswind diffusion). Streamline diffusion is added everywhere, while crosswind diffusion is added only near discontinuities. To determine characteristic curves, space-time elements must be used, which increases the number of unknowns and results in an implicit time scheme. See [12, 13, 22].
3 MODEL PROBLEMS

In this chapter, we present the model problems we will use to test our FE scheme. We also present their analytic solutions.

3.1 Burgers’ equation

The quintessential scalar equation in one dimension is Burgers’ equation, in which \( f(q) = \frac{q^2}{2} \). We take the domain to be \( U = (-1, +1) \). Equation (1.1) becomes

\[
\frac{\partial q}{\partial t} + \frac{\partial}{\partial x} \left( \frac{q^2}{2} \right) = 0.
\]

3.1.1 Problem with steady state solution for long time

Let us first consider Burgers’ equation with Dirichlet boundary conditions:

\[
\begin{align*}
\frac{\partial q}{\partial t} + \frac{\partial}{\partial x} \left( \frac{q^2}{2} \right) &= 0 \text{ on } (-1, +1) \times (0, \infty) \\
q(\pm1, t) &= \mp1 \text{ for all } t \in (0, \infty) \\
q(x, 0) &= -\cos \left[ \frac{\pi (x - 1)}{2} \right].
\end{align*}
\]

As \( t \to \infty \),

\[
q(x, t) \to \begin{cases} 
+1, & x < 0 \\
-1, & x > 0.
\end{cases}
\]
3.1.2 Periodic problem

Let us now consider the periodic Burgers' equation:

\[
\begin{aligned}
\frac{\partial q}{\partial t} + \frac{\partial}{\partial x} \left( \frac{q^2}{2} \right) &= 0 \text{ on } (-1, +1) \times (0, T) \\
q (-1, t) &= q (+1, t) \text{ for all } t \in (0, T) \\
q (x, 0) &= 1 + \frac{1}{2} \sin (\pi x).
\end{aligned}
\] (3.1)

See Figure 3.1 for solutions at \( t = 0.25 \) and \( t = 1 \).

It is shown in [15, 7] that the entropy solution will be

\[ q (x, t) = \frac{x - y}{t}, \]

where \( y = y (x, t) \) minimizes

\[ G (x, y, t) = \frac{t}{2} \left( \frac{x - y}{t} \right)^2 + \int q (y, 0) \, dy. \]
If $x$ is not the point of discontinuity, then $G$ will have only one global minimum, although it can have several local minima.

The simplest strategy of solving for $q$ is, for each $x$, to find a point $y_0$ near the global minimum of $G(x, y, t)$, then use any nonlinear solver to find the solution, $y$, of

$$-\frac{1}{t} (x - y) + q(y, 0) = 0,$$

then set $q(x, t) = \frac{x - y_0}{t}$.

### 3.2 Euler's equations

An important example of a system of conservation laws is Euler's equations

$$\begin{align*}
\frac{\partial \rho}{\partial t} + \sum_{j=1}^{d} \frac{\partial}{\partial x_j} (\rho u_j) &= 0 \\
\frac{\partial}{\partial t} (\rho u_i) + \sum_{j=1}^{d} \frac{\partial}{\partial x_j} (\rho u_i u_j) + \frac{\partial P}{\partial x_i} &= 0 \quad \text{for } 1 \leq i \leq d \\
\frac{\partial}{\partial t} (\rho e) + \sum_{j=1}^{d} \frac{\partial}{\partial x_j} (\rho e u_j + P u_j) &= 0.
\end{align*}$$

$p, u \in \mathbb{R}^d$, and $e$ are the density, velocity, and the specific (per unit mass) total energy respectively of the fluid. $P$ is the pressure. $\rho u$ is the momentum of the fluid, and $\rho e$ is the total energy density. Euler's equations are a statement of conservation of mass, momentum, and energy.

Euler's equations consist of three equations with four unknowns. An equation of state is required to relate $P$, $\rho$, and the specific internal energy, $\varepsilon = e - \frac{|u|^2}{2}$. For a polytropic ideal gas, $\rho e = \frac{P}{\gamma - 1}$, where $\gamma$ is the ratio of specific heats. The value of $\gamma$ for a diatomic gas, such as air (nitrogen and oxygen), is $\gamma = 1.4$. Our equation of state becomes

$$P = (\gamma - 1) \left( \rho e - \rho \frac{|u|^2}{2} \right).$$

Two other quantities will be used in our study of Euler's equations. The speed of sound in the fluid, $a$, and the Mach number. With the assumptions we have already made about the nature of the fluid, $a = \sqrt{\frac{\gamma P}{\rho}}$. The Mach number is defined as $\frac{|u|}{a}$.

The above definitions are taken from [16, 9, 1].
3.2.1 Shock tube problem

Let us consider a tube, closed at both ends, which is divided into two sections. On the left is a high pressure region. On the right is a low pressure region. At time $t = 0$, the divider is removed, and the gases in the two regions are allowed to mix.

If the variables are constant in the two regions, then this is a Riemann problem, and so at most four distinct regions should form. A boundary between two regions is either an expansion wave, a contact discontinuity, or a shock discontinuity. In this particular case, all three boundaries are represented.

Figure 3.2 is the solution of the shock tube problem at time $t = 0.287$ with initial and boundary conditions:

$$(\rho, u, P)(x, 0) = \begin{cases} (1, 0, 1), & x < 0 \\ (0.125, 0, 0.1), & x > 0 \end{cases}$$

$$(\rho, u, P)(-1, t) = (1, 0, 1)$$

$$(\rho, u, P)(+1, t) = (0.125, 0, 0.1).$$

Let us label the four regions as in Figure 3.2. Let $(\rho_i, u_i, P_i)$ be the variables in the $i$-th region. $(\rho_4, u_4, P_4)$ and $(\rho_1, u_1, P_1)$ are known from the initial data. In [1], it is shown how the initial data can be used to calculate $(\rho, u, P)$ in the entire domain. The contact discontinuity separates the two gases and moves with a velocity $u_p$. The shock moves to the right with speed $W$. $\frac{P_4}{P_1}$ can be calculated from the equation

$$\frac{P_4}{P_1} = \frac{P_2}{P_1} \left\{ 1 - \frac{(\gamma_4 - 1) \left( \frac{u_1}{u_4} \right) \left( \frac{P_4}{P_1} - 1 \right)}{\sqrt{2\gamma_1 \left[ 2\gamma_1 + (\gamma_1 + 1) \left( \frac{P_4}{P_1} - 1 \right) \right]}} \right\} \frac{1}{\gamma_4 - 1}. $$
We can now easily calculate the following:

\[ u_3 = u_p \]

\[ u_2 = u_p \]

\[ u_p = \frac{a_1}{\gamma_1} \left( \frac{P_2}{P_1} - 1 \right) \left( \frac{2\gamma_1}{\gamma_1+1} \right)^{\frac{1}{2}} \]

\[ P_3 = P_2 \]

\[ P_2 = \frac{P_1 P_2}{P_1} \]

\[ \rho_3 = \rho_4 \left( \frac{P_3}{P_4} \right)^{\frac{1}{4}} \]

\[ \rho_2 = \rho_1 \frac{1 + \left( \frac{\gamma_1+1}{\gamma_1-1} \right) \left( \frac{P_2}{P_1} \right)}{\frac{\gamma_1+1}{\gamma_1-1} + \frac{P_2}{P_1}} \]

\[ W = a_1 \sqrt{\left( \frac{\gamma_1+1}{2\gamma_1} \right) \left( \frac{P_2}{P_1} - 1 \right) + 1}. \]
The left and right ends of the expansion wave travel with velocities $-a_4$ and $u_3-a_3$ respectively. Inside the expansion wave,

$$u = \left( \frac{2}{\gamma_4 + 1} \right) \left( a_4 + \frac{x}{t} \right)$$
$$\rho = \rho_4 \left[ 1 - \left( \frac{\gamma_4 - 1}{2} \right) \left( \frac{u}{a_4} \right) \right]^{\frac{2}{\gamma_4 - 1}}$$
$$P = P_4 \left[ 1 - \left( \frac{\gamma_4 - 1}{2} \right) \left( \frac{u}{a_4} \right) \right]^{\frac{2\gamma_4}{\gamma_4 - 1}}.$$

### 3.2.2 Supersonic flow over a wedge

Let us consider a uniform horizontal flow traveling with Mach number $M_1 > 1$ which hits a wedge at an angle $\theta$ with the flow. As long as $\theta$ is not too large, a shock at an angle $\beta$ will form which is attached to the wedge. The flow will be uniform on either side of the shock.

Figure 3.3 is the steady state solution to the wedge problem with $\theta = 15^\circ$. On the left boundary, we have the boundary conditions

$$\rho = 1.4$$
$$M = 2.5$$
$$P = 1.$$

$|u|$ can be calculated from $M$, and $u$ is in the $x$ direction. On the top and bottom boundary, we have the boundary conditions

$$u \cdot \hat{n} = 0,$$

where $\hat{n}$ is the outward unit normal. No boundary conditions are prescribed on the right boundary.

Let $(\rho_i, u_i, P_i)$ be the constant values on either side of the shock, $i = 1, 2$ on the left and right of the shock respectively. In [1], it is shown that $(\rho_1, u_1, P_1)$ and $\theta$ determines $(\rho_2, u_2, P_2)$ and $\beta$. $\beta$ can be determined by the $\theta-\beta-M$ relation

$$\tan \theta = 2 (\cot \beta) \left\{ \frac{M_1^2 \sin^2 \beta - 1}{M_1^2 \left[ \gamma + \cos (2\beta) \right] + 2} \right\}.$$
If \( \theta \) is too large, then the \( \theta-\beta-M \) does not have any solutions. Physically, this corresponds to the shock forming in front of the wedge and becoming a detached shock. For a given \( \theta \), there can be two solutions of the \( \theta-\beta-M \) relation. If there is no additional downstream pressure, then the smaller value of \( \beta \) is the physically correct one.

\((\rho_2, u_2, P_2)\) are determined by \( \gamma \) and the normal component, \( M_n \), of the Mach number relative to the shock.

\[
\begin{align*}
M_{n_1} &= M_1 \sin \beta \\
\rho_2 &= \rho_1 \frac{(\gamma + 1) M_{n_1}^2}{(\gamma - 1) M_{n_1}^2 + 2} \\
P_2 &= P_1 \left[ 1 + \left( \frac{2\gamma}{\gamma + 1} \right) \left( M_{n_1}^2 - 1 \right) \right] \\
M_{n_2} &= \sqrt{\frac{M_{n_1}^2 + \frac{2}{\gamma - 1}}{\frac{2\gamma}{\gamma - 2} M_{n_2}^2 - 1}} \\
M_2 &= \frac{M_{n_2}}{\sin (\beta - \theta)}.
\end{align*}
\]

The direction of \( \mathbf{u} \) is parallel to the wedge angle.
4 BACKGROUND

In this chapter, we present the tools needed for our FE formulation.

4.1 Spectral viscosity methods

In [23], the SV method was introduced as a solution to Equation (3.1) using Fourier spectral basis functions. The theory was further refined and extended in a series of papers [20, 19, 4, 25, 24, 8, 11]. Of particular importance are [11, 19], which use Legendre polynomials. The variational formulation of the Legendre SV method is closest to our FE formulation.

4.1.1 Overview of SV method

We present the most basic SV method, which uses Fourier spectral basis functions. Using standard notation from Fourier Spectral Methods, we define:

\[ u_N = P_N u(x, t), \]
\[ P_N u = \sum_{|k| \leq N} \hat{u}_k(t) e^{ik\pi x}, \quad \text{and} \]
\[ \hat{u}_k(t) = \frac{1}{2} \int_{-1}^{+1} u(x, t) e^{-ik\pi x} dx. \]

We seek \( u_N \) such that:

\[ \frac{\partial u_N}{\partial t} + \frac{\partial}{\partial x} \left( P_N u_N^2 \right) = \frac{\partial}{\partial x} \left( Q_N \frac{\partial u_N}{\partial x} \right). \]
$Q_N$ is the spectral viscosity operator defined as a convolution with the viscosity kernel $Q_N(x)$, so

$$
Q_N \frac{\partial u_N}{\partial x} = Q_N(x) * \frac{\partial u_N(x,t)}{\partial x} \quad \text{and} \quad (4.6)
$$

$$
Q_N(x) = \sum_{|k| \leq N} \tilde{Q}_k \tilde{e}^{ik\pi x}, \text{ with} \quad (4.7)
$$

$$
0 \leq \tilde{Q}_k \leq 1 \text{ and} \quad (4.8)
$$

$$
\tilde{Q}_k = 0 \text{ for small } |k|. \quad (4.9)
$$

It is easy to see the effect of $Q_N$ if we write the diffusion term in the Fourier space:

$$
\epsilon \frac{\partial}{\partial x} \left( Q_N \frac{\partial u_N}{\partial x} \right) = -\epsilon \sum_{|k| \leq N} \left( k^2 \frac{\pi^2}{2} \tilde{Q}_k \tilde{u}_k e^{ik\pi x} \right). \quad (4.10)
$$

Since $0 \leq \tilde{Q}_k < 1$ for all but large $|k|$, $Q_N$ dampens or eliminates the low frequency modes of $u_N$ in the diffusion term. We can now see that the SV diffusion term is a compromise between not adding diffusion, which is unstable, and adding full diffusion, which limits the convergence rate and smears the solution.

Ideally, we would like to add diffusion only in the area around a shock. The global nature of the basis functions, however, does not allow for an adaptive viscosity kernel.

### 4.1.2 Post-processing of SV solution

The SV solution, $u_N$, does not converge to the exact solution, $u$, at the optimal rate because of the poor convergence of $P_N u$. $P_N u$ is limited to first order convergence in the smooth regions and has $O(1)$ Gibbs oscillations near a discontinuity. Post processing $u_N$ recovers spectral convergence.

The post-processing scheme can be enhanced by knowing the locations of discontinuities, as in [8]. Because of the global nature of basis functions, this edge detection is not a trivial task.
4.2 Hierarchical finite element basis functions

The usual (nodal) basis functions in the FE method all have the same frequency. In order to have multi-frequency basis functions, we use hierarchical basis functions. In the elliptic partial differential equation setting, an early analysis of hierarchical basis functions, especially for one dimension, can be found in [30]. For two dimensions, see [28]. A good overview of hierarchical basis functions can be found in [29].

4.2.1 Hierarchical bases on polygonal domains

Let us first consider a polygonal domain, $U$. Let $T_0$ be a coarse grid approximation of $U$. The $n$-th level triangulation, $T_n$, is obtained by subdividing the elements of $T_{n-1}$.

Let $S^N$ be the space of continuous functions which are polynomials of degree $p$ on the elements of $T_N$. Let $N_N \subseteq U$ be the nodes of the elements of $T_N$. The nodal basis functions of $S^N$, $\{\phi_i\}_i$, are defined as

$$\phi_i \in S^N \text{ such that } \phi_i(x_j) = \delta_{ij} \forall x_j \in N_N.$$

It is well known that $S^N = \text{span} \{\phi_i\}_i$. The use of the nodal basis leads to many nice numerical properties, such as sparse matrices and local assembly of matrices. We cannot use the nodal basis for our purposes because, as we noted earlier, the elements of $\{\phi_i\}_i$ all have the same frequencies.

Let $B^n$ be the nodal basis of $S^n$. Let us define the hierarchical basis functions, $\{\psi_{n,i}\}_{n,i}$, as

$$\psi_{n,i} \in B^n \text{ such that } \psi_{n,i}(x_j) = 0 \forall x_j \in N_{n-1}.$$

For $0 \leq n \leq N$, $\{\psi_{n,i}\}_{n,i} \subseteq S^N$ is a linearly independent set with the same dimension as $S^N$, so $S^N = \text{span} \{\psi_{n,i}\}_{n,i}$. See Figure 4.1 for a comparison of the two bases for linear elements in one dimension. As we can see from the picture, $\psi_{n,i}$ is a low frequency function for small $n$ and a high frequency function for large $n$.

The strategy outlined above works for polynomials of any degree. The first column of Figure 4.2 consists of quadratic hierarchical basis functions. An alternate strategy is, for
Linear Hierarchical Basis Functions
level 0: 2 functions
level 1: 1 function
level 2: 2 functions
level 3: 4 functions

Standard Linear Basis Functions
9 functions

n < N, to use linear hierarchical basis functions, as in the second column of Figure 4.2.

In order to determine $T_{n+1}$ from $T_n$, we must decide, for a given $T \in T_n$, how many sub-elements to divide $T$ into. For linear and quadratic elements, the natural choice is $2^d$ sub-elements, where the domain is in $\mathbb{R}^d$. For cubic elements, the natural choice is $3^d$ since the vertices of an element will then be a subset of the vertices of its parent. We will limit our attention to linear and quadratic basis functions.

4.2.2 Hierarchical bases on general domains

For domains with curved boundaries, the situation is more complicated. Let $U$ be our potentially complicated domain. One strategy is to use the hierarchical decomposition of some polygonal domain $U'$ such that $U \subseteq U'$, as in [14]. Another strategy is to try and impose a hierarchical structure on an unstructured mesh, as in [3].
4.2.3 Properties of hierarchical bases

Let $x^D$ and $x^H$ be the nodal and hierarchical coefficients respectively of $u_N \in S^N$.

A FE discretization will result in a system of ordinary differential equations. After choosing an ODE solver to discretize time, we are left with a system of equations, possibly non-linear. Let $J^D$ and $J^H$ be the nodal and hierarchical Jacobian matrices respectively of the discretization. Let $R^D$ and $R^H$ be the residuals of the nonlinear solver. For simple problems, $J$ could be the mass or stiffness matrix. Let us first consider residuals of the form

\[ [R^D]_i = \int_U g_i(u_N) L_i(\phi_i) \, dx \]  
\[ [R^H]_i = \int_U g_i(u_N) L_i(\psi_i) \, dx \]  
\[ (4.11) \quad (4.12) \]
and Jacobian matrices of the form

\[ [J^D]_{i,j} = \int_U f_{i,j} (u_N) L_i (\phi_i) \tilde{L}_j (\phi_j) \, dx, \]  
\[ (4.13) \]

\[ [J^H]_{i,j} = \int_U f_{i,j} (u_N) L_i (\psi_i) \tilde{L}_j (\psi_j) \, dx \]  
\[ (4.14) \]

for some linear operators (such as partial derivatives) \( \{L_i\} \) and \( \{\tilde{L}_j\} \) and some functions \( \{g_i\} \) and \( \{f_{i,j}\} \). Due to the small support of all nodal basis functions, \( J^D \) is much more sparse than \( J^H \). It is also much easier to assemble \( J^D \) than \( J^H \). It is therefore best to work in the nodal basis as much as possible, then translate to the hierarchical basis only when needed.

Let \( S \) be the matrix such that

\[ x^D = S x^H. \]

A simple but tedious calculation gives

\[ J^H = S^T J^D S \]
\[ R^H = S^T R^D. \]

The nonlinear solver requires the solution of

\[ (J^H)^{-1} R^H = (S^T J^D S)^{-1} S^T R^D \]
\[ = (J^D S)^{-1} R^D. \]

So in an iterative linear solver, we need to calculate \( J^D S x \) and perhaps \( (J^D S)^T x \) for some vector \( x \). Compared to \( J^D \), \( S \) is not sparse, so we do not want to actually construct \( S \). Making the linear solver efficient requires being able to calculate \( Sx \) and \( S^T x \) quickly. In Section 4.2.4, we will present algorithms for calculating \( Sx \), \( S^T x \), and \( S^{-1} x \).

When the Jacobian matrix is in the form of Equations (4.13)–(4.14), the matrix \( S \) can be removed entirely from the linear solver by noticing that \( (J^H)^{-1} R^H = S^{-1} (J^D)^{-1} R^D \). As we will see in Section 5.1, this simplification is not possible in our FE formulation.

### 4.2.3.1 Condition numbers of standard matrices in hierarchical bases

Let \( h_N \) be maximum mesh size of the finest triangulation \( T_N \). For a matrix \( J \), let \( \kappa (J) \) be the condition number. Let \( r \) be defined by \( \kappa = O (h_N^{-r}) \). Let \( M \) and \( K \) be the mass
and stiffness matrices respectively in the standard basis. It is well known $\kappa(M) = O(1)$ and $\kappa(K) = O\left(h_N^{-2}\right)$. In the hierarchical basis, the condition number of the stiffness matrix, $K^H$, gets better, while the condition number of the mass matrix, $M^H$, is worse.

Let us first consider some numerical experiments. We calculate the condition numbers of $J$, $J S$, and $S^T J S$ for $J = K, M$ and for first degree polynomials. The results are in Figure 4.3 and Tables 4.1-4.4. The numerical evidence indicates that $\kappa\left(M^H\right) = \kappa\left(K^H\right) = O\left(h_N^{-1}\right)$.

We can prove that in one dimension, we can not do better than $\kappa\left(M^H\right) = \kappa\left(K^H\right) = O\left(h_N^{-1}\right)$ for linear and quadratic basis functions. Let $H^1(U)$ and $H^1_0(U)$ be the standard notations for Sobolev spaces with inner product, norm, and semi-norm $(\cdot, \cdot)_1$, $\|\cdot\|_1$, and $|\cdot|_1$. 
respectively. Let \( p \in \{1, 2\} \) be the degree of the polynomial. For

\[
v_N = \sum_{i=0}^{p+1} \beta_{0,i} \psi_{0,i} + \sum_{k=1}^{N} \sum_{i=0}^{2^{k+p-2}-1} \beta_{k,i} \psi_{k,i} \in S_N(U),
\]

let us define the norm

\[
\|v_N\|^2 = \sum_{i=0}^{p+1} |\beta_{0,i}|^2 + \sum_{k=1}^{N} \sum_{i=0}^{2^{k+p-2}-1} |\beta_{k,i}|^2.
\]

For a lower bound of the condition number of the mass matrix, we will need to determine how small and how large \( \|v_N\|^2 \) can be. For \( p = 1 \),

\[
\|\psi_{1,0}\|^2 = 1
\]
\[
\|\psi_{1,0}\|_0^2 = \frac{|U|}{3}.
\]
Table 4.3 Comparison of mass matrix condition numbers in one dimension (polynomial degree=1)

<table>
<thead>
<tr>
<th>N</th>
<th>M</th>
<th>MS</th>
<th>(S^T MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\kappa)</td>
<td>(r)</td>
<td>(\kappa)</td>
</tr>
<tr>
<td>3</td>
<td>3.9e+00</td>
<td>-</td>
<td>1.1e+01</td>
</tr>
<tr>
<td>4</td>
<td>3.9e+00</td>
<td>0.0</td>
<td>1.6e+01</td>
</tr>
<tr>
<td>5</td>
<td>4.0e+00</td>
<td>0.0</td>
<td>2.4e+01</td>
</tr>
<tr>
<td>6</td>
<td>4.0e+00</td>
<td>0.0</td>
<td>3.6e+01</td>
</tr>
<tr>
<td>7</td>
<td>4.0e+00</td>
<td>0.0</td>
<td>5.4e+01</td>
</tr>
<tr>
<td>8</td>
<td>4.0e+00</td>
<td>0.0</td>
<td>7.9e+01</td>
</tr>
<tr>
<td>9</td>
<td>4.0e+00</td>
<td>0.0</td>
<td>1.2e+02</td>
</tr>
</tbody>
</table>

Table 4.4 Comparison of mass matrix condition numbers in one dimension (polynomial degree=2)

<table>
<thead>
<tr>
<th>N</th>
<th>M</th>
<th>MS</th>
<th>(S^T MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\kappa)</td>
<td>(r)</td>
<td>(\kappa)</td>
</tr>
<tr>
<td>3</td>
<td>5.4e+00</td>
<td>-</td>
<td>1.7e+01</td>
</tr>
<tr>
<td>4</td>
<td>5.4e+00</td>
<td>0.0</td>
<td>2.5e+01</td>
</tr>
<tr>
<td>5</td>
<td>5.4e+00</td>
<td>0.0</td>
<td>3.7e+01</td>
</tr>
<tr>
<td>6</td>
<td>5.4e+00</td>
<td>0.0</td>
<td>5.5e+01</td>
</tr>
<tr>
<td>7</td>
<td>5.4e+00</td>
<td>0.0</td>
<td>8.1e+01</td>
</tr>
<tr>
<td>8</td>
<td>5.4e+00</td>
<td>0.0</td>
<td>1.2e+02</td>
</tr>
<tr>
<td>9</td>
<td>5.4e+00</td>
<td>0.0</td>
<td>1.7e+02</td>
</tr>
</tbody>
</table>

For \(p = 2\),

\[
\|\psi_{0,1}\|_0^2 = 1 \\
\|\psi_{0,1}\|_0^2 = C |U| ,
\]

where \(C\) can be \(\frac{8}{15}\) or \(\frac{1}{3}\) depending on which quadratic hierarchical representation we use.

\[
\|\psi_{N,0}\|_0^2 = 1 , \\
\|\psi_{N,0}\|_0^2 = \begin{cases} 
\frac{|U|}{3} 2^{-N+1} & p = 1 \\
\frac{16}{15} \frac{|U|}{2} 2^{-N} & p = 2 .
\end{cases}
\]
Therefore for all $N$, there exist $v_N, u_N \in S^N(U) \cap H^1_0(U)$ such that

$$\frac{\|v_N\|_0^2}{\|v_N\|^2} = C_1$$

$$\frac{\|u_N\|_0^2}{\|u_N\|^2} = C_2 |U| 2^{-N}.$$

For a lower bound of the condition number of the stiffness matrix, we will need to determine how small and how large $\frac{|\psi|}{\|\psi\|}$ can be. For $p = 1$,

$$\|\psi_{1,0}\|^2 = 1$$

$$|\psi_{1,0}|_1^2 = \frac{4}{|U|}.$$

For $p = 2$,

$$\|\psi_{0,1}\|^2 = 1$$

$$|\psi_{0,1}|_1^2 = \frac{C}{|U|},$$

where $C$ can be $\frac{16}{3}$ or 4 depending on which quadratic hierarchical representation we use. We then have

$$\|\psi_{N,0}\|^2 = 1,$$

$$|\psi_{N,0}|_1^2 = \begin{cases} \frac{4}{|U|} 2^{N-1} & p = 1 \\ \frac{16}{3} \frac{1}{|U|} 2^N & p = 2 \end{cases}$$

Therefore for all $N$, there exist $v_N, u_N \in S^N(U) \cap H^1_0(U)$ such that

$$\frac{|v_N|_1^2}{\|v_N\|^2} = C_3$$

$$\frac{|u_N|_1^2}{\|u_N\|^2} = \frac{C_4}{|U|} 2^N.$$

From now on, $C$ will denote a positive constant, not necessarily the same from line to line.

Let $\lambda_{N,\text{min}}$ and $\lambda_{N,\text{max}}$ be the minimum and maximum eigenvalues, respectively, for a matrix.

**Proposition 1.** In the hierarchical basis, the condition number of the mass matrix grows at least exponentially with the number of levels.
Proof.

\[
\lambda_{N,\text{min}}(M^H) = \inf_x \frac{x^T M^H x}{x^T x} \quad \text{(Rayleigh quotient)}
\]
\[
= \inf_{v_N \in \mathcal{S}_N} \frac{\|v_N\|_0^2}{\|v_N\|^2} \leq C 2^{-N}.
\]

\[
\lambda_{N,\text{max}}(M^H) = \sup_x \frac{x^T M^H x}{x^T x} \quad \text{(Rayleigh quotient)}
\]
\[
= \sup_{v_N \in \mathcal{S}_N} \frac{\|v_N\|_0^2}{\|v_N\|^2} \geq C.
\]

\[
\kappa_N(M^H) = \frac{\lambda_{N,\text{max}}(M^H)}{\lambda_{N,\text{min}}(M^H)} \geq C 2^N
\]

\[\square\]

Proposition 2. In the hierarchical basis, the condition number of the stiffness matrix grows at least exponentially with the number of levels.

Proof.

\[
\lambda_{N,\text{min}}(K^H) = \inf_x \frac{x^T K^H x}{x^T x} \quad \text{(Rayleigh quotient)}
\]
\[
= \inf_{v_N \in \mathcal{S}_N} \frac{|v_N|_1^2}{\|v_N\|^2} \leq C.
\]

\[
\lambda_{N,\text{max}}(K^H) = \sup_x \frac{x^T K^H x}{x^T x} \quad \text{(Rayleigh quotient)}
\]
\[
= \sup_{v_N \in \mathcal{S}_N} \frac{|v_N|_1^2}{\|v_N\|^2} \geq C 2^N.
\]

\[
\kappa_N(K^H) = \frac{\lambda_{N,\text{max}}(K^H)}{\lambda_{N,\text{min}}(K^H)} \geq C 2^N
\]

\[\square\]
4.2.4 Implementation issues of hierarchical bases

Let us now present algorithms for calculating $Sx$, $S^T x$, and $S^{-1} x$. All the notation is the same as in previous sections. $i$ and $j$ represent global node numbers.

Let $X$ be the vector representing some $u_N \in S^N$ in the hierarchical basis. $S X$ represents $u_N$ in the nodal basis. To calculate $S X$, we evaluate $u_N$ at the nodes. $X$ is overwritten by $S X$.

\begin{itemize}
\item for level $n = 0$ to $N - 1$
\item for element $T \in T_n$
\item for $T_c \in T_{n+1}$ and $T_c \subseteq T$
\item for $x_i \in (N_{n+1} \setminus N_n) \cap T_c$
\item for nodal basis function $\phi_j \in B^n$ with support in $T$
\item $X(i) = X(i) + X(j) \phi_j(x_i)$
\item end for $\phi_j$
\item end for $x_i$
\item end for $T_c$
\item end for $T$
\item end for $n$
\end{itemize}

Now let $X$ be the vector representing some $u_N \in S^N$ in the standard basis. $S^{-1} X$ represents $u_N$ in the hierarchical basis. To calculate $S^{-1} X$, at each level, we must use the coarser mesh to subtract away the global behavior of $u_N$. $X$ is overwritten by $S^{-1} X$.

\begin{itemize}
\item for level $n = N - 1$ to $0$
\item for element $T \in T_n$
\item for $T_c \in T_{n+1}$ and $T_c \subseteq T$
\item for $x_i \in (N_{n+1} \setminus N_n) \cap T_c$
\item for nodal basis function $\phi_j \in B^n$ with support in $T$
\item $X(i) = X(i) - X(j) \phi_j(x_i)$
\item end for $\phi_j$
\item end for $x_i$
\item end for $T_c$
\item end for $T$
\item end for $n$
\end{itemize}
end for \( x_i \)

end for \( T_c \)

end for \( T \)

end for \( n \)

Unlike \( S \) and \( S^{-1} \), \( S^T \) does not have a clear geometric interpretation. As noted in [28], \( S \) can be represented as \( S = S_0 S_1 \cdots S_{N-1} S_N \), where \( S_n \) represents the matrix multiplication in the outer loop of the \( S \times X \) algorithm. This gives us \( S^T = S_N^T S_{N-1}^T \cdots S_1^T S_0^T \).

for level \( n = N - 1 \) to 0

for element \( T \in T_n \)

for \( T_c \in T_{n+1} \) and \( T_c \subseteq T \)

for \( x_i \in (N_{n+1} \setminus N_n) \cap T_c \)

for nodal basis function \( \phi_j \in B^n \) with support in \( T \)

\[
X(j) = X(j) + X(i) \phi_j(x_i)
\]

end for \( \phi_j \)

end for \( x_i \)

end for \( T_c \)

end for \( T \)

end for \( n \)
5 FINITE ELEMENT MULTI-RESOLUTION VISCOSITY METHOD

Let us assume we have a hierarchical sequence of partitions \( \{ T_n \}_{n=0}^N \) of \( U \subseteq \mathbb{R}^d \), which is an open bounded set. Let \( S^N \) be the space of continuous vector valued functions whose components are in \( S^N \). We seek an approximate solution to the hyperbolic conservation law

\[
\begin{cases}
\frac{\partial q}{\partial t} + \sum_{j=1}^d \frac{\partial}{\partial x_j} f_j(q) = 0 \text{ in } U \times (0, \infty) \\
q(\cdot, 0) = g \text{ in } U
\end{cases}
\] (5.1)

with appropriate boundary conditions. The notation is the same as in Chapters 1 and 4.

Our FE solution of Equation (5.1) is seek \( q^N \in S^N \) such that

\[
\begin{align*}
\frac{d}{dt} \int_U q^N \cdot \nu \, dx + \int_U \sum_{j=1}^d \frac{\partial}{\partial x_j} f_j(q^N) \cdot \nu \, dx + \epsilon_N \sum_{i=1}^p \sum_{j,k=1}^d \int_U \frac{\partial}{\partial x_j} \left( Q_N^{i,k} q_i^N \right) \frac{\partial v_i}{\partial x_k} \\
&- \epsilon_N \sum_{i=1}^p \sum_{j,k=1}^d \int_{\partial U} \frac{\partial}{\partial x_j} \left( Q_N^{i,k} q_i^N \right) v_i \hat{n}_k \, ds = 0 \ \forall \ \nu \in S^N. 
\end{align*}
\] (5.2)

\( \hat{n} \) is the unit normal to the boundary, \( \partial U \), of \( U \). As in the SV method, \( Q_N^{i,k} \) dampens or eliminates the low frequency modes of a function:

\[
Q_N^{i,k} : S^N \rightarrow S^N
\]

\[
\sum_{n=0}^N \sum_{\iota} \beta_{n,\iota} \psi_{n,\iota} \rightarrow \sum_{n=0}^N \sum_{\iota} Q_{N;n,\iota}^{i,k} \beta_{n,\iota} \psi_{n,\iota}, \text{ where}
\] (5.4)

\[
0 \leq Q_{N;n,\iota}^{i,k} \leq 1 \text{ and}
\]

\[
Q_{N;n,\iota}^{i,k} = \begin{cases} 
0 & \text{for small } n \ (n \text{ near } 0) \\
1 & \text{for large } n \ (n \text{ near } N).
\end{cases}
\]
To account for the boundary conditions of Equation (5.1), a subspace of $S_N$ might be used (for essential boundary conditions) or the boundary integral in Equation (5.2) might not be over all of $\partial U$ (for natural boundary conditions).

Equation (5.2) is a weak formulation of

$$\frac{\partial \mathbf{q}}{\partial t} + \sum_{j=1}^{d} \frac{\partial}{\partial x_j} f_j (\mathbf{q}) - \epsilon_N \sum_{j,k=1}^{d} \frac{\partial^2}{\partial x_j \partial x_k} (Q_{N}^{j,k} \mathbf{q}) = 0, \quad \text{(5.5)}$$

where $[Q_{N}^{j,k} q]_i = Q_{N}^{j,k} q_i$ for $1 \leq i \leq p$. The dependence of $Q_{N}^{j,k}$ on $j$ and $k$ allows for the possibility of directional bias in the diffusion, which could reduce crosswind diffusion. As in the streamline diffusion method, this would probably require the use of entropy variables. We can simplify our formulation by using an isotropic diffusion term, $Q_N$, such that

$$Q_{N,i,j}^{j,k} = Q_{N,n} \delta_{j,k}. \quad \text{(5.6)}$$

Equations (5.2) and (5.5) reduce to

$$\frac{d}{dt} \int_U \mathbf{q} \cdot \mathbf{v} \, dx + \int_U \sum_{j=1}^{d} \frac{\partial}{\partial x_j} f_j (\mathbf{q}) \cdot \mathbf{v} \, dx + \epsilon_N \int_U \nabla (Q_N \mathbf{q}) : \nabla \mathbf{v} \, dx$$

$$- \epsilon_N \int_{\partial U} \frac{\partial}{\partial n} (Q_N \mathbf{q}) \cdot \mathbf{v} \, ds = 0 \quad \forall \mathbf{v} \in S^N \quad \text{and} \quad \text{(5.7)}$$

$$\frac{\partial \mathbf{q}}{\partial t} + \sum_{j=1}^{d} \frac{\partial}{\partial x_j} f_j (\mathbf{q}) - \epsilon_N \Delta (Q_N \mathbf{q}) = 0. \quad \text{(5.8)}$$

## 5.1 Implementation of the method

Once we choose a time discretization technique, we have a nonlinear system of equations. The Jacobian will have the form

$$J^H = J^H + K^H Q,$$

where $J^H$ is the Jacobian of the time dependent and flux terms, and $Q$ is a diagonal matrix whose nonzero elements are $\{Q_{N,n,i}^{j,k}\}$. For ease of notation, we ignore the boundary term.
Translating to the nodal basis, we get

\[ J^H = S^T \left( J^D S + K^H S Q \right), \]

\[ R^H = S^T R^D, \]

\[ (J^H)^{-1} R^H = \left( J^D S + K^H S Q \right)^{-1} R^D. \]

Therefore our iterative linear solver requires us to calculate, for a given vector \( x \),

\[ \left( J^D S + K^H S Q \right) x \text{ and} \]

\[ \left( J^D S + K^H S Q \right)^T x. \]

### 5.2 Advantages of hierarchical bases over spectral bases

In the SV method, there is only one function, with global support, at a given frequency. In our FE formulation, there are many functions, with local support, at a given frequency. The hierarchical basis offers two advantages. Diffusion can be added locally, and edge detection is trivial.

#### 5.2.1 Adaptive diffusion

For large values of \( n \) the hierarchical basis function \( \psi_{n,t} \) has local support, so \( Q_{N,n,t}^{1,k} \) only has a local effect. We can therefore add more diffusion near a discontinuity and less diffusion in the smooth regions. This should improve the accuracy of the method. As we are about to see, the size of \( |\beta_{n,t}| \) can be used to determine if the support of \( \psi_{n,t} \) resides in a smooth region or is near a discontinuity.

#### 5.2.2 Edge detection

Using the hierarchical basis, edge detection is a trivial task. Near a discontinuity the high frequency hierarchical coefficients are order one. In a smooth region, they shrink exponentially. Figure 5.1 is a hierarchical decomposition of a piecewise smooth function with a discontinuity. We can easily determine the location of discontinuities by looking at the magnitude of the high frequency coefficients.
5.2.2.1 Edge detection for piecewise linear polynomials

Let us prove the behavior of the hierarchical coefficients for linear basis functions. As we can see in Figure 5.2, $\beta_{k+1,j}$ can be calculated from the value of the function $u$ at $x_{k+1,j}$ and the node points of the parent cell. For a uniform partition, the cell size at level $k$ is $\Delta x_k = \frac{|k|}{2} 2^{-k+1}$. So $x_{k+1,j} - x_{k,i} = x_{k,i+1} - x_{k+1,j} = \Delta x_{k+1}$.

$$\beta_{k+1,j} = \frac{u(x_{k+1,j}) - u(x_{k,i+1}) + u(x_{k,i})}{2} = \frac{u(x_{k+1,j}) - u(x_{k,i})}{2} - \frac{u(x_{k,i+1}) - u(x_{k+1,j})}{2}.$$ (5.9)

Let $T_{k,i} = (x_{k,i}, x_{k,i+1})$.

$u$ is discontinuous: Let us assume that $u$ has a discontinuity in $T_{k,i}$. At least one of the two terms in Equation (5.9) will have a relatively large value, and so $|\beta_{k+1,j}|$ will be the
same order as the jump of \( u \) independent of \( k \).

\( u \) is continuously differentiable: Let us now assume that \( u \in C^1(T_{k,i}) \). Equation (5.9) can be written as

\[
\beta_{k+1,j} = \frac{|U|}{4} 2^{-k} \left( \frac{u(x_{k+1,j}) - u(x_{k,i})}{x_{k+1,j} - x_{k,i}} - \frac{u(x_{k,i+1}) - u(x_{k+1,j})}{x_{k,i+1} - x_{k+1,j}} \right)
\]

By the mean value theorem,

\[
\beta_{k+1,j} = \frac{|U|}{4} 2^{-k} [u'(\bar{x}_1) - u'(\bar{x}_2)] ,
\]

for some \( \bar{x}_1, \bar{x}_2 \in T_{k,i} \). Therefore \( \beta_{k+1,j} \) is of order \( 2^{-k} \):

\[
|\beta_{k+1,j}| \leq |U| \|u'\|_{L^\infty(T_{k,i})} 2^{-k-1}.
\]

\( u \) is twice continuously differentiable: Finally, let us assume \( u \in C^2(T_{k,i}) \). By the mean value theorem, Equation (5.10) can be written as

\[
\beta_{k+1,j} = -\frac{|U|}{4} 2^{-k} (\bar{x}_2 - \bar{x}_1) u''(\bar{x}).
\]

for some \( \bar{x} \in T_{k,i} \). Therefore \( \beta_{k+1,j} \) is of order \( 4^{-k} \):

\[
\beta_{k+1,j} = \frac{|U|}{4} |\bar{x}_2 - \bar{x}_1| |u''(\bar{x})| 2^{-k}
\leq \frac{|U|^2}{2} \|u''\|_{L^\infty(T_{k,i})} 4^{-k-1}.
\]

5.2.2.2 Edge detection for piecewise quadratic polynomials

We can produce similar results as Section 5.2.2.1 for quadratic basis functions by using:

\[
\beta_{k+1,j} = u(x_{k+1,j}) - \left( \frac{3}{8} u(x_{k,i-1}) + \frac{3}{4} u(x_{k,i}) - \frac{1}{8} u(x_{k,i}) \right)
= \frac{3}{8} [u(x_{k+1,j}) - u(x_{k,i-1})] - \frac{5}{8} [u(x_{k,i}) - u(x_{k+1,j})] \\
+ \frac{1}{8} [u(x_{k,i+1}) - u(x_{k,i})],
\]

as in Figure 5.3. We can further show that the hierarchical coefficients of a smooth function shrink more quickly in the quadratic case.
Figure 5.2 Linear hierarchical coefficient from function values

Figure 5.3 Quadratic hierarchical coefficient from function values

\textbf{u is three times continuously differentiable:} We can use an interpolation argument from [18]. Let

\[ F(t) := u(t) - u_k(t) - RL(t), \]

where \( u_k \in S^k \) is the interpolation of \( u \),

\[
R := \frac{u(x_{k+1,j}) - u_k(x_{k+1,j})}{(x_{k+1,j} - x_{k,i-1})(x_{k+1,j} - x_{k,i})(x_{k+1,j} - x_{k,i+1})}
\]

\[ = \frac{64}{3|U|^3} S^k \beta_{k+1,j}, \text{ and} \]

\[ L(t) := (t - x_{k,i-1})(t - x_{k,i})(t - x_{k,i+1}). \]
Since \( u_k(x) = u(x) \) for \( x \in \{x_{k,i}, x_{k,i+1}\} \), we have \( F(x_{k,i-1}) = F(x_{k,i}) = F(x_{k,i+1}) = 0 \). From the construction of \( F(t) \), we also have \( F(x_{k+1,i}) = 0 \). By repeated applications of Rolle’s theorem, \( F'''(t) \) has at least one zero \( \eta \in (x_{k,i-1}, x_{k,i+1}) \). \( u_k \) is a quadratic polynomial on \((x_{k,i-1}, x_{k,i+1})\), so \( u_k'''(t) = 0 \). So we have

\[
\begin{align*}
0 &= F(\eta) \\
&= u'''(\eta) - u'''_k(\eta) - RL'''(\eta) \\
&= u'''(\eta) - 6R \\
&= u'''(\eta) - \frac{128}{|U|} 8^k \beta_{k+1,i}.
\end{align*}
\]

Therefore, \( \beta_{k+1,i} = \frac{|U|}{128} 8^{-k} u'''(\eta) \) and

\[
|\beta_{k+1,i}| \leq \frac{|U|}{16} 8^{-k-1} \|u'''\|_{L^\infty(x_{k,i-1},x_{k,i+1})}.
\]

### 5.3 Comparison of hierarchical and nodal bases

We use the hierarchical bases because they are multi-frequency. As an added bonus, the condition number of the stiffness matrix is smaller, although the condition number of the mass matrix increases. The nodal bases, however, have important computational advantages. The matrices are much more sparse and can be locally assembled.

As we showed in Section 5.1, we can retain most of the advantages of the nodal bases. We assemble and store all of the system matrices in a nodal basis and use the translation matrix \( S \) in an iterative linear solver. \( S \) does not even have to be stored by using the algorithms in Section 4.2.4.
6 PROOF OF CONVERGENCE

In this chapter, we will prove that our FE formulation converges to the entropy solution of the one dimensional periodic Burgers' equation. We will make use of the method of compensated compactness. In particular, we will use the Div-Curl lemma and Murat's lemma.

6.1 The Div-Curl lemma and Murat's lemma

For the sake of completeness, we state the Div-Curl \[27, 6\] and Murat’s lemma \[21, 27, 6\]. The notation used in this section is taken from \[6\]. It is limited to this section and conflicts with other sections.

Let \( n \geq 2 \) be an integer. Let \( U \subseteq \mathbb{R}^n \) be open, bounded, and smooth. Let \( w \in L^2(U; \mathbb{R}^n) \) with \( w = (w^1, \cdots, w^n) \). We define the divergence, \( \text{div} \ w \in H^{-1}(U; \mathbb{R}) \) and the curl, \( \text{curl} \ w \in H^{-1}(U; \mathbb{R}^{n \times n}) \) as:

\[
\text{div} \ w := \sum_{i=1}^{n} \frac{\partial w^i}{\partial x_i} \quad \text{and} \\
(\text{curl} \ w)_{i,j} := \frac{\partial w^i}{\partial x_j} - \frac{\partial w^j}{\partial x_i} \quad (1 \leq i, j \leq n).
\]

For \( n = 2 \),

\[
\text{curl} \ w = \begin{bmatrix}
0 & \frac{\partial w^1}{\partial x_2} - \frac{\partial w^2}{\partial x_1} \\
-\frac{\partial w^1}{\partial x_1} + \frac{\partial w^2}{\partial x_2} & 0
\end{bmatrix}. \tag{6.3}
\]

Since there is only one independent quantity in Equation 6.3, it is a common practice to define, for \( n = 2 \), \( \text{curl} \ w \in H^{-1}(U; \mathbb{R}) \) as

\[
\text{curl} \ w = \frac{\partial w^1}{\partial x_2} - \frac{\partial w^2}{\partial x_1} \quad (n = 2). \tag{6.4}
\]
**Div-Curl lemma:** Let \( \{v_k\}_{k=1}^{\infty}, \{w_k\}_{k=1}^{\infty} \subseteq L^2(U; \mathbb{R}^n) \) be bounded sequences such that

- \( v_k \rightharpoonup v \) weakly in \( L^2(U; \mathbb{R}^n) \) for some \( v \in L^2(U; \mathbb{R}^n) \),
- \( w_k \rightharpoonup w \) weakly in \( L^2(U; \mathbb{R}^n) \) for some \( w \in L^2(U; \mathbb{R}^n) \),
- \( \{\text{div } v_k\}_{k=1}^{\infty} \) is precompact in \( H^{-1}(U; \mathbb{R}) \), and
- \( \{\text{curl } w_k\}_{k=1}^{\infty} \) is precompact in \( H^{-1}(U; \mathbb{R}^{n \times n}) \).

Then \( v_k \cdot w_k \rightharpoonup v \cdot w \) in the sense of distributions.

For our application of the Div-Curl lemma, we will be able to construct a sequence \( \{v_k\}_{k=1}^{\infty} \) such that \( \text{div } v_k \) converges in \( H^{-1}(U; \mathbb{R}) \) and so is trivially precompact in \( H^{-1}(U; \mathbb{R}) \). Showing that our choice of \( \{\text{curl } w_k\}_{k=1}^{\infty} \) is precompact in \( H^{-1}(U; \mathbb{R}) \) is much more difficult and requires Murat’s lemma.

**Murat’s lemma:** Let \( \{f_k\}_{k=1}^{\infty} \subseteq W^{-1,p}(U) \) be a bounded sequence for some \( p > 2 \). Let \( \{g_k\}_{k=1}^{\infty} \) be precompact in \( H^{-1}(U) \) and \( \{h_k\}_{k=1}^{\infty} \) be bounded in \( L^1(U) \). If \( f_k = g_k + h_k \) for all \( 1 \leq k < \infty \), then \( \{f_k\}_{k=1}^{\infty} \) is precompact in \( H^{-1}(U) \).

### 6.2 Problem

Let \( U = (a, b) \) be an open bounded interval. We seek a Finite Element (FE) approximation to \( u(x, t) \), the entropy solution of the periodic hyperbolic conservation law on some finite time interval \((0, T)\):

\[
\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left( \frac{u^2}{2} \right) = 0 \quad (6.5)
\]

\[
\frac{\partial}{\partial t} \left( \frac{u^2}{2} \right) + \frac{\partial}{\partial x} \left( \frac{u^3}{3} \right) \leq 0 \quad (6.6)
\]

\[
u(a, \cdot) = u(b, \cdot) \quad (6.7)
\]

\[
u(\cdot, 0) = g. \quad (6.8)
\]

Equations (6.5) and (6.6) are in the distributional sense. We will assume the \( g \in H^1(U) \) and that \( g \) is periodic.
The entropy solution of Burgers' equation can be found using Equation (6.6) instead of the more general entropy condition Equation (1.9). This greatly simplifies our proof by requiring an entropy type inequality for one entropy/entropy flux pair, \( \left( \frac{u^2}{2}, \frac{u^3}{3} \right) \), instead of all of them.

6.2.1 Notation

To formulate the FE approximation, we need some notation. Let \( \mathcal{T}_0 = U \). \( \mathcal{T}_N \) is obtained by subdividing the elements of \( \mathcal{T}_{N-1} \) into \( M \) distinct elements. Let \( |U| h_N \) be the max diameter of the elements of \( \mathcal{T}_N \). Since the partition is quasi-uniform, there exists a positive constant \( \nu \) such that \( M^{-N} \leq h_N \leq \nu M^{-N} \) for all \( N \).

Let \( \{\psi_{k,i}\}_{k,i} \) be a hierarchical basis of \( S_p^N \). Let us define \( Q_N : S_p^N \rightarrow S_p^N \) as a damping operator \( Q_N u_N = \sum_{k,i} (Q_{k,i} \alpha_{k,i} \psi_{k,i}) \) for \( u = \sum_{k,i} (\alpha_{k,i} \psi_{k,i}) \in S_p^N \), where \( 0 \leq Q_{k,i} \leq 1 \) and \( Q_{k,i} = 1 \) for \( k > m_H \). So \( Q_N \) dampens (or eliminates) the low frequencies of a function while keeping the high frequencies above the level \( m_H \). Occasionally, when the level of a basis function is unimportant, we will switch to the less cumbersome notation \( \{\psi_i\}_i \) and \( \{Q_i\}_i \) for the basis functions and damping coefficients respectively.

We will also use the following convention: \( C \) will denote any positive constant which depends on known quantities, and is independent of any indexing variables. Let \( U_T = U \times (0,T) \).

6.2.2 FE Formulation

Let \( g_N \in S_p^N \) be the interpolant of \( g \). The FE approximation of Equations (6.5)-(6.8) is:

Seek \( u_N(x,t) \) with \( u_N(\cdot,0) = g_N \) such that for all \( v \in S_p^N \),

\[
\int_U \left[ \frac{\partial u_N}{\partial t} + \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) \right] v \, dx + \epsilon_N \int_U \left[ \frac{\partial}{\partial x} (Q_N u_N) \frac{\partial v}{\partial x} \right] \, dx = 0. \tag{6.9}
\]
6.3 Sufficient conditions for convergence

Sufficient conditions for proving that $u_N$ converges to a solution of Burgers’ equation are:

\[ \|u_N\|_{L^\infty(U_T)} \leq C, \]  \hspace{1cm} (6.10)

\[ \epsilon_N, h_N \to 0 \text{ as } N \to \infty, \]  \hspace{1cm} (6.11)

\[ \frac{\epsilon_N}{h_N} \geq C, \]  \hspace{1cm} (6.12)

\[ \frac{\sqrt{\epsilon_N}}{d_x} \left[ (I - Q_N) u_N \right]_{L^2(U)} \leq C \|u_N\|_{L^2(U)} \]  \hspace{1cm} for \( u_N \in \left\{ u_N, \frac{\partial u_N}{\partial t} \right\}, \] and (6.13)

\[ \frac{d}{dx} (Q_N g_N) \leq C \frac{dg_N}{dx} \]  \hspace{1cm} (6.14)

To prove that $u_N$ converges to an entropy solution of Burgers’ equation, we must strengthen Equations (6.12) and (6.13):

\[ \frac{\epsilon_N}{h_N} \to \infty \text{ as } N \to \infty \]  \hspace{1cm} (6.15)

\[ \frac{\sqrt{\epsilon_N}}{d_x} \left[ (I - Q_N) u_N \right]_{L^2(U)} \to 0 \text{ as } N \to \infty. \]  \hspace{1cm} (6.16)

For the moment, we will assume that Equations (6.10)-(6.16) are true. In Section 6.9, we will show how to satisfy the convergence conditions.

6.4 Properties of the FE system of ODEs

The FE approximation is equivalent to: seek $\alpha : (0, T) \to \mathbb{R}^d$ such that

\[ \alpha' + M^{-1} F(\alpha) + M^{-1} K Q \alpha = 0, \]  \hspace{1cm} (6.17)

where $M$ is the mass matrix, $K$ is the stiffness matrix, $Q$ is a diagonal matrix whose elements are $\{Q_i\}$, and $F$ is the flux term:

\[ [F(\alpha)]_i = \int_U \frac{\partial}{\partial x} \left[ \left( \sum_j \alpha_j \psi_j \right)^2 \right] \psi_i \, dx. \]  \hspace{1cm} (6.18)
It is evident that the diffusion term is globally Lipschitz continuous. We shall show that the flux term is locally Lipschitz continuous.

\[
[F(\alpha)]_i = \int_U \frac{\partial}{\partial x} \left( \sum_j \sum_k \alpha_j \alpha_k \psi_j \psi_k \right) \psi_i \, dx = \alpha^T A_i \alpha, \tag{6.19}
\]

where \( A_i \) is the symmetric matrix

\[
[A_i]_{j,k} = \int_U \psi_i \frac{\partial}{\partial x} (\psi_j \psi_k) \, dx. \tag{6.21}
\]

For all \( \alpha \) and \( \beta \in \mathbb{R}^2 \) and all \( i \),

\[
||F(\beta) - F(\alpha)||_i = ||\beta^T A_i \beta - \alpha^T A_i \alpha|| \tag{6.22}
\]

\[
= ||\beta^T A_i \beta - \alpha^T A_i \alpha - \beta^T A_i \alpha + \alpha^T A_i \beta|| \tag{6.23}
\]

\[
= ||(\beta + \alpha)^T A_i (\beta - \alpha)|| \tag{6.24}
\]

\[
\leq ||\beta + \alpha||_2 ||A_i||_2 ||\beta - \alpha||_2 \tag{6.25}
\]

\[
\leq d ||\beta + \alpha||_2 ||A_i||_2 ||\beta - \alpha||_\infty. \tag{6.26}
\]

\[
\Rightarrow ||F(\beta) - F(\alpha)||_\infty \leq d ||\beta + \alpha||_2 ||A_i||_2 ||\beta - \alpha||_\infty. \tag{6.27}
\]

\( ||A_i||_2 \) and \( d \) are independent of \( \alpha \) and \( \beta \), thus the flux term is locally Lipschitz continuous for any \( T \).

ODE theory says there exists a unique \( C^1 [0, T] \) solution of Equation (6.17).

### 6.5 Estimates of \( u_N \)

#### 6.5.1 Estimates from the FE formulation

In our FE approximation (6.9), let us choose \( v = u_N \), then

\[
\int_U \left[ \frac{\partial}{\partial t} \left( \frac{u_N^2}{2} \right) + \frac{\partial}{\partial x} \left( \frac{u_N^3}{3} \right) \right] \, dx + \epsilon_N \int_U \frac{\partial}{\partial x} (Q_N u_N) \frac{\partial u_N}{\partial x} \, dx = 0. \tag{6.28}
\]

Since \( u_N \) is periodic, \( \int_U \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) \, dx = 0 \), which gives us

\[
\frac{1}{2} \frac{d}{dt} \|u_N\|_{L^2(U)}^2 + \epsilon_N \int_U \frac{\partial}{\partial x} (Q_N u_N) \frac{\partial u_N}{\partial x} \, dx = 0. \tag{6.29}
\]
6.5.2 $H^1(U_T)$ estimates of $u_N$ for linear polynomials

The piecewise linear hierarchical basis is orthogonal with respect to the $H^1(U)$ semi-norm, so

$$\int_U \frac{\partial}{\partial x} (Q_N u_N) \frac{\partial u_N}{\partial x} \, dx = \sum_i \sum_j Q_i \alpha_i \alpha_j \int_U \psi_i' \psi_j' \, dx \tag{6.30}$$

$$= \sum_i Q_i \alpha_i^2 \int_U (\psi_i')^2 \, dx \tag{6.31}$$

$$\geq \sum_i Q_i^2 \alpha_i^2 \int_U (\psi_i')^2 \, dx \tag{6.32}$$

$$= \sum_i \sum_j Q_i Q_j \alpha_i \alpha_j \int_U \psi_i' \psi_j' \, dx \tag{6.33}$$

$$= \left\| \frac{\partial}{\partial x} (Q_N u_N) \right\|_{L^2(U)}^2. \tag{6.34}$$

Integrating Equation (6.29) over time, we get

$$C \| g \|_{L^2(U)}^2 \geq \| g_N \|_{L^2(U)}^2$$

$$= \| u_N (\cdot, t) \|_{L^2(U)}^2 + 2\epsilon_N \int_0^t \int_U \frac{\partial}{\partial x} (Q_N u_N) \frac{\partial u_N}{\partial x} \, dx \, dt \tag{6.35}$$

$$\geq \| u_N (\cdot, t) \|_{L^2(U)}^2 + 2\epsilon_N \int_0^t \left\| \frac{\partial}{\partial x} (Q_N u_N) \right\|_{L^2(U)}^2 \, dt. \tag{6.36}$$

Therefore, we have

$$\left\{ \begin{array}{l}
\| u_N \|_{L^2(U_T)} \leq C \sqrt{T} \| g \|_{L^2(U)} \\
\sqrt{\epsilon_N} \left\| \frac{\partial}{\partial x} (Q_N u_N) \right\|_{L^2(U_T)} \leq C \| g \|_{L^2(U)}.
\end{array} \right. \tag{6.37}$$
6.5.3 $H^1(U_T)$ estimates of $u_N$ for higher degree polynomials

The quadratic hierarchical basis functions are not orthogonal, but we can still get an estimate similar to Equation (6.37):

$$
\int_U \frac{\partial}{\partial x} (Q_N u_N) \frac{\partial u_N}{\partial x} \, dx = \int_U \frac{\partial}{\partial x} (Q_N u_N) \frac{\partial (Q_N u_N)}{\partial x} \, dx + \int_U \frac{\partial}{\partial x} (Q_N u_N) \frac{\partial}{\partial x} [(I - Q_N) u_N] \, dx \tag{6.38}
$$

$$
\geq \int_U \left| \frac{\partial}{\partial x} (Q_N u_N) \right|^2 \, dx
- \frac{1}{2} \int_U \left| \frac{\partial}{\partial x} [(I - Q_N) u_N] \right|^2 \, dx \tag{6.39}
$$

$$
= \frac{1}{2} \left\| \frac{\partial}{\partial x} (Q_N u_N) \right\|^2_{L^2(U)}
- \frac{1}{2} \left\| \frac{\partial}{\partial x} [(I - Q_N) u_N] \right\|^2_{L^2(U)} \tag{6.40}
$$

Substituting Equation (6.41) into Equation (6.29), we get

$$
\frac{d}{dt} \left\| u_N \right\|^2_{L^2(U)} \leq C \left\| u_N \right\|^2_{L^2(U)} - \epsilon_N \left\| \frac{\partial}{\partial x} (Q_N u_N) \right\|^2_{L^2(U)} \tag{6.42}
$$

We require a non-standard formulation of the differential form of Gronwall’s inequality. For the sake of completeness, we present the proof here. Let $\eta(t)$ be an absolutely continuous function on $[0, T]$ such that for a.e. $t \in [0, T]$,

$$
\eta'(t) \leq \phi(t) \eta(t) + \psi(t),
$$

where $\phi(t)$ and $\psi(t)$ are summable functions on $[0, T]$.

**Proposition 3.** $\eta(t) \leq e^{\int_0^t \phi(r) \, dr} \left[ \eta(0) + \int_0^t e^{-\int_0^r \phi(s) \, ds} \psi(s) \, ds \right] \forall t \in [0, T]$. 
Proof. This proof is taken from [7]. For a.e. \(0 \leq s \leq T\),
\[
\frac{d}{ds} \left( \eta(s) e^{-\int_0^s \phi(r) \, dr} \right) = e^{-\int_0^s \phi(r) \, dr} \left( \eta'(s) - \phi(s) \eta(s) \right) \\
\leq e^{-\int_0^s \phi(r) \, dr} \psi(s). \tag{6.43}
\]
Integrating Equation (6.43) over the variable \(s\) over \([0, t]\), we get, for all \(0 < t < T\),
\[
\eta(t) e^{-\int_0^t \phi(r) \, dr} - \eta(0) e^{-\int_0^0 \phi(r) \, dr} \leq \int_0^t e^{-\int_0^s \phi(r) \, dr} \psi(s) \, ds. \tag{6.44}
\]
From which we get the assertion.

Let us now assume \(\phi = C\) is a positive constant, and \(\psi \leq 0\) is never positive. We then have
\[
\eta(t) \leq e^{Ct} \left[ \eta(0) + \int_0^t e^{-Cs} \psi(s) \, ds \right] \\
\leq e^{Ct} \left[ \eta(0) + \int_0^t e^{-Cs} \psi(s) \, ds \right] \\
= e^{Ct} \eta(0) + \int_0^t \psi(s) \, ds. \tag{6.45}
\]
Using Equation (6.45) with Equation (6.42), we get
\[
\sum \nu_N^2 \left( u_N \right)_{L^2(U)} \leq e^{Ct} \|g_N\|_{L^2(U)}^2 - \epsilon_N \int_0^t \left\| \frac{\partial}{\partial x} (Q_N u_N) \right\|_{L^2(U)}^2 \, ds. \tag{6.46}
\]
Since \(\|g_N\|_{L^2(U)} \leq C \|g\|_{L^2(U)}\),
\[
\sum \nu_N^2 \left( u_N \right)_{L^2(U)} + \epsilon_N \int_0^t \left\| \frac{\partial}{\partial x} (Q_N u_N) \right\|_{L^2(U)}^2 \, ds \leq C e^{Ct} \|g\|^2_{L^2(U)}. \tag{6.47}
\]
Therefore, we have
\[
\begin{cases} 
\|u_N\|_{L^2(U)} \leq C \sqrt{\epsilon \|g\|_{L^2(U)}} \\
\sqrt{\epsilon \|\frac{\partial}{\partial x} (Q_N u_N)\|_{L^2(U, T)}} \leq C \sqrt{\epsilon \|g\|_{L^2(U)}}. \tag{6.48}
\end{cases}
\]

### 6.6 Strong Convergence of \(\{u_N\}\)

Let us define \(v_N = \left( \frac{u_N^2}{2}, u_N \right)\) and \(w_N = \left( \frac{u_N^2}{2}, -\frac{u_N^3}{3} \right)\), so
\[
\text{div } v_N = \frac{\partial u_N}{\partial t} + \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) \quad \text{and} \\
\text{curl } w_N = \frac{\partial}{\partial t} \left( \frac{u_N^2}{2} \right) + \frac{\partial}{\partial x} \left( \frac{u_N^3}{3} \right).
\]
6.6.1 \( L^2(\Omega_T) \) bound on \( \{\nabla v_N \} \)

In our FE approximation (6.9), let us choose \( v = \frac{\partial u_N}{\partial t} \), then

\[
\int_U \left| \frac{\partial u_N}{\partial t} \right|^2 dx + \int_U \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) \frac{\partial u_N}{\partial t} \, dx
\]

\[
= - \epsilon_N \int_U \frac{\partial}{\partial x} \left( Q_N u_N \right) \frac{\partial^2 u_N}{\partial x^2} \, dx
\]

\[
= - \epsilon_N \int_U \frac{\partial}{\partial x} \left( Q_N u_N \right) \frac{\partial^2 u_N}{\partial x \partial t} \, dx
\]

\[
- \epsilon_N \int_U \frac{\partial}{\partial x} \left( Q_N u_N \right) \frac{\partial^2 u_N}{\partial x \partial t} [(I - Q_N) u_N] \, dx
\]

\[
= - \frac{\epsilon_N}{2} \int_U \frac{\partial}{\partial t} \left\{ \left( \frac{\partial}{\partial x} (Q_N u_N) \right)^2 \right\} \, dx
\]

\[
- \epsilon_N \int_U \frac{\partial}{\partial x} \left( Q_N u_N \right) \frac{\partial^2 u_N}{\partial x \partial t} [(I - Q_N) u_N] \, dx
\]

\[
\leq - \frac{\epsilon_N}{2} \int_U \frac{\partial}{\partial t} \left\{ \left( \frac{\partial}{\partial x} (Q_N u_N) \right)^2 \right\} \, dx
\]

\[
+ \epsilon_N \left\| \frac{\partial}{\partial x} (Q_N u_N) \right\|_{L^2(U)} \left\| \frac{\partial^2 u_N}{\partial x \partial t} [(I - Q_N) u_N] \right\|_{L^2(U)}
\]

\[
\leq - \frac{\epsilon_N}{2} \int_U \frac{\partial}{\partial t} \left\{ \left( \frac{\partial}{\partial x} (Q_N u_N) \right)^2 \right\} \, dx
\]

\[
+ C \sqrt{\epsilon_N} \left\| \frac{\partial}{\partial x} (Q_N u_N) \right\|_{L^2(U)} \left\| \frac{\partial u_N}{\partial t} \right\|_{L^2(U)}
\]

\[
\leq - \frac{\epsilon_N}{2} \int_U \frac{\partial}{\partial t} \left\{ \left( \frac{\partial}{\partial x} (Q_N u_N) \right)^2 \right\} \, dx + C \left\| \frac{\partial u_N}{\partial t} \right\|_{L^2(U)}
\]

\[
\leq - \frac{\epsilon_N}{2} \int_U \frac{\partial}{\partial t} \left\{ \left( \frac{\partial}{\partial x} (Q_N u_N) \right)^2 \right\} \, dx + C^2 + \frac{1}{4} \left\| \frac{\partial u_N}{\partial t} \right\|_{L^2(U)}^2.
\]

Rearranging the terms of Equation (6.55), we get

\[
\frac{3}{4} \left\| \frac{\partial u_N}{\partial t} \right\|_{L^2(U)}^2 + \epsilon_N \frac{d}{dt} \left\| \frac{\partial}{\partial x} (Q_N u_N) \right\|_{L^2(U)}^2 \leq - \int_U \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) \frac{\partial u_N}{\partial t} \, dx
\]

\[
\leq - \left\| \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) \right\|_{L^2(U)} \left\| \frac{\partial u_N}{\partial t} \right\|_{L^2(U)}
\]

\[
\leq \left\| \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) \right\|_{L^2(U)}^2 + \frac{1}{4} \left\| \frac{\partial u_N}{\partial t} \right\|_{L^2(U)}^2.
\]
Rearranging the terms of Equation (6.58), we get
\[
\left\| \frac{\partial u_N}{\partial t} \right\|_{L^2(U)}^2 \leq C - \epsilon_N \frac{d}{dt} \left\| \frac{\partial}{\partial x} \left( Q_N u_N \right) \right\|_{L^2(U)}^2 + 2 \left\| \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) \right\|_{L^2(U)}^2. \tag{6.59}
\]

Integrating Equation (6.59) over time
\[
\left\| \frac{\partial u_N}{\partial t} \right\|_{L^2(U_T)}^2 \leq C + 2 \int_{U_T} \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) dx dt - \epsilon_N \left( Q_N u_N (\cdot, T) \right)_{L^2(U)}^2 + \epsilon_N \left( \frac{d}{dx} (Q_N g_N) \right)_{L^2(U)}^2 \tag{6.60}
\]
\[
\leq C + 2 \int_{U_T} \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) dx dt + \epsilon_N \left( \frac{d}{dx} (Q_N g_N) \right)_{L^2(U)}^2 \tag{6.61}
\]
\[
\leq C + 2 \int_{U_T} \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) dx dt + C \epsilon_N \left( \frac{d}{dx} g_N \right)_{L^2(U)}^2 \tag{6.62}
\]
\[
\leq C + 2 \int_{U_T} \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) dx dt + C \epsilon_N \left( \frac{d}{dx} \right)_{L^2(U)}^2. \tag{6.63}
\]

The second term in Equation (6.63) can be estimated by:
\[
\int_{U_T} \left| \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) \right|^2 dx dt = \int_{U_T} |u_N|^2 \left| \frac{\partial u_N}{\partial x} \right|^2 dx dt \tag{6.64}
\]
\[
\leq \left\| u_N \right\|_{L^\infty(U_T)}^2 \left\| \frac{\partial u_N}{\partial x} \right\|_{L^2(U_T)}^2 \tag{6.65}
\]
\[
\leq \frac{C}{\epsilon_N}. \tag{6.66}
\]

Equation (6.63) now becomes
\[
\epsilon_N \left\| \frac{\partial u_N}{\partial t} \right\|_{L^2(U)}^2 \leq C \left( 1 + \epsilon_N + \epsilon_N^2 \right). \tag{6.67}
\]

Therefore,
\[
\sqrt{\epsilon_N} \left\| \frac{\partial u_N}{\partial t} \right\|_{L^2(U_T)} \leq C. \tag{6.68}
\]

Combining Equations (6.66) and (6.68), we get
\[
\sqrt{\epsilon_N} \left\| \text{div } v_N \right\|_{L^2(U_T)} = \sqrt{\epsilon_N} \left\| \frac{\partial u_N}{\partial t} + \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) \right\|_{L^2(U_T)} \tag{6.69}
\]
\[
\leq \sqrt{\epsilon_N} \left\| \frac{\partial u_N}{\partial t} \right\|_{L^2(U_T)} + \sqrt{\epsilon_N} \left\| \frac{\partial}{\partial x} \left( \frac{u_N^2}{2} \right) \right\|_{L^2(U_T)} \tag{6.70}
\]
\[
\leq C. \tag{6.71}
\]
6.6.2 $H^{-1}(U_T)$ bound on $\{\text{div } v_N\}$

Let $\bar{\varphi} \in H^1_0(U_T)$. For all $t \in (0, T)$, let $\bar{\varphi}_N(\cdot, t) \in S^N_p \cap H^1_0(U)$ be the $H^1(U)$ projection of $\bar{\varphi}$, so for all $v \in S^N_p \cap H^1_0(U)$,

$$
\int_U \frac{\partial \bar{\varphi}_N(\cdot, t)}{\partial x} \frac{\partial v}{\partial x} \, dx = \int_U \frac{\partial \bar{\varphi}(\cdot, t)}{\partial x} \frac{\partial v}{\partial x} \, dx.
$$

We need the $H^1(U)$ projection into $S^N_p$ of an arbitrary $\bar{\varphi} \in H^1_0(U_T)$ in order to use our FE formulation:

$$
\int_{U_T} (\text{div } v_N) \bar{\varphi} \, dx \, dt = \int_{U_T} (\text{div } v_N) \bar{\varphi}_N \, dx \, dt \\
+ \int_{U_T} (\text{div } v_N) (\bar{\varphi} - \bar{\varphi}_N) \, dx \, dt \\
= -\epsilon_N \int_{U_T} \frac{\partial}{\partial x} (Q_N v_N) \frac{\partial \bar{\varphi}_N}{\partial x} \, dx \, dt \\
+ \int_{U_T} (\text{div } v_N) (\bar{\varphi} - \bar{\varphi}_N) \, dx \, dt \\
= -\epsilon_N \int_{U_T} \frac{\partial}{\partial x} (Q_N v_N) \frac{\partial \bar{\varphi}}{\partial x} \, dx \, dt \\
+ \int_{U_T} (\text{div } v_N) (\bar{\varphi} - \bar{\varphi}_N) \, dx \, dt \\
\leq \epsilon_N \left\| \frac{\partial}{\partial x} (Q_N v_N) \right\|_{L^2(U_T)} \left\| \frac{\partial \bar{\varphi}}{\partial x} \right\|_{L^2(U_T)} \\
+ \|\text{div } v_N\|_{L^2(U_T)} \|\bar{\varphi} - \bar{\varphi}_N\|_{L^2(U_T)} \\
\leq \epsilon_N \left\| \frac{\partial}{\partial x} (Q_N v_N) \right\|_{L^2(U_T)} \left\| \frac{\partial \bar{\varphi}}{\partial x} \right\|_{L^2(U_T)} \\
+ C h_N \|\text{div } v_N\|_{L^2(U_T)} \left\| \frac{\partial \bar{\varphi}}{\partial x} \right\|_{L^2(U_T)} \\
\leq C \left(\sqrt{\epsilon_N} + \frac{h_N}{\sqrt{\epsilon_N}}\right) \left\| \frac{\partial \bar{\varphi}}{\partial x} \right\|_{L^2(U_T)} \\
= C \sqrt{\epsilon_N} \left(1 + \frac{h_N}{\epsilon_N}\right) \left\| \frac{\partial \bar{\varphi}}{\partial x} \right\|_{L^2(U_T)}.
$$

Requirement (6.12) says that $\frac{h_N}{\epsilon_N} \leq C$, so

$$
\int_{U_T} (\text{div } v_N) \bar{\varphi} \, dx \, dt \leq C \sqrt{\epsilon_N} \left\| \frac{\partial \bar{\varphi}}{\partial x} \right\|_{L^2(U_T)}.
$$
6.6.3 \{\text{div } v_N\} \text{ lies in a compact subset of } H^{-1}(U_T)

Let \( \varphi \in H_0^1(U_T) \) with \( \|\varphi\|_{H^1(U_T)} \leq 1 \). From Section 6.6.2, we know that

\[
\int_{U_T} (\text{div } v_N) \varphi \, dx \, dt \leq C \sqrt{\varepsilon_N},
\]

which gives us

\[
\|\text{div } v_N\|_{H^{-1}(U_T)} \leq C \sqrt{\varepsilon_N} \rightarrow 0 \text{ as } N \rightarrow \infty.
\]

Thus \( \{\text{div } v_N\} \) lies in a compact subset of \( H^{-1}(U_T) \).

6.6.4 \{\text{curl } w_N\} \text{ lies in a compact subset of } H^{-1}(U_T)

Let \( \varphi \in C_0^\infty(U_T) \) be a test function. \( u_N \varphi \in H_0^1(U_T) \), so we can use the result from Section 6.6.2, with \( \psi = u_N \varphi \).

\[
\int_{U_T} (\text{curl } w_N) \varphi \, dx \, dt = \int_{U_T} (\text{div } v_N) u_N \varphi \, dx \, dt
\]

\[
\leq C \sqrt{\varepsilon_N} \left\| \frac{\partial}{\partial x} (u_N \varphi) \right\|_{L^2(U_T)}
\]

\[
= C \sqrt{\varepsilon_N} \left( \left\| u_N \frac{\partial \varphi}{\partial x} + \varphi \frac{\partial u_N}{\partial x} \right\|_{L^2(U_T)} \right)
\]

\[
\leq C \sqrt{\varepsilon_N} \left( \left\| u_N \frac{\partial \varphi}{\partial x} \right\|_{L^2(U_T)} + \left\| \varphi \frac{\partial u_N}{\partial x} \right\|_{L^2(U_T)} \right)
\]

\[
\leq C \sqrt{\varepsilon_N} \left( \left\| u_N \right\|_{L^\infty(U_T)} \left\| \frac{\partial \varphi}{\partial x} \right\|_{L^2(U_T)}
\right.
\]

\[
+ \left\| \varphi \right\|_{L^\infty(U_T)} \left\| \frac{\partial u_N}{\partial x} \right\|_{L^2(U_T)} \right)
\]

\[
\leq C \left( \sqrt{\varepsilon_N} \left\| \frac{\partial \varphi}{\partial x} \right\|_{L^2(U_T)} + \left\| \varphi \right\|_{L^\infty(U_T)} \right).
\]

We can not use Murat’s lemma directly. Instead, we use a variational form of Murat’s lemma used in [19, 12] which gives us, from Equation (6.88), that \( \{\text{curl } w_N\} \) lies in a compact subset of \( H^{-1}(U_T) \).
6.6.5 Subsequence of $u_N$ converges strongly in $L^2(UT)$

Since $\|u_N\|_{L^\infty(UT)} \leq C$, there exists a subsequence $\{u_{N'_k}\}$ of $\{u_N\}$ such that for $1 \leq p \leq 4$, 
$\{u^{p}_{N'_k}\}$ converges weakly in $L^2(UT)$. Let $\overline{u^{(p)}} \in L^2(UT)$ be the weak limit of $u^{p}_{N'_k}$. So $u_{N'_k}$ and $w_{N'_k}$ converge weakly:

\[
\begin{align*}
  v_{N'_k} & \to \left( \frac{u^{(2)}_{N'_k}}{2}, \frac{u^{(1)}_{N'_k}}{3} \right) \\
  w_{N'_k} & \to \left( \frac{u^{(2)}_{N'_k}}{2}, \frac{u^{(3)}_{N'_k}}{3} \right).
\end{align*}
\] (6.89) (6.90)

Using the Div-Curl lemma, we have

\[
\lim_{k \to \infty} \int_{UT} (v_{N'_k} \cdot w_{N'_k}) \varphi \, dx \, dt = \int_{UT} (\overline{\nabla \cdot w}) \varphi \, dx \, dt \quad \forall \varphi \in C_0^\infty (UT). \quad (6.91)
\]

For all $\varphi \in C_0^\infty (UT),$

\[
\begin{align*}
  \lim_{k \to \infty} \int_{UT} (v_{N'_k} \cdot w_{N'_k}) \varphi \, dx \, dt & = \lim_{k \to \infty} \int_{UT} \left( \frac{u_{N'_k}^4}{4} - \frac{u_{N'_k}^4}{3} \right) \varphi \, dx \, dt \\
  & = \lim_{k \to \infty} \int_{UT} -\frac{u_{N'_k}^4}{12} \varphi \, dx \, dt \\
  & = \int_{UT} -\frac{u^{(4)}}{12} \varphi \, dx \, dt. \quad (6.92) \quad (6.93) \quad (6.94)
\end{align*}
\]

The right hand side of Equation (6.91) is:

\[
\int_{UT} (\overline{\nabla \cdot w}) \varphi \, dx \, dt = \int_{UT} \left( \frac{u^{(2)}}{4} - \frac{u^{(1)} \cdot u^{(3)}}{3} \right) \varphi \, dx \, dt. \quad (6.95)
\]

Since Equations (6.94) and (6.95) are equal, we have

\[
\overline{u^{(4)}} = 4 \overline{u^{(1)}} \overline{u^{(3)}} - 3 \left( \overline{u^{(2)}} \right)^2 \text{ a.e.} \quad (6.96)
\]

We can use Equation (6.96) to show that $u_{N'_k}$ converges strongly to $\overline{u^{(1)}}$ in $L^2(UT)$:

\[
\left( u_{N'_k} - \overline{u^{(1)}} \right)^4 = u_{N'_k}^4 - 4 u_{N'_k}^3 \overline{u^{(1)}} + 6 u_{N'_k}^2 \left( \overline{u^{(1)}} \right)^2 - 4 u_{N'_k} \left( \overline{u^{(1)}} \right)^3 + \left( \overline{u^{(1)}} \right)^4. \quad (6.97)
\]
Taking the weak limit of both sides of Equation (6.97), we have

\[
\lim_{k \to \infty} \left( u_{N_k} - u^{(1)} \right)^4 = u^{(4)} - 4 u^{(3)} u^{(1)}
\]

\[
+ 6 u^{(2)} u^{(1)} - 4 u^{(1)} + u^{(1)}^4
\]

\[
= 4 u^{(1)} u^{(3)} - 3 u^{(2)} - 4 u^{(3)} u^{(1)}
\]

\[
+ 6 u^{(2)} u^{(1)} - 4 u^{(1)} + u^{(1)}^4
\]

\[
= -3 \left[ u^{(2)} - u^{(1)} \right]^2
\]

\[
\leq 0.
\]

So we have

\[
0 \leq \lim_{k \to \infty} \int_{U_T} \left( u_{N_k} - u^{(1)} \right)^4 dx dt
\]

\[
= \int_{U_T} -3 \left[ u^{(2)} - u^{(1)} \right]^2 dx dt
\]

\[
\leq 0.
\]

We now have \( \overline{u^{(2)}} = \left( \overline{u^{(1)}} \right)^2 \) a.e., which gives us

\[
\left\| \overline{u^{(1)}} \right\|_{L^2(U_T)}^2 = \int_{U_T} \overline{u^{(2)}} dx dt
\]

\[
= \lim_{k \to \infty} \int_{U_T} u_{N_k}^2 dx dt
\]

\[
= \lim_{k \to \infty} \left\| u_{N_k} \right\|_{L^2(U_T)}^2.
\]

Therefore, \( u := \overline{u^{(1)}} \) is the strong limit of \( u_{N_k} \) in \( L^2(U_T) \).
6.7 Convergence to a solution

We will now show that \( \bar{u} \) is a solution of the conservation law. For all test functions \( \varphi \in C_0^\infty(U_T) \),

\[
\int_{U_T} \left[ \frac{\partial \bar{u}}{\partial t} \varphi + \frac{\partial}{\partial x} \left( \frac{\bar{u}^2}{2} \right) \varphi \right] \, dx \, dt = - \int_{U_T} \left[ \frac{\partial \varphi}{\partial t} + \frac{\bar{u}^2}{2} \frac{\partial \varphi}{\partial x} \right] \, dx \, dt \tag{6.109}
\]

\[
= - \int_{U_T} \left[ u^{(1)} \frac{\partial \varphi}{\partial t} + \frac{u^{(2)}}{2} \frac{\partial \varphi}{\partial x} \right] \, dx \, dt \tag{6.110}
\]

\[
= - \lim_{k \to \infty} \int_{U_T} \left[ u_{N_k} \frac{\partial \varphi}{\partial t} + \frac{u_{N_k}^2}{2} \frac{\partial \varphi}{\partial x} \right] \, dx \, dt \tag{6.111}
\]

\[
= \lim_{k \to \infty} \int_{U_T} \left[ \frac{\partial u_{N_k}}{\partial t} \varphi + \frac{\partial}{\partial x} \left( \frac{u_{N_k}^2}{2} \right) \varphi \right] \, dx \, dt \tag{6.112}
\]

\[
= \lim_{k \to \infty} \int_{U_T} (\text{div} \, v_{N_k}) \varphi \, dx \, dt. \tag{6.113}
\]

The right hand side of Equation (6.113) is zero since

\[
0 \leq \int_{U_T} (\text{div} \, v_{N_k}) \varphi \, dx \, dt \leq \| \text{div} \, v_{N_k} \|_{H^{-1}(U_T)} \| \varphi \|_{H^1(U_T)} \tag{6.114}
\]

\[
\leq C \sqrt{\epsilon_N} \| \varphi \|_{H^1(U_T)} \tag{6.115}
\]

\[
\to 0 \text{ as } k \to \infty \tag{6.116}
\]

6.8 Convergence to the entropy solution

We showed in Sections 6.6 and 6.7 that \( \{u_{N_k}\} \) converges strongly to a solution \( \bar{u} \) of the conservation law. We can show that \( \bar{u} \) is the physically relevant entropy solution with the strengthened requirements:

\[
\frac{\epsilon_N}{h_N} \to \infty \text{ as } N \to \infty, \text{ and} \tag{6.117}
\]

\[
\sqrt{\epsilon_N} \left\| \frac{\partial}{\partial x} \left[ (I - Q_N) u_N \right] \right\|_{L^2(U)} \to 0 \text{ as } N \to \infty. \tag{6.118}
\]
Let $\varphi \in C_0^\infty (U_T)$.

\[
0 \leq \left| \int_{U_T} (\bar{u}^3 - u_{N_k}^3) \varphi \, dx \, dt \right|
\leq \left| \int_{U_T} (\bar{u}^2 + u_{N_k} + u_{N_k}^2 - 2 \bar{u} - u_{N_k}) \varphi \, dx \, dt \right|
\leq \left\| \bar{u} - u_{N_k} \right\|_{L^2(U_T)} \left\| (\bar{u}^2 + u_{N_k} + u_{N_k}^2) \varphi \right\|_{L^2(U_T)}
\leq \left\| \varphi \right\|_{L^\infty(U_T)} \left\| \bar{u} - u_{N_k} \right\|_{L^2(U_T)} \left\| \bar{u}^2 + u_{N_k} + u_{N_k}^2 \right\|_{L^2(U_T)}
\leq \left\| \varphi \right\|_{L^\infty(U_T)} \left\| \bar{u} - u_{N_k} \right\|_{L^2(U_T)}
\left( \left\| \bar{u}^2 \right\|^2_{L^2(U_T)} + \left\| \bar{u} u_{N_k} \right\|^2_{L^2(U_T)} + \left\| u_{N_k}^2 \right\|^2_{L^2(U_T)} \right)
\leq \left\| \varphi \right\|_{L^\infty(U_T)} \left\| \bar{u} - u_{N_k} \right\|^2_{L^2(U_T)}
\left( \left\| \bar{u} \right\|^2_{L^4(U_T)} + \left\| u_{N_k} \right\|^2_{L^\infty(U_T)} \left\| \bar{u} \right\|_{L^2(U_T)} + \left\| u_{N_k} \right\|^2_{L^\infty(U_T)} \sqrt{|U_T|} \right).
\] (6.124)

\{$u_N\}$ is uniformly bounded, so $\left\| u_{N_k} \right\|_{L^\infty(U_T)} \leq C$. $\bar{u}$ is in $L^4 (U_T)$ since $\bar{u}^2 = \bar{u} (2) \in L^2 (U_T)$.

Since $\lim_{k \to \infty} \left\| \bar{u} - u_{N_k} \right\|_{L^2(U_T)} = 0$, we have

\[
\lim_{k \to \infty} \int_{U_T} u_{N_k}^3 \varphi \, dx \, dt = \int_{U_T} \bar{u}^3 \varphi \, dx \, dt.
\] (6.125)

Let $\varphi \in C_0^\infty (U_T)$ with $\varphi \geq 0$, then

\[
\int_{U_T} \left[ \frac{\partial}{\partial t} \left( \frac{\bar{u}^2}{2} \right) + \frac{\partial}{\partial x} \left( \frac{\bar{u}^3}{3} \right) \right] \varphi \, dx \, dt = - \int_{U_T} \left( \frac{\bar{u}^2}{2} \frac{\partial \varphi}{\partial t} + \frac{\bar{u}^3}{3} \frac{\partial \varphi}{\partial x} \right) \, dx \, dt
\]
\[
= - \lim_{k \to \infty} \int_{U_T} \left( \frac{u_{N_k}^2}{2} \frac{\partial \varphi}{\partial t} + \frac{u_{N_k}^3}{3} \frac{\partial \varphi}{\partial x} \right) \, dx \, dt
\] (6.126)
\[
= \lim_{k \to \infty} \int_{U_T} \left[ \frac{\partial}{\partial t} \left( \frac{u_{N_k}^2}{2} \right) + \frac{\partial}{\partial x} \left( \frac{u_{N_k}^3}{3} \right) \right] \varphi \, dx \, dt
\] (6.127)
\[
= \lim_{k \to \infty} \int_{U_T} \left( \frac{\partial}{\partial t} \left( \frac{u_{N_k}^2}{2} \right) + \frac{\partial}{\partial x} \left( \frac{u_{N_k}^3}{3} \right) \right) \varphi \, dx \, dt
\] (6.128)
\[
= \lim_{k \to \infty} \int_{U_T} \left( \text{div} u_{N_k} \right) u_{N_k} \varphi \, dx \, dt
\] (6.129)

Let $z_N := u_N \varphi$. For all $t \in (0, T)$, let $z_N^k (\cdot, t) \in S_p^N \cap H_0^1 (U)$ be the $H^1 (U)$ projection of $z_N$, so for all $v \in S_p^N \cap H_0^1 (U)$,

\[
\int_U \frac{\partial z_N^k (\cdot, t)}{\partial x} \frac{\partial v}{\partial x} \, dx = \int_U \frac{\partial z_N (\cdot, t)}{\partial x} \frac{\partial v}{\partial x} \, dx.
\] (6.130)
We can show that as $k \to \infty$, the right hand side of Equation (6.129) is non-positive.

\[
\int_{U_T} (\text{div } u_N) u_N \varphi \, dx \, dt = \int_{U_T} (\text{div } u_N) z_N \, dx \, dt \tag{6.131}
\]

\[
= \int_{U_T} (\text{div } u_N) z_N^h \, dx \, dt
\]

\[
+ \int_{U_T} (\text{div } u_N_k) \left( z_N - z_N^h \right) \, dx \, dt \tag{6.132}
\]

\[
= -\epsilon_N \int_{U_T} \frac{\partial z_N^h}{\partial x} \frac{\partial}{\partial x} (Q_N u_N) \, dx \, dt
\]

\[
+ \int_{U_T} (\text{div } u_N) \left( z_N - z_N^h \right) \, dx \, dt \tag{6.133}
\]

\[
= -\epsilon_N \int_{U_T} \frac{\partial z_N}{\partial x} \frac{\partial u_N}{\partial x} \, dx \, dt
\]

\[
+ \epsilon_N \int_{U_T} \frac{\partial z_N}{\partial x} \frac{\partial}{\partial x} [(I - Q_N) u_N] \, dx \, dt
\]

\[
+ \int_{U_T} (\text{div } u_N) \left( z_N - z_N^h \right) \, dx \, dt \tag{6.134}
\]

\[
= -\epsilon_N \int_{U_T} \frac{\partial}{\partial x} (u_N \varphi) \frac{\partial u_N}{\partial x} \, dx \, dt
\]

\[
+ \epsilon_N \int_{U_T} \frac{\partial z_N}{\partial x} \frac{\partial}{\partial x} [(I - Q_N) u_N] \, dx \, dt
\]

\[
+ \int_{U_T} (\text{div } u_N) \left( z_N - z_N^h \right) \, dx \, dt \tag{6.135}
\]

\[
= -\epsilon_N \int_{U_T} \varphi \left| \frac{\partial u_N}{\partial x} \right|^2 \, dx \, dt
\]

\[
- \epsilon_N \int_{U_T} u_N \frac{\partial}{\partial x} \frac{\partial u_N}{\partial x} \, dx \, dt
\]

\[
+ \epsilon_N \int_{U_T} \frac{\partial z_N}{\partial x} \frac{\partial}{\partial x} [(I - Q_N) u_N] \, dx \, dt
\]

\[
+ \int_{U_T} (\text{div } u_N) \left( z_N - z_N^h \right) \, dx \, dt. \tag{6.137}
\]

The first term on the right hand side of Equation (6.137) is non-positive, while the other terms go to zero as $N \to \infty$. 
For the second term on the right hand side of Equation (6.137):

\[
0 \leq | -\epsilon_N \int_{U_T} u_N \frac{\partial \varphi}{\partial x} \frac{\partial u_N}{\partial x} \, dx \, dt | \tag{6.138}
\]

\[
\leq \epsilon_N \| u_N \|_{L^\infty(U_T)} \left\| \frac{\partial u_N}{\partial x} \right\|_{L^2(U_T)} \left\| \frac{\partial \varphi}{\partial x} \right\|_{L^2(U_T)}
\]

\[
\leq C \sqrt{\epsilon_N} \left\| \frac{\partial \varphi}{\partial x} \right\|_{L^2(U_T)}
\tag{6.140}
\]

\[
\rightarrow 0 \text{ as } N \rightarrow \infty.
\tag{6.141}
\]

For the third term on the right hand side of Equation (6.137):

\[
0 \leq | \epsilon_N \int_{U_T} \frac{\partial z_N}{\partial x} \frac{\partial}{\partial x} \left( (I - Q_N) u_N \right) \, dx \, dt |
\tag{6.142}
\]

\[
\leq \epsilon_N \left\| \frac{\partial z_N}{\partial x} \right\|_{L^2(U_T)} \left\| \frac{\partial}{\partial x} \left( (I - Q_N) u_N \right) \right\|_{L^2(U_T)}
\tag{6.143}
\]

\[
= \epsilon_N \left\| \frac{\partial}{\partial x} (u_N \varphi) \right\|_{L^2(U_T)} \left\| \frac{\partial}{\partial x} \left( (I - Q_N) u_N \right) \right\|_{L^2(U_T)}
\tag{6.144}
\]

\[
\leq \epsilon_N \left\| \frac{\partial}{\partial x} \left[ (I - Q_N) u_N \right] \right\|_{L^2(U_T)}
\tag{6.145}
\]

\[
\leq \epsilon_N \left\| \frac{\partial}{\partial x} \left[ (I - Q_N) u_N \right] \right\|_{L^2(U_T)} \left( \| \varphi \|_{L^\infty(U_T)} \left\| \frac{\partial u_N}{\partial x} \right\|_{L^2(U_T)} + \| u_N \|_{L^\infty(U_T)} \left\| \frac{\partial \varphi}{\partial x} \right\|_{L^2(U_T)} \right)
\tag{6.146}
\]

\[
\rightarrow 0 \text{ as } N \rightarrow \infty.
\tag{6.147}
\]
For the fourth term on the right hand side of Equation (6.137):

\[
0 \leq \left| \int_{U_T} \left( \text{div } u_N \right) \left( z_N - z_N^h \right) \, dx \, dt \right| \leq \|\text{div } u_N\|_{L^2(U_T)} \left\| z_N - z_N^h \right\|_{L^2(U_T)} \leq C h_N \|\text{div } u_N\|_{L^2(U_T)} \left\| \frac{\partial z_N}{\partial x} \right\|_{L^2(U_T)} \leq C \frac{h_N}{\sqrt{\epsilon_N}} \left\| \frac{\partial z_N}{\partial x} \right\|_{L^2(U_T)} = C \left\| \frac{\partial (u_N \varphi)}{\partial x} \right\|_{L^2(U_T)} = C \left( \frac{h_N}{\sqrt{\epsilon_N}} \left\| \frac{\partial \varphi}{\partial x} \right\|_{L^2(U_T)} + \frac{h_N}{\epsilon_N} \|\varphi\|_{L^\infty(U_T)} \right) \leq C \left( \frac{h_N}{\sqrt{\epsilon_N}} \left\| \frac{\partial \varphi}{\partial x} \right\|_{L^2(U_T)} + \frac{h_N}{\epsilon_N} \|\varphi\|_{L^\infty(U_T)} \right) \rightarrow 0 \quad \text{as } N \to \infty.
\]

Therefore

\[
\int_{U_T} \left[ \frac{\partial}{\partial t} \left( \frac{u^2}{2} \right) + \frac{\partial}{\partial x} \left( \frac{u^3}{3} \right) \right] \varphi \, dx \, dt = \liminf_{k \to \infty} \int_{U_T} \left( \text{div } u_{N_k} \right) u_{N_k} \varphi \, dx \, dt \leq 0.
\]

6.9 Free Parameters

In the FE formulation, we have to choose \( \epsilon_N, m, \) and the form of  \( Q_N \) for \( k \leq m. \) Let

\( 0 < \delta \leq \theta \leq 1. \) We choose \( \epsilon_N \) and \( m \) as

\[
\epsilon_N = C h_N^\theta \quad \text{and} \quad m_H \leq \frac{\delta N}{2}.
\]

\( Q_{k,i} \) is chosen to be either 1 or 0:

\[
Q_{k,i} = \begin{cases} 
0 & k \leq m_H \\
1 & k > m_H.
\end{cases}
\]
These definitions satisfy the requirements for convergence, Equations (6.10)-(6.14). To satisfy the requirements that the approximations converge to the entropy solution, Equations (6.15) and (6.16), we take \(0 < \delta < \theta < 1\).

### 6.9.1 Satisfying the convergence criteria

Showing that \(\{u_N\}\) is uniformly bounded is a difficult task. It is a common practice in compensated compactness arguments to take it as “God given” as in [23, 19, 12, 6]. We shall therefore assume Equation (6.10) is true.

It is evident from our definitions that \(\epsilon_N, h_N \rightarrow 0\) as \(N \rightarrow \infty\), so Equation (6.11) is true.

If \(0 < \theta \leq 1\),

\[
\frac{\epsilon_N}{h_N} = C h_N^{\theta-1} \geq C,
\]

so Equation (6.12) is satisfies.

If \(0 < \theta < 1\),

\[
\frac{\epsilon_N}{h_N} = C h_N^{\theta-1} \rightarrow \infty \text{ as } N \rightarrow \infty,
\]

so Equation (6.15) is satisfied.

### 6.9.2 Inverse estimate of the low frequencies

Let \(v \in S_p^N\). \((I - Q_N)\) is simply an interpolation operator on a coarse grid and so

\[
\|(I - Q_N) v\|_{L^2(U)} \leq C \|v\|_{L^2(U)}.
\] (6.161)

\(Q_N\) keeps the high frequencies of a function, so \((I - Q_N)\) eliminates them.

\[
Q_{k,i} = 1 \Rightarrow 1 - Q_{k,i} = 0 \text{ for } k > m_H,
\] (6.162)
so \((I - Q_N) v \in S^m_p\). Using a standard inverse estimate,

\[
\left\| \frac{\partial}{\partial x} [(I - Q_N) u_N] \right\|_{L^2(U)} \leq C \left( h_m \right)^{-1} \| (I - Q_N) v \|_{L^2(U)}
\]

\[
\leq C M^{m_H} \| (I - Q_N) v \|_{L^2(U)}
\]

\[
\leq C M^{2 \delta} \| (I - Q_N) v \|_{L^2(U)}
\]

\[
\leq C \left( \frac{\nu}{h_N} \right)^{\frac{\delta}{2}} \| (I - Q_N) v \|_{L^2(U)}
\]

\[
\leq C \sqrt{\nu} h_N^{-\delta} \| v \|_{L^2(U)}
\]

\[
\sqrt{\epsilon N} \left\| \frac{\partial}{\partial x} [(I - Q_N) u_N] \right\|_{L^2(U)} \leq C \sqrt{\nu} \sqrt{\frac{\epsilon N}{h_N^3}} \| v \|_{L^2(U)}
\]

\[
= C \sqrt{h_N^{3-\delta}} \| v \|_{L^2(U)}.
\]

If \(\delta \leq \theta\), then \(h_N^{\theta-\delta} \leq C\), and Equation (6.13) is satisfied. If \(\delta < \theta\), then \(h_N^{\theta-\delta} \rightarrow 0\) as \(N \rightarrow \infty\), and Equation (6.16) is satisfied.

### 6.9.3 \(H^1\) error estimates of the low frequencies

Since \((I - Q_N)\) is an interpolation operator on a coarse grid, for all \(v \in S^N_p\),

\[
\left\| \frac{d}{dx} (Q_N v) \right\|_{L^2(U)} = \left\| \frac{dv}{dx} - \frac{d}{dx} [(I - Q_N) v] \right\|_{L^2(U)}
\]

\[
\leq C \left\| \frac{dv}{dx} \right\|_{L^2(U)}.
\]

Therefore, Equation (6.14) is satisfied.
7 NUMERICAL RESULTS

All of our numerical results were generated using the finite element library deal.II [2].

7.1 Free parameters

In our finite element formulation, Equation (5.2), we have several free parameters which we must choose.

7.1.1 Choosing $\epsilon_N$

A natural choice for $\epsilon_N$ is $\epsilon_N = h_N$.

7.1.2 Choosing $m$ so that $Q_{N,n,m} = 0$ for $n < m$

We were able to prove convergence when $Q_{N,n,m} = 1$ for $n > \frac{N}{2}$. Our numerical experiments indicate that we can add much less diffusion, but we still choose $m$ to be a fixed fraction of $N$: $m = \frac{3}{4}N$.

7.1.3 Choosing the form of $Q_N$

In our two dimensional calculations, we have chosen an isotropic diffusion term, as in Equation (5.6). As we can see in Equation (5.7), this greatly simplifies the formulation.

As we mentioned in Section 5.2.1, one of the potential advantages of this method is the ability to add diffusion only near a discontinuity. We were not able to find a form of $Q_N$ that gave better results with adaptive diffusion.

The SV method gives better results if $Q_{N,n,m}$ is smooth with respect to the level number $n$. Spectral basis functions have many more levels than hierarchical basis functions to achieve
Numerical evidence indicates that a sharp jump in \( Q_{N_{i};m,i} \) with respect to \( n \) does not adversely affect the calculation. Slightly better results were attained with a continuous \( Q_{N_{i};m,i} \):

\[
Q_{N_{i};m,i} = \frac{n - m + 1}{N - m}.
\]

### 7.1.4 Choosing the quadratic hierarchical structure

As we can see in Figure 4.2, there are two possible hierarchical structures for piecewise quadratic polynomials. In the first column, all the basis functions are quadratic. In the second column, the low frequency basis functions are linear. The major difference between the two is that for the same \( Q_N \) and \( m \), the linear low frequency structure will add more diffusion. For a piecewise quadratic function \( v \), \( Q_N v = \sum_{n,i} Q_{N_{i};m,i} \beta_{n,i} \psi_{n,i} \) with \( Q_{N_{i};m,i} = 0 \) for small \( n \). If all of the basis functions are quadratic, then in the smooth regions, \( \beta_{n,i} = O(8^{-n}) \), as we showed in Section 5.2.2.2. If only the highest frequency basis functions are quadratic, then for \( n < N \), \( \beta_{n,i} = O(4^{-n}) \) in the smooth regions, as we showed in Section 5.2.2.1. Therefore \( Q_N v \) will be larger for the linear low frequency basis functions.

### 7.1.5 Ordinary differential equation solver

Our FE formulation turns the partial differential equation into a system of ordinary differential equations. We must choose an ODE solver. We implemented several solvers: Crank-Nicholson, second through fourth order backward difference formulae, and second through fourth Adams-Moulton methods. All the methods are implicit with time step the same as the grid size \( \Delta t = h_N \). For our FE formulation to be competitive with other methods, we would need an explicit time stepping scheme or a much larger time step. For our initial calculation however we wanted to minimize the error from the ODE solver and only study the error from the spatial discretization.

For the solution of Burgers' equation, we took a ODE solver that was an order of magnitude more accurate than the best possible spatial error. For an arbitrary smooth function, the best piecewise linear approximating polynomial has an error of \( O\left( h_N^2 \right) \) so we chose a third order
ODE solver. For the piecewise quadratic polynomials, we chose a fourth order ODE solver.

For the shock tube problem, the ODE solvers that were more than second order accurate were either unstable and did not converge or resulted in a large number of oscillations between the contact and shock discontinuities as in Figure 7.22.

For the steady state problems, the ODE solver was unimportant, and we were able to take a much larger time step: $\Delta t = 250 \ h_N$.

7.1.6 Post-processing strategy

Because we are approximating a discontinuous solution with continuous piecewise polynomials, we see Gibbs oscillations near the discontinuity. A simple strategy to remove the oscillations is to set the coefficients of the hierarchical expansion to zero around the discontinuity. The question then becomes the location of the discontinuity. Let $\beta_{n+1,i}$ be a high frequency hierarchical coefficient. Let $\beta_{n,J}$ be the parent hierarchical coefficient, so the support of $\psi_{n+1,i}$ is a subset of $\psi_{n,J}$. If the solution is continuously differentiable in the region of the support of $\psi_{n,J}$, then

$$\frac{\beta_{n,J}}{\beta_{n+1,i}} = \frac{C_1 2^{-n}}{C_2 2^{-n-1}} \approx 2.$$  

Therefore, our strategy is: for the highest four frequencies, if a hierarchical coefficient is larger than half the value of its parent, then it is set to zero. The choice of the highest four levels was chosen through experimentation.

Our simple strategy only affects the region around a discontinuity, but it has the disadvantage of smoothing across the discontinuity. Our solution therefore becomes more smeared.

We apply the post-processing strategy to solutions of the two variants of Burgers’ equation, although it could also be applied to solutions of Euler’s equations.

7.2 Convergence rates

All our errors are given in the $L^1$ norm. There seems to be some consensus that this is a natural norm for hyperbolic conservation laws.
Near a discontinuity, we are limited to how well a piecewise polynomial can approximate a solution. We are more interested in the convergence rates in the smooth regions. We therefore exclude a region of length 0.2 around all the discontinuities in our error calculations.

For some regions where the solution is constant, the convergence rates are erratic because the $L^1$ error is almost zero. In the domain of the steady state Burgers' equation and in regions 4 and 1 of the shock tube problem, the exact solution is constant, and the error is less than $10^{-13}$.

For the shock tube problem, we observed different convergence rates in the five distinct regions depicted in Figure 3.2. We therefore present the five different errors and convergence rates. Similarly for the wedge problem, we present the errors and convergence rates for the regions to the left and right of the shock.

### 7.3 Boundary conditions for the wedge problem

As we saw in Section 3.2.2, the top and bottom of the wedge have a no penetration boundary condition $\mathbf{u} \cdot \mathbf{n} = 0$. As we can see in the solution, Figure 3.3, we can over specify the boundary conditions by setting $\rho$, $\rho e$, and $\rho \mathbf{u}$ on the boundary segments $(x, 0)$ and $(x, 2)$ for $0 < x < 2$.

The latter boundary condition produces much better numerical results, as we can see by comparing Figures 7.23 and 7.24.

Because of our diffusion scheme, information is sent upstream, which is necessary but unphysical. The altered lower boundary condition for all variables (except $u_2$) has a disastrous affect on the solution near the lower boundary. This does not happen when all variables are specified on the lower boundary.

### 7.4 Numerical simulations

#### 7.4.1 Steady state Burgers' equation

In the smooth regions, the solution is constant, and the convergence rates are erratic because the $L^1$ error is almost zero. The discontinuity is better resolved by quadratic polynomials
than linear polynomials. The post-processing strategy removed the oscillations near the discontinuity.

Figure 7.1 Solution of steady state Burgers' equation with linear polynomials (without post-processing)
Steady State Burgers’ Equation

\[ m = \lfloor 3N/4 \rfloor \quad (75\% \text{ of the diff. levels} = 0) \]
\[ Q_{k,j} = (k-m+1)(N-m) \]

poly. degree: 1

4 levels were removed (post-processing)

Figure 7.2 Solution of steady state Burgers’ equation with linear polynomials (with post-processing)
Figure 7.3 Solution of steady state Burgers' equation with quadratic linear low frequency polynomials (without post-processing)
Figure 7.4  Solution of steady state Burgers' equation with quadratic linear low frequency polynomials (with post-processing)
Steady State Burgers' Equation

$m = \lfloor 3N/4 \rfloor$ (75% of the diff. levels = 0)

$Q_{i,j} = (k-m+1)(N-m)$

poly. degree: 2

0 levels were removed (post-processing)

Figure 7.5 Solution of steady state Burgers' equation with quadratic polynomials (without post-processing)
Figure 7.6 Solution of steady state Burgers' equation with quadratic polynomials (with post-processing)
Table 7.1 Convergence rate of steady state Burgers’ equation with linear polynomials

<table>
<thead>
<tr>
<th>levels</th>
<th>without post-processing</th>
<th>with post-processing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$L^1$ error</td>
<td>rate</td>
</tr>
<tr>
<td>8</td>
<td>1.918e-04</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>2.370e-05</td>
<td>3.02</td>
</tr>
<tr>
<td>10</td>
<td>2.109e-08</td>
<td>10.13</td>
</tr>
<tr>
<td>11</td>
<td>8.708e-13</td>
<td>14.56</td>
</tr>
<tr>
<td>12</td>
<td>5.524e-13</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Table 7.2 Convergence rate of steady state Burgers’ equation with quadratic linear low frequency polynomials

<table>
<thead>
<tr>
<th>levels</th>
<th>without post-processing</th>
<th>with post-processing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$L^1$ error</td>
<td>rate</td>
</tr>
<tr>
<td>8</td>
<td>2.849e-06</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>2.683e-07</td>
<td>3.41</td>
</tr>
<tr>
<td>10</td>
<td>6.202e-12</td>
<td>15.40</td>
</tr>
<tr>
<td>11</td>
<td>9.421e-13</td>
<td>2.72</td>
</tr>
<tr>
<td>12</td>
<td>5.621e-13</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Table 7.3 Convergence rate of steady state Burgers’ equation with quadratic polynomials

<table>
<thead>
<tr>
<th>levels</th>
<th>without post-processing</th>
<th>with post-processing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$L^1$ error</td>
<td>rate</td>
</tr>
<tr>
<td>8</td>
<td>1.485e-04</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>1.351e-05</td>
<td>3.46</td>
</tr>
<tr>
<td>10</td>
<td>1.159e-08</td>
<td>10.19</td>
</tr>
<tr>
<td>11</td>
<td>1.229e-12</td>
<td>13.20</td>
</tr>
<tr>
<td>12</td>
<td>9.498e-13</td>
<td>0.37</td>
</tr>
</tbody>
</table>
7.4.2 Periodic Burgers’ equation

In the smooth regions, we see quasi-optimal convergence rates using linear and quadratic polynomials. Using linear low frequency quadratic polynomial, the convergence is only second order, but this is explained in Section 7.1.4. The discontinuity is better resolved by quadratic polynomials than linear polynomials. The post-processing strategy removed the oscillations near the discontinuity.

Figure 7.7 Solution of periodic Burgers’ equation with linear polynomials (without post-processing)
Figure 7.8 Solution of periodic Burgers' equation with linear polynomials (with post-processing)
Periodic Burgers' Equation

\[ m = \left\lfloor 3 \frac{N}{4} \right\rfloor \quad (75\% \text{ of the diff. levels} = 0) \]

\[ Q_{k,j} = (k - m + 1)(N - m) \]

polynomial degree: 2 (linear low frequency)
time = 1.00 (4th Order Back. Diff. Form.)

0 levels were removed (post-processing)

Figure 7.9 Solution of periodic Burgers' equation with quadratic linear low frequency polynomials (without post-processing)
Figure 7.10 Solution of periodic Burgers' equation with quadratic linear low frequency polynomials (with post-processing)
Figure 7.11  Solution of periodic Burgers' equation with quadratic polynomials (without post-processing)
Figure 7.12 Solution of periodic Burgers' equation with quadratic polynomials (with post-processing)
Table 7.4  Convergence rate of periodic Burgers' equation with linear polynomials

<table>
<thead>
<tr>
<th>levels</th>
<th>$L^1$ error without post-processing</th>
<th>rate</th>
<th>$L^1$ error with post-processing</th>
<th>rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>2.637e-04</td>
<td>-</td>
<td>1.909e-03</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>2.126e-05</td>
<td>3.63</td>
<td>7.171e-04</td>
<td>1.41</td>
</tr>
<tr>
<td>10</td>
<td>4.874e-06</td>
<td>2.13</td>
<td>1.984e-04</td>
<td>1.85</td>
</tr>
<tr>
<td>11</td>
<td>1.273e-06</td>
<td>1.94</td>
<td>5.098e-05</td>
<td>1.96</td>
</tr>
<tr>
<td>12</td>
<td>3.242e-07</td>
<td>1.97</td>
<td>1.260e-05</td>
<td>2.02</td>
</tr>
</tbody>
</table>

Table 7.5  Convergence rate of periodic Burgers' equation with quadratic linear low frequency polynomials

<table>
<thead>
<tr>
<th>levels</th>
<th>$L^1$ error without post-processing</th>
<th>rate</th>
<th>$L^1$ error with post-processing</th>
<th>rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1.000e-03</td>
<td>-</td>
<td>1.610e-03</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>7.478e-06</td>
<td>7.06</td>
<td>1.854e-04</td>
<td>3.12</td>
</tr>
<tr>
<td>10</td>
<td>1.867e-06</td>
<td>2.00</td>
<td>4.874e-05</td>
<td>1.93</td>
</tr>
<tr>
<td>11</td>
<td>4.653e-07</td>
<td>2.00</td>
<td>1.206e-05</td>
<td>2.01</td>
</tr>
<tr>
<td>12</td>
<td>1.162e-07</td>
<td>2.00</td>
<td>2.991e-06</td>
<td>2.01</td>
</tr>
</tbody>
</table>

Table 7.6  Convergence rate of periodic Burgers' equation with quadratic polynomials

<table>
<thead>
<tr>
<th>levels</th>
<th>$L^1$ error without post-processing</th>
<th>rate</th>
<th>$L^1$ error with post-processing</th>
<th>rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>8.724e-04</td>
<td>-</td>
<td>1.100e-02</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>1.140e-05</td>
<td>6.26</td>
<td>1.806e-03</td>
<td>2.61</td>
</tr>
<tr>
<td>10</td>
<td>1.580e-05</td>
<td>6.17</td>
<td>4.326e-06</td>
<td>8.71</td>
</tr>
<tr>
<td>11</td>
<td>1.764e-08</td>
<td>3.16</td>
<td>5.473e-07</td>
<td>2.98</td>
</tr>
<tr>
<td>12</td>
<td>2.209e-09</td>
<td>3.00</td>
<td>6.796e-08</td>
<td>3.01</td>
</tr>
</tbody>
</table>
7.4.3 Shock tube problem

In the smooth regions where the solution is constant, the convergence rates are erratic, but the $L^1$ error is almost zero only in regions four and one. In regions three and four, the $L^1$ error is not as small as in regions four and one, but it is between 10 and 100 times smaller than the error in the expansion wave. The $L^1$ error in the expansion wave seems to be limited to first order. This is most likely an example of the phenomenon known as “downstream pollution” which similarly affects other numerical solvers of hyperbolic conservation laws.

Figure 7.13 Solution of shock tube problem with linear polynomials ($\rho$)
Figure 7.14 Solution of shock tube problem with linear polynomials (pu)
Shock Tube Problem
Variable: \(p_e\)

\[m = \lfloor 3 \times N / 4 \rfloor \quad (75\% \text{ of the diff. levels} = 0)\]

\[Q_{k,i} = (k - m + 1) / (N - m)\]

poly. degree: 1

time = 0.29 (Crank–Nicholson)

0 levels were removed (post-processing)

Figure 7.15  Solution of shock tube problem with linear polynomials \((p_e)\)
Figure 7.16 Solution of shock tube problem with linear low frequency polynomials ($\rho$)
Figure 7.17  Solution of shock tube problem with linear low frequency polynomials ($\rho u$)
Figure 7.18 Solution of shock tube problem with linear low frequency polynomials ($\rho e$)
Figure 7.19 Solution of shock tube problem with quadratic polynomials (ρ)
Figure 7.20 Solution of shock tube problem with quadratic polynomials

\( \rho u \)
Shock Tube Problem
Variable: \( \rho e \)

\[ m = \lfloor 3N/4 \rfloor \text{ (75\% of the diff. levels = 0)} \]
\[ Q_{k,i} = (k-m+1)(N-m) \]

poly. degree: 2
time=0.29 (Crank-Nicholson)

0 levels were removed (post-processing)

Figure 7.21 Solution of shock tube problem with quadratic polynomials \((\rho e)\)
Shock Tube Problem
Variable: \( p \)

\[ m = \lfloor 3N/4 \rfloor \quad (75\% \text{ of the diff. levels} = 0) \]

\[ Q_{k,j} = (k-m+1)(N-m) \]

poly. degree: 2

time = 0.25 (3rd Order Back. Diff. Form.)

0 levels were removed (post-processing)

Figure 7.22 Solution of shock tube problem with quadratic polynomials and 3rd order ODE solver (unstable)
<table>
<thead>
<tr>
<th>Levels</th>
<th>Region 4</th>
<th>Expansion Wv.</th>
<th>Region 3</th>
<th>Region 2</th>
<th>Region 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$L^1$ error</td>
<td>rate</td>
<td>$L^1$ error</td>
<td>rate</td>
<td>$L^1$ error</td>
</tr>
<tr>
<td>$\rho$</td>
<td>8</td>
<td>6.9e-05</td>
<td>-</td>
<td>3.0e-03</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.8e-05</td>
<td>1.7</td>
<td>1.7e-03</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>6.7e-07</td>
<td>4.7</td>
<td>8.8e-04</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1.6e-09</td>
<td>8.7</td>
<td>4.4e-04</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>5.1e-14</td>
<td>14.9</td>
<td>2.2e-04</td>
<td>0.9</td>
</tr>
<tr>
<td>$\mu$</td>
<td>8</td>
<td>7.1e-05</td>
<td>-</td>
<td>1.9e-03</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>2.1e-05</td>
<td>1.7</td>
<td>1.0e-03</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>8.0e-07</td>
<td>4.7</td>
<td>5.4e-04</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1.9e-09</td>
<td>8.7</td>
<td>2.7e-04</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>4.0e-14</td>
<td>15.5</td>
<td>1.3e-04</td>
<td>0.9</td>
</tr>
<tr>
<td>$\rho e$</td>
<td>8</td>
<td>2.1e-04</td>
<td>-</td>
<td>7.9e-03</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>6.3e-05</td>
<td>1.7</td>
<td>4.4e-03</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>2.3e-06</td>
<td>4.7</td>
<td>2.2e-03</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>5.6e-09</td>
<td>8.7</td>
<td>1.1e-03</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>9.6e-14</td>
<td>15.8</td>
<td>5.7e-04</td>
<td>0.9</td>
</tr>
</tbody>
</table>
Table 7.8  Convergence rate of shock tube problem with quadratic linear
low frequency polynomials

<table>
<thead>
<tr>
<th></th>
<th>Region 4</th>
<th>Expansion Wv.</th>
<th>Region 3</th>
<th>Region 2</th>
<th>Region 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>levels</td>
<td>$L^1$ error</td>
<td>rate</td>
<td>$L^1$ error</td>
<td>rate</td>
</tr>
<tr>
<td>$\rho$</td>
<td>8</td>
<td>2.6e-06</td>
<td>-</td>
<td>1.3e-03</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.9e-06</td>
<td>0.4</td>
<td>8.3e-04</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>7.4e-09</td>
<td>8.0</td>
<td>4.2e-04</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>6.0e-13</td>
<td>13.5</td>
<td>2.1e-04</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>4.2e-14</td>
<td>3.8</td>
<td>1.0e-04</td>
<td>0.9</td>
</tr>
<tr>
<td>$\mu$</td>
<td>8</td>
<td>3.1e-06</td>
<td>-</td>
<td>8.3e-04</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>2.2e-06</td>
<td>0.4</td>
<td>5.1e-04</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>8.8e-09</td>
<td>8.0</td>
<td>2.5e-04</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>4.8e-13</td>
<td>14.1</td>
<td>1.2e-04</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>3.4e-14</td>
<td>3.7</td>
<td>6.4e-05</td>
<td>1.0</td>
</tr>
<tr>
<td>$\rho_e$</td>
<td>8</td>
<td>9.3e-06</td>
<td>-</td>
<td>3.4e-03</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>6.6e-06</td>
<td>0.4</td>
<td>2.1e-03</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>2.6e-08</td>
<td>8.0</td>
<td>1.0e-03</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1.3e-12</td>
<td>14.2</td>
<td>5.4e-04</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>1.8e-13</td>
<td>2.8</td>
<td>2.7e-04</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Table 7.9  Convergence rate of shock tube problem with quadratic polynomials

<table>
<thead>
<tr>
<th></th>
<th>Region 4</th>
<th>Expansion Wv.</th>
<th>Region 3</th>
<th>Region 2</th>
<th>Region 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>levels</td>
<td>$L^1$ error</td>
<td>rate</td>
<td>$L^1$ error</td>
<td>rate</td>
</tr>
<tr>
<td>$\rho$</td>
<td>8</td>
<td>5.8e-06</td>
<td>-</td>
<td>1.3e-03</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>7.4e-06</td>
<td>-0.3</td>
<td>8.8e-04</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1.3e-07</td>
<td>6.0</td>
<td>4.4e-04</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>4.1e-11</td>
<td>11.4</td>
<td>2.2e-04</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>6.8e-14</td>
<td>9.2</td>
<td>1.1e-04</td>
<td>0.9</td>
</tr>
<tr>
<td>$\mu$</td>
<td>8</td>
<td>6.8e-06</td>
<td>-</td>
<td>8.4e-04</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>8.7e-06</td>
<td>-0.3</td>
<td>5.4e-04</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1.3e-07</td>
<td>6.0</td>
<td>2.7e-04</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>4.9e-11</td>
<td>11.4</td>
<td>1.3e-04</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>2.3e-13</td>
<td>7.7</td>
<td>6.8e-05</td>
<td>1.0</td>
</tr>
<tr>
<td>$\rho_e$</td>
<td>8</td>
<td>2.0e-05</td>
<td>-</td>
<td>3.5e-03</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>2.5e-05</td>
<td>-0.3</td>
<td>2.2e-03</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>4.0e-07</td>
<td>6.0</td>
<td>1.1e-03</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1.4e-10</td>
<td>11.4</td>
<td>5.7e-04</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>1.5e-13</td>
<td>9.8</td>
<td>2.8e-04</td>
<td>0.9</td>
</tr>
</tbody>
</table>
7.4.4 Wedge problem

To the right of the shock, we have first order convergence. This is most likely another example of downstream pollution, which we saw in our solution of the shock tube problem. To the left of the shock, we see second order convergence, when we should see an error close to zero since the exact solution is constant. Although not as disastrous as the situation discussed in Section 7.3, this is another example of excessive propagation of information upstream for the wedge problem.

Figure 7.23 Poor approximation of wedge problem ($\rho$) because of slip boundary conditions
Wedge Problem

Variable: \( p \)

\[ m = \left\lfloor 3^*N/4 \right\rfloor \] (75% of the diff. levels = 0)

\[ Q_{ji} = (k-m+1)(N-m) \]

poly. degree: 1

0 levels were removed (post-processing)

Figure 7.24 Solution of wedge problem with linear polynomials (\( \rho \))
Figure 7.25 Solution of wedge problem with linear polynomials ($|\mu u|$)

Wedge Problem
Variable: $|\mu u|

m = \left\lfloor 3^*N/4 \right\rfloor \ (75\% \text{ of the diff. levels} = 0)

Q_{k,i} = (k-m+1)(N-m)

poly. degree: 1

0 levels were removed (post-processing)
Figure 7.26 Solution of wedge problem with linear polynomials ($\frac{u}{|u|}$)
Wedge Problem
Variable: pe

\[ m = \lfloor 3N/4 \rfloor \quad (75\% \text{ of the diff. levels} = 0) \]
\[ Q_{k,i} = (k-m+1)(N-m) \]

poly. degree: 1

0 levels were removed (post-processing)

Figure 7.27 Solution of wedge problem with linear polynomials (pe)
Wedge Problem

Variable: $pe$

$m = \lfloor 3N/4 \rfloor$ (75% of the diff. levels = 0)

$Q_{kj} = (k-m+1)/(N-m)$

poly. degree: 1

0 levels were removed (post-processing)

Figure 7.28 Solution of wedge problem with linear polynomials ($pe$ along curves)
Figure 7.29  Grids for the wedge problem
Table 7.10 Convergence rate of wedge problem with linear polynomials

<table>
<thead>
<tr>
<th>levels</th>
<th>$L^1$ error left of shock</th>
<th>rate</th>
<th>$L^1$ error right of shock</th>
<th>rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>7.7e-02</td>
<td>-</td>
<td>4.9e-02</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>3.1e-02</td>
<td>1.3</td>
<td>2.9e-02</td>
<td>0.7</td>
</tr>
<tr>
<td>7</td>
<td>7.0e-03</td>
<td>2.1</td>
<td>1.5e-02</td>
<td>0.9</td>
</tr>
<tr>
<td>$\rho u_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.4e-01</td>
<td>-</td>
<td>8.8e-02</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>5.9e-02</td>
<td>1.3</td>
<td>5.5e-02</td>
<td>0.6</td>
</tr>
<tr>
<td>7</td>
<td>1.3e-02</td>
<td>2.1</td>
<td>3.1e-02</td>
<td>0.8</td>
</tr>
<tr>
<td>$\rho u_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.4e-01</td>
<td>-</td>
<td>8.8e-02</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>5.9e-02</td>
<td>1.3</td>
<td>5.5e-02</td>
<td>0.6</td>
</tr>
<tr>
<td>7</td>
<td>1.3e-02</td>
<td>2.1</td>
<td>3.1e-02</td>
<td>0.8</td>
</tr>
<tr>
<td>$\rho e$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>7.1e-02</td>
<td>-</td>
<td>6.3e-02</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>2.7e-02</td>
<td>1.3</td>
<td>3.6e-02</td>
<td>0.8</td>
</tr>
<tr>
<td>7</td>
<td>5.9e-03</td>
<td>2.2</td>
<td>1.6e-02</td>
<td>1.0</td>
</tr>
</tbody>
</table>
The initial results for this new method are promising. We have a stable finite element method which, in some cases, attains quasi-optimal convergence rates in the smooth regions. We also have developed a theoretical foundation for understanding why this method works. These results, however, are preliminary. There are potential pitfalls awaiting in more complicated problems, but there is also untapped potential in the framework.

**Find an efficient adaptive diffusion form of** $Q_N$: As we mentioned in Section 7.1.3, we were not able to find a diffusion operator $Q_N$ which explicitly added more diffusion near a discontinuity. In other words, our $Q_{N,m,s}^{i,k}$ in Equations (5.3)-(5.4) does not depend on $i$. It was done implicitly since for large $n$, $\beta_{n,s}$ (and so $Q_{N,m,s}^{i,k}$) is small in the smooth regions and large near a discontinuity. Nevertheless, an adaptive diffusion operator should improve the convergence rates in the smooth regions.

**Run two and three dimensional test problems, especially on curved domains:** The problems we have chosen are good initial test problems, but more realistic problems are required. We would like to implement this method on curved domains (such as airplane wings) and in three dimensions.

**Perform comparisons with other numerical methods:** We would like to perform rigorous comparisons with other methods to see if our method is competitive.

**Extend the theory:** In the SV framework, convergence results are available for multi-dimensional scalar conservation laws [4]. We would like to extend the theory of Chapter 6 to
multidimensional problems. We would also like to extend the theory to include convergence rates.

**Implement an anisotropic diffusion term to minimize crosswind diffusion:** As we mentioned in Section 7.1.3, the diffusion operator in our calculations is isotropic. In other words, our $Q_{N^1}^{j,k}$ in Equations (5.3)-(5.4) is zero if $j \neq k$. We would like to implement an anisotropic diffusion term to minimize crosswind diffusion.

**Implement adaptive refinement:** In all of our calculations, we used uniform grids. Clearly, we would like to have a finer grid near a discontinuity. The ease of edge detection gives us a good way of knowing where more elements are needed. Hierarchical bases are well suited to adaptive refinement because adding another level does not affect the coefficient values of the previous levels.

**Implement discontinuous basis functions:** In Section 2.4.1, we listed some of the advantages of discontinuous Galerkin (DG) methods. Multi-resolution viscosity ideas can be incorporated into the DG framework by using discontinuous hierarchical basis functions, as depicted in Figure 8.1.

![Discontinuous Linear Hierarchical Basis Functions](image)

Figure 8.1 Discontinuous linear hierarchical basis functions
Find a better post-processing strategy: As we mentioned in Section 7.1.6, setting hierarchical high frequency coefficients to zero near a discontinuity removes the oscillations, but it also smoothes across the discontinuity, adding more diffusion. A better strategy would be to dampen the high frequency coefficient instead of setting them to zero. We would like to find an automatic process which finds what values to set the high frequency coefficients to near the discontinuity.
BIBLIOGRAPHY


ACKNOWLEDGEMENTS

First and foremost, I would like to thank Prof. Max Gunzburger and Prof. Janet Peterson. Before I met them, I had not even heard of the finite element method. Their help and guidance throughout my graduate education have been invaluable. I feel privileged to have been their student.

Thanks to Prof. Eitan Tadmor for his helpful discussions, especially concerning the convergence theory.

Thanks to my committee members Prof. Scott Hansen, Prof. L. Steven Hou, and Prof. Sunder Sethuraman for their time and input.

I am grateful for the support I have received from Sandia National Labs. Dr. Curtis Ober has generously provided his time expertise in hyperbolic conservation laws. Dr. Pavel Bochev and Dr. Rich Lehoucq have been of great help during my stay at Sandia. David Womble and the members of the Computational Mathematics and Algorithms, Computational Sciences Research Institute (CSRI) have provided an excellent learning environment.

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000

Last but not least, I would like to thank Liane Hansen, whose math discussions were always helpful and whose friendship was never failing.