On the Advective Component of Active Flux Schemes for Nonlinear Hyperbolic Conservation Laws

by

Jungyeoul Brad Maeng

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Aerospace Engineering)
in the University of Michigan
2017

Doctoral Committee:

Professor Philip L. Roe, Chair
Associate Professor Karthik Duraisamy
Associate Professor Krzysztof J. Fidkowski
Professor Robert Krasny

Aequam memento rebus in arduis servare mentem

Jungyeoul Brad Maeng

jmaeng@umich.edu

ORCID iD: 0000-0002-7798-7865

For my family

Acknowledgments

I express my deepest gratitude to my advisor Professor Philip Roe who has provided an immense amount of support and guidance through every step along the Ph.D. journey. What makes this journey memorable is the process of working out the challenges we had encountered along the way. I am especially grateful for the patience and the sense of optimism that Professor Roe has always demonstrated throughout many challenging, and at times frustrating, experiences. I have learned a lot from him.

I thank Professor Karthik Duraisamy, Professor Krzysztof Fidkowski, and Professor Robert Krasny for taking their precious time to serve as committee members. Their insights and constructive criticisms have positively shaped the outcome of the thesis. I also extend my sincere gratitude to Professor Bram van Leer, who has expressed a special interest in my research and given me thoughtful comments.

I would like to thank my friends at Michigan for reminding me that there is a whole world out there besides the graduate school. Special thanks go out to the original 2211 FXB colleagues, Yuntao Chen, Kentaro Hara, Yee Chee See, Anand Singh, and Rohan Morajkar for great memories we had together. I also thank the new 2211 FXB mates, Chris Petersen, Dominic Liao-McPherson, Will Dunham, AJ Berning, Nan Li, and Dongkun Han, for opening their hearts to the only CFD student in the office. A very special thanks go to Wai Lee Chan for the positive influence he had on me. Among many personal qualities, his kindness, compassion, and caring towards friends and family have shaped my life. I am very fortunate to call him my true friend.

It would have been difficult to navigate through the preliminary examination process without the help of Kyle Ding, Tyler Lung, Kevin Neitzel, and Rohan Morajkar. Their willingness to help and motivate each other had given me a clear direction through the process. I would also like to thank Doreen Fan, Tyler Lung, Ju Wang, and Tim Eymann for their willingness to give and share thoughtful discussions on research. Their perspectives and insights have expanded my view on many things.

I could not have reached this point without the most important people in my life. I am grateful to Sara for being next to me through thick and thin, and for steering me in the right directions. Her dedication and support undoubtedly made this process much better

than I would have imagined. Most importantly, I am very thankful to my parents who have provided me with this amazing life-changing opportunity. And I thank my brother for making himself available to our parents when I couldn't. I am forever indebted to them for their unwavering support so I could find my own path.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

AF Active Flux

CABARET Compact Accurately Boundary-Adjusting high-REsolution Technique

CFD computational fluid dynamics

CFL Courant-Friedrichs-Lewy

DG discontinuous Galerkin

DOF degree(s) of freedom

ENO Essentially Non-Oscillatory

FCT Flux Corrected Transport

FD finite difference

FS fluctuation splitting

FV finite volume

MUSCL Monotonic Upstream-centered Scheme for Conservation Laws

RD residual distribution

SE spectral element

SV spectral volume

TVB total variation bounded

TVD total variation diminishing

WENO Weighted Essentially Non-Oscillatory

ABSTRACT

A new class of numerical methods called Active Flux (AF) is investigated for nonlinear hyperbolic conservation laws. The AF method is designed specifically to address the aspect that most modern compressible flow methods fail to do; the multidimensionality aspect. It addresses the shortcoming by employing a two stage update process. In the first stage, a nonconservative form of the system is introduced to provide the flexibility to pursue distinct numerical approaches for flow processes with differing physics. Because each process is treated separately, the numerical method can be appropriately formed to reflect each type of physics and to provide the maximal stability. The method is completed with the conservation update to produce a third-order accurate scheme.

The AF advection scheme is founded on the characteristic tracing method, a semi-Lagrangian method, which has long been used for developing numerical methods for hyperbolic problems. The first known AF method for advection, Scheme V by van Leer, is revisited as a part of the development of the scheme. Details of Scheme V are examined closely, and new improvements are made for the multidimensional nonlinear advection scheme.

A detailed study of the nonlinear system of equations is made possible by the pressureless Euler system, which is the advective component of the Euler system. It serves as a stepping stone for the Euler system, and all necessary details of the nonlinear system are explored. Lastly, an extension to the Euler system is presented where a novel nonlinear operator splitting method is introduced to correctly blend the contributions of the nonlinear advection and acoustic processes. The AF method, as a result, produces a maximally stable, third-order accurate method for the multidimensional Euler system.

Some guiding principles of limiting are presented. Because two types of flow feature are separately treated, the limiting process must also be kept separate. Advective problems obeying natural bounding principles are treated differently from acoustic problems with no explicit bounding principles. Distinct limiting approaches are explored along with discussions.

CHAPTER 1

Introduction

In some of the first attempts to computationally obtain solutions to governing equations of fluid dynamics problems, finite difference (FD) methods have provided a simple and effective way to discretize and model partial differential equations. In FD, approximations to the derivatives are made with difference equations, which are simple to repeat and easy to produce. And higher-order approximations are achieved by reducing truncation errors of the numerical formula. Subsequently, the development of mathematical foundations used in FD has changed the immediate extent to which computational fluid dynamics (CFD) shapes the modern scientific and engineering communities for years to come.

The aerospace engineering community has been at the forefront of modern CFD developments since the 1970s. It has been continually providing some of the state of the art techniques utilized across all disciplines of CFD today. Some of the most important developments are the advent of *high-resolution* and *upwind* methods. These developments along with the use of *Riemann solvers* in finite volume (FV) methods have enabled substantial improvements to be made towards modeling nonlinear systems of conservation laws. The core principles of high-resolution upwind methods still have a lasting influence on the way that numerical methods are designed to date.

Although high-resolution upwind methods have been the bedrock on which most current industry standard compressible CFD solvers are designed, it is our belief that these methods often do not provide an adequate basis that modern complex scientific and engineering problems demand. There have been many theoretically intriguing developments in modern CFD and numerical methods. Many new methods now aim to address these demands: having a high-order accurate property, having a compact numerical stencil for a suitable parallel computing framework, being able to properly model multidimensional processes, having a maximum stability requirement, and being suitable for unstructured grid types. Despite all these efforts to develop as many of the new improvements in CFD research communities, the current trend of industrial and commercial flow solvers is still influenced largely by the fundamental developments made in the 1980s. The main reason

for high-resolution upwind methods still being the workhorse in modern CFD solvers is the *robustness* property. These methods are not only robust but also easily adapted across a wide variety of purposes; making them quite versatile when a quick turnaround time is essential.

With the knowledge and experience learned from the past three decades, the time is now ripe to culminate the best practices of the past time successes in high-order numerical methods. We introduce the Active Flux (AF) method for nonlinear hyperbolic conservation laws in this thesis. In the new method, we specifically address the new features that modern scientific and engineering problems demand. Achieving all of the good properties is made possible by two main features of AF. One is its unique two stage update process; the other is the operator splitting approach. As a result, the AF method is a third-order compact numerical scheme whose design principles are rooted in proper 'multidimensional' physics processes. Furthermore, it is specifically designed for unstructured grids which align well with the future trend of massively parallel computing. In this dissertation, some specific developments of the AF method are introduced such as nonlinear advection, a conservation treatment, and a physically motivated operator splitting method for nonlinear conservation systems.

1.1 Thesis Overview

A framework for the multidimensional AF method for hyperbolic conservation laws will be established throughout this dissertation. In Chapter 2, important design principles of AF are presented with respect to some past and current examples of high-order numerical methods for hyperbolic conservation laws. In Chapter 3, we lay the groundwork for advective AF schemes in one- and two-dimensional linear problems. Some detailed linear analyses are also presented. In Chapter 4, we present the extension of AF for nonlinear scalar conservation laws where one- and two-dimensional variations of Burgers' equation are considered. In Chapter 5, we present new approaches to limiting. We recommend that the limiting procedures be based on the property of problems and provide a few simple examples. In Chapter 6, the AF method for nonlinear advective conservation system is presented. We employ a pressureless version of the Euler equations to develop and extend the advective scheme for multidimensional conservation systems. In Chapter 7, the AF method for conservation systems with different physical processes is presented. The Euler system is composed of advective and acoustic processes. Although these physical processes behave similarly in one dimension, they pose different disturbance patterns in multidimensions. We develop distinct numerical methods for the nonlinear advective and

acoustic processes and present novel operator splitting methods that combine these processes to solve the Euler equations.

CHAPTER 2

Active Flux Method

2.1 Relevant Schemes

We begin by describing some related schemes in the context of the AF advection scheme.

2.1.1 Scheme V

In the penultimate paper of the five-part series titled "Toward the ultimate conservative difference scheme" [84], van Leer demonstrated six different ways to achieve better than first-order accurate finite difference schemes for the one-dimensional convection problem. One of these, Scheme III, became the foundation for what is now known as Monotonic Upstream-centered Scheme for Conservation Laws (MUSCL) [85], and have been popularized for many years since. On the other hand, third-order finite difference schemes, Schemes IV-VI, introduced at the same time have not had an as significant role in modern compressible flow solvers.

Of these three third-order schemes, Scheme V features some noteworthy properties. It introduced additional evolution variables at element interfaces to increase the solution approximation polynomial order, and to obtain higher numerical accuracy. The additional degrees of freedom are shared with the neighboring elements to produce a *compact* and *continuous* numerical scheme for advection problems. The compactness property of the scheme, in the sense of reduced stencil size, reduced the number of total degrees of freedom in computation. And the continuous solution reconstruction meant that the Riemann solver was no longer needed. The scheme still remained upwind because a semi-Lagrangian method was used to advance the solution.

During recent years, high-order numerical schemes have gained steady attention in the CFD communities due to proclaimed efficiency. And the current trend of massively parallel computations favors numerical schemes that are compact in nature. These factors have

renewed some interest in the potentials of Scheme V. Some of the closely related schemes are discussed next. In fact, the topic of this dissertation, the Active Flux method, is derived from the original Scheme V.

2.1.2 CABARET

The Compact Accurately Boundary-Adjusting high-REsolution Technique (CABARET) was proposed in 1998 by Goloviznin and Samarskii [27] for linear advection. It was originally introduced as an explicit three-time level difference scheme and is closely related to the second-order upwind Leapfrog scheme. The explicit three-time level scheme, like the upwind Leapfrog, requires one initialization step. A new formulation has since been introduced by Karabasov [41] in a predictor-corrector form. This form of CABARET effectively removes the need for an initialization process. It has also made the extension to nonlinear hyperbolic conservation laws easier by enabling a variation in time step sizes in predictor-corrector stages.

In the predictor-corrector form of the CABARET scheme, two sets of distinct values are utilized: conservative and interface flux variables. In the predictor stage, the conservative variables are updated to an intermediate time step using the available interface flux values. Before the corrector stage, the flux values are extrapolated to the full-time step. Then, the extrapolated interface flux values are used to update the conservative variables to the full-time step in the correction stage. The CABARET scheme achieves second-order accuracy. Because of the close resemblance in update procedures, the predictor-corrector form of CABARET is regarded as a close second-order cousin of the AF scheme. Eymann showed in [21] the existence of a class of second-order schemes, of which CABARET is the unique member with no dissipation and Scheme V is the unique member with third-order accuracy.

2.1.3 Multi-Moment Finite Volume Method

A class of methods called the multi-moment finite volume method has been introduced for nonlinear hyperbolic conservation laws by numerous authors Akoh *et al.* [5, 6], Ii and Xiao [38, 37], Yabe *et al.* [94], and Takewaki *et al.* [82], and see the references therein. These authors are all part of groups from Osaka University and Tokyo Institute of Technology. Alternatively, the multi-moment finite volume method is known as the CIP method where the initials CIP are applied to a number of different but somewhat related methods: ‘constrained interpolated profile’, ‘cubic-interpolated propagation’ or ‘cubic interpolated pseudo-particle’. It was originally proposed as a third-order nonconservative semi-Lagrangian scheme for the one-dimensional wave propagation problem. It introduced

a cubic solution interpolation function with one of its constraints defined by a gradient parameter that doubles as a parameter to control numerical oscillations [82]. In the most recent form, the multi-moment finite volume method has defined itself as a third-order and compact numerical method for nonlinear hyperbolic problems in one and two dimensions using unstructured grids. Two moments are introduced in this method: the volume integrated average (VIA) and the point value (PV) moment. Although these are introduced with different terminologies, they are in essence the conservative cell average and interface values used in finite volume methods. Importantly, solution gradients are additionally introduced as a way to control numerical oscillations and to define interpolation functions.

In one-dimensional problems, the interface variables, or PV moments, are obtained from a point-wise Riemann solver. A cubic solution reconstruction is constructed within each cell and provides necessary data for a point-wise Riemann solver. The cubic constraints are provided within each cell: left and right interface values, a cell average variable, and the first-order derivative at a cell centroid. This cubic reconstruction is compactly defined in each element but may be discontinuous across cell interfaces because interface values are not shared. Conservative fluxes are evaluated from previously updated interface variables. The cell average is conservatively evaluated by a finite volume scheme. The first-order derivative at the centroid, one of the cubic constraints, serves as a parameter to control and remove numerical oscillations in the numerical scheme.

The two-dimensional update process for interface and cell average variables remains largely the same. However, a quadratic reconstruction is introduced instead of a cubic reconstruction. The quadratic reconstruction is provided by three triangle vertex values, a cell average and two first-order derivatives at a cell centroid. The quadratic reconstruction is discontinuous across all neighboring elements. The local Lax-Friedrichs Riemann solver is used to find point value moments. The point value moments are used to find interface fluxes to be used in conservation. Two first-order derivatives defined at cell centroids are used to control and remove numerical oscillations. The first-order derivatives are defined by simple difference equations of point values.

It, however, seems highly improbable to achieve a third-order accurate version of the multi-moment finite volume scheme in three dimensions in its current formulation. The number of constraints required to achieve a piecewise quadratic interpolation function in three-dimensional simplex is ten. A straightforward extension of the two-dimensional design principle gives: four vertex interface variables, a cell average, and three first-order derivatives at cell centroid. These are two constraints short of achieving a quadratic interpolation function for a three-dimensional quadratic polynomial.

Because the two-dimensional multi-moment finite volume method is based on a two-

dimensional quadratic interpolation function to achieve third-order accuracy, it draws some interesting comparison to the AF method, despite the shortcoming in the three-dimensional application. Furthermore, the introduction of derivatives in reconstruction suggests a very interesting prospect for limiting.

2.1.4 Discontinuous Galerkin Method

The original discontinuous Galerkin (DG) finite element method was introduced by Reed and Hill [65] for solving neutron transport equation. It had demonstrated the property of DG methods that the approximation solution could be achieved on an element by element basis when they are suitably ordered in the characteristic direction. DG methods have since been developed and widely used in many fields of computational science including the compressible and incompressible flows in CFD, and a variety of other applications, see [13, 14].

DG methods have gained a wide reception for hyperbolic problems for the following favorable properties: locally conservative, stable, and high-order accurate. They have made DG methods suitable for handling complex geometry, irregular meshes. Moreover, local solution adaptability by having different discretization orders in different elements has been an attractive feature of DG methods.

The notion of conservation is deeply embedded in the space discretization process called the discontinuous Galerkin method. A discontinuous approximate solution is sought such that when it is restricted to one element, it belongs to the finite dimensional space by use of test functions. The locally accurate approximate solutions are discontinuous across elements, and it conveniently leads to the use of approximate Riemann solvers for upwinding. The stability of DG methods depends closely on the stability of time discretization method which is advantageous for extending the method to high-order accuracy. For time discretization methods, either explicit multistage time stepping methods such as the Runge-Kutta or implicit pseudo time stepping such as the backward Euler, Adam-Bashforth and implicit Runge-Kutta methods are commonly used.

Despite these features that enable users to pursue high-order accurate and locally adaptable numerical simulations, there are largely two issues that need resolving before DG methods become a widely accepted practice in aerospace industrial applications. The main issue is the lack of robustness for nonlinear problems to reach steady state solutions [7]. The second issue is the computational expense related to high memory requirements. For nonlinear problems, DG requires enough degree(s) of freedom (DOF) in each element for a sufficient order of approximation. This is true of any high-order methods, but it is worth

noting that sharing of DOF between neighboring elements would greatly benefit large-scale problems. Hybridizable DG methods [60] and embedded DG methods [61] were introduced to address the high memory requirements. However, these DG schemes still require either a costly pseudo time stepping and/or an inversion of mass matrices for unsteady flow simulations.

2.1.5 Flux Reconstruction Method

The flux reconstruction (FR), or correction procedure using reconstruction, approach was introduced by Huynh [35] as a new high-order accurate method that serves as a bridge between two existing approaches of discontinuous Galerkin (DG) and staggered-grid [42]. Because it employs differential forms of the governing equations instead of integral forms, it is able to provide a simple and economical framework to derive arbitrarily high-order numerical schemes for conservation laws. It is considered to be a unifying framework for many similar high-order methods such as spectral volume (SV), spectral difference (SD) and DG by changing the way continuous reconstructions are constructed.

The one-dimensional approach for the method is described here. Two-dimensional extensions for quadrilaterals [35] and simplex elements [36] are referred to the respective references. A solution reconstruction is described by some polynomial of order p. This solution reconstruction, in general, does not need to be continuous across adjacent cells, therefore, the flux function constructed from it is also discontinuous. The differential form of conservation laws cannot be solved from this discontinuous flux reconstruction as it ignores the interactions with neighboring data. For proper physics, an upwind flux must be obtained at all interfaces. To obtain a *continuous* p^{th} order flux reconstruction so that the resulting derivative retains an order of $p-1$, correction functions g are introduced. These correction functions seek to correct the left and right flux interface points by introducing jumps (the difference between the flux evaluated from the discontinuous flux reconstruction and the common upwind flux from the Riemann problem), but they do not modify any other solution points in the process. Once the continuous flux reconstruction and its derivative are known, the differential form of conservation laws can be evaluated at solution points using some explicit time stepping scheme. As a result of its differential formulation, no mass matrix is created, and it leads to a more economical scheme. Different correction functions g are formed based on various orthogonal polynomials and they define other high-order schemes, see [35] for the list of schemes and g.

For a specific application in DG, the FR approach results in a quadrature free formulation because the continuous flux reconstruction depends only on correction functions

defined on solution points. It is especially advantageous because fluxes are only evaluated at solutions points, and there is no need to evaluate the mass matrix or quadrature points within elements. Also, certain choices of the correction function g can lead to a larger stability limit than the traditional DG scheme. However, it is observed that the increase in stability limit is accompanied by one order reduction in accuracy. While there are many choices available for g, there doesn't seem to be a guideline for the best choice; prompting a problem dependent nature of g. In addition, steep correction functions used to achieve higher-order schemes introduce a trade-off between the accuracy and CFL limit; the higher the accuracy, the smaller the CFL limit.

The FR method provides a unifying framework to bridge the gap between a finite element method to a spectral difference scheme. In doing so, the favorable aspects of both methods were retained while most of the unfavorable aspects were discarded; namely increased stability limits and reduction in computational cost. While not perfect, the FR scheme exemplifies the future direction of high-order methods that aim to achieve all the virtues while selectively discarding the vices.

2.2 Active Flux Design Philosophy

We have seen a number of good numerical methods such as those discussed earlier in this section, yet they have not sufficiently answered our quest to a satisfactory multidimensional fluid dynamics method. An apt excerpt from Richtmyer and Morton [66] in 1967 foreshadowed the current trend in multidimensional CFD methods.

> We seem to be far from having any universal methods for multidimensional fluid dynamics, and it seems likely that for some time there will be nearly as many methods as problems. . . . It is our opinion that we are more often limited by inadequacies of the mathematical (including numerical) methods than by inadequacies of the computers.

However, we believe that we can address the inadequacies of past numerical methods with AF. Here is the list of properties that we believe are missing in many numerical methods so far, and the AF method aims to deliver. New numerical methods should

1. have a compact supporting stencil
2. have high-order accuracy
3. be designed to have a maximally stable property that the Courant-Friedrichs-Lewy (CFL) condition would allow

4. reflect multidimensional physics processes properly

Now, the design philosophy of AF is described.

2.2.1 Compact Support

Some early finite difference methods to solve partial differential equations to higher than the first-order accuracy involved expanding the numerical domain of dependence. For time explicit numerical schemes, it meant larger numerical stencil sizes. It was soon realized that high-order accurate methods could also be obtained by making smart decisions using locally available information such as high-order moments. These modern high-order *compact* schemes were pioneered by van Leer [84].

There were some notable finite volume method extensions of van Leer's Scheme V. The piecewise parabolic method (PPM) by Colella and Woodward [16] was introduced for gasdynamic systems, and it involved a semi-Lagrangian evolution of interface values. Discontinuous interface solutions were introduced across interfaces and the Riemann solver was used to reconcile the discontinuity. Also, it was primarily used for structured Cartesian grids due to the complex nature of the algorithm. The piecewise-parabolic dual-mesh method by Huynh [34] achieved an extension of van Leer's Scheme V to the one-dimensional Euler equations using a dual-mesh configuration. The dual-mesh property produced a twice as large error as the corresponding single mesh upwind scheme (Scheme V). In addition, the dual-mesh, or staggered mesh, configuration made it difficult to extend to unstructured multidimensional schemes. As a result, it was developed for one-dimensional hyperbolic problems.

A widely regarded high-order finite element method in the compressible flow community is the DG method [13]. It has become a popular numerical method for hyperbolic conservation laws because of its ability to deliver a high-order approximation of a solution in a local and compact sense. Because the high-order approximation solution polynomials are available in individual elements, DG methods can be used with unstructured grid types. Other high-order compact stencil supporting numerical methods have been developed using various spectral methods such as spectral element (SE) and spectral volume (SV) [87, 88, 80]. Discretization methods differ, however, the core concept is to provide a highly accurate yet compact numerical method. Despite their theoretical popularity, they have been slow to be industrially accepted due to two main issues: the robustness for reaching steady state in nonlinear problems and high memory requirement from a large number of independent DOF per element.

Some contrasting points to compactness property are made from going expansive in

stencil support. There are many numerical methods based on the Essentially Non-Oscillatory (ENO) notion. ENO [29] and Weighted Essentially Non-Oscillatory (WENO) [40] schemes are well known in the shock capturing community for their property to produce essentially non-oscillatory shock capturing solutions. While the numerical results produced by these schemes are truly admirable, they are subject to some contentions of their own. The reason is a threefold one. Firstly, the expansive stencil footprint makes it harder to implement numerical boundary conditions. The main principle used in ENO and WENO methods to construct the least oscillatory reconstruction leads to forming a large numerical stencil. Secondly, referring from an extensive reference of the solution does not agree well with the physical observation that hyperbolic problems are 'local' problems. Lastly and most importantly, achieving a large-scale algorithm would be challenging because of the amount of data transfer needed.

The notion of compact supporting schemes also places an interesting perspective on how a slope/flux limiting procedure must take place. The current inadequacies in limiting can be seen across many high-order and multidimensional numerical methods. This is perhaps due to the lack of good understanding of multidimensional methods. While it may seem easier to use ENO or WENO methods to construct the smoothest non-oscillatory solution reconstructions, a counterpoint is made that the limiting process is just as much part of a numerical method and deserves the same level of attention. As a consequence, the new kind of limiting process should be rooted within the confines of the compact numerical stencil.

The future outlook of CFD and the computational science at large depend heavily on a massively scalable and parallel computing paradigm. One of the main requirements of any good numerical methods for the future engineering and scientific endeavor is an ability to provide a locally manageable algorithmic framework. The compact stencil is one of the core aspects to be considered in numerical methods development.

2.2.2 Third-Order Accuracy

High-order numerical methods, usually defined as methods that achieve third- or higher-order accuracy, have received unwavering attention from all aspects of computational sciences in recent years. They are more efficient than the lower-order counterpart based on the computational cost to achieve the *same level of error*. Roughly speaking, reducing the level of error by a factor of 16 requires a first-order method to increase its number of degrees of freedom (DOF) by a factor of 16^4 and a third-order method to increase its DOF by a factor of 2^4 in a full three-dimensional (plus time dimension) simulation. The number

of DOF one scheme must increase in order to reach the same level of error is staggering if the problem at hand is a large-scale complex engineering project. This demonstrates a quantitative case for using high-order methods.

However, a more fitting question to ask is one regarding the efficiency of high-order schemes. For a fixed order of accuracy, the effectiveness of the scheme can be measured based on the *measure of work* to achieve the desired accuracy. For a third-order accurate two-dimensional solver such as DG with spatial polynomial order $p = 2$ (DG2), the number of DOF per element needed to achieve that level is 6. For the AF scheme, it would only require 3 due to the sharing of DOF between immediate neighboring elements. While the measure of work cannot tell the whole story nor be used universally for different schemes, it provides us with some sense of computational cost.

A more pragmatic point of view is presented in the context of the state of the art production codes available in aeronautics and automotive industries alike. Many production level codes are largely limited to second-order numerical schemes. In many cases, second-order methods can be used to model smooth flows with few nonlinear features. For flows with strong nonlinearity or around complex geometry, however, second-order methods may present inadequate performance to capture these features. The shortfall in second-order schemes is often resolved via mesh refinements in regions of trouble. However, we claim that the practice of employing mesh refinements with low-ordered schemes is not only cost ineffective but also quite illogical on the fundamental level.

Fourier analysis of differently ordered numerical schemes reveals that the truncation errors from even-ordered schemes tend to accentuate oscillatory behaviors in the solution much more than odd-ordered schemes [10]. In certain circumstances, even unlimited results from third-order schemes (anecdotal evidence from AF) are quite acceptable compared to the limited results from second-order schemes. Superior dispersion and dissipation error properties of odd-ordered methods have also been previously observed and enjoyed by DG schemes. Since the error properties of third-order schemes are superior to those of second- or lower-ordered schemes, the amount of computational effort needed in limiting is theoretically lesser as well. This means that limiters for third-order schemes don't have to be subject to restrictive principles such as total variation diminishing (TVD) or bounded (TVB).

High-order numerical methods are on their rise to become the new norm in CFD. From the practical standpoint, third-order numerical schemes like the AF and DG2 methods provide the right amount of compromise between methods of low- and high-order accuracy. They provide a more efficient way to achieve the desired outcome at a slight increase in the computational cost.

2.2.3 Maximal Time Accuracy and Stability

In the early development of finite difference schemes for multidimensional conservation laws, many numerical methods were based on the work of Lax and Wendroff [46]. These schemes were not without merits: they are time explicit, and the second-order discretization based on central differencing allowed for a systematic extension of the one-dimensional discretization to two or three dimensions. The original one-step two-dimensional Lax-Wendroff scheme had very restrictive stability limits, which soon were superseded by two-step variations. The two-step Lax-Wendroff schemes of Richtmyer [66] and MacCormack [56] presented relaxed stability limits. The two-step structure also made the extension to nonlinear problems much easier.

The one- and two-step versions of dimensionally unsplit Lax-Wendroff schemes suffered from some accuracy loss for nonlinear problems. The non-commutative nature of operators in nonlinear problems meant these schemes could not take into account of some nonlinear terms without special mathematical or numerical procedures. The restrictive nature of dimensionally unsplit Lax-Wendroff schemes was resolved by the introduction of the fractional step method by Janenko [39] and the dimensional splitting by Strang [79]. But it was soon found that these new forms of Lax-Wendroff schemes failed to produce marked resolution of multidimensional features [71].

More recently, Lax-Wendroff rotated Richtmyer schemes [58] were shown to preserve discrete vorticity in acoustic problems. These schemes have achieved the discrete vorticity preservation along with the maximal stability that CFL would allow by exploiting the ambiguous formulation of the Lax-Wendroff scheme in multidimensions. In two dimensions, the Lax-Wendroff scheme has nine stencil points but only six are required for a second-order discretization. Therefore, a new scheme could form by taking an advantage of the property of the Lax-Wendroff scheme. And this is of particular interest for the AF scheme. The Lax-Wendroff scheme, precisely because of its ambiguity in multidimensional formulations, provides a surprisingly solid basis for constructing time explicit numerical schemes. In our treatment of the Euler equations, the acoustic process is treated with a symmetric stencil, but the advective system, which is the focus of this thesis, uses a method with an asymmetric numerical stencil. One of the challenges in the design of a stable and high-order numerical method is how to provide a suitable numerical framework on which physically accurate numerical models are built. The Lax-Wendroff scheme provides that.

2.2.4 Multidimensional Solver

Numerical methods for hyperbolic conservation laws have followed a somewhat narrow path for the past five decades since the invention of Godunov scheme [26]. Riemann solvers, which are inherently one-dimensional by construct, have been the dominant method of choice for solving hyperbolic problems including for many well-regarded multidimensional numerical methods such as DG and WENO. Despite having been widely used for multidimensional problems, there are fundamental objections to Riemann solvers and the type of data they use to represent multidimensional hyperbolic problems. For instance, taking discontinuous reconstruction data to be a genuine feature in multidimensional problems introduces strong one-dimensional waves that travel perpendicular to the arbitrary cell boundary [71] which may present a strong dependency on the computational grid rather than the solution.

There have been many numerical methods motivated to better represent the multidimensional physics processes, that are not based on the dimensional splitting method or the Riemann problem. These methods were based on the 'multidimensional Riemann problem' concept in which the multidimensional upwind process is applied to some form of the Euler equations. Among many, there are three notable multidimensional upwind methods. In [15], Colella introduced the multidimensional corner treatment (MCT) method. The MCT method aimed to introduce the multidimensional upwinding principle by introducing a corner transport feature, which was missing from many methods employing a dimension-by-dimension approach. In [48], LeVeque suggested the 'multidimensional' method that solved different Riemann problems for each direction of the problem. This was accomplished by decomposing the multidimensional data into different one-dimensional initial data for each direction. In [24], Fey introduced the method of transport (MOT). In The MOT, multidimensional fluxes were decomposed into vectors of fluxes in different directions. Then, these fluxes were transported according to an upwind transport method. In a sense, the MOT resembled flux vector splitting methods [86, 78]. Although these numerical methods have aimed to address the multidimensional aspect of problems more rigorously by implementing some multidimensional processes, they have not proved to perform significantly better than dimensional split methods such as MUSCL, DG, and WENO. In fact, these aforementioned multidimensional methods are computationally complex, which only adds to the overall cost for a marginal improvement in performance [73].

It was discussed by Roe [73, 74] that multidimensional physical processes, especially the acoustic process, in nonlinear systems are not suitably modeled by an inherently one-dimensional process such as the Riemann problem. The acoustic process, instead, must be modeled by a numerical method that considers a proper domain of dependence, which is

a symmetric circular domain in two dimensions. Therefore, we must revise the notion of upwinding in multidimensional problems.

In the early 1980's, a steadfast movement toward genuinely multidimensional solvers was initiated by Roe [69]. A class of schemes called residual distribution (RD) or the fluctuation splitting (FS) schemes were introduced to provide multidimensional steady state solutions of various systems [19]. The RD scheme tried to seek genuinely multidimensional numerical methods without resorting to one-dimensional Riemann solvers [2, 3]. However, the RD scheme in some sense relied on 'patterns' that exist in flows [74] and didn't target the significance of each flow process. One particularly intriguing development was by Rad [64] where the Euler equations were decomposed into an elliptic and a hyperbolic part. This decomposition method produced remarkable numerical solutions for various flow regimes, most notably a well behaved numerical entropy. This was attributed to correctly assigning appropriate numerical methods for different flow processes. However, the method was short lived because it could not be extended to three dimensions and the unsteady formulation was not available. Nevertheless, the RD scheme has certainly provided inspirations for multidimensional flow solvers, and offered a tantalizing outlook for 'going multidimensional'.

2.3 Active Flux Method Description

The Active Flux (AF) "method" is a class of methods, just as the FV or the discontinuous Galerkin (DG) method is. We describe the main components of the AF method for hyperbolic conservation laws.

In the AF method, we solve the hyperbolic conservation laws in two stages. In the first stage, the nonconservative formulation of the conservation law system is introduced to provide more meaningful numerical methods that reflect proper physical processes. The advantage of introducing the first stage is that the numerical method does not need to be conservative. This opens up new opportunities for flexible numerical schemes. In the second stage, individually updated interface flux values are gathered and conservation is enforced by a finite volume method.

2.3.1 Nonconservative Formulation

The hyperbolic conservation system expressed in conservation form bears little relevance for how numerical methods should be applied. The nonconservative form of a system can present a clear composition of the governing equation, and it is an integral part of the AF

method.

There are many ways to consider the nonconservative form of governing system of equations. The one that makes the most sense is the form in which the physical process of physical variables is revealed. Physical, or primitive, variables refer to quantities that appeal to our physical intuition. We will adopt the physical, primitive or nonconservative variables throughout the dissertation, choosing, in particular $(\rho, \mathbf{v}, p)$. Governing equations written in these variables make the physics especially transparent.

In scalar conservation laws, the physical variables are also the conservative variables. But the distinction becomes more important when nonlinear systems are considered. For example, the one-dimensional Euler equations can be written in the conservative form.

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0 \tag{2.1}$$

where

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho u \\ \rho E \end{pmatrix}, \qquad \mathbf{F} = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho u H \end{pmatrix}, \tag{2.2}$$

and $E = e + \frac{1}{2}\mathbf{v}\cdot\mathbf{v}$ and $H = E + \frac{p}{\rho}$ are the specific total energy and enthalpy. We propose to split the Euler equations into advective and acoustic components. The approach, presented in [100, 92, 45], has been to write the flux as a sum of an advective and an acoustic part.

$$\mathbf{F} = \mathbf{F}_{\text{AD}} + \mathbf{F}_{\text{AC}} \tag{2.3}$$

where

$$\mathbf{F}_{\text{AD}} = \begin{pmatrix} \rho u \\ \rho u^2 \\ \rho u E \end{pmatrix}, \qquad \mathbf{F}_{\text{AC}} = \begin{pmatrix} 0 \\ p \\ up \end{pmatrix}. \tag{2.4}$$

Although this makes intuitive sense, the flux Jacobian matrices of $\mathbf{F}_{\text{AD}}$ and $\mathbf{F}_{\text{AC}}$ have eigenstructures that are not closely related to the physics. The advective flux Jacobian matrix is analyzed first.

$$\mathbf{A}_{\text{AD}} = \frac{\partial \mathbf{F}_{\text{AD}}}{\partial \mathbf{u}} = \begin{pmatrix} 0 & 1 & 0 \\ -u^2 & 2u & 0 \\ -uE & E & u \end{pmatrix} \tag{2.5}$$

It has three eigenvalues equal to $\lambda^{\text{AD}}_{1,2,3} = u$ and is rank deficient having only two right

eigenvectors.

$$\mathbf{r}_1^{\text{AD}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad \mathbf{r}_2^{\text{AD}} = \begin{pmatrix} 1 \\ u \\ 0 \end{pmatrix} \tag{2.6}$$

The acoustic flux Jacobian matrix is shown as follows.

$$\mathbf{A}_{\text{AC}} = \frac{\partial \mathbf{F}_{\text{AC}}}{\partial \mathbf{u}} = \begin{pmatrix} 0 & 0 & 0 \\ \frac{1}{2}(\gamma-1)u^2 & -(\gamma-1)u & \gamma-1 \\ -\frac{pu}{\rho} + \frac{\gamma-1}{2}u^3 & \frac{p}{\rho} - (\gamma-1)u^2 & (\gamma-1)u \end{pmatrix} \tag{2.7}$$

whose eigenvalues are $\lambda_{1,2,3}^{\text{AC}} = 0, \pm\sqrt{\frac{(\gamma-1)p}{\rho}}$ with eigenvectors

$$\mathbf{r}_1^{\text{AC}} = \begin{pmatrix} 1 \\ u \\ \frac{1}{2}u^2 \end{pmatrix}, \qquad \mathbf{r}_{2,3}^{\text{AC}} = \begin{pmatrix} 0 \\ \frac{(\gamma-1)\rho u \pm \sqrt{(\gamma-1)p\rho}}{(\gamma-1)\rho u^2 - p} \\ 1 \end{pmatrix}. \tag{2.8}$$

The eigensystem analysis of the acoustic Jacobian reveals some physically correct observations. A purely advective disturbance, $\mathbf{r}_1^{\text{AC}}$, does not lead to an acoustic response and will propagate at the correct wave speed $\lambda_1^{\text{AC}} = 0$. The acoustic disturbances, $\mathbf{r}_{2,3}^{\text{AC}}$, however, do not propagate at correct wave speeds but the density is not affected by them.

In fact, there is no satisfactory splitting of the fluxes into advective and acoustic components. The problem lies in the energy equation. Using the flux vector split approach, the energy equation is shown in divergence form.

$$\frac{\partial \rho E}{\partial t} + \nabla \cdot (\rho u E) + \nabla \cdot (up) = 0 \tag{2.9}$$

where the term corresponding to the acoustic flux divergence $\nabla \cdot (up)$ splits into two nonconservative terms $u \cdot \nabla p$ and $p \nabla \cdot u$. The first is an advective term, and the second describes an acoustic effect. For this reason, this particular flux vector splitting approach cannot result in a conservation formulation. Our method, instead, employs a nonconservative formulation to evaluate the fluxes. The Euler equations, Eq. (2.1), are written in nonconservative form.

$$\frac{\partial \mathbf{q}}{\partial t} + \mathbf{A}^q \frac{\partial \mathbf{q}}{\partial x} = 0 \tag{2.10}$$

where

$$\mathbf{q} = \begin{pmatrix} \rho \\ u \\ p \end{pmatrix}, \qquad \mathbf{A}^q = \begin{pmatrix} u & \rho & 0 \\ 0 & u & 1/\rho \\ 0 & \rho a^2 & u \end{pmatrix}, \tag{2.11}$$

and a is the speed of sound. Further decomposition of $\mathbf{A}^q$ leads to the following.

$$\begin{aligned} \mathbf{A}^q &= \mathbf{A}^q_{\text{AD}} + \mathbf{A}^q_{\text{AC}} \\ &= \begin{pmatrix} u & 0 & 0 \\ 0 & u & 0 \\ 0 & 0 & u \end{pmatrix} + \begin{pmatrix} 0 & \rho & 0 \\ 0 & 0 & 1/\rho \\ 0 & \rho a^2 & 0 \end{pmatrix} \end{aligned} \tag{2.12}$$

This process of decomposing the Euler equations reveals some details about the evolutionary properties of the physical variables. For each physical variable, there are two distinct processes: an advective $\mathbf{A}^q_{\text{AD}}$ and an acoustic $\mathbf{A}^q_{\text{AC}}$ component.

$\mathbf{A}^q_{\text{AC}}$ is a singular matrix, which has made some authors avoid this decomposition choice. We take the position that the null vector $\mathbf{x} = (1, 0, 0)^T$ is a physically correct assertion that acoustic disturbances do not directly affect the density.

2.3.2 Conservation Laws

The second stage of the AF method is the conservation stage. A conservation law is a mathematical statement concerning conservation of quantities such as mass, momentum, and energy. Conservation is an important notion that allows us to accurately model flows with nonlinear features like shocks or discontinuities. It is particularly useful to consider an integral form of conservation laws.

$$\int_\Omega \frac{\partial \mathbf{u}}{\partial t} dV + \int_\Omega \nabla \cdot \mathbf{F} dV = 0 \tag{2.13}$$

$$\frac{d}{dt} \int_\Omega \mathbf{u} dV + \oint_{\partial\Omega} \mathbf{F} \cdot \mathbf{n} dS = 0 \tag{2.14}$$

where the conservative variable vector and flux tensors are denoted as

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{pmatrix}, \qquad \mathbf{F} = \begin{pmatrix} F_{1,1} & \dots & F_{1,d} \\ F_{2,1} & \dots & F_{2,d} \\ \vdots & \ddots & \vdots \\ F_{N,1} & \dots & F_{N,d} \end{pmatrix}, \tag{2.15}$$

and N is the number of variables, d is the number of dimensions. The significance of conservation law formulation Eq. (2.14) lies in the presence of the surface integral. Changes to the conserved variables, $\mathbf{u}$, inside Ω only depend on the surface flux values across $\partial\Omega$.

Consistency is an important notion in designing conservation schemes, especially for conservative fluxes. The conservative fluxes in AF are derived from a preliminary nonconservative computation. Therefore, fluxes are consistent in the sense of the Lax Wendroff Theorem, that is, if all input states are equal, the numerical flux corresponds to that state.

2.3.3 Active Flux Algorithm

For some simple problems, scalar conservation laws, certain AF stages are not needed. Where they are needed, there are usually options available. In this thesis, the AF method for conservation laws of the advective nature is primarily developed. But the details of the acoustic aspect will be given in Section 3.1 for the basic theory and in Section 7.2.2 for basic numerical techniques. The general algorithm for the AF method is presented for an easy understanding.

Algorithm 2.1 Active Flux Method

procedure RECONSTRUCTION
    for all Conservation element j **do**:
        Define high-order spatial reconstruction coefficients
    end for
end procedure
procedure EVOLUTION(Nonconservation Update)
    for all Solution update on quadrature point i **do**:
        Evolve physical variables according to *physically* distinct numerical schemes:
            • Advective process
            • Acoustic process
    end for
end procedure
procedure CONSERVATION UPDATE
    for all Conservation element j **do**:
        for all Element interface l **do**:
            Evaluate conservative and consistent flux
        end for
        Evaluate conservation law
    end for
end procedure
if Nonlinear System **then**
    procedure RECONCILIATION
        for all Conservation element j **do**:
            Reconcile between conservation and physical variables
        end for
    end procedure
end if
Store data for next iteration

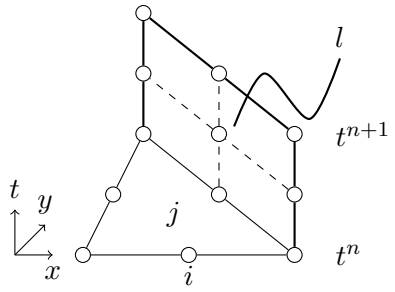


Figure 2.1: Active Flux element

A two-dimensional element is shown to describe the concepts necessary in the AF method in Fig. 2.1. The concept is shown for a two-dimensional case but can be easily extended to other dimensions. The element j is surrounded by quadrature points i defined on the Lagrange basis locations. There are six points in two dimensions. An element interface l is defined by total nine quadrature points: three quadrature points from each time step (t^n, $t^{n+1/2}$ and t^{n+1}).

CHAPTER 3

Linear Advection Problem

We discuss an initial value problem for a first-order linear partial differential equation.

$$\begin{aligned} &\frac{\partial u}{\partial t} + \mathbf{a} \cdot \frac{\partial u}{\partial \mathbf{x}} = 0 \\ &u(\mathbf{x}, 0) = u_0(\mathbf{x}) \end{aligned} \tag{3.1}$$

where $\mathbf{a}$ is a constant coefficient vector and $\mathbf{x}$ is a linearly independent vector. This linear partial differential equation, known as linear advection, has traditionally been used to understand and design numerical schemes for hyperbolic partial differential equations [47]. In particular, this chapter is devoted to a type of numerical scheme for advection called the upwind scheme that mimics both the physical and mathematical character of the problem.

The AF method foundations of the one- and two-dimensional linear advection problems have been previously established by Eymann [21] in his Ph.D. dissertation, nevertheless, we revisit them for the sake of continuity. The specifics of Eymann's contributions are detailed in Sections 3.2 and 3.3.

The current author's contribution builds on the foundations of Eymann to develop the AF method for general nonlinear multidimensional problems, including nonlinear systems as shown throughout the dissertation. In this chapter, noteworthy contributions of the current author include, but are not limited to, the identification of potentially unstable behavior of the multidimensional advection scheme featured in Section 3.3.7 and some remedies for it.

3.1 Characteristics

The theory of characteristics plays an important role in the development of numerical methods for hyperbolic conservation laws. The most important aspect of the theory is the equivalence of partial differential equations and characteristic equations.

We first consider the one-dimensional linear advection equation.

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0 \tag{3.2}$$

for definiteness we define $a > 0$. We begin by introducing a parameter, s, in Eq. (3.2) so that we form characteristic lines on which solutions are constant.

$$\frac{du}{ds} = \frac{\partial u}{\partial t}\frac{dt}{ds} + \frac{\partial u}{\partial x}\frac{dx}{ds} \tag{3.3}$$

We define the slope as $\frac{dx}{dt} = a$, or by the chain rule $\frac{dx}{ds} = a$ and $\frac{dt}{ds} = 1$, then we obtain the following.

$$\frac{du}{ds} = \frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0 \tag{3.4}$$

Thus, we see that every line $u(x,t)$ generated by a parameter family of characteristic lines is an integral line of the partial differential equation, Eq. (3.2). And it establishes a special relation between the partial differential equation and ordinary differential equation, from which the characteristic tracing method is born.

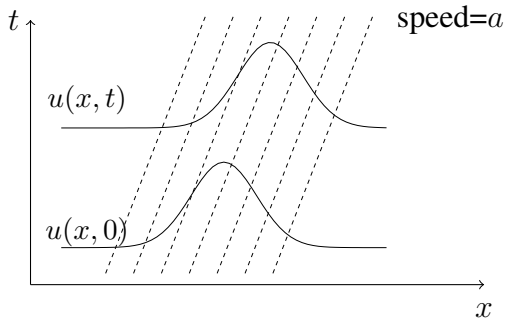


Figure 3.1: Characteristics of linear advection problem in one dimension. An initial solution translates along the same family of characteristics with the same speed $a > 0$.

In Fig. 3.1, an initially prescribed solution simply translates along characteristics. The newly advanced solution is just a translation of the solution from the "upwind" direction, and numerical methods that are physically motivated from this observation are called upwind methods. The exact solution of Eq. (3.2) is easily defined as a linear translation of the initial condition.

$$u(x,t) = u(x - at, 0) \tag{3.5}$$

where $x - at$ refers to the origin of characteristics.

In general, for a $n+1$-dimensional scalar partial differential equation ($+1$ including the

time dimension),

$$\frac{\partial u}{\partial t} + \sum_{i}^{n} a_i \frac{\partial u}{\partial x_i} = 0 \tag{3.6}$$

an envelope is formed conjoining n-parameter family of solution lines. The characteristics manifold is formed in the characteristic space $\mathbf{x} - t$. The n-parameter family of characteristics is an integral manifold of the $n+1$-dimensional partial differential equation along which the solution is constant. Then, the scalar partial differential equation is reduced to a characteristic equation, which is now an ordinary differential equation. This is the theory used in developing multidimensional advection numerical schemes for partial differential equations, see [28, 17] for more information. For $n = 2$, the characteristic manifold would be a surface, and the solution is found from integrating along the characteristics.

$$u(x, y, t) = u(x - a_1 t, y - a_2 t, 0) \tag{3.7}$$

The discussion below is not pertinent to the core of the thesis, and may be skipped for uninterested readers. However, the distinction must be made for methodologies in acoustics or wave systems, which involve infinitely many dimensions, unlike advection. This will be relevant in Chapter 7.

For a general $n+1$-dimensional system of hyperbolic partial differential equations, such as wave systems, the method of characteristics offers a different strategy.

$$\frac{\partial \mathbf{u}}{\partial t} + \sum_{i}^{n} \mathbf{A}_i \frac{\partial \mathbf{u}}{\partial x_i} = 0 \tag{3.8}$$

where $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{A}_i \in \mathbb{R}^{m \times m}$ and m is the number of equations in the system. There are m n-parameter family of characteristic equations for the system. We take $\mathbf{l}_1$, which is a left eigenvector of $\mathbf{A}_1$ with eigenvalue λ_1, and multiply it to all terms in Eq. (3.8) to obtain a characteristic equation.

$$\left(\frac{\partial}{\partial t} + \lambda_1 \frac{\partial}{\partial x_1} \right) \mathbf{l}_1 \mathbf{u} + \sum_{i=2}^{n} \mathbf{l}_1 \mathbf{A}_i \frac{\partial \mathbf{u}}{\partial x_i} = 0 \tag{3.9}$$

The characteristic equation, defined in an n-dimensional manifold of the $n+1$-dimensional space-time, are allowed to be discontinuous across the manifold. It can be then identified as a wavefront. A special subset of the characteristic equation called the bicharacteristic equation is obtained by selecting a special element of the basis set so that the characteristic manifold intersects with a neighboring characteristic manifold. The bicharacteristic equation, as a result, has a form of an ordinary differential equation along the intersection

together with source terms.

We assume for the moment that the system is linear, and choose a unit basis vector in some direction $\mathbf{e} = (e_1, \ldots, e_n)$. The characteristic variables are defined as follows.

$$\mathbf{w}(\mathbf{e}) = \mathcal{L}(\mathbf{e})\mathbf{u} \tag{3.10}$$

where $\mathcal{L}(\mathbf{e}) = (\mathbf{l}_1, \ldots, \mathbf{l}_m)^T$ is the left eigenvector matrix. The characteristic system is

$$\left(\frac{\partial}{\partial t} + \lambda_k \frac{\partial}{\partial x_k}\right) \mathbf{l}_k \mathbf{u} = -\sum_{i \neq k}^{n} \mathbf{l}_k e_i \mathbf{A}_i \frac{\partial \mathbf{u}}{\partial x_i}, \qquad k = 1, \ldots, m \tag{3.11}$$

$$\left(\frac{\partial}{\partial t} + \lambda_k \frac{\partial}{\partial x_k}\right) w_k = S_k(x_k(t, e_k), t, e_k), \qquad k = 1, \ldots, m \tag{3.12}$$

Along the bicharacteristics, we have total derivatives of w_k but this is not constant as in one dimension because the right hand side is generally not zero. Integration along the bicharacteristics introduces a formula for the characteristic variables.

$$w_k(\mathbf{x}, t, \mathbf{e}) = w_k(\mathbf{x} - \lambda_k t, 0, \mathbf{e}) + \int_0^t S_k(\mathbf{x}_k(\tau, \mathbf{e}), \tau, \mathbf{e}) d\tau, \qquad k = 1, \ldots, m \tag{3.13}$$

Or, for the bicharacteristic of a linear system shown in Fig. 3.2,

$$w_k(P, \mathbf{e}) = w_k(Q_k(\mathbf{e}), \mathbf{e}) + \int_0^t S_k(\mathbf{x}_k(\tau, \mathbf{e}), \tau, \mathbf{e}) d\tau, \qquad k = 1, \ldots, m \tag{3.14}$$

The exact integral equation in the original variable $\mathbf{u}$ is obtained after multiplying by a column vector matrix, $\mathcal{R}(\mathbf{e}) = \mathcal{L}^{-1}(\mathbf{e}) = (\mathbf{r}_1, \ldots, \mathbf{r}_m)$, then integrating around the characteristic cone for which $\mathbf{e} = (\cos\theta, \sin\theta)$.

$$\mathbf{u}(P) = \int_0^{2\pi} \left(\mathcal{R}(\mathbf{e}) \begin{bmatrix} w_1(Q_1(\mathbf{e}), \mathbf{e}) \\ \vdots \\ w_m(Q_m(\mathbf{e}), \mathbf{e}) \end{bmatrix} + \int_0^t \mathcal{R}(\mathbf{e}) \mathbf{S}(\tau, \mathbf{e}) d\tau \right) d\theta \tag{3.15}$$

$$\mathbf{u}(P) = \int_0^{2\pi} \left(\sum_{k=1}^{m} w_k(Q_k(\mathbf{e}), \mathbf{e}) \mathbf{r}_k + \int_0^t \mathcal{R}(\mathbf{e}) \mathbf{S}(\tau, \mathbf{e}) d\tau \right) d\theta \tag{3.16}$$

The solution $\mathbf{u}(P)$ is composed of the integral along the characteristic ray and the mantle integral, which is the surface of the characteristic cone between 0 and t. The second term

arises due to the source term $\mathbf{S}$. In a general form,

$$\mathbf{u}(\mathbf{x},t) = \frac{1}{|O|}\int_O \left(\sum_{k=1}^{m} w_k(\mathbf{x}-\lambda_k t,0,\mathbf{e})\mathbf{r}_k + \int_0^t \mathcal{R}(\mathbf{e})\mathbf{S}(\tau,\mathbf{e})d\tau\right) dO \tag{3.17}$$

where O is a sphere in $\mathbb{R}^n$. This is an implicit formula where $\mathbf{S}(\tau,\mathbf{e})$ depends on both time levels, and a suitable approximation is needed.

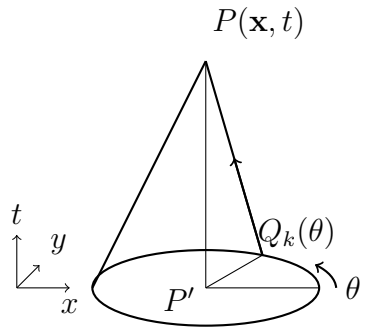


Figure 3.2: A Bicharacteristic PQ of a linear wave system for $n = 2$. The characteristic cone has a circular cross-section for a linear wave system with a constant speed of sound. The solution at P is found by integrating along all characteristic rays and the mantle of the characteristic cone.

Therefore, in a system of hyperbolic partial differential equations, the proper domain of dependence includes the cone surface, and it differs from scalar hyperbolic partial differential equations. More discussion on the AF method for acoustic systems will be presented in detail in Section 7.2.2. Other numerical methods using bicharacteristics equations are found [72, 50, 51]. These numerical methods usually make a selection of which bicharacteristics to employ, and are usually tied to the grid.

In this dissertation, characteristic equations of scalar hyperbolic partial differential equations are used throughout to develop numerical methods for linear and nonlinear advection problems.

3.2 Linear Advection in One Dimension

The AF method for one-dimensional linear advection is summarized in Algorithm 2.1. A scalar conservation law such as linear advection does not necessitate the last step reconciliation in the algorithm. Details regarding each algorithm step will be discussed in the order of execution.

3.2.1 Reconstruction

The reconstruction of AF is discussed.

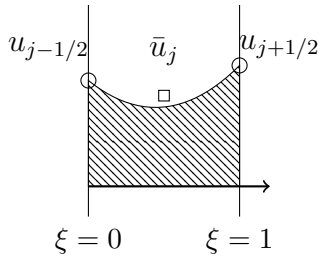


Figure 3.3: Numerical stencil in one-dimensional reference coordinate.

In a one-dimensional reconstruction, an element is defined by two interface values denoted as $u_{j-1/2}$ and $u_{j+1/2}$ at either boundary and an additional degree of freedom $\bar{u}_j$ at the centroid of the element, see Fig. 3.3. A quadratic reconstruction polynomial is constructed using the three pieces of DOF available. The interface variables are shared between the immediate neighbors, and the solution reconstruction is continuous across interfaces. As a result of the node sharing, the effective number of DOF reduces to two in the one-dimensional case.

The cell average $\bar{u}_j$ is an area averaged solution of an individual element. The definition of cell average is consistent with the definition found in standard finite volume schemes.

$$\bar{u}_j = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} u(\xi) d\xi \tag{3.18}$$

where ξ is introduced as a reference coordinate, and the mapping from a physical to a reference coordinate is achieved by the following relation.

$$\xi = \frac{x - x_{j-1/2}}{x_{j+1/2} - x_{j-1/2}} \tag{3.19}$$

Depending on the piecewise approximation of solution $u(\xi)$, the overall accuracy of the numerical scheme can be varied. In [84], van Leer introduced six schemes from which second- and third-order solutions can be obtained by augmenting the piecewise constant solution by high-order approximations. He introduced the Legendre polynomial in order to demonstrate different forms of piecewise polynomials.

$$u(\xi) = \bar{u}_j + \overline{\Delta}_j u \cdot (\xi - \xi_j) + \frac{1}{2}\overline{\Delta^2}_j u \cdot \left\{ (\xi - \xi_j)^2 - \frac{1}{12} \right\}, \quad \xi \in [0, 1] \tag{3.20}$$

where coefficients $\overline{\Delta}_j u$ and $\overline{\Delta^2}_j u$ correspond to approximate averages of gradient and curvature of the solution. A piecewise linear polynomial is obtained by selecting some approximation to $\overline{\Delta}_j u$. Likewise, a piecewise quadratic polynomial is obtained by supplying an additional approximation to $\overline{\Delta^2}_j u$. The list of coefficients from van Leer's schemes is summarized in Table 3.1.

Table 3.1: Summary of van Leer's schemes

Scheme	$\overline{\Delta}_j u$	$\overline{\Delta^2}_j u$
I	$\frac{1}{2}\left(\Delta_j u + \Delta_{j+1} u\right)$	0
II	$u_{j+1/2} - u_{j-1/2}$	0
III	$12\int_0^1 u(\xi)\cdot(\xi-\xi_j)d\xi$	0
IV	$\frac{1}{2}\left(\Delta_j u + \Delta_{j+1} u\right)$	$\Delta_{j+1}\bar{u} - \Delta_j \bar{u}$
V	$u_{j+1/2} - u_{j-1/2}$	$6\left(u_{j-1/2} - 2\bar{u}_j + u_{j+1/2}\right)$
VI	$12\int_0^1 u(\xi)\cdot(\xi-\xi_j)d\xi$	$360\int_0^1 u(\xi)\cdot\left\{\left(\frac{\xi-\xi_j}{\Delta\xi}\right)^2 - \frac{1}{12}\right\}d\xi$

In particular, the quadratic reconstruction for Scheme V is explicitly shown.

$$\begin{aligned} u(\xi) = \bar{u}_j &+ \left(u_{j+1/2} - u_{j-1/2}\right)\cdot(\xi-\xi_j) \\ &+ 3\left(u_{j-1/2} - 2\bar{u}_j + u_{j+1/2}\right)\cdot\left\{(\xi-\xi_j)^2 - \frac{1}{12}\right\}, \quad \xi\in[0,1] \end{aligned} \tag{3.21}$$

Note that this piecewise quadratic polynomial is continuous across neighboring elements. Equivalently, the Lagrange basis functions can be employed to write the piecewise quadratic reconstruction.

$$u(\xi) = \sum_{i=1}^{3} c_i \phi_i(\xi) \tag{3.22}$$

Reconstruction coefficients c_i and basis functions ϕ_i are summarized in Table 3.2.

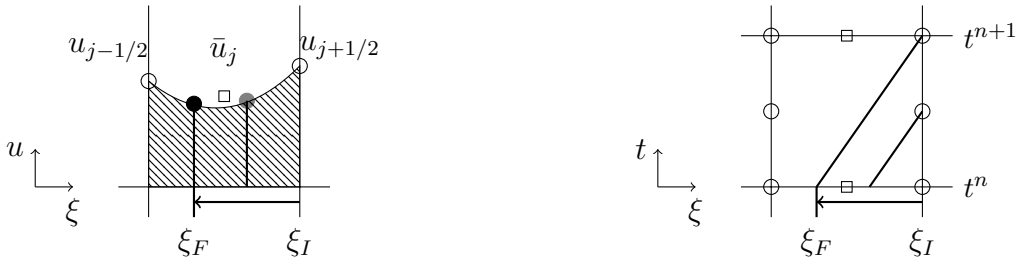
Table 3.2: Lagrange basis functions and reconstruction coefficients for 1D element

Index	c_i	ϕ_i
1	$u_{j-1/2}$	$(\xi-1)(2\xi-1)$
2	$\frac{1}{4}\left(6\bar{u}_j - u_{j-1/2} - u_{j+1/2}\right)$	$4\xi(1-\xi)$
3	$u_{j+1/2}$	$\xi(2\xi-1)$

The piecewise quadratic polynomial is unique to each element and is convenient to interpolate from. Hence, the Lagrange interpolation polynomial is primarily used.

3.2.2 Evolution

The evolution step of AF is described.



(a) Solution reconstruction

(b) Characteristic tracing for $a > 0$

Figure 3.4: Description of linear advection characteristic tracing in 1D

The evolution step is based on a characteristic tracing method. It is a semi-Lagrangian scheme that involves updating the fixed interface variables according to interpolations at characteristic origins. The use of characteristic tracing in numerical methods has a long history [1]. The work of Courant, Isaacson, and Rees [18] first suggested the method of characteristics in finite difference methods. Since then, many numerical methods were born out of this methodology including in some finite element methods [57, 12, 59].

When Scheme V was first introduced, the interface update/evolution was based on a semi-Lagrangian method. And the conservation update was achieved by exactly integrating solutions that have been evolved by the semi-Lagrangian process. In AF, we also use the method of characteristics for the evolution of interface variables, but a slightly more refined approach is introduced for the conservation update. In addition to updating a full-time step interface variable, we also introduce an additional interface value at the half-time step $t^{n+1/2}$. This would allow us to formulate an equivalent conservation update by means of a numerically accurate integration rule, the Simpson's quadrature rule. In linear problems, both approaches are exactly equivalent. However, the new approach used in AF will prove to be very useful in nonlinear problems.

The solution u of the linear advection equation, Eq. (3.2), is constant along characteristics $\frac{dx}{dt} = a$. By tracing the characteristics, advanced solutions at t^{n+1} and $t^{n+1/2}$ are found at respective characteristic origins, see Fig. 3.4b. With a provided solution reconstruction polynomial at t^n,

$$u^n(\xi) = \sum_{i=1}^{3} c_i \phi_i(\xi) \tag{3.23}$$

the new interface solution $u^{n+1}(\xi_I)$, located at the interface ξ_I, is the value interpolated at

the characteristic origin ξ_F.

$$u^{n+1}(\xi_I) = u^n(\xi_F) \tag{3.24}$$

where the characteristic origin ξ_F is defined as

$$\xi_F = \xi_I - \frac{a\Delta t}{\Delta x} = \xi_I - \nu \tag{3.25}$$

The Courant number ν is restricted by the usual CFL condition for stability, Eq. (3.26). For stability, the interpolation must take place within the proper numerical domain of dependence.

$$0 \leq \nu \leq 1 \tag{3.26}$$

As mentioned before, we now require the half-time step interface value for the conservation update stage. It is quite trivial to obtain the extra interface value defined at $t^{n+1/2}$. The corresponding half-time value and the characteristic origin are shown.

$$u^{n+1/2}(\xi_I) = u^n(\xi_H) \tag{3.27}$$

$$\xi_H = \xi_I - \frac{a\Delta t}{2\Delta x} = \xi_I - \frac{\nu}{2} \tag{3.28}$$

Fully explicit expressions at the right interface location are shown.

$$\begin{aligned} u^{n+1}(\xi_I) &= u^{n+1}_{j+1/2} \\ &= \nu(3\nu-2)u^n_{j-1/2} + 6\nu(1-\nu)\bar{u}^n_j + (1-\nu)(1-3\nu)u^n_{j+1/2} \end{aligned} \tag{3.29}$$

$$\begin{aligned} u^{n+1/2}(\xi_I) &= u^{n+1/2}_{j+1/2} \\ &= \frac{\nu}{2}\left(3\frac{\nu}{2}-2\right)u^n_{j-1/2} \\ &\qquad + 6\frac{\nu}{2}\left(1-\frac{\nu}{2}\right)\bar{u}^n_j + \left(1-\frac{\nu}{2}\right)\left(1-3\frac{\nu}{2}\right)u^n_{j+1/2} \end{aligned} \tag{3.30}$$

3.2.3 Conservation

The conservation step is discussed next. Both the original approach of Scheme V and the new approach of AF are presented.

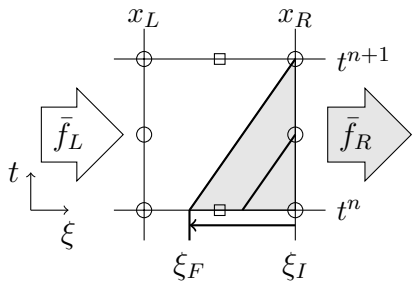


Figure 3.5: Conservation in $[x_L, x_R] \times [t^n, t^{n+1}]$

For the conservation step, we consider the integral form of the conservation law Eq. (3.2) in a control volume $[x_L, x_R] \times [t^n, t^{n+1}]$.

$$\int_{t^n}^{t^{n+1}} \int_{x_L}^{x_R} \left(\frac{\partial}{\partial t} u(x,t) + \frac{\partial}{\partial x} f(x,t) \right) dx dt = 0 \tag{3.31}$$

$$\int_{x_L}^{x_R} \left[u(x, t^{n+1}) - u(x, t^n) \right] dx + \int_{t^n}^{t^{n+1}} \left[f(x_R, t) - f(x_L, t) \right] dt = 0 \tag{3.32}$$

where the analytic flux is $f = au(x,t)$. The first term in Eq. (3.32) is simply the change of conservative quantity u in time. It is equal in magnitude and opposite in sign to the net conservative flux of the control volume. The flux across an interface can be evaluated in a couple of different, but equivalent, ways. We consider the flux across the right interface x_R.

$$\bar{f}_R = \int_{t^n}^{t^{n+1}} f(x_R, t) dt \equiv \int_{\xi_F}^{\xi_I} f(u^n(\xi)) d\xi \tag{3.33}$$

The first way is to integrate the flux function in time analytically, which is shown as the second term in Eq. (3.33). This was the approach that van Leer presented, and is easy for linear problems. However, it becomes difficult for nonlinear problems.

The second way utilizes the property of characteristics that solutions are constant along characteristics. The amount of conservative quantity displaced across the interface x_R between $[t^n, t^{n+1}]$ is equivalent to the quantity emerging from $[\xi_F, \xi_I]$ at t^n. The equivalence relation in Eq. (3.33) can be *numerically* satisfied by an accurate numerical quadrature rule.

The Simpson's rule is employed.

$$\begin{aligned}\bar{f}_R &= \int_{\xi_F}^{\xi_I} f(u^n(\xi))d\xi \\ &= \frac{1}{6}\bigg(f(u^n(\xi_I)) + 4f(u^n(\xi_H)) + f(u^n(\xi_F)) \bigg) + \mathcal{O}(\Delta\xi^4) \qquad (3.34) \\ &= \frac{a}{6}\bigg(u^n(\xi_I) + 4u^n(\xi_H) + u^n(\xi_F) \bigg) + \mathcal{O}(\Delta\xi^4) \qquad (3.35)\end{aligned}$$

where $u^n(\xi)$ is the solution reconstruction available at t^n, ξ_H and ξ_F indicate the characteristic origins at half- and full-time steps. Since the numerical flux is averaged from individually upwind fluxes, the resulting flux is upwind by default. The general expression, Eq. (3.34), is easily extended to nonlinear problems, in Section 4.2.

Once consistent interface fluxes are obtained, a change in the conservative variable is evaluated according to the conservative discretization of Eq. (3.32). A fully discrete form of the conservation law with the forward-Euler time stepping is

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \frac{\Delta t}{\Delta x_j}\left(\bar{f}_{j+1/2} - \bar{f}_{j-1/2}\right) \qquad (3.36)$$

where $\bar{f}_{j-1/2}$ and $\bar{f}_{j+1/2}$ are the left and right interface fluxes of the cell j. By the virtue of telescoping property of fluxes, the change in conservation in the domain only depends on the net flux change across the left and right most interfaces.

$$\begin{aligned}\sum_{j=1}^{N}(\bar{u}_j^{n+1} - \bar{u}_j^n)\Delta x_j &= -\Delta t \sum_{j=1}^{N}\left[\left(\bar{f}_{j+1/2} - \bar{f}_{j-1/2}\right)\right] \\ &= -\Delta t\left(\bar{f}_{N+1/2} - \bar{f}_{1/2}\right) \qquad (3.37)\end{aligned}$$

The explicit cell average value for linear advection is shown below.

$$\begin{aligned}\bar{u}_j^{n+1} = \nu^2(\nu-1)u_{j-3/2}^n + \nu^2(3-2\nu)\bar{u}_{j-1}^n + \nu(1-\nu)u_{j-1/2}^n \\ + (1-\nu)^2(1+2\nu)\bar{u}_j^n - \nu(1-\nu)^2 u_{j+1/2}^n \qquad (3.38)\end{aligned}$$

3.2.4 Stability Analysis

In the AF method, a scalar conservation law is solved by first updating interface fluxes by the characteristic tracing method, Eq. (3.24), then updating conservative variables by a finite volume method, Eq. (3.36). Because the interface variables are also evolving quanti-

ties, the entire update procedure for the AF method can be regarded as a system of updates.

$$\begin{pmatrix} \bar{u}_j \\ u_{j+1/2} \end{pmatrix}^{n+1} = \mathbf{G}_{\text{1D}} \begin{pmatrix} \bar{u}_j \\ u_{j+1/2} \end{pmatrix}^{n} \tag{3.39}$$

where $\mathbf{G}_{\text{1D}}$ is the amplification matrix.

In order to analyze the stability of the system, Von Neumann analysis, or Fourier analysis, is performed. Fourier analysis assumes the solution vector takes the following form.

$$\begin{pmatrix} \bar{u}_j^n \\ u_{j+1/2}^n \end{pmatrix} = g^n \begin{pmatrix} e^{\mathbf{i}j\phi_x} \\ e^{\mathbf{i}i\phi_x} \end{pmatrix} \tag{3.40}$$

where $\mathbf{i} = \sqrt{-1}$ and indices $i = j + 1/2$ and j are associated with interface and cell average values, see Fig. 3.3. By performing Fourier analysis, the amplification matrix for one-dimensional linear advection is obtained.

$$\mathbf{G}_{\text{1D}} = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \tag{3.41}$$

$$\begin{aligned} G_{11} &= \nu^2(3-2\nu)e^{-\mathbf{i}\phi_x} + (1-\nu)^2(1+2\nu) \\ G_{12} &= \nu^2(\nu-1)e^{-2\mathbf{i}\phi_x} + \nu(1-\nu)e^{-\mathbf{i}\phi_x} - \nu(1-\nu)^2 \\ G_{21} &= 6\nu(1-\nu) \\ G_{22} &= \nu(3\nu-2)e^{-\mathbf{i}\phi_x} + (1-\nu)(1-3\nu) \end{aligned}$$

Analyses of the amplification matrix lead to some important stability criteria of the scheme. The characteristic polynomial of $\mathbf{G}_{\text{1D}}$ is quadratic, and there are two eigenvalues as a result. Two eigenvalues and the corresponding eigenvectors are presented explicitly for $\nu = 0.5$.

$$g_{1,2} = \frac{1}{8} e^{-\mathbf{i}\frac{\phi_x}{2}} \left(2\cos\frac{\phi_x}{2} \mp \sqrt{6}\sqrt{7-\cos\phi_x} \right) \tag{3.42}$$

$$\mathbf{v}_{1,2} = \begin{pmatrix} \frac{1}{12} e^{-\mathbf{i}\frac{\phi_x}{2}} \left(6\cos\frac{\phi_x}{2} \mp \sqrt{6}\sqrt{7-\cos\phi_x} \right) \\ 1 \end{pmatrix} \tag{3.43}$$

The eigenvalues can be analyzed for the stability property of the numerical scheme. The amplification error, also known as the dissipation error, defines the amount of numerical dissipation per time step.

$$\epsilon_D = 1 - |g| \tag{3.44}$$

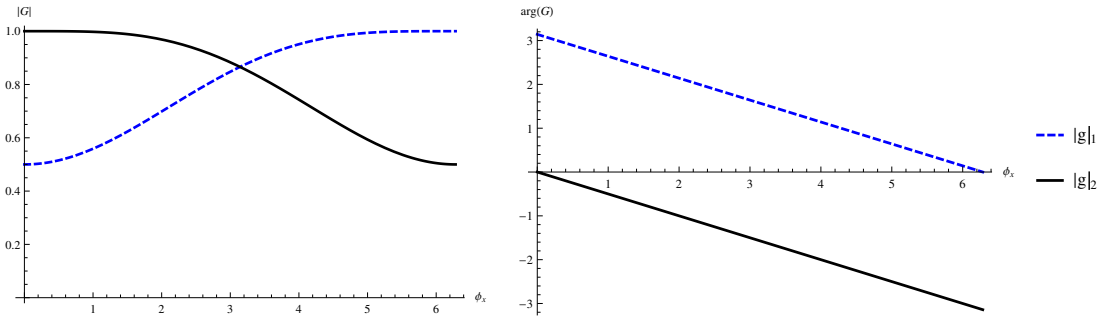
The dispersion error, or the phase error, is defined as the ratio of numerical and physical propagation speeds, and describes either a lagging or a leading nature of the numerical solution.

$$\epsilon_P = 1 - \frac{\text{Arg}(g)}{-\nu\phi} \tag{3.45}$$

Before providing the results of AF stability properties, we first shed some light on some general properties of the amplification matrix $\mathbf{G}_{\text{1D}}$.

For Scheme V, van Leer initially reported that only one eigenvalue is relevant to the stability of the system by studying the eigenvectors associated with the eigenvalues [84]. He described the inappropriate mode as the "stegosaur" bias mode and disregarded it from the analysis. Since then, a new analysis [21] has shown that both eigenvalues, in fact, contribute to the stability of the system. Of the two eigenvalues in Eq. (3.42), one describes the stability property of the scheme in the low-frequency domain $[0, \pi]$ while the other describes the stability property in the high-frequency domain $[\pi, 2\pi]$. And this increases the frequency domain to $[0, 2\pi]$; twice that of a traditional range. The outcome is that numerical solutions with twice as high-frequency components can be resolved by Scheme V and the AF scheme.

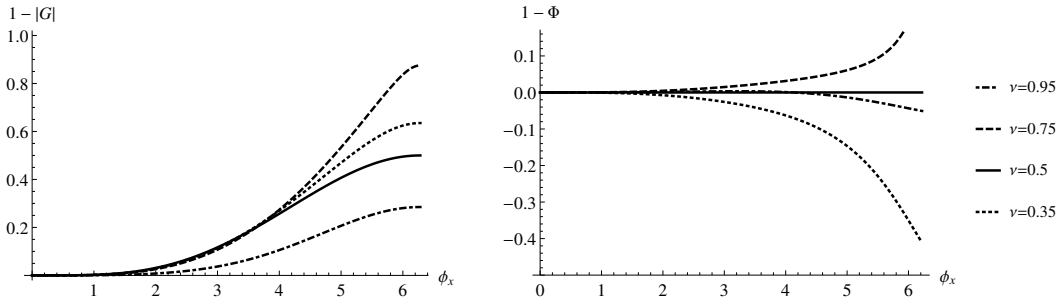
Thus far, we have only considered the low-frequency component of one eigenvalue and the high-frequency component of the other eigenvalue for describing the stability of the scheme. The other components of eigenvalues need some mentioning even though they do not bear any physical meaning. All eigenvalues in Eq. (3.42) for $\nu = 0.5$ are plotted in Fig. 3.6. There are two curves in magnitude and argument plots of the eigenvalues. We use the physical observation that an explicit upwind scheme is only correctly described by a forward moving numerical solution, *i.e.* black curves. The remaining solution is of only mathematical significance and bears little physical significance in the context of upwind numerical schemes. Therefore, we only consider the physical mode for the stability analysis.



(a) Magnitudes of eigenvalues (b) Arguments of eigenvalues

Figure 3.6: Magnitudes and arguments of two eigenvalues at $\nu = 0.5$. The eigenvalue with negative argument values indicates the physical solution. The eigenvalue with positive argument values is only mathematically correct.

Dissipation and dispersion properties of the AF method for one-dimensional linear advection are shown in Fig. 3.7.



(a) Dissipation error (b) Dispersion error

Figure 3.7: AF dissipation and dispersion errors at various ν

We want to make a comparison between AF and a similar third-order scheme. For this, we compare the stability properties of AF to those of the discontinuous Galerkin (DG) method. A spatial discretization order $p = 1$ DG with a third-order Runge-Kutta method (RKDG1) time discretization is chosen. For linear problems, RKDG1 can achieve third-order accuracy by superconvergence.

To determine the stability of RKDG1, we need to establish a stable time discretization method for the spatial discretization method of DG1. To do this, we make sure that the time integration scheme is compatible with the range of the eigenvalue spectrum of the spatial discretization method. The eigenvalue spectrum of DG1 is found from the matrix of the explicit formulation of the spatial discretization [101]. For $p = 1$, two linear basis functions describe the spatial discretization for DG1. The semi-discrete system can be written as follows. The linear basis functions are defined at $-1/4$ and $+1/4$ locations in

the reference domain $\xi \in [-1, 1]$.

$$\frac{\partial}{\partial t}\begin{pmatrix} u_{-1/4}(t) \\ u_{1/4}(t) \end{pmatrix} = \mathbf{G}_{\text{DG1}}(h, \phi_x)\begin{pmatrix} u_{-1/4}(0) \\ u_{1/4}(0) \end{pmatrix} \tag{3.46}$$

where the spatial discretization matrix is

$$\mathbf{G}_{\text{DG1}}(h, \phi_x) = \frac{a}{h}\begin{pmatrix} -\frac{5}{4}e^{-\mathbf{i}\phi_x} - \frac{7}{4} & \frac{15}{4}e^{-\mathbf{i}\phi_x} - \frac{3}{4} \\ \frac{1}{4}e^{-\mathbf{i}\phi_x} + \frac{11}{4} & -\frac{3}{4}e^{-\mathbf{i}\phi_x} - \frac{9}{4} \end{pmatrix}, \tag{3.47}$$

a is the advection speed and h is the reference length or the mesh spacing. The spatial discretization matrix $\mathbf{G}_{\text{DG1}}(h, \phi_x)$ has two eigenvalues.

$$\Omega_{1,2} = \frac{a}{h}\left(-e^{-\mathbf{i}\phi_x} - 2 \mp \sqrt{e^{-2\mathbf{i}\phi_x} + 10e^{-\mathbf{i}\phi_x} - 2}\right) \tag{3.48}$$

The consistency of the scheme requires that one of the values represent the physical behavior. This physical solution is recognized by the property that it approaches 1 in the limit of an exact solution (*i.e.* $\Delta t \to 0$). The other solution represents a spurious solution. In the RKDG1 case, the spurious solution damps out quickly and doesn't affect the stability.

Once the eigenvalue spectrum of the spatial discretization scheme is known, the stability region of a time discretization method must include the whole spectrum for the overall stability of the scheme [30]. A general system of ordinary differential equations is written as follows.

$$\frac{d\mathbf{u}}{dt} = \mathbf{H}(\mathbf{u}) \tag{3.49}$$

where $\mathbf{H}$ is a general form of the spatial discretized matrix operator of a numerical scheme. Once the Fourier mode eigenvalues $\mathbf{\Omega} = \text{diag}(\Omega_i)$ of the spatial discretization matrix operator $\mathbf{H}$ are found, a canonical modal equation can replace Eq. (3.49). For the mode i having Ω_i as its eigenvalue,

$$\frac{dw_i}{dt} = \Omega_i \cdot w_i \tag{3.50}$$

Then, a time integration method, operator P, is selected to advance the solution w_i.

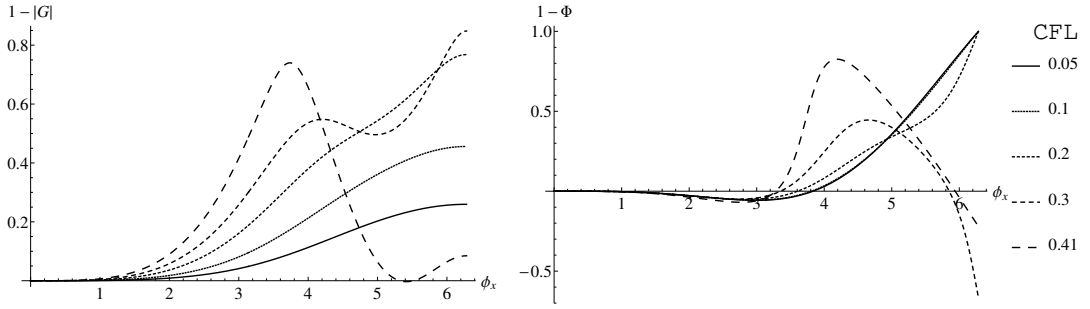
$$w_i^{n+1} = P(E, \Omega_i \Delta t) \cdot w_i^n \tag{3.51}$$

where E is a time shift operator. In relation to the stability analysis of fully discrete schemes, the stability condition of a semi-discrete scheme is defined by $|P| \leq 1$. For the third-order Runge-Kutta method with the eigenvalues of DG1 space discretization, P

is defined as follows.

$$P(E, \Omega_i \Delta t) = P(\nu \Omega_i) = 1 + (\nu \Omega_i) + \frac{1}{2}(\nu \Omega_i)^2 + \frac{1}{6}(\nu \Omega_i)^3 \tag{3.52}$$

where $\Omega_i \Delta t \equiv \nu \Omega_i$ because of the coefficient in the Fourier spectrum of DG1 in Eq. (3.48). The stability properties, dissipation and dispersion errors, of RKDG1 are described by Eq. (3.52), using the definitions previously stated in Eqs. (3.44) and (3.45). Dissipation and dispersion errors of RKDG1 are shown in Fig. 3.8 for several Courant numbers. Note that RKDG1 is stable for $\nu \leq 0.41$.

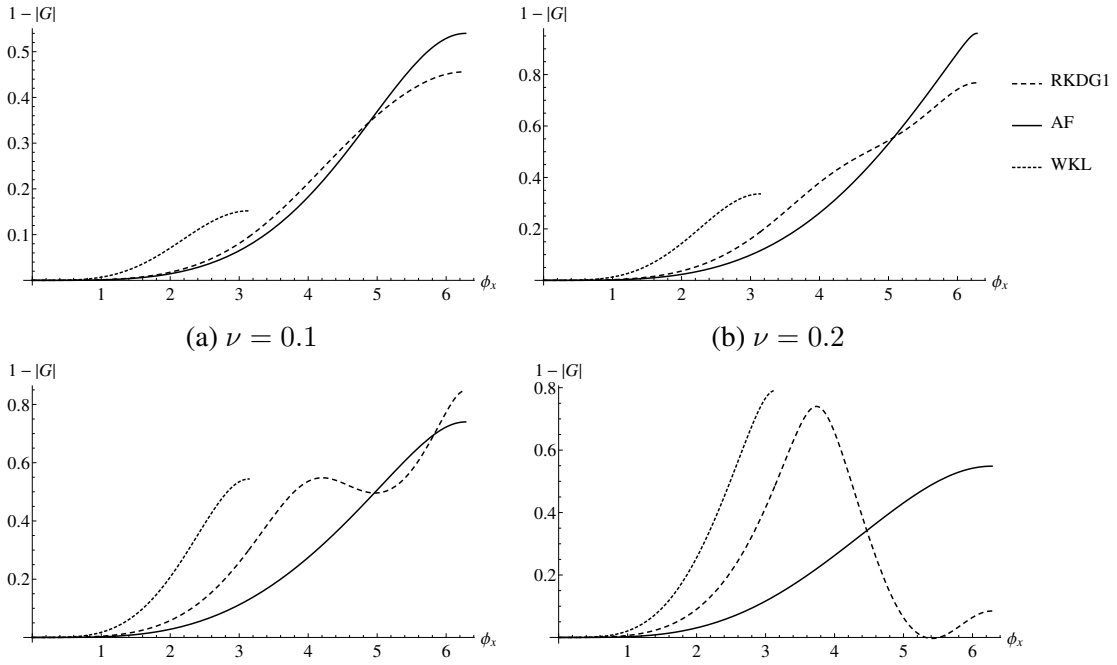


(a) Dissipation error (b) Dispersion error

Figure 3.8: RKDG1 dissipation and dispersion errors at various ν. The stability limit for RKDG1 is $\nu \leq 0.41$.

In addition, we consider the stability properties of Warming Kutler Lomax (WKL), a third-order finite volume scheme [91]. The WKL scheme and its stability properties which have been studied in Eymann [21].

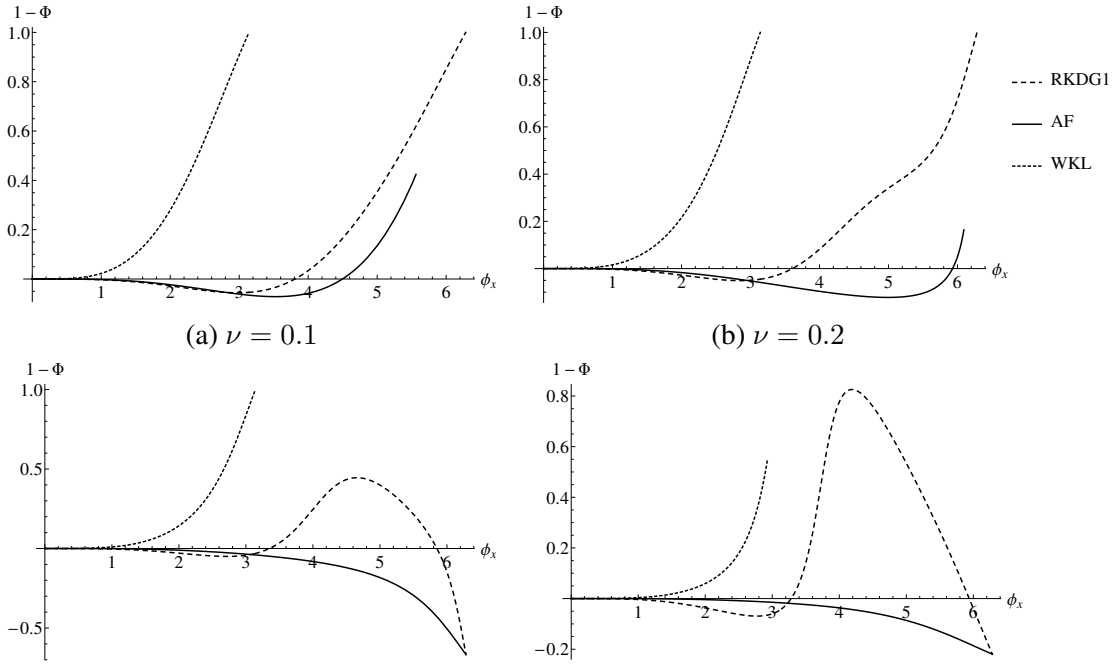
Figs. 3.9 and 3.10 show the dissipation and dispersion errors of the third-order schemes: AF, the conventional finite volume scheme WKL, and the finite element scheme RKDG1. Note that the comparison is made for Courant numbers less than or equal to $\nu = 0.41$ because of the stability limitation of RKDG1. Fully explicit schemes like AF and WKL can reach Courant numbers up to 1. While AF and RKDG1 can resolve frequencies in the range $[0, 2\pi]$, WKL only resolves a half of that. The dissipation and dispersion property of AF are well-behaved compared to the other two schemes. This is more evident in the high-frequency range, and as larger time steps are taken (*i.e.* larger ν).



(a) $\nu = 0.1$ (b) $\nu = 0.2$

(c) $\nu = 0.3$ (d) $\nu = 0.41$

Figure 3.9: Dissipation error comparison between various third-order schemes



(a) $\nu = 0.1$ (b) $\nu = 0.2$

(c) $\nu = 0.3$ (d) $\nu = 0.41$

Figure 3.10: Dispersion error comparison between various third-order schemes

The local truncation error is an estimate of the error introduced in a single iteration of a numerical method. It determines how well the difference equation, or the method, models the differential equation locally. The order of local truncation errors can determine the local order of the method. But for smooth problems, the global error is of the same order provided the method is stable [47].

The local truncation error analysis is performed by substituting an exact solution of the partial differential equation into numerical schemes, then performing Taylor expansions. A leading error term corresponds to the local truncation error and consequently defines the accuracy. For third-order accurate schemes, leading errors are dissipative and proportional to Δx^3.

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = -F(\nu)\frac{a\Delta x^3}{24}\frac{\partial^4 u}{\partial x^4} + \mathcal{O}(\Delta x^4) \tag{3.53}$$

where the left hand side is the exact partial differential equation and right hand side consists of the error terms. $F(\nu)$ is the magnitude of the truncation error as a function of the Courant number. The magnitude of truncation error can be used to determine the amount of error incurred in one iteration of a method.

We compare local truncation errors of AF, Scheme III of van Leer, RKDG1, and WKL third-order finite volume schemes. Some analyses between RKDG1 and Scheme III were made in [67]. We note that RKDG1 is a semi-discrete limit of Scheme III. The summary of local truncation errors is listed in Table 3.3. At Courant numbers $\nu = 0, 0.5, 1$, the AF scheme is exact. Although WKL is exact for Courant numbers $\nu = -1, 0, 1, 2$, it is only stable if $0 < \nu < 1$. On the other hand, Scheme III and RKDG1 have non-zero truncation errors at $\nu = 0$. Since RKDG1 is a semi-discrete version of Scheme III, the truncation error should approach that of Scheme III as $\nu \to 0$.

Table 3.3: Summary of local truncation error magnitude of various third-order schemes

Scheme	$F(\nu)$
AF	$\nu(1-\nu)(1-2\nu)$
Scheme III	$\frac{1}{3}(1-\nu)(1-\nu(1-\nu))$
RKDG1	$\frac{1}{3}+\nu^3$ for $\nu < 0.41$
WKL	$\nu(2-\nu)(1-\nu^2)$

Local truncation errors as functions of the Courant number are shown in Fig. 3.11. In fact, the magnitude of local truncation error of AF illustrates the superior error property exhibited by it compared to other third-order schemes.

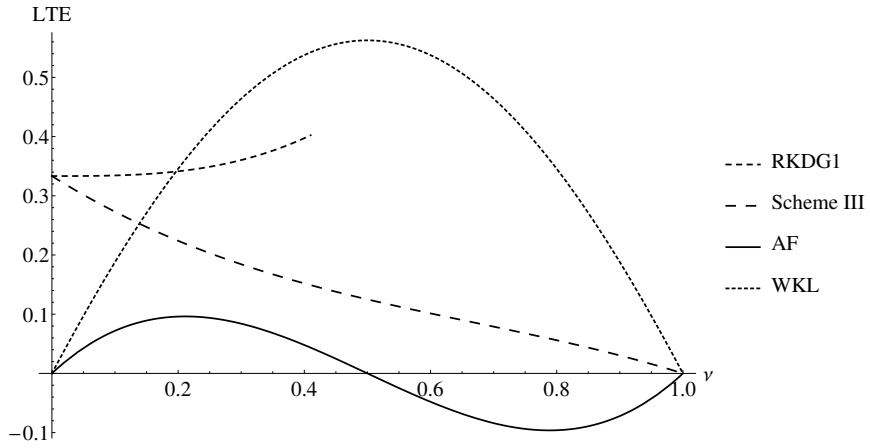


Figure 3.11: Local truncation error magnitude, $F(\nu)$, of various third-order schemes

3.2.5 Numerical Accuracy

In order to define the accuracy of numerical methods, we must define the measure of error estimates. An error norm typically used to measure a solution accuracy at a given time level in finite volume schemes is the $L_2(\overline{u})$ norm.

$$L_2(\overline{u}) = \left\{ \frac{1}{\Omega} \sum_{j=1}^{N} \left| \overline{u}_j \Omega_j - \overline{u}_j^{\text{exact}} \Omega_j \right|^2 \right\}^{1/2} \tag{3.54}$$

where N is the total number of cells, $\sum_j^N \Omega_j = \Omega$ is the total volume of a domain, and $\overline{u}$, $\overline{u}^{\text{exact}}$ are numerical and exact solutions in conservative variables $\overline{u}$. This norm definition is used to define errors associated with conservative variables, or cell averages. Because this norm measures an error of *averaged* quantities, it may not be suitable to measure the projection of the exact solution onto the interpolant of the numerical solution.

Another norm is defined to account for variations in the numerical solution reconstruction.

$$L_2(q(\mathbf{x})) = \left\{ \frac{1}{\Omega} \sum_{j=1}^{N} \int_{\Omega_j} \left| q_j(\mathbf{x}) - q_j^{\text{exact}}(\mathbf{x}) \right|^2 dV \right\}^{1/2} \tag{3.55}$$

where $q(\mathbf{x})$ and $q^{\text{exact}}(\mathbf{x})$ are the numerical and exact reconstruction solutions. This norm, unlike the $L_2(\overline{u})$ norm, does not ignore the projection of the exact solution onto the interpolant of the numerical solution, and can be a good measure of the error for reconstruction variables that are continuously defined in a domain.

Throughout the thesis, both definitions of norms, $L_2(\overline{u})$ and $L_2(q(\mathbf{x}))$, are used when some distinction is required. As it turns out, two definitions of norms make very little

difference in one-dimensional scalar problems. However, for problems in two or higher dimensions, it becomes important to address variations in solution reconstructions for an accurate description of the numerical solution. This becomes especially crucial for nonlinear systems of equations.

A sine wave initial condition is considered for the numerical accuracy of the AF scheme.

$$u_0(x) = \frac{1}{2\pi} \sin(2\pi x) \tag{3.56}$$

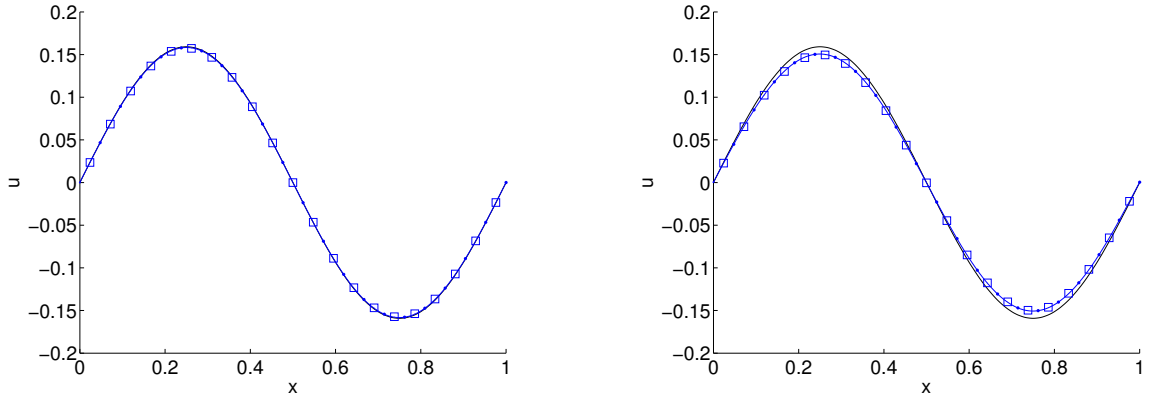
with the advection speed $a = 1$ and $x \in [0, 1]$. A periodic boundary condition is imposed. After advecting the initial solution for one period for a various number of elements, errors in conservative variables are measured and summarized in Table 3.4. N is the number of elements and $h = 1/N$ is a reference length. The conservative variable errors in $L_2(\overline{u})$ converge as $\mathcal{O}(h^3)$ in one-dimensional linear advection.

Table 3.4: 1D linear advection of Eq. (3.56) $L_2(\overline{u})$ norm error convergence at $t = 1$ and $\nu = 0.7$

Level	N	$L_2(\overline{u})$	Order
1	20	7.30e-05	
2	40	8.92e-06	3.03
3	80	1.12e-06	2.99
4	160	1.41e-07	2.99
5	320	1.76e-08	3.01
6	640	2.20e-09	3.00
7	1280	2.75e-10	3.00
8	2560	3.44e-11	3.00

3.2.6 Numerical Test

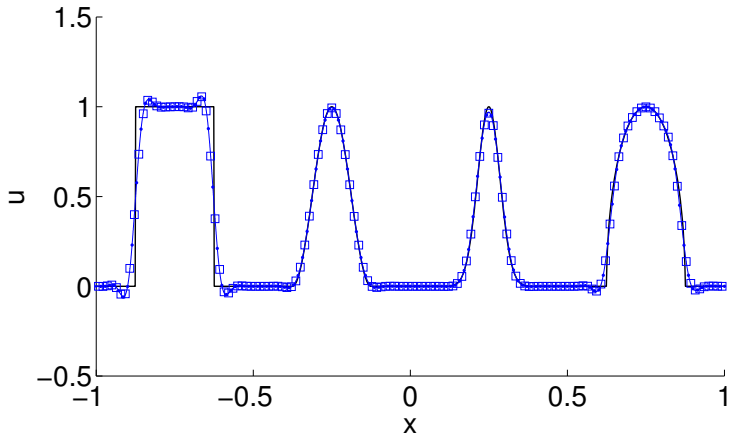
Additional linear advection numerical tests are presented. The sine wave problem, Fig. 3.12, considered for the numerical accuracy study is presented after 10 and 100 cycles. Even after 100 cycles, the sine wave shows little dissipation in comparison to the exact solution.



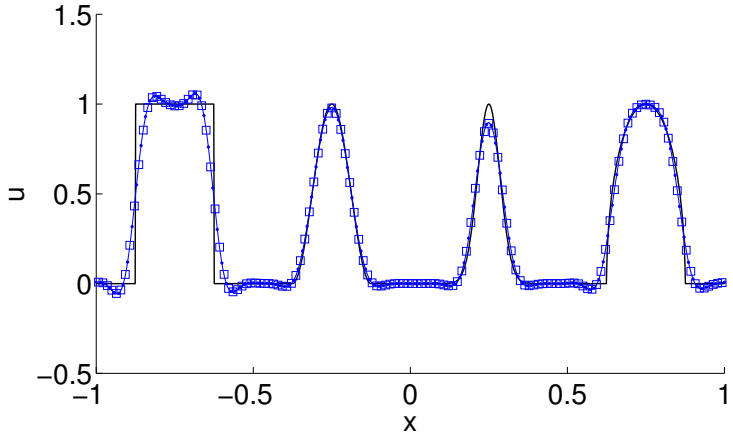
(a) $t = 10$ (b) $t = 100$

Figure 3.12: Sine wave problem, Eq. (3.56), at various times, $N = 21$, $\nu = 0.7$. The exact solution is indicated by the solid black line.

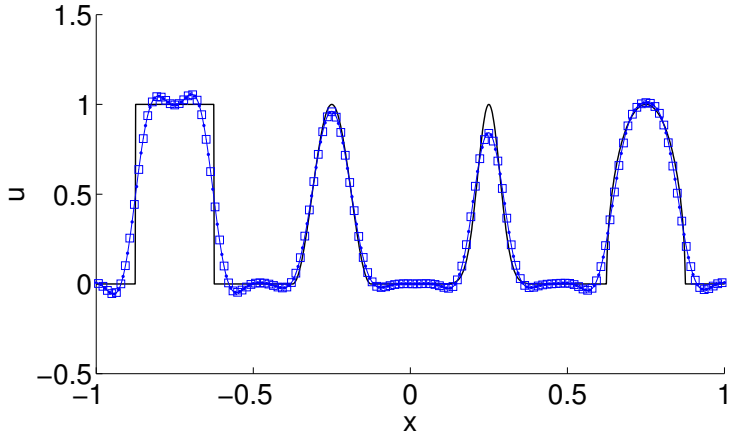
The Zalesak wave suite test problem demonstrates the numerical scheme's performance over a variety of solutions; a square, a cosine, a Gaussian, and an elliptic wave [98]. The AF results after multiple cycles are presented in Fig. 3.13. The smooth wave solutions like the cosine, the Gaussian, and the elliptic waves, are well-preserved with approximately the same number of cells even after ten cycles. The peak of the Gaussian wave shows some noticeable dissipation after five cycles. The square wave, however, exhibits some noticeable dissipation which translates to a smearing of discontinuities.



(a) One cycle, $t = 2$



(b) Five cycles, $t = 10$



(c) Ten cycles, $t = 20$

Figure 3.13: Zalesak wave suite with $N = 140$ and $\nu = 0.7$. The exact solution is indicated by the solid black line.

3.3 Linear Advection in Two Dimensions

The multidimensional linear advection problem is shown in a vector form.

$$\frac{\partial u}{\partial t} + \mathbf{a} \cdot \nabla u = 0 \tag{3.57}$$

where $\mathbf{a}$ is a constant vector of length d. In this section, the AF scheme for a two-dimensional case is presented, however, an extension to higher dimensions is straightforward.

3.3.1 Spatial Discretization – Choice of Reconstruction

The AF method is designed to be genuinely multidimensional. One attribute is that the spatial discretization can easily be varied between the most commonly used geometries in CFD simulations. Typical unstructured grids consisting of simplex elements are a popular choice for the spatial discretization. Simplex elements can vary in shape and size to conform around complex objects where structured grids may not be well suited.

Structured, Cartesian, grids perhaps date back as far as the history of CFD itself. Even though the implementation is simple, it becomes difficult to manage complex geometries that modern engineering problems often require.

An example of a two-dimensional triangular element is presented in Fig. 3.14. Since physical elements may vary in orientation and size, we introduce an element in the reference space where all numerical computation tasks are performed.

(a) Physical space

(b) Reference space defined in $\xi, \eta \in [0, 1]$

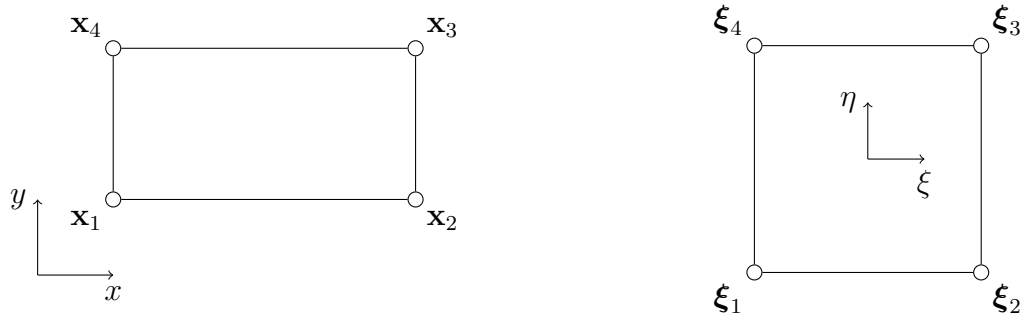
Figure 3.14: Triangular elements

A linear mapping Jacobian is introduced to map between the physical and the reference element spaces.

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{J}\boldsymbol{\xi}, \qquad \boldsymbol{\xi} = \mathbf{J}^{-1}\mathbf{x} \tag{3.58}$$

$$\mathbf{J} = \frac{\partial x_i}{\partial \xi_j} = \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix} \tag{3.59}$$

Likewise, a two-dimensional quadrilateral element is presented in Fig. 3.15.



(a) Physical space

(b) Reference space in $\xi, \eta \in [-1, 1]$

Figure 3.15: Quadrilateral elements

Assuming straight sided quadrilateral elements, a nonlinear mapping relation is given.

$$\mathbf{x}(\xi, \eta) = \sum_{i=1}^{4} \phi_i(\xi, \eta) \mathbf{x}_i \tag{3.60}$$

The vertex basis functions are defined as follows.

$$\begin{aligned} \phi_1 &= 0.25\,(\xi - 1)\,(\eta - 1)\,\xi\eta \\ \phi_2 &= 0.25\,(\xi + 1)\,(\eta - 1)\,\xi\eta \\ \phi_3 &= 0.25\,(\xi + 1)\,(\eta + 1)\,\xi\eta \\ \phi_4 &= 0.25\,(\xi - 1)\,(\eta + 1)\,\xi\eta \end{aligned}$$

$$\mathbf{J} = \frac{\partial x_i}{\partial \xi_j} = \begin{pmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} \end{pmatrix} \tag{3.61}$$

This mapping applies to any convex arbitrary quadrilaterals.

$$\mathbf{x} = \mathbf{x}_c + \mathbf{J}\boldsymbol{\xi}, \qquad \boldsymbol{\xi} = \mathbf{J}^{-1}\mathbf{x} \tag{3.62}$$

where $\mathbf{x}_c$ is the centroid coordinate in the physical space.

A compact quadratic reconstruction is obtained in two dimensions by Lagrange basis polynomials. For triangular elements, a fully defined quadratic reconstruction is defined by six Lagrange basis functions located along the boundary of the element. However, the con-

servation update, by discrete Gauss theorem, produces a third-order conservative scheme. We introduce an additional basis function, called the bubble function, at the centroid of the element, ϕ_7, so that the scheme maintains the third-order accurate property. The bubble function is a third-order basis function that vanishes on element boundaries, so it does not influence the reconstruction coefficients. As a result, solution reconstructions between elements are kept continuous. The detailed explanation of the bubble function and its use is discussed Section 3.3.4. Table 3.5 lists all reconstruction coefficients and basis functions. The two-dimensional element reconstruction is simply constructed by a sum of Lagrange basis functions and a bubble function.

$$u\left(\boldsymbol{\xi}\right) = \sum_{i=1}^{7} c_i \phi_i \left(\xi, \eta\right) \tag{3.63}$$

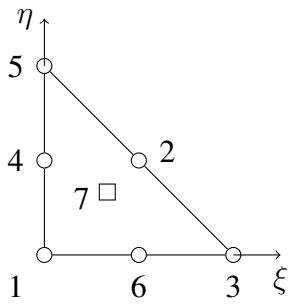


Figure 3.16: Locations of Lagrange basis functions for reference triangular elements

Table 3.5: Lagrange basis functions and coefficients for reference triangular elements

Index	c_i	ϕ_i
1	u_1	$(1-\xi-\eta)(1-2\xi-2\eta)$
2	u_2	$4\xi\eta$
3	u_3	$\xi(2\xi-1)$
4	u_4	$4\eta(1-\xi-\eta)$
5	u_5	$\eta(2\eta-1)$
6	u_6	$4\xi(1-\xi-\eta)$
7	$\frac{20}{9}\left(\bar{u}-\frac{1}{3}(u_2+u_4+u_6)\right)$	$27\xi\eta(1-\xi-\eta)$

For quadrilateral elements, a bi-quadratic polynomial is used instead. This is due to additional interface nodes introduced in the geometry. The bubble function in a quadrilateral is defined at the centroid by ϕ_9, and the purpose of it is the same as in the triangular element

case. Table 3.6 lists all reconstruction coefficients and basis functions. The bi-quadratic reconstruction is constructed by a sum of Lagrange basis functions and the bubble function.

$$u\left(\boldsymbol{\xi}\right) = \sum_{i=1}^{9} c_i \phi_i \left(\xi, \eta\right) \tag{3.64}$$

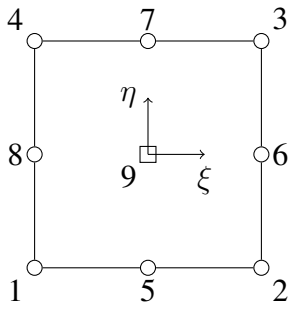


Figure 3.17: Locations of Lagrange basis functions for reference quadrilateral elements

Table 3.6: Lagrange basis functions and coefficients for reference quadrilateral elements

Index	c_i	ϕ_i
1	u_1	$0.25\left(\xi-1\right)\left(\eta-1\right)\xi\eta$
2	u_2	$0.25\left(\xi+1\right)\left(\eta-1\right)\xi\eta$
3	u_3	$0.25\left(\xi+1\right)\left(\eta+1\right)\xi\eta$
4	u_4	$0.25\left(\xi-1\right)\left(\eta+1\right)\xi\eta$
5	u_5	$0.5\left(1-\xi^2\right)\left(\eta-1\right)\eta$
6	u_6	$0.5\left(1-\eta^2\right)\left(\xi+1\right)\xi$
7	u_7	$0.5\left(1-\xi^2\right)\left(\eta+1\right)\eta$
8	u_8	$0.5\left(1-\eta^2\right)\left(\xi-1\right)\xi$
9	$\frac{1}{16}\left(36\bar{u}-\left(u_1+u_2+u_3+u_4\right)-4(u_5+u_6+u_7+u_8)\right)$	$\left(1-\xi^2\right)\left(1-\eta^2\right)$

3.3.2 Evolution

The evolution process for two-dimensional linear advection is described.

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} + b\frac{\partial u}{\partial y} = 0 \tag{3.65}$$

The characteristic tracing method developed in one-dimensional linear advection is used. All characteristics in two-dimensional linear advection are straight and parallel to each

other, and it is trivial to extend the general methodology from the one-dimensional characteristic tracing method.

In order to simplify the conservation update in one-dimensional linear advection in Section 3.2, we introduced an additional interface variable at the half-time step. This allowed us to evaluate numerically accurate interface fluxes, therefore avoiding costly exact flux integrations. The same approach is pursued in multidimensional advection problems.

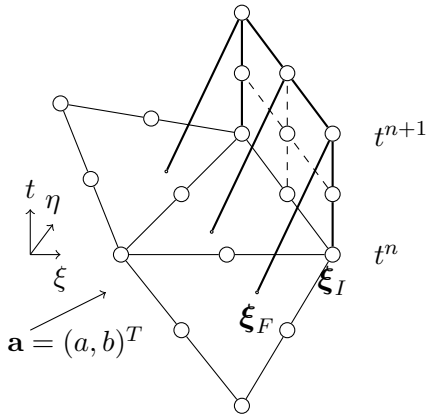


Figure 3.18: Description of linear advection characteristic tracing in 2D

An example of the characteristic tracing is presented in Fig. 3.18. As a consequence of multidimensionality, three quadrature points belonging to a common interface can have characteristic origins lying on different upwind elements. The upwind direction can be determined by tracing the characteristic vector, which is just the constant velocity vector $\mathbf{a}$ in linear advection.

The solution u of Eq. (3.65) is constant along characteristics. The time advanced interface solution is evaluated at the characteristic origin $\boldsymbol{\xi}_F$.

$$u^{n+1}(\boldsymbol{\xi}_I) = u^n(\boldsymbol{\xi}_F) \tag{3.66}$$

where the two-dimensional characteristic origin $\boldsymbol{\xi}_F$ is defined as follows.

$$\boldsymbol{\xi}_F = \boldsymbol{\xi}_I - \Delta t \mathbf{J}^{-1}\mathbf{a} \tag{3.67}$$

where $\mathbf{J}^{-1}$ is the inverse coordinate mapping Jacobian matrix to transform the vector $\mathbf{a}$ in the physical space to the reference space.

The half-time step value at the same interface $\boldsymbol{\xi}_I$ is also easily evaluated by interpolating

at the characteristic origin corresponding to half-time.

$$u^{n+1/2}(\boldsymbol{\xi}_I) = u^n(\boldsymbol{\xi}_F) \tag{3.68}$$

$$\boldsymbol{\xi}_F = \boldsymbol{\xi}_I - \frac{\Delta t}{2}\mathbf{J}^{-1}\mathbf{a} \tag{3.69}$$

3.3.3 Conservation

The integral form of the conservation law for two-dimensional linear advection, Eq. (3.65), is written as follows.

$$\begin{aligned} \int_\Omega \frac{\partial u}{\partial t} dV + \int_\Omega \nabla \cdot \mathbf{f} dV = 0 \\ \frac{d}{dt}\int_\Omega u dV + \oint_{\partial\Omega} \mathbf{f}\cdot\mathbf{n} dS = 0 \end{aligned} \tag{3.70}$$

where the two-dimensional linear advection flux vector is

$$\mathbf{f} = \begin{pmatrix} au \\ bu \end{pmatrix} \tag{3.71}$$

The fully discrete conservation law is presented below.

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \frac{\Delta t}{\Omega_j}\sum_{l=1}^{\text{sides}} \bar{\mathbf{f}}_l \cdot \mathbf{n}_l = 0 \tag{3.72}$$

where $\mathbf{n}_l$ is a scaled normal vector belonging to an interface edge l.

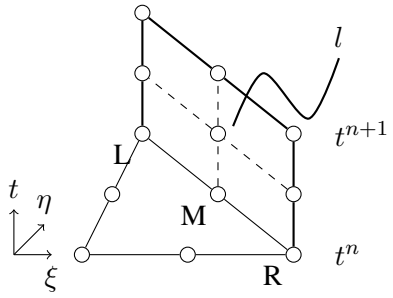


Figure 3.19: Flux interface l with space-time coordinate of $[\mathrm{L}, \mathrm{M}, \mathrm{R}] \times [t^n, t^{n+1}]$.

The two-dimensional interface flux is evaluated as a direct extension of the one-dimensional flux. The particular flux interface l with space-time coordinate of $[\mathrm{L}, \mathrm{M}, \mathrm{R}] \times [t^n, t^{n+1}]$ is presented as an example in Fig. 3.19. The interface flux is evaluated by the two-dimensional

Simpson's composite rule, which produces a second-order accurate numerical integration.

$$\begin{aligned}\bar{\mathbf{f}}_l &= \bar{\mathbf{f}}_{[\mathrm{L,M,R}]\times[t^n,t^{n+1}]} \\ &= \frac{1}{36}\Big[\mathbf{f}_L^n + \mathbf{f}_R^n + \mathbf{f}_L^{n+1} + \mathbf{f}_R^{n+1} + \\ &\quad 4\left(\mathbf{f}_M^n + \mathbf{f}_M^{n+1} + \mathbf{f}_L^{n+1/2} + \mathbf{f}_R^{n+1/2}\right) + 16\mathbf{f}_M^{n+1/2}\Big] + \mathcal{O}(\Delta\xi^3, \Delta\eta^3)\end{aligned} \tag{3.73}$$

where

$$\mathbf{f}_i^n = \begin{pmatrix} au_i^n \\ bu_i^n \end{pmatrix}, \quad \mathbf{f}_i^{n+1/2} = \begin{pmatrix} au_i^{n+1/2} \\ bu_i^{n+1/2} \end{pmatrix}, \quad \mathbf{f}_i^{n+1} = \begin{pmatrix} au_i^{n+1} \\ bu_i^{n+1} \end{pmatrix}, \quad i = L, M, R \tag{3.74}$$

3.3.4 Bubble Function

Previously, it was explained that the bubble function was introduced to maintain conservation in the two-dimensional AF scheme. We now describe why the bubble function is required in two-dimensional schemes. In order to do that, we first discuss why one-dimensional problems do not require a separate bubble function.

In one-dimensional problems, a continuous and compact quadratic reconstruction is constructed from two interface variables and one cell average value. From the conservative update, the cell average is approximated to $\mathcal{O}(\Delta\xi^3)$ by a finite volume update. In order to produce and maintain a third-order accurate scheme, the reconstruction used for next iteration updates must be accurate enough to continue the trend. The one-dimensional quadratic reconstruction has a built-in 'bubble function' property. The midpoint value in the quadratic reconstruction polynomial can freely be defined in terms of the conservative cell average. This is evident in the one-dimensional Simpson's quadrature formula.

$$\overline{u}_j^{n+1} = \frac{1}{\Delta\xi}\int_{\xi=0}^{1} u(\xi)d\xi = \frac{1}{6}\left(u_i^{n+1} + 4u_{i+1/2}^{n+1} + u_{i+1}^{n+1}\right) + \mathcal{O}(\Delta\xi^4) \tag{3.75}$$

where $u_{i+1/2}^{n+1} = u_j^{n+1}$ is the solution value at the midpoint $\xi = 0.5$, not the cell average. Indeed, the solution value at the midpoint corresponds to the Lagrange basis coefficient c_2 in one-dimensional reconstructions, see Table 3.2. This means that the quadratic reconstruction, constructed with already known three numerical solution values (two interface values and a cell average), is sufficient to completely convey the third-order accurate approximation of the solution.

The situation is a bit different in two-dimensional reconstructions. A two-dimensional triangular reconstruction is considered, but the same argument can be had for higher-

dimensional or two-dimensional quadrilateral reconstructions. A two-dimensional triangular quadratic reconstruction, $u_Q(x, y)$, is integrated in the physical space to find an area averaged value of the reconstruction.

$$\overline{u}_Q = \frac{1}{\Omega} \int_{x=0}^{1} \int_{y=0}^{1-x} u_Q(x, y) dy dx \tag{3.76}$$

where Ω is the element area. Because physical elements can vary in size and shape, we define the integration in the reference space. The mapping between the physical and reference element is exact in the case of linear simplex elements, so two integration processes are equivalent. Then, the area averaged reconstruction quantity, $\overline{u}_Q$, is obtained using a suitable numerical quadrature rule.

$$\begin{aligned} \overline{u}_Q &= \frac{\text{det}\mathbf{J}}{\Omega} \int_{\xi=0}^{1} \int_{\eta=0}^{1-\xi} u_Q(\xi, \eta) d\eta d\xi \\ &= 2 \int_{\xi=0}^{1} \int_{\eta=0}^{1-\xi} \sum_{i=1}^{6} c_i \phi_i(\xi, \eta) d\eta d\xi \\ &= \frac{1}{3} (u_2 + u_4 + u_6) \end{aligned} \tag{3.77}$$

The area integral of the two-dimensional quadratic reconstruction is only expressed as a sum of edge coefficients, u_2, u_4, u_6, of the reconstruction, and the inclusion of vertex values only increases the truncation error. It turns out this quadrature rule accurately integrates only up to quadratic polynomials [55]. If the cell average is approximated to third-order accuracy, then one needs to sample more points in the quadrature rule.

Without modifying the quadratic Lagrange basis functions so that solution reconstructions are C^0 continuous across elements, we introduce a *cubic* bubble function at a cell centroid. By defining the additional cubic coefficient in addition to the quadratic reconstruction, third-order data can be accurately approximated. The bubble function is

$$\phi_7 = 27\xi\eta(1 - \xi - \eta) \tag{3.78}$$

and it can also be found in Table 3.5. The bubble function, ϕ_7, vanishes on element boundaries.

$$u(\xi, \eta) = u_Q(\xi, \eta) + u_7\phi_7(\xi, \eta) \tag{3.79}$$

The integration of this new reconstruction polynomial is now third-order accurate.

$$\begin{aligned}\overline{u} &= \frac{\text{detJ}}{\Omega}\int_{\xi=0}^{1}\int_{\eta=0}^{1-\xi} u(\xi,\eta)d\eta d\xi \\ &= 2\int_{\xi=0}^{1}\int_{\eta=0}^{1-\xi}\left[\sum_{i=1}^{6} c_i\phi_i(\xi,\eta) + c_7\phi_7(\xi,\eta)\right]d\eta d\xi \\ &= \underbrace{\frac{1}{3}\left(u_2+u_4+u_6\right)}_{=\overline{u}_Q} + \frac{9}{20}u_7 \end{aligned} \tag{3.80}$$

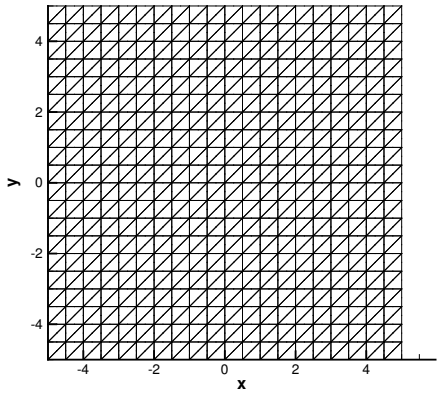
The coefficient associated with the bubble function, u_7, is easily found from Eq. (3.80), and it is also found in Table 3.5.

So far, we have been very insistent on **not** modifying the coefficients of Lagrange basis functions defined on element boundaries. This is to stay true to the AF philosophy of keeping a continuous solution reconstruction, and it is accomplished by the bubble function. As a result, we do not need a Riemann solver.

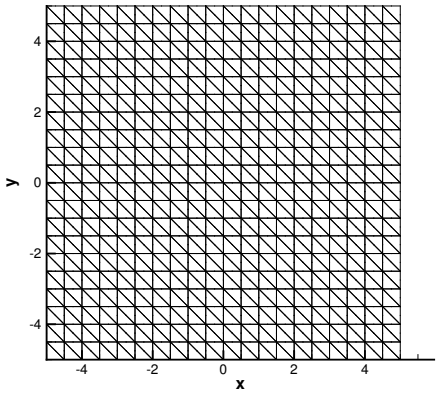
The bubble function in a quadrilateral element is derived using the same reasoning. However, the cubic function defined at the cell centroid is a natural part of the bi-quadratic reconstruction polynomial, unlike the cubic bubble function for a triangle. Table 3.6 summarizes the bubble coefficient and function.

3.3.5 Numerical Accuracy

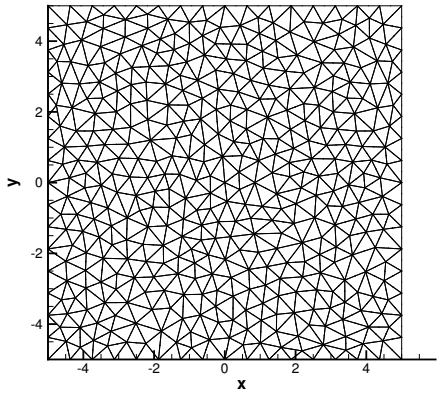
Computational grids used for numerical accuracy tests are illustrated in Fig. 3.20. For each grid type, there are four additional refinements. Triangular structured grids are uniformly refined by bisecting element edges. Unstructured grids have no definite structure to be refined uniformly. A Delaunay triangulation, provided by the mesh-generation software PointWise, is used to generate basic unstructured grids. A quadrilateral grid type is used for linear advection and nonlinear advection cases.



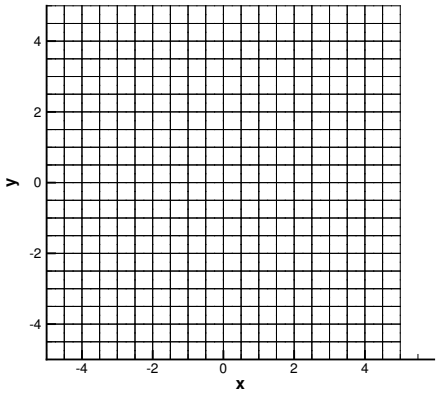
(a) Triangular structured grid type A. $N = 800$



(b) Triangular structured grid type B. $N = 800$



(c) Triangular unstructured grid. $N = 814$



(d) Quadrilateral grid. $N = 400$

Figure 3.20: Two-dimensional computational grid types used in numerical tests. The domain is restricted to $[x, y] \in [-5, 5]$. There are five grids for each type.

In AF, DOF for a triangle is defined as follows.

$$\text{DOF} = N_{\text{Cell}} + \frac{1}{2} N_{\text{Face}} + \frac{1}{6} N_{\text{Vertex}} \tag{3.81}$$

For a quadrilateral, DOF is defined as follows.

$$\text{DOF}_{\text{Quadrilateral}} = N_{\text{Cell}} + \frac{1}{2} N_{\text{Face}} + \frac{1}{4} N_{\text{Vertex}} \tag{3.82}$$

And the reference length of a grid, h, is defined as follows.

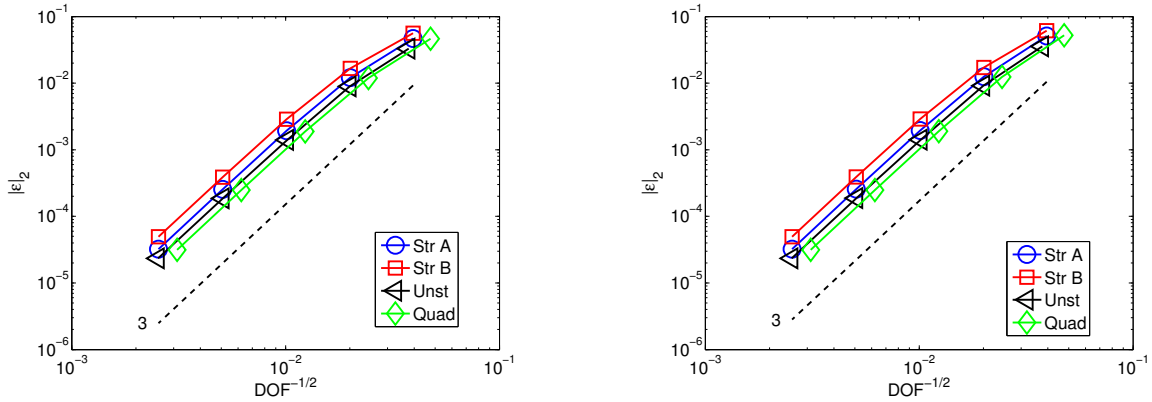
$$h = \text{DOF}^{-1/2} \tag{3.83}$$

A Gaussian initial condition is considered for the numerical accuracy test.

$$u_0(x, y) = \exp\left(-0.5(x^2 + y^2)\right) \tag{3.84}$$

First, we present a case where the advection direction is arbitrarily chosen. The advection speed vector is set to $\mathbf{a} = (1, 0.5)^T$. Periodic boundary conditions are imposed on both directions. The initial condition is advanced to final time $t = 20$ so that it travels around the computational domain two cycles and arrives at the original location. We verify the numerical accuracy using all grid configurations in Fig. 3.20.

The L_2 norm errors of the conservative variables and the solution reconstructions are shown in Fig. 3.21 for the Gaussian initial condition, Eq. (3.84). The conservative variables and solution reconstructions errors all exhibit convergence rates of $\mathcal{O}(h^3)$.



(a) $L_2(\overline{u})$ norm error

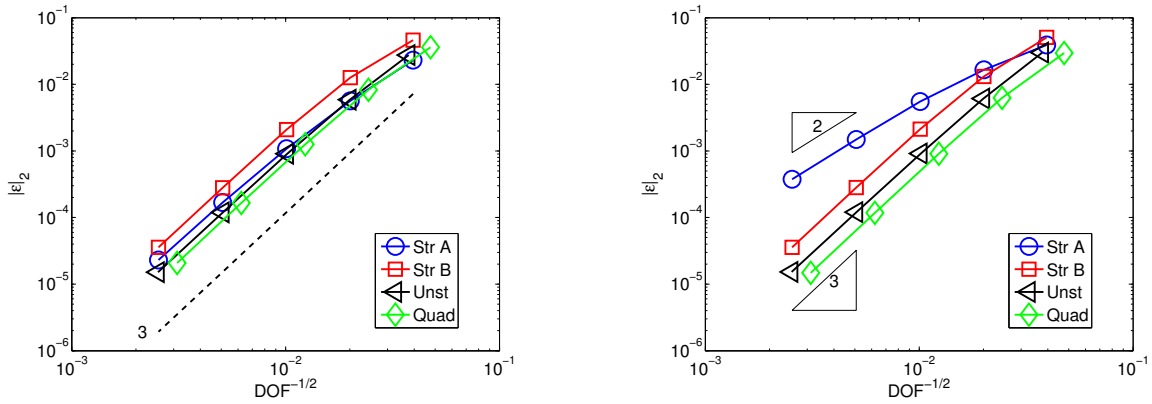
(b) $L_2(u(\mathbf{x}))$ norm error

Figure 3.21: Linear advection of Eq. (3.84) conservative variable error $L_2(\overline{u})$ and reconstruction error $L_2(u)$ at $t = 20$, $\nu = 0.7$, and $\mathbf{a} = (1, 0.5)^T$. Grid labels are: triangular structured grid A (Str A), B (Str B), unstructured grid (Unst), and quadrilateral grid (Quad).

Now, the advection velocity vector is set to $\mathbf{a} = (1, 1)^T$ so that it is aligned with one set of mesh edges for the triangular grid A. It had been expected that the special orientation of the mesh in the advection problem would prove favorable, and perhaps even provide an objective for a mesh adaptation. The L_2 norm errors of conservative variables and solution reconstructions are shown in Fig. 3.22 for the Gaussian initial condition, Eq. (3.84). The $L_2(\overline{u})$ norm errors of conservative variables show a convergence rate of $\mathcal{O}(h^3)$. However, the solution reconstruction $L_2(u(\mathbf{x}))$ norm error convergence rate for the triangular structured grid A is $\mathcal{O}(h^2)$.

This odd outcome of AF illustrates a possible shortcoming of the characteristic tracing method in very isolated conditions. In order to understand the cause and devise a remedy

for it, some analysis is due. The stability analysis is discussed in Section 3.3.7. In fact, convergence rates corresponding to a range of advection velocity are affected by the same phenomena. For more detail, see Section 3.4.



(a) $L_2(\overline{u})$ error (b) $L_2(u(\mathbf{x}))$ error

Figure 3.22: Linear advection of Eq. (3.84) conservative variable error $L_2(\overline{u})$ and reconstruction error $L_2(u)$, at $t = 10$, $\nu = 0.7$, and $\mathbf{a} = (1, 1)^T$. Grid labels are: triangular structured grid A (Str A), B (Str B), unstructured grid (Unst) and quadrilateral grid (Quad).

3.3.6 Numerical Comparison

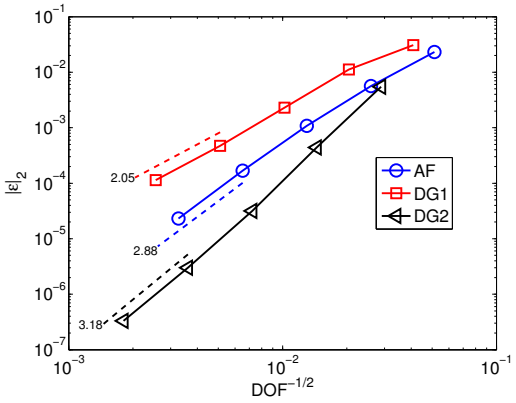
The AF error convergence results of the Gaussian wave initial condition, Eq. (3.84), are compared to the DG scheme error convergence results. DG schemes considered in this comparison are time explicit Runge-Kutta DG schemes. For spatial discretization, a linear ($p = 1$) and a quadratic ($p = 2$) basis functions are used, and are respectively referred to as DG1 and DG2. A time explicit Runge-Kutta (RK) method is employed for time discretization. Two- and three-stage RK methods are used for DG1 and DG2, respectively. Computational grids for the DG results are based on the same grids shown in Fig. 3.20, except for the quadrilateral grid. For DG schemes, the CFL stability is restricted by

$$\nu < \frac{1}{2p+1} \tag{3.85}$$

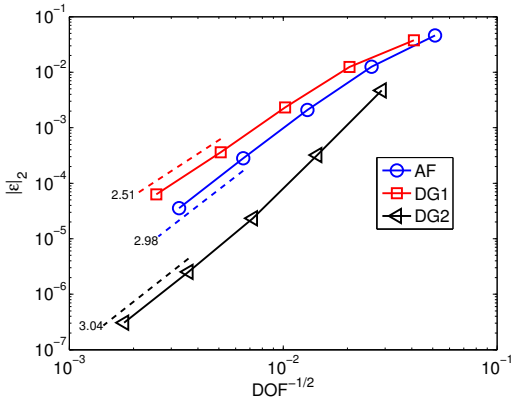
where p is the order of spatial discretization [14].

We first compare the $L_2(\overline{u})$ norm errors of the conservative variables in Fig. 3.23. DG schemes for linear problems can result in superconvergent accuracy. In this study, however, we only observe the standard convergence rate of $p + 1$ for DG schemes because global errors are measured. The superconvergence in DG is, however, is not generally observed for nonlinear problems. Certain scalar outputs, such as pressure and entropy, can reveal

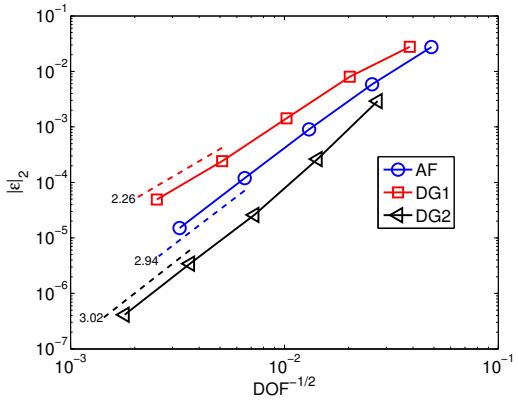
superconvergent behaviors. The superconvergence of errors can be observed when an integrated output error is instead measured. Alternatively, a special set of basis functions may also reveal the superconvergence of errors [90].



(a) Triangular structured grid A



(b) Triangular structured grid B



(c) Triangular unstructured grid

Figure 3.23: Comparisons of the $L_2(\overline{u})$ norm errors of conservative variables at $t = 10$ and $\mathbf{a} = (1, 0)^T$. Courant numbers are $\nu = 0.7$, $\nu = 0.3$, and $\nu = 0.2$ for AF, DG1, and DG2.

Since AF and DG schemes are very different by design, we must introduce a more appropriate standard to compare the schemes. The number of DOF would be appropriate for similarly based numerical schemes. Instead of using the DOF, we devise a common measure called a "work unit" (WU). The WU is defined as follows.

$$\mathrm{WU} = \mathrm{DOF}\frac{\text{Computation stages per iteration}}{\nu_{\max}} \tag{3.86}$$

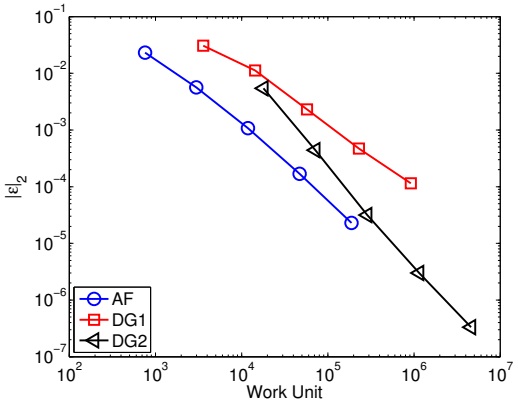
The WU introduces both spatial and temporal discretization parameters in order to compare a fully discrete scheme to a semi-discrete one. The parameters for evaluating the WU are listed in Table 3.7 for both schemes.

Table 3.7: Summary of parameters used for WU evaluations in AF and DG

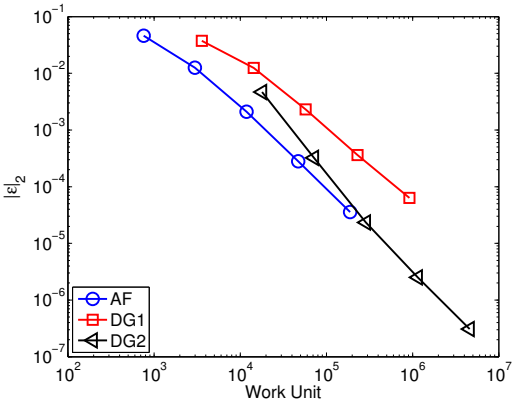
Scheme	DOF	Computation stages per iteration	ν_{max}
AF	$N_{\text{cell}} + \frac{1}{2}N_{\text{face}} + \frac{1}{6}N_{\text{vertex}}$	2	1
DG1	$3 \times N_{\text{cell}}$	2 (RK 2)	$\frac{1}{3}$
DG2	$6 \times N_{\text{cell}}$	3 (RK 3)	$\frac{1}{5}$

Even though the WU provides a theoretical standard to compare the efficiency of different schemes, it makes no allowance for the arithmetical complexity of schemes. From the perspective of each scheme, the efficiency depends largely on the way that conservation laws are solved. In DG, conservation laws must be formulated in the weak form. This prompts two residual terms to be evaluated using numerically accurate quadrature integration procedures. Because the way in which conservation laws are solved is already constrained by a set of procedures and rules, the DG method does not leave a lot of choice for variation. However, it creates a mathematically tidy framework for solving conservation laws. Unlike DG, the AF method takes a two stage approach with the last step being the conservation stage. The first stage, the nonconservative stage, a semi-Lagrangian method is used to update the variables in advection problems but other equivalent methods can be used. This aspect of AF provides many choices for improvements. Given the perspectives of each method, we understand the WU is neither the best nor the universal standard, but it can be understood as a *good* measure of the efficiency with regard to the theoretical construction of each scheme.

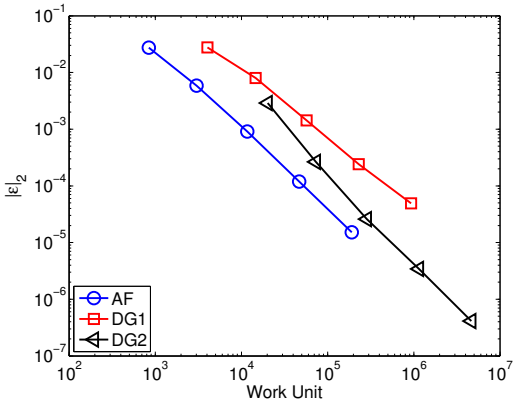
Using the WU, $L_2(\overline{u})$ norm errors of the conservative variables for all schemes are re-examined in Fig. 3.24. The results indicate that AF is nearly as effective as DG2 scheme. In fact, there is a marginal advantage for AF in the triangular unstructured grid case, Fig. 3.24c, see also Tables 3.8, 3.10 for tabulated data. Since the problem is smooth and well behaved, both AF and DG schemes would result in good error convergence. More interesting comparisons will be made for nonlinear problems, in Chapter 4 for Burgers' equation, Chapter 6 for the pressureless Euler equations and Chapter 7 for the Euler equations.



(a) Triangular structured grid A



(b) Triangular structured grid B



(c) Triangular unstructured grid

Figure 3.24: Comparisons of the $L_2(\overline{u})$ norm errors of conservative variables measured as a function of WU at $t = 10$ and $\mathbf{a} = (1, 0)^T$. Courant numbers are $\nu = 0.7$, $\nu = 0.3$ and $\nu = 0.2$ for AF, DG1, and DG2.

Table 3.8: L_2 norm errors of conservative variable and WU for AF on unstructured grid at $t = 10$ and $\nu = 0.7$

DOF	Work Unit	$L_2(\overline{u})$	Order
4.21e+02	8.42e+02	2.74e-02	
1.51e+03	3.03e+03	5.88e-03	2.41
5.86e+03	1.17e+04	9.05e-04	2.76
2.34e+04	4.69e+04	1.19e-04	2.93
9.52e+04	1.90e+05	1.51e-05	2.94

Table 3.9: L_2 norm errors of conservative variable and WU for DG1 on unstructured grid at $t = 10$ and $\nu = 0.3$

DOF	Work Unit	$L_2(\overline{u})$	Order
6.72e+02	4.03e+03	2.76e-02	
2.44e+03	1.47e+04	8.00e-03	1.92
9.53e+03	5.72e+04	1.42e-03	2.53
3.82e+04	2.29e+05	2.41e-04	2.56
1.56e+05	9.33e+05	4.93e-05	2.26

Table 3.10: L_2 norm errors of conservative variable and WU for DG2 on unstructured grid at $t = 10$ and $\nu = 0.2$

DOF	Work Unit	$L_2(\overline{u})$	Order
1.34e+03	2.02e+04	2.90e-03	
4.88e+03	7.33e+04	2.64e-04	3.71
1.91e+04	2.86e+05	2.60e-05	3.41
7.64e+04	1.15e+06	3.44e-06	2.91
3.11e+05	4.67e+06	4.14e-07	3.02

3.3.7 Stability Analysis

The stability analysis of the two-dimensional AF for linear advection is discussed. The triangular structured grid configuration does not make a convenient analysis candidate due to the placement of the bubble function for the conservation property of the scheme. Therefore, we introduce a quadrilateral structured stencil to perform Fourier analysis. Although Fourier analysis is applied on the quadrilateral grid structure, the analysis is still suggestive of the stability of the AF method in other grid configurations.

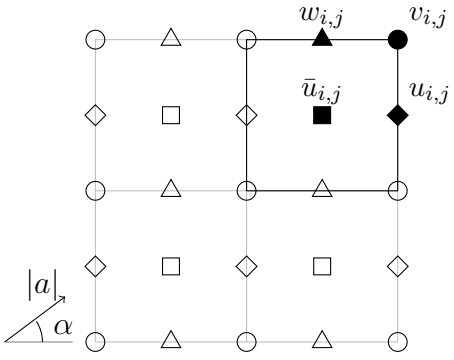


Figure 3.25: Quadrilateral grid for 2D stability analysis. $\mathbf{a} = |\mathbf{a}|(\cos\alpha, \sin\alpha)^T$

We introduce four degrees of freedom, see Fig. 3.25. Because interface variables evolve independently of the cell average, the AF method for two-dimensional linear advection is a system comprised of the vector $\mathbf{u}_{i,j}$. The system in a compact form is shown.

$$\mathbf{u}_{i,j}^{n+1} = \mathbf{G}_{\text{2D}}\mathbf{u}_{i,j}^{n} \tag{3.87}$$

where

$$\mathbf{u}_{i,j} = \begin{pmatrix} \bar{u}_{i,j} \\ u_{i,j} \\ v_{i,j} \\ w_{i,j} \end{pmatrix} = \begin{pmatrix} \blacksquare \\ \blacklozenge \\ \bullet \\ \blacktriangle \end{pmatrix} \tag{3.88}$$

and $\mathbf{G}_{\text{2D}}$ is a 4×4 matrix describing the discrete update nature of the system.

Von Neumann analysis in two dimensions assumes the solution vector takes the form, as a superposition of x and y harmonics.

$$\mathbf{u}_{i,j}^{n} = g^{n} e^{\mathbf{i}(i\phi_x + j\phi_y)} \tag{3.89}$$

where $\mathbf{i} = \sqrt{-1}$. The amplification matrix $\mathbf{G}_{\text{2D}}$ can be decomposed to produce four sets of eigensystem, the eigenvalue and eigenvector pair. Among the four eigensystems, two systems compose a solution mode as will be described soon.

3.3.8 Mesh Alignment Problem

Heuristic evidence for mesh alignment problem is presented. We will take a look at the special case in which $\alpha = 0$ and $\nu = 0.5$. Under this condition, we have two solution modes that should propagate in the x-direction. One solution mode involves the corner ($\bullet$) and the horizontal edge ($\blacktriangle$). And the other involves the vertical edge ($\blacklozenge$) and the cell average ($\blacksquare$). The complete amplification matrix for this special case is presented.

$$\mathbf{G}_{\text{2D}} = \begin{pmatrix} \frac{1}{2} + \frac{1}{2}e^{-\mathbf{i}\phi_x} & -\frac{1}{12} + \frac{1}{6}e^{-\mathbf{i}\phi_x} - \frac{1}{12}e^{-2\mathbf{i}\phi_x} & G_{1,3} & 0 \\ \frac{9}{4} & -\frac{1}{4} - \frac{1}{4}e^{-\mathbf{i}\phi_x} & G_{2,3} & -\frac{1}{4} - \frac{1}{4}e^{-\mathbf{i}\phi_y} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & e^{-\mathbf{i}\phi_x} & 0 \end{pmatrix} \tag{3.90}$$

where $G_{1,3}$ and $G_{2,3}$ are listed separately below.

$$G_{1,3} = -\frac{1}{48} + \frac{1}{24}e^{-\mathbf{i}\phi_x} - \frac{1}{48}e^{-2\mathbf{i}\phi_x} + \frac{1}{24}e^{-\mathbf{i}(\phi_x+\phi_y)} - \frac{1}{48}e^{-\mathbf{i}(2\phi_x+\phi_y)} - \frac{1}{48}e^{-\mathbf{i}\phi_y}$$
$$G_{2,3} = -\frac{1}{16} - \frac{1}{16}e^{-\mathbf{i}\phi_x} - \frac{1}{16}e^{-\mathbf{i}(\phi_x+\phi_y)} - \frac{1}{16}e^{-\mathbf{i}\phi_y}$$

In order to understand the mesh alignment problem further, we first find the eigenvalues and eigenvectors of the matrix $\mathbf{G}_{\text{2D}}$. The matrix is 4×4 so there are four eigenvalue and eigenvectors. While the eigenvalues describe the stability property of the mode, we may not be able to distinguish the behavior of the mode in relation to the solution vector. So, it is instructive to show the corresponding eigenvectors to determine what that solution mode it describes.

The first pair is shown.

$$g_{1,2} = \mp e^{-\mathbf{i}\phi_x} \tag{3.91}$$

$$\mathbf{v}_{1,2} = \begin{pmatrix} 0 \\ \pm\frac{1}{2}\cos\frac{\phi_x}{2} \\ \mp e^{\mathbf{i}\frac{\phi_x}{2}} \\ 1 \end{pmatrix} \tag{3.92}$$

What is quite noticeable is the first component of these eigenvectors being zero. This means that this solution mode is decoupled from the cell average (■). Likewise, we take a look at the remaining pair.

$$g_{3,4} = \frac{1}{8}e^{-\mathbf{i}\frac{\phi_x}{2}}\left(2\cos\frac{\phi_x}{2} \mp \sqrt{6}\sqrt{7-\cos\phi_x}\right) \tag{3.93}$$

$$\mathbf{v}_{3,4} = \begin{pmatrix} \frac{1}{18}e^{-\mathbf{i}\phi_x}\left(3 \mp \sqrt{6}\sqrt{-e^{-\mathbf{i}\phi_x}(-7+\cos\phi_x)} + e^{3\mathbf{i}\phi_x}\right) \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{3.94}$$

In this case, the eigenvectors have zero entries in the last two rows corresponding to (●) and (▲) spaces. Therefore, this mode $g_{3,4}$ belongs to the solution that consists of the cell average (■) and the vertical edge (◆). In fact, this mode corresponds exactly to the one-dimensional scheme, Eq. (3.42).

The Taylor expansions of two solution modes $g_{1,2}$ and $g_{3,4}$ are provided to gain some insights of the low-frequency behavior. It turns out that only g_2 and g_4 are relevant for the

low-frequency range.

$$g_2 = 1 - \frac{1}{2}\mathbf{i}\phi_x - \frac{1}{8}\phi_x^2 + \frac{1}{48}\mathbf{i}\phi_x^3 + \frac{1}{384}\phi_x^4 + \mathcal{O}(\phi_x^5) \tag{3.95}$$

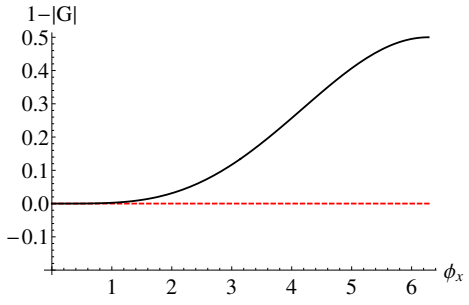
$$g_4 = 1 - \frac{1}{2}\mathbf{i}\phi_x - \frac{1}{8}\phi_x^2 + \frac{1}{48}\mathbf{i}\phi_x^3 + 0\phi_x^4 + \mathcal{O}(\phi_x^5) \tag{3.96}$$

The g_4 expansion is accurate up to, but not including, the $\mathcal{O}(\phi_x^4)$ term. This would indicate that the truncation error of the mode involving the cell average is of order $\mathcal{O}(\phi_x^4)$, so the numerical accuracy of the cell average is third-order. The eigenvalue expansion g_2 is an exact expansion of $e^{-\mathbf{i}\frac{\phi_x}{2}}$. It does not seem obvious at first that this should happen, but the AF scheme is exact at Courant numbers of $\nu = 0, 0.5, 1$. Since the analysis is done for $\nu = 0.5$, the solution mode along the corner (●) and the horizontal edge (▲) is exact.

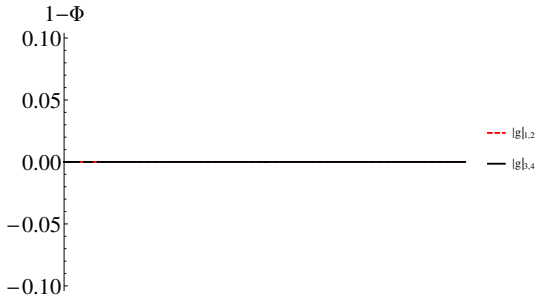
From all these, we have successfully determined the following.

- $g_{1,2}$ mode: the solution along the mesh geometry involving the corner (●) and the horizontal edge (▲). This mode corresponds to the nonconservative solution mode.
- $g_{3,4}$ mode: the solution along the interior of element involving the cell average (■) and the vertical edge (◆). This mode corresponds to the conservative solution mode.

The dissipation and dispersion error properties are shown in Fig. 3.26. The red dashed curve describes the stability properties of one mode $g_{1,2}$. The black curve describes the stability properties of the other mode $g_{3,4}$. The solution mode $g_{1,2}$ shows zero dissipation while the solution mode $g_{3,4}$ shows exactly the same behavior as one-dimensional AF scheme, for reference see Fig. 3.7 and compare Eqs. (3.42) and (3.93). Both solution modes exhibit zero dispersion error because the phase property is exact for $\nu = 0.5$.



(a) Dissipation error. Mode $g_{1,2}$ shows zero dissipation error.

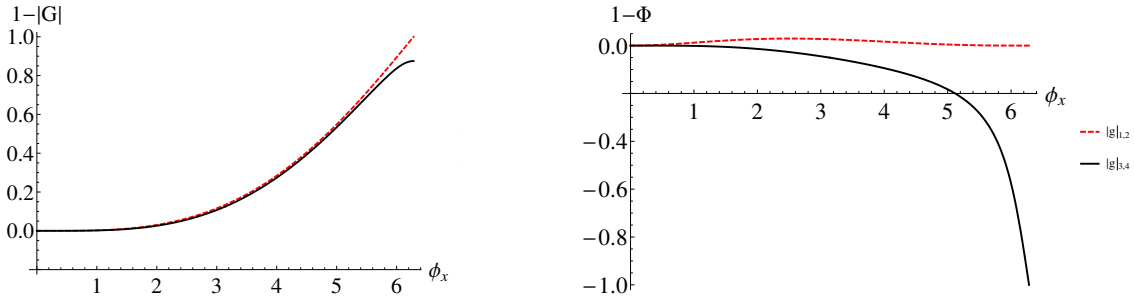


(b) Dispersion error showing zero for both modes

Figure 3.26: Dissipation and dispersion error properties of mesh aligned problem modes for $\alpha = 0, \nu = 0.5$, and $\phi_y = \phi_x$

The error properties shown under the special condition, $\alpha = 0$ and $\nu = 0.5$, definitely illustrate two distinct solution modes in AF. In fact, the solution along the mesh geometry never diffuses. This lack of cross-wind diffusion between two solution modes explains the mesh alignment problem in two dimensions.

It is also instructive to consider a case at an arbitrary Courant number. We consider changing the Courant number to $\nu = 0.25$ while leaving the flow angle unchanged. The dissipation and dispersion errors are shown in Fig. 3.27. The dissipation errors of the two modes are very similar. In fact, the only noticeable difference comes out at high-frequency ranges. However, the dispersion errors are very distinct. The nonconservative solution mode $g_{1,2}$ has an overall better phase property than the conservative solution mode $g_{3,4}$. Also, we note that the conservative solution mode $g_{3,4}$ corresponds exactly to the one-dimensional AF scheme.



(a) Dissipation error

(b) Phase error

Figure 3.27: Dissipation and dispersion error properties of mesh aligned problem modes for $\alpha = 0, \nu = 0.25$, and $\phi_y = \phi_x$

The current hypothesis is as follows. In the case of mesh alignment, two solution modes are decoupled and can behave in two distinct ways. This may be unnoticed over a short time period but will be definitely noticeable for long periods of time. Although both modes have the same order of accuracy, they will behave a little differently because there is, by design, no cross-wind diffusion between them. We now present some ways we could add some numerical diffusion to the two-dimensional scheme in the following section.

3.4 Mesh Alignment Problem in Multidimensional Advection

We have found some evidence of the mesh alignment problem from the stability analysis and numerical accuracy test. From the stability analysis in Section 3.3.7, we have dis-

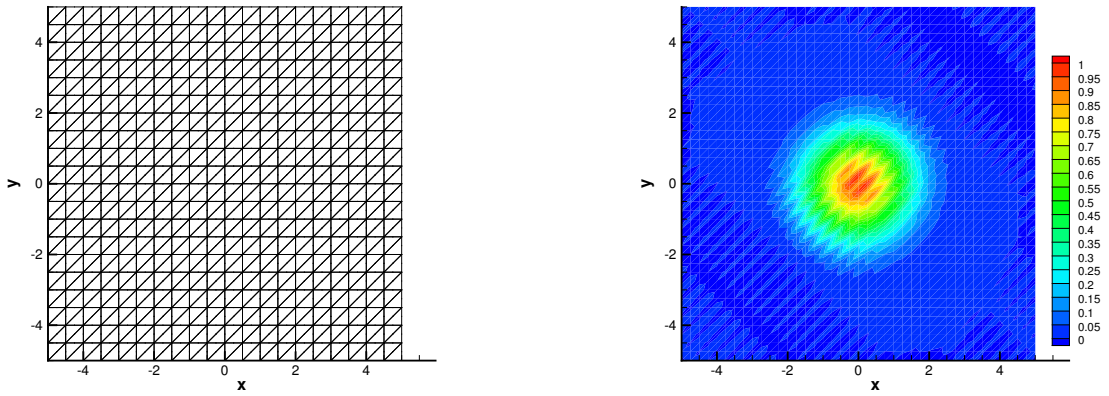
covered two distinct solution modes which may behave differently in time. From the numerical accuracy test in Section 3.3.5, the error associated with the solution reconstruction converged at a lower convergence rate.

Here, we try to identify the mesh alignment problem in detail. Specifically, we try to identify to what extent the mesh alignment problem is an issue by quantifying some error behavior. Also, we propose a couple remedies for the mesh alignment problem. These remedies are derived from observations that two distinct solution modes evolve without significant coupling between them. And these remedies will involve some sort of crosswind diffusion to introduce some artificial coupling between the two solution modes identified in the stability analysis.

First, a Gaussian wave initial data is subjected to linear advection to present the problem we face.

$$u(x, y) = \exp\left(-0.5(x^2 + y^2)\right) \tag{3.97}$$

with $\mathbf{a} = (1, 1)^T$ so that one solution mode is aligned with the diagonals of the mesh, see Fig. 3.28 for the mesh. The solution is advanced to $t = 10$, at which time the solution has moved around the domain one cycle. The final solution is presented to have the distinct mesh alignment problem in Fig. 3.28.



(a) Triangular structured grid type A

(b) Solution at $t = 10$

Figure 3.28: Linear advection of the Gaussian wave, Eq. (3.97), at $t = 10$, $\nu = 0.8$, and $\mathbf{a} = (1, 1)^T$

3.4.1 Identifying Mesh Alignment Problem

In order to identify to what extent the mesh alignment problem affects the numerical solution, a simple test is carried out.

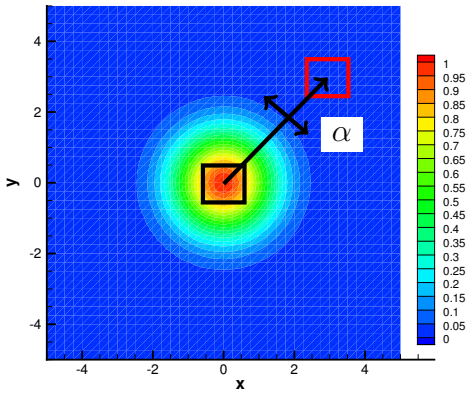
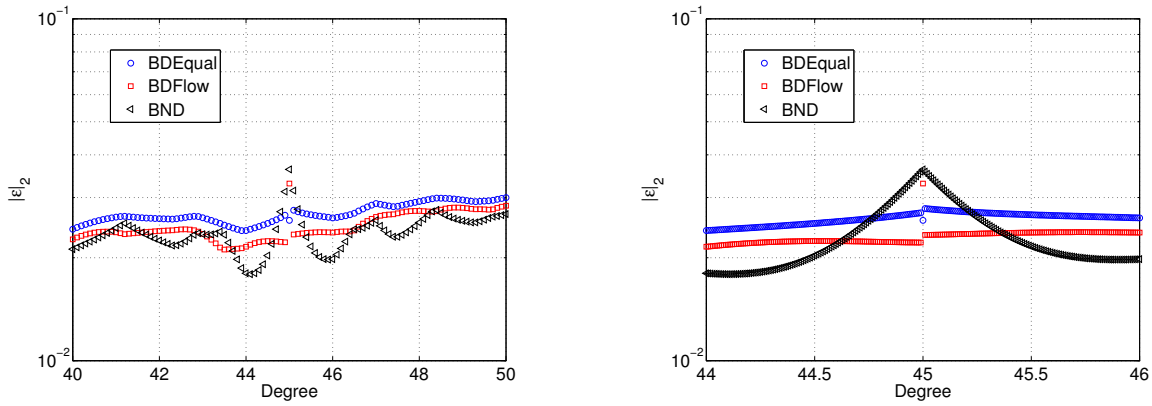


Figure 3.29: Initial (black) and final (red) regions from which the solution error is computed. Each region size is $[-0.25, 0.25] \in [x, y]$. The advection velocity vector is $\mathbf{a} = |\mathbf{a}|(\cos\alpha, \sin\alpha)^T$ where the angle α is varied around $45°$.

The initial solution in Fig. 3.29 is advected for a set number of time steps. The advection velocity is varied such that the advection angle, α, changes while the magnitude remains constant. The solution at the initial time (region shown in black) is compared to the solution at the final time (red box) and the relative error is measured by comparing solutions at two times. Prior to discussing the details of remedies for the mesh alignment problem, the error measurements are displayed in Fig. 3.30. The error measurement for the original solution without the bubble function treatment, labeled (BND), shows a drastic increase in the relative error at $45°$. The solutions treated with two bubble function remedies studied in the following section have more well-behaved errors near $45°$, solutions labeled (BDEqual) and (BDFlow).



(a) Angle range $40° - 50°$ (b) Angle range $44° - 46°$

Figure 3.30: $L_2(u(\mathbf{x}))$ norm error of the solutions between the initial and final regions shown in Fig. 3.29. The errors are shown for the original solution without any bubble function treatments (BND), the solution treated with remedy 1 (BDEqual) and the solution treated with remedy 2 (BDFlow).

We identified the mesh alignment problem occurs when the advection direction is aligned, or very nearly aligned, with the mesh geometry by measuring the numerical errors before and after some times. This is in accordance with what we have observed from the previous stability analysis in Section 3.3.7 and the numerical error analysis in Section 3.3.5. With these supporting arguments, we now propose two remedies that aim to reduce the mesh alignment problem.

3.4.2 Mesh Alignment Problem Remedy – Bubble Function Damping

As previously suggested, the mesh alignment problem is closely associated with two distinct solution modes. The two solution modes decouple from each other when the characteristic interpolation process occurs near mesh geometry. The solution mode of the conservative scheme, which we regard as the solution that approximates the numerical scheme closely, involves the bubble function. The bubble function, therefore, will take a central role in the design and implementation of the mesh alignment problem remedy.

The bubble function, discussed in Section 3.3.4, has two main attributes in the AF scheme. The first, and obvious, one is the compact conservation property. The second is the anti-diffusive effect. The second effect is partly due to the compact design of the bubble function. The bubble function does not have any influence on the element boundaries. This means that when the solution interpolation occurs along the boundary of an element, the interpolated solution does not contain the conservative property of solution. This can be mitigated by introducing some crosswind diffusion in the scheme. One particular way

to introduce the crosswind diffusion is to apply some damping on the bubble coefficient. Then, the same amount of bubble function that was damped is redistributed to the immediate outflow edges to maintain conservation.

This so-called the bubble function damping idea is similar to smoothing process exhibited by many elliptic problems. However, unlike in elliptic problems, this damping process must take place concurrently with the conservation update because we are solving unsteady flow problems.

The bubble function damping procedure which respects the conservation property is laid out.

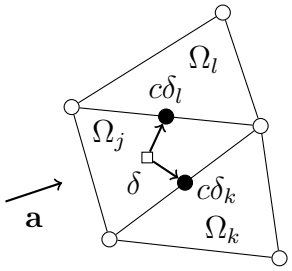


Figure 3.31: Bubble function damping problem setup. Assume here that there are two outflow edges to which the bubble function damping δ is distributed.

Suppose that the cell j has one inflow edge, Fig. 3.31. A portion of the bubble function in cell j, δ, is to be spread out to outflow edges of the element by $c\delta_l$ and $c\delta_k$. Here, c is a coefficient that makes sure the damping process occurs without loss of conservation in the control volume encompassing cells Ω_j, Ω_k, and Ω_l. The coefficient c will be evaluated soon.

Total conservation in cell j is

$$\Omega_j \bar{u}_j = \Omega_j \left[\frac{9}{20}(u_7) + \frac{1}{3}\left(u_2 + u_4 + u_6\right) \right]. \tag{3.98}$$

where the indices for the standard triangular element are found in Fig. 3.16. Conservation in each element is obtained by the sum of a bubble function and edge values.

If the bubble function, u_7, is damped by the amount δ and the outflow edges receive amounts $c\delta_{l,k}$, the change in conservation in element j and its adjacent elements k, and l

are established as

$$\Delta u_j = \Omega_j \left[\frac{9}{20}(-\delta) + \frac{1}{3}\left(c\delta_l + c\delta_k\right) \right] \tag{3.99}$$

$$\Delta u_k = \Omega_k \left[\frac{1}{3} c\delta_k \right] \tag{3.100}$$

$$\Delta u_l = \Omega_l \left[\frac{1}{3} c\delta_l \right] \tag{3.101}$$

where the coefficient c is required for the following reason. In the general unstructured grid setup, every cell has different areas. Consequently, the effect of adjusting edge values is felt differently for elements with different sizes. So, we introduce the coefficient c to accurately reflect this nature of the unstructured grid. The coefficient c is solved by correctly assuming that $\sum_{i=j,k,l} \Delta u_i = 0$, is shown.

$$c = \frac{\frac{27}{20}\delta}{\left(1 + \frac{\Omega_k}{\Omega_j}\right)\delta_k + \left(1 + \frac{\Omega_l}{\Omega_j}\right)\delta_l} \tag{3.102}$$

The choice of $c\delta_l$ and $c\delta_k$ determines how the bubble function is damped. Two choices are described in following subsections.

One may find that conservation in cell j has not changed because of the damping. However, the total conservation in adjacent elements is changed because one of the shared edge values have changed. In order to avoid this, conservation is enforced in the adjacent elements by re-evaluating a new bubble coefficient in those cells.

So far, we have only considered a case where cell j has one inflow edge and two outflow edges. The other case where cell j has two inflow edges is a simple special case. In this special case, we damp the bubble function to the only outflow edge available. In the following sections, the specifics of the bubble damping remedies are discussed. They only differ in how the distribution is made.

3.4.2.1 Equal Distribution Remedy

In this method, the bubble function is distributed equally to all edges, excluding the inflow edge. Because all outflow edges get equal amounts, the following relation for δ is established.

$$\delta = \delta_l = \delta_k \tag{3.103}$$

The coefficient c, Eq. (3.102), is simplified as follows.

$$c = \frac{\frac{27}{20}}{\left(1 + \frac{\Omega_k}{\Omega_j}\right) + \left(1 + \frac{\Omega_l}{\Omega_j}\right)} \tag{3.104}$$

The streamline edges in mesh flow aligned cases are treated the same way as the outflow edges, and this introduces a crosswind diffusion to the scheme.

3.4.2.2 Flow Direction Dependent Distribution Remedy

In this method, the bubble function is distributed to edges according to the flow direction variable k_i. For an edge i, we define

$$k_i = \frac{1}{2}\mathbf{a} \cdot \mathbf{n}_i \tag{3.105}$$

where $\mathbf{n}_i$ is scaled outward normal vector and $\mathbf{a}$ is the linear advection velocity vector. Additionally, the flow direction parameter for each outflow edge is defined as follows.

$$\gamma_k = \frac{k_k}{|k_{\text{inflow}}|}, \quad \gamma_l = \frac{k_l}{|k_{\text{inflow}}|} \tag{3.106}$$

Note that $\sum k_i = k_{\text{inflow}} + k_l + k_k = 0$ due to the closure of the triangle. Then,

$$\delta_l = 2\gamma_l \delta \tag{3.107}$$

$$\delta_k = 2\gamma_k \delta \tag{3.108}$$

The coefficient c, Eq. (3.102), is evaluated as according to the original general form.

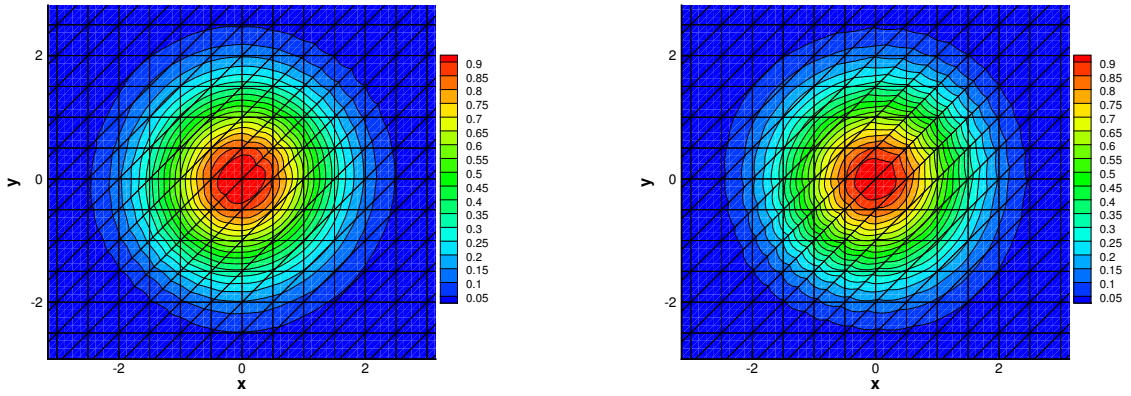
$$c = \frac{\frac{27}{20}\delta}{\left(1 + \frac{\Omega_k}{\Omega_j}\right)\delta_k + \left(1 + \frac{\Omega_l}{\Omega_j}\right)\delta_l} \tag{3.102}$$

where δ_l and δ_k are now known. By employing this method, we can distribute the bubble function to the outflow edges only. However, it creates a strong bias to damp the bubble function towards outflow edges in case of a perfect alignment.

3.4.2.3 Results

The linear advection results after the bubble damping procedures have been applied are shown. In Fig. 3.32, the solutions have improved and no longer exhibit very strong mesh

alignment problem. The close up solutions of each remedy display some qualitative measures of the bubble function damping procedures.



(a) Equal distribution

(b) Flow direction dependent distribution

Figure 3.32: Comparison of two bubble function damping mechanisms on the triangular structured grid type A at $t = 10$, $\nu = 0.8$, and $\mathbf{a} = (1, 1)^T$

For the two bubble function damping mechanisms compared, the numerical error convergence study is performed and the summary plot is shown in Fig. 3.33. The reconstruction error convergence for the original scheme, $\mathcal{O}(h^2)$ due to mesh alignment problem, is shown as a reference. After applying the bubble function damping method where the damped amount is equally distributed to all outflow edges, the error in reconstructions converges at the rate of $\mathcal{O}(h^3)$. However, the flow direction dependent bubble damping does not show a measurable gain as the mesh is refined.

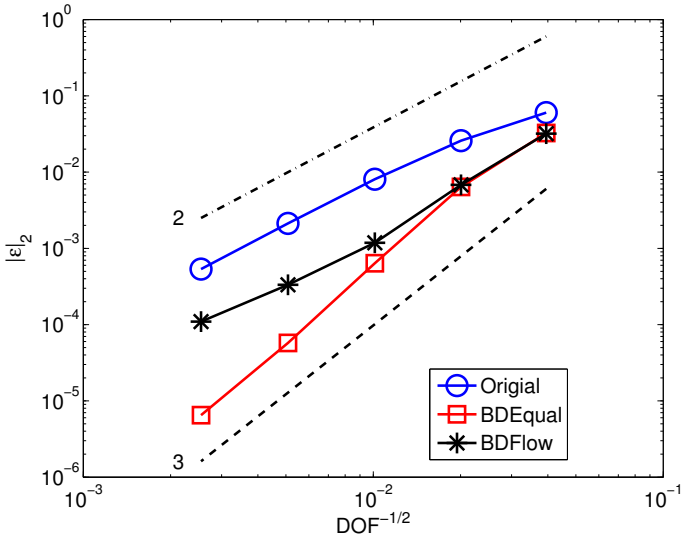


Figure 3.33: Gaussian problem $L_2(u(\mathbf{x}))$ norm error in solution reconstruction comparison between bubble function damping mechanisms with $\mathbf{a} = (1,1)^T$, $t = 10$, $\nu = 0.8$ on the triangular structured grid type A. Original scheme (Org), bubble damping by equal distribution (BDEqual) and flow direction dependent distribution (BDFlow).

3.4.3 Assessment

One of the remedies presented here, the equal distribution method, has successfully removed noticeable mesh alignment issues in linear advection. In the equal distribution method, bubble function coefficients were distributed to triangle edges to ensure that the conservation requirement was not violated. We recognize that vertex points do not affect the overall conservation within the element. So, an alternative form of the remedy can be made to include vertex points of triangles to improve the smoothness of solution reconstructions.

However, the bubble function damping approaches we developed here to resolve the mesh alignment problems might pose some practical limitations. The computational cost for distributing and damping bubble functions may become a deciding point. In scalar conservation laws, only one bubble function exists. However, there are multiple bubble functions to damp in systems. In addition, these remedies will need to be applied at every iteration. Since bubble functions are third-order quantities, one must also consider the potential increase in cost. Furthermore, the characteristic tracing method works well on unstructured grids. However, working only with unstructured grids may take away a possibility of a useful mesh adaptation property. We would like to work with unstructured grids that are locally structured for the mesh adaption.

For these reasons, it is beneficial to devise an economical advection algorithm that

works well with all mesh types. As we have discovered, the characteristic tracing method introduces no numerical diffusion. Some clever introduction of numerical diffusion that does not interfere with the accuracy of the scheme would be one way to avoid the mesh alignment issue. For this, the residual distribution method may be a suitable candidate. The residual distribution method is less intuitive than the characteristic tracing method, but it is exact on aligned grids.

CHAPTER 4

Nonlinear Scalar Conservation Laws

A nonlinear scalar conservation law is presented.

$$\frac{\partial u}{\partial t} + \frac{\partial \mathbf{f}(u)}{\partial \mathbf{x}} = 0 \tag{4.1}$$

where $\mathbf{f}(u)$ is a nonlinear flux function. Nonlinear scalar conservation laws have traditionally been used as stepping stones toward the treatment of a realistic system.

The AF method for nonlinear scalar conservation laws is a simple extension of the method for the linear problem. The first update stage is achieved by the characteristic tracing method, or a semi-Lagrangian method, to advance interface solutions fixed at mesh points. In it, a simple yet powerful nonlinear correction is introduced to achieve an accurate interpolation in nonlinear advection. This certainly offers some difference to many numerical methods based on Riemann solvers. The second update stage, a finite volume method, provides conservation and accuracy to the final solution. The conservative flux used in the conservation update stage is obtained from the Simpson's rule, previously developed in linear scalar conservation laws in Chapter 3.

4.1 Nonlinear Characteristics

We begin with the one-dimensional inviscid Burgers' equation with initially smooth data. The nonconservative, or the quasi-linear, form of the nonlinear conservation law Eq. (4.1) is

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial u}\frac{\partial u}{\partial x} = 0 \tag{4.2}$$

where $f(u) = \frac{1}{2}u^2$. It is further reduced to the following form which reveals the advective property of the conservation law.

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0 \tag{4.3}$$

One notices straightaway that this equation draws a striking similarity with linear advection Eq. (3.2). The method of characteristics established for linear hyperbolic partial differential equations still hold, but on a more restricted basis. The parametrization of Burgers' equation gives analogous results to the linear equation case.

$$\frac{du}{ds} = \frac{\partial u}{\partial t}\frac{dt}{ds} + \frac{\partial u}{\partial x}\frac{dx}{ds} = \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0 \tag{4.4}$$

A notable difference comes from the observation that each characteristic line may have a different slope or speed. Varying slopes of characteristics motivate mathematically and physically interesting phenomena. As characteristics converge to or diverge from each other, two analogous features, a shock or an expansion wave, can be produced in Burgers' equation. Burgers' equation is a remarkable model problem which allows us to closely study the nonlinear scalar conservation laws.

The method of characteristics is not valid when the solution contains discontinuities. These nonlinear features in flows are resolved through the conservation update stage in the AF advection scheme.

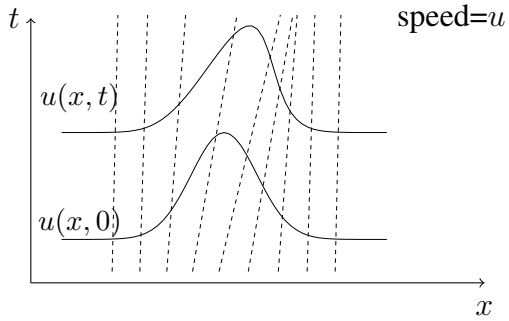


Figure 4.1: Characteristics of Burgers' equation in one dimension. The initial solution translates along the same family of characteristics with varying speed u.

The exact solution of Burgers' equation is obtained by the implicit formula.

$$u(x,t) = u(x - ut, 0) \tag{4.5}$$

where the characteristic speed is not constant anymore in contrast to Eq. (3.5).

One version of Burgers' equation is considered for the developing a multidimensional method.

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} + \frac{\partial g(u)}{\partial y} = 0 \tag{4.6}$$

where $f(u) = g(u) = \frac{1}{2}u^2$. This is a straightforward generalization of the one-dimensional characteristic equation. It follows immediately that the nonconservative form of this Burg-

ers' equation is the characteristic equation.

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + u\frac{\partial u}{\partial y} = 0 \tag{4.7}$$

The exact solution is found by integrating characteristic equations.

$$u(x, y, t) = u(x - ut, y - ut, 0) \tag{4.8}$$

The general approach described here for solving nonlinear scalar conservation laws presents some contrast with the historical achievement made by van Leer [84]. In the reference, van Leer presented Scheme V to solve the one-dimensional inviscid Burgers' equation. A quadratic solution was advanced in time using a semi-Lagrangian method. Once advanced, the quadratic solution was exactly integrated to obtain the analytic fluxes for conservation. It showed an effective and accurate way to solve nonlinear advection, but it proved to be difficult to achieve for more complex problems or in multidimensional problems. In the AF method, we introduce a simple and elegant extension to the characteristic tracing method for nonlinear advection.

4.2 Inviscid Burgers' Equation in One Dimension

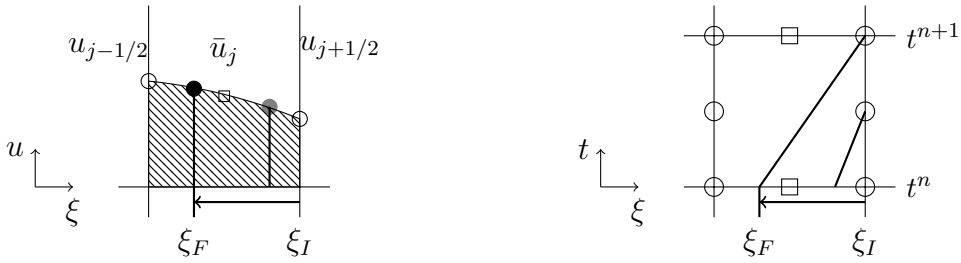
An advection scheme for one-dimensional nonlinear scalar conservation laws is described. It is important to consider both the conservative and nonconservative forms of Burgers' equation. In the evolution update stage, the nonlinear characteristic tracing method is used to evolve the interface variables to half- and full-time steps. These are used to evaluate conservative fluxes for the conservation update. Conservation is achieved by a finite volume method. The one-dimensional reconstruction procedure remains the same as the procedure described in linear advection and is omitted. Refer to Section 3.2 for details.

4.2.1 Evolution

We recall the nonconservative form of Burgers' equation.

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0 \tag{4.3}$$

Characteristics in Burgers' equation are straight. However, they no longer share the same slope. As a consequence, the advection scheme introduced in linear advection must be modified to address the nonlinear property.



(a) Solution reconstruction

(b) Nonlinear characteristic tracing

Figure 4.2: Description of nonlinear advection characteristic tracing in 1D

Recalling the exact solution of Burgers' equation, Eq. (4.5), it is apparent that one cannot find the exact solution without either solving the exact implicit formula or solving it iteratively. Instead, we present a nonlinear variation of the characteristic tracing method. The nonlinear characteristic tracing method requires a correction to the characteristic speed.

For a third-order conservative scheme, a second-order numerical flux is required. And the second-order accurate flux requires at least a first-order characteristic wave speed. The characteristic speed provided at the interface is only an $\mathcal{O}(1)$ value. So, a Taylor series approximation of the characteristic speed can provide an added degree of accuracy. We seek to find a Taylor expansion of the solution u about a cell interface since it is where a conservative flux is evaluated.

For the moment, we lay out the mathematical derivation with no regard to the numerical implementation aspect. Not considering the dependency of the solution in space, the abbreviation is introduced $u(x,t) = \tilde{u}(t)$ to indicate we are interested in the temporal expansion of the solution $u(x,t)$ at some location x. Then, a Taylor expansion of solution $\tilde{u}(t)$ in time is written as follows.

$$\tilde{u}(t) = \tilde{u}(0) + t\frac{\partial \tilde{u}}{\partial t} + \mathcal{O}(t^2) \tag{4.9}$$

where $\tilde{u}(0)$ is the solution at $t = 0$. By a Lax-Wendroff procedure, a temporal derivative is replaced by a space derivative term, $\frac{\partial \tilde{u}}{\partial t} = -\tilde{u}(t) \cdot \frac{\partial \tilde{u}}{\partial x}$. This substitution returns the first-order time accurate $\tilde{u}(t)$ in terms of the solution at $\tilde{u}(0)$.

$$\tilde{u}(t) = \tilde{u}(0) - t\tilde{u}(t) \cdot \frac{\partial \tilde{u}}{\partial x} \tag{4.10}$$

$$\Rightarrow \tilde{u}(t) = \frac{\tilde{u}(0)}{1 + t\frac{\partial \tilde{u}}{\partial x}} \tag{4.11}$$

$$= \tilde{u}(0) \cdot \left(1 - t\frac{\partial \tilde{u}}{\partial x}\right) + \mathcal{O}(t^2) \tag{4.12}$$

With Eq. (4.12), a first-order accurate characteristic speed can be evaluated at any x location

of one's choice only with a knowledge of the solution and its first derivative.

Now, we shift our focus back to the numerical implementation aspect. At the time of interpolation, the solution reconstruction at t^n is known.

$$u^n(\xi) = \sum_{i=1}^{3} c_i \phi_i(\xi) \tag{4.13}$$

where $\xi \in [0, 1]$ is the local coordinate and ϕ_i represent the basis functions. Solution reconstruction coefficients c_i are defined in Table 3.2. A full-time solution at ξ_I is obtained by the characteristic tracing method.

$$u^{n+1}(\xi_I) = u^n(\xi_F) \tag{4.14}$$

The new characteristic origin ξ_F for nonlinear problem is defined as follows.

$$\xi_F = \xi_I - \tilde{u}^n(\Delta t; \xi_I)\frac{\Delta t}{\Delta x} \tag{4.15}$$

where $\tilde{u}^n(\Delta t; \xi_I)$ is the first-order accurate nonlinear characteristic speed evaluated at $\xi = \xi_I$. Note that the characteristic origin in nonlinear advection is very similar to the linear advection characteristic origin shown in Eq. (3.25). We can rewrite the nonlinear characteristic origin, Eq. (4.15), as

$$\xi_F = \xi_I - \tilde{u}^n(\Delta t; \xi_I)\frac{\Delta t}{\Delta x} = \xi_I - \tilde{\nu} \tag{4.16}$$

where $\tilde{\nu}$ is the Courant number in nonlinear advection. This shows a simple yet powerful extension of the one-dimensional linear advection scheme. In fact, the introduction of the first-order estimate of the local wave speed is the only device needed to give the AF method a third-order accuracy in any problem encountered to date.

The conservative flux requires three interface fluxes at equally spaced time steps, so a half-time step solution is acquired.

$$u^{n+1/2}(\xi_I) = u^n(\xi_H) \tag{4.17}$$

$$\xi_H = \xi_I - \tilde{u}^n\left(\frac{\Delta t}{2}; \xi_I\right)\frac{\Delta t}{2\Delta x} \tag{4.18}$$

A characteristic origin at half-time step, ξ_H, may have a different characteristic speed since characteristics can vary. It is easy to see that characteristic speeds and origins at half- and full-time steps will have different values in our approximation because the first-order

terms vary in Eq. (4.12).

4.2.2 Conservation

The conservative form of Burgers' equation Eq. (4.3) is

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0 \tag{4.19}$$

where $f(u) = \frac{1}{2}u^2$. The integral form of Eq. (4.19) is considered for the conservative discretization.

$$\frac{d}{dt}\int_{\Omega} u dV + \oint_{\partial\Omega} f \cdot n dS = 0 \tag{4.20}$$

And the fully discrete finite volume formulation is shown as follows.

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \frac{\Delta t}{\Delta x_j}\left(\bar{f}_{j+1/2} - \bar{f}_{j-1/2}\right) \tag{4.21}$$

where $\bar{f}_{j-1/2}$ and $\bar{f}_{j+1/2}$ are left and right interface flux values.

The conservation step requires consistent interface fluxes. As it was established in Section 3.2.3, the numerical flux for the inviscid Burgers' equation, or for nonlinear problems in general, is evaluated from averaging interface fluxes at three equally spaced time steps using the Simpson's rule. The general expression for the interface flux formula Eq. (3.34) is reiterated for nonlinear fluxes.

$$\bar{f}_{j+1/2} = \int_{t^n}^{t^{n+1}} f(t)dt \equiv \int_{\xi_F}^{\xi_I} f(u^n(\xi))d\xi \tag{4.22}$$

$$\approx \frac{1}{6}\left(f(u^n(\xi_I)) + 4f(u^n(\xi_H)) + f(u^n(\xi_F))\right) + \mathcal{O}(\Delta\xi^4) \tag{4.23}$$

4.2.3 Shocks and Rarefaction Waves

It is a property of nonlinear hyperbolic problems that even initially smooth solutions may develop discontinuities, or shocks, at a later time. Shocks are formed when characteristics converge into each other. The conservation stage automatically takes care of the formation of shock.

Another nonlinear feature can develop in nonlinear problems. When characteristics diverge from each other, rarefaction waves are formed. Characteristic based advection schemes handle rarefaction cases with ease, unlike certain Godunov-type schemes where sonic point treatment is necessary [70].

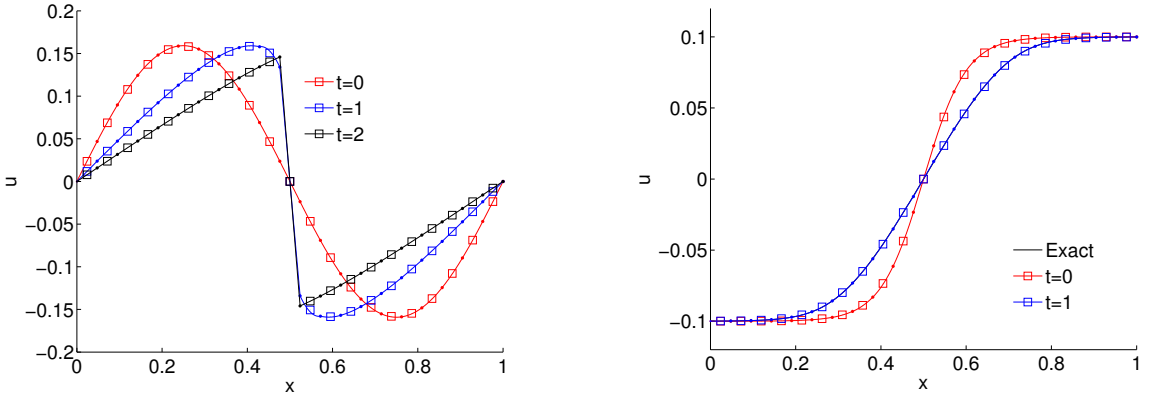
We look at two examples of nonlinearities found in Burgers' equation. The first example is a shock formation problem.

$$u_0(x) = \frac{1}{2\pi t_s} \sin(2\pi x) \tag{4.24}$$

where $t_s = 1$ is the shock formation time. And the rarefaction wave problem is described by the following initial condition.

$$u_0(x) = 0.1 \tanh(10(x - 0.5)) \tag{4.25}$$

The results are shown in Fig. 4.3. In the case of a shock, characteristics collide at $x = 0.5$ and form into a shock. The shock cell is initialized to be stationary $u = 0$. In the rarefaction wave, characteristics diverge away from the sonic cell.



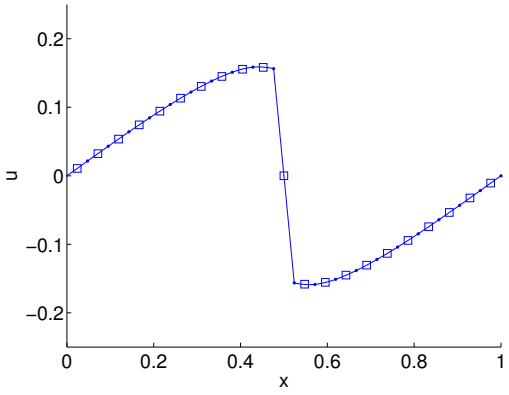
(a) Shock formation

(b) Expansion wave

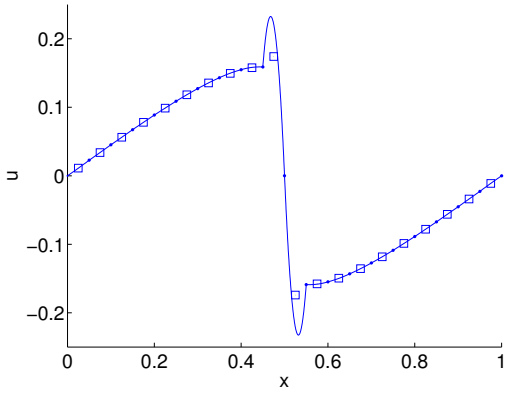
Figure 4.3: Two nonlinear phenomena in Burgers' equation with $N = 21$ and $\nu = 0.7$. Square symbols indicate cell averages and dots indicate interface values. They are accompanied by internal reconstructions shown in solid lines.

Solving Burgers' equation using the characteristic tracing method may present some isolated issues with certain initial conditions for shocks as shown in Fig. 4.4. In one case, Fig. 4.4a, a proper solution initialization leads to a normal shock capturing behavior. In the other case, Fig. 4.4b, the solution is initialized with one of its interface values at the stationary point, and it shows a sign of the shock accumulation issue. If a shock is initialized so that one of its interface values coincide with the stationary point of the solution, then the solution/mass accumulates at the shock front without any means to dissipate away. This occurs in the AF method for scalar conservation laws because a stationary interface is not updated in the characteristic tracing method due to its zero advection speed. This may concern some readers. However, we believe this will have lessened effect, or be eliminated,

when governing equations have other sources of physics such as diffusion. Furthermore, any higher than first-order accurate methods need to be properly limited. Results presented here are unlimited, so some oscillations are expected.



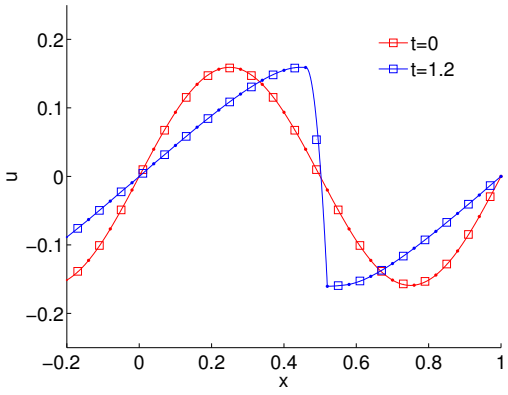
(a) Initialized with $N = 21$ elements. The stationary shock point is at the center of the cell in the middle. The shock is well captured and behaved.



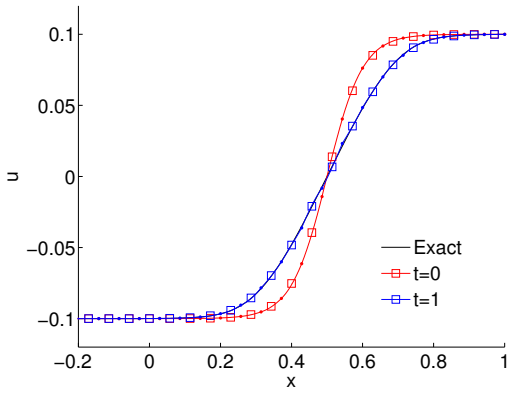
(b) Initialized with $N = 20$ elements. The stationary shock point coincides with one interface value of the cell. The shock accumulates at the interface.

Figure 4.4: Stationary shock with two different initializations at $t = 1.2$ and $\nu = 0.7$. Square symbols indicate cell averages and dots indicate interface values. They are accompanied by internal reconstructions shown in solid lines.

To demonstrate that the above-mentioned shock accumulation issue is a special case based on initialization, we show slightly off-centered initial conditions so that both nonlinear phenomena are no longer symmetrically initialized, in Fig. 4.5. The shock is captured without any numerical oscillations in the solution, or in the internal reconstruction shown by solid lines. The exact solution for the shock problem isn't available. But the shock stays stationary according to the pre- and post-shock states. The expansion wave is well described. The sonic point shows no deviation from the exact solution at $t = 1$.



(a) Asymmetrically initialized shock wave formation with $N = 21$



(b) Asymmetrically initialized expansion wave with $N = 21$

Figure 4.5: Asymmetrically initialized shock and expansion wave problems with $\nu = 0.7$. Square symbols indicate cell averages and dots indicate interface values. They are accompanied by internal reconstructions shown in solid lines.

4.2.4 Numerical Accuracy

A sine wave initial condition is considered for the numerical accuracy test.

$$u_0(x) = \frac{1}{2\pi t_s} \sin(2\pi x) \tag{4.26}$$

where $t_s = 1$ is the shock formation time. The exact solution can be iteratively evaluated for times before the shock formation. We take errors in cell averages at $t = 0.9$.

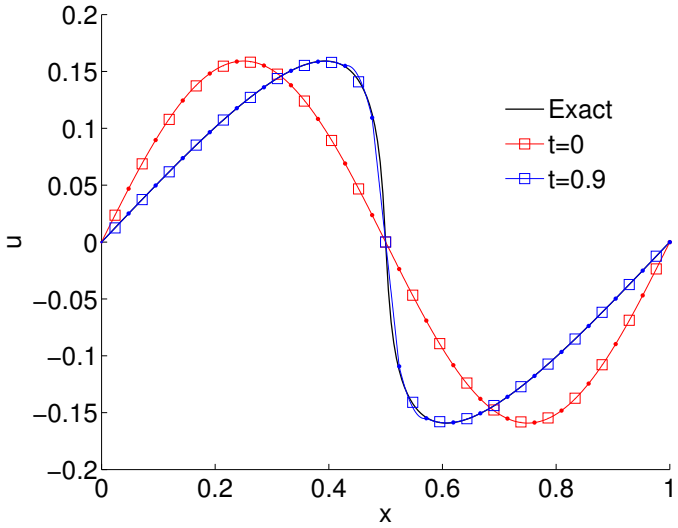


Figure 4.6: Burgers' equation solution of Eq. (4.26) at $t = 0.9$, $\nu = 0.7$ and $N = 20$. Square symbols indicate cell averages and dots indicate interface values. They are accompanied by internal reconstructions shown in solid lines.

Errors in conservative variables are measured by the $L_2(\overline{u})$ norm definition. It is formally used for error measurements in finite volume methods.

$$L_2(\overline{u}) = \left\{ \frac{1}{\Omega} \sum_{j=1}^{N} \left| \overline{u}_j \Omega_j - \overline{u}_j^{\text{exact}} \Omega_j \right|^2 \right\}^{1/2} \tag{4.27}$$

Table 4.1 summarizes the error in conservative variables for the sine wave problem at $t = 0.9$. The AF scheme for Burgers' equation converges at third-order but not until the mesh is fine enough to capture the steep solution near the center.

Table 4.1: Burgers' equation solution of Eq. (4.26) $L_2(\overline{u})$ norm error convergence at $t = 0.9$ and $\nu = 0.7$

Level	DOF	$L_2(\overline{u})$	Order
1	20	1.63e-04	
2	40	1.30e-04	0.33
3	80	5.46e-05	1.25
4	160	1.12e-05	2.28
5	320	1.73e-06	2.70
6	640	2.51e-07	2.78
7	1280	3.32e-08	2.92
8	2560	4.23e-09	2.98

4.3 Inviscid Burgers' Equation in Two Dimensions

The two-dimensional inviscid Burgers' equation is considered. The nonlinear characteristic tracing method for two-dimensional nonlinear advection is established in the evolution stage. The conservation stage is briefly discussed to highlight the procedure of AF scheme for two-dimensional scalar conservation laws. Although a two-dimensional case is discussed exclusively, the extension to three-dimensional problems is straightforward. The two-dimensional reconstruction remains the same as the procedure discussed in linear advection, refer to Section 3.3 for details.

4.3.1 Evolution

A nonconservative form of a two-dimensional generalization of Burgers' equation is considered.

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + u\frac{\partial u}{\partial y} = 0 \tag{4.28}$$

The scalar version of Burgers' equation has both of its advection speed components defined by its solution u. In Chapter 6, a genuinely multidimensional Burgers' equation system is presented in the context of nonlinear conservation systems.

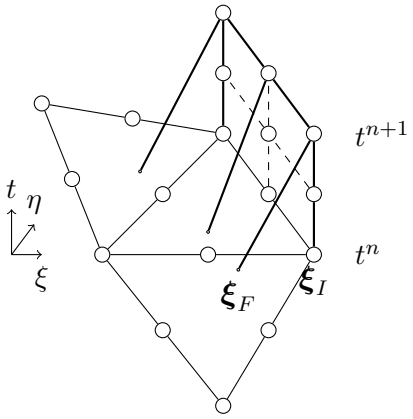


Figure 4.7: Description of nonlinear characteristic tracing in 2D. Not all characteristics have the same speed.

As illustrated in Fig. 4.7, the nonlinear variation in the advection speed, characteristic speed, from one node to another gives non-parallel characteristics in the space-time coordinate. Unlike linear advection in two dimensions, the advection speed for each node may be different and this suggests that the characteristic interpolation process also reflects this property.

As previously discussed, second-order accurate interface fluxes are constructed from at least first-order accurate interface variables. Therefore, a nonlinear wave speed correction is required. The nonlinear wave speed is approximated from a Taylor series approximation, similar to the one-dimensional procedure. We introduce a solution at some location $u(\mathbf{x}, t) = \tilde{u}(t)$

$$\tilde{u}(t) = \tilde{u}(0) + t\frac{\partial \tilde{u}}{\partial t} + \mathcal{O}(t^2) \tag{4.29}$$

where $\tilde{u}(0)$ is the solution at $t = 0$. A Lax-Wendroff procedure is applied to replace the temporal derivative with spatial derivatives.

$$\tilde{u}(t) = \tilde{u}(0) - t\tilde{u}(t) \cdot \left(\frac{\partial \tilde{u}}{\partial x} + \frac{\partial \tilde{u}}{\partial y}\right) \tag{4.30}$$

$$\Rightarrow \tilde{u}(t) = \frac{\tilde{u}(0)}{1 + t\left(\frac{\partial \tilde{u}}{\partial x} + \frac{\partial \tilde{u}}{\partial y}\right)} \tag{4.31}$$

$$= \tilde{u}(0) \cdot \left(1 - t\left(\frac{\partial \tilde{u}}{\partial x} + \frac{\partial \tilde{u}}{\partial y}\right)\right) + \mathcal{O}(t^2) \tag{4.32}$$

This approximation evidently illustrates the multidimensional nature of the scheme. Because we consider the velocity gradients in both directions at the same time, it adheres to the design principles of multidimensionality closely. More importantly, the first-order nonlinear wave speed is obtained only with a knowledge of the solution and its first derivatives. The extension to three dimensions can be made easily.

A two-dimensional solution reconstruction is recalled.

$$u^n(\boldsymbol{\xi}) = \sum_{i=1}^{7} c_i \phi_i(\boldsymbol{\xi}) \tag{4.33}$$

where coefficients c_i and the basis functions ϕ_i are found in Table 3.5 for triangles and Table 3.6 for quadrilaterals. A full-time solution at interface $\boldsymbol{\xi}_I$ is evaluated at a characteristic origin $\boldsymbol{\xi}_F$.

$$u^{n+1}(\boldsymbol{\xi}_I) = u^n(\boldsymbol{\xi}_F) \tag{4.34}$$

$$\boldsymbol{\xi}_F = \boldsymbol{\xi}_I - \Delta t \mathbf{J}^{-1} \left[\tilde{u}^n(\Delta t; \boldsymbol{\xi}_I)\hat{\mathbf{e}}\right] \tag{4.35}$$

where $\mathbf{J}$ is a mapping Jacobian of the element and $\hat{\mathbf{e}} = (1_x, 1_y)^T$ is a unit direction vector.

And a half-time step solution is found trivially.

$$u^{n+1/2}(\boldsymbol{\xi}_I) = u^n(\boldsymbol{\xi}_H) \tag{4.36}$$

$$\boldsymbol{\xi}_H = \boldsymbol{\xi}_I - \frac{\Delta t}{2}\mathbf{J}^{-1}\left[\tilde{u}^n\left(\frac{\Delta t}{2};\boldsymbol{\xi}_I\right)\hat{\mathbf{e}}\right] \tag{4.37}$$

Note that the first-order corrected characteristic speed $\tilde{u}^n(t)$ directly depends on time, and takes different values for Δt and $\frac{\Delta t}{2}$, see Eq. (4.32).

4.3.2 Conservation

The conservative form of the inviscid two-dimensional Burgers' equation Eq. (4.28) is shown as follows.

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} + \frac{\partial g(u)}{\partial y} = 0 \tag{4.38}$$

where the flux vector is

$$\mathbf{f}(u) = \begin{pmatrix} f(u) \\ g(u) \end{pmatrix} = \frac{1}{2}u^2. \tag{4.39}$$

The integral form of Burgers' equation is shown as follows.

$$\frac{d}{dt}\int_\Omega u dV + \oint_{\partial\Omega} \mathbf{f}\cdot\mathbf{n} dS = 0 \tag{4.40}$$

A fully discrete form of the conservation law is presented.

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \frac{\Delta t}{\Omega_j}\sum_{l=1}^{\text{sides}} \bar{\mathbf{f}}_l \cdot \mathbf{n}_l \tag{4.41}$$

where $\mathbf{n}_l$ is a scaled normal vector of flux interface l. The flux vector on each interface $\bar{\mathbf{f}}_l$ is evaluated numerically according to the composite Simpson's formula described in Eq. (3.73).

4.3.3 Numerical Accuracy

Errors in conservative variables are measured by the $L_2(\overline{u})$ norm definition. It is formally used for error measurements in finite volume methods.

$$L_2(\overline{u}) = \left\{\frac{1}{\Omega}\sum_{j=1}^{N}\left|\overline{u}_j\Omega_j - \overline{u}_j^{\text{exact}}\Omega_j\right|^2\right\}^{1/2} \tag{4.42}$$

Errors in solution reconstruction variables are measured by the $L_2(q(\mathbf{x}))$ norm definition. This error norm measure provides an additional perspective in quantifying the solution quality such as the variation of solutions in reconstructions.

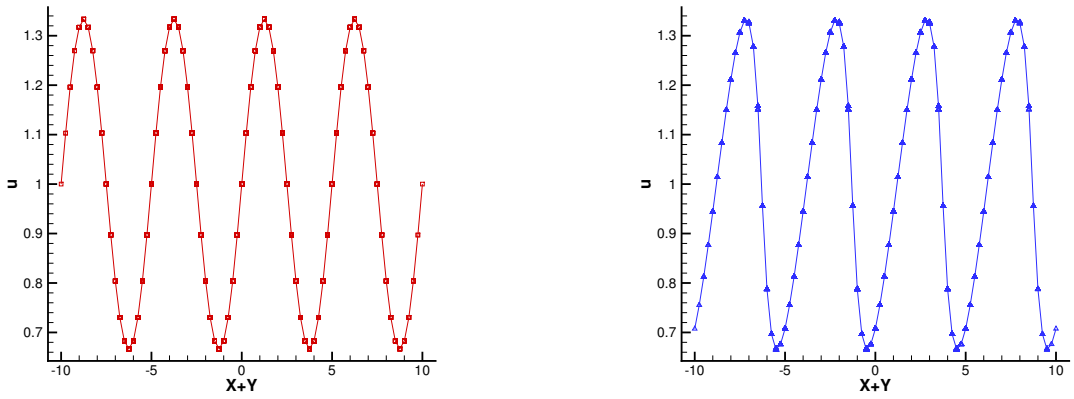
$$L_2(q(\mathbf{x})) = \left\{ \frac{1}{\Omega} \sum_{j=1}^{N} \int_{\Omega_j} \left| q_j(\mathbf{x}) - q_j^{\text{exact}}(\mathbf{x}) \right|^2 dV \right\}^{1/2} \tag{4.43}$$

A sine wave initial condition is considered for the numerical test.

$$u_0(x, y) = 1 + \frac{1}{3} \sin\left(\frac{2\pi}{5} (x + y) \right) \tag{4.44}$$

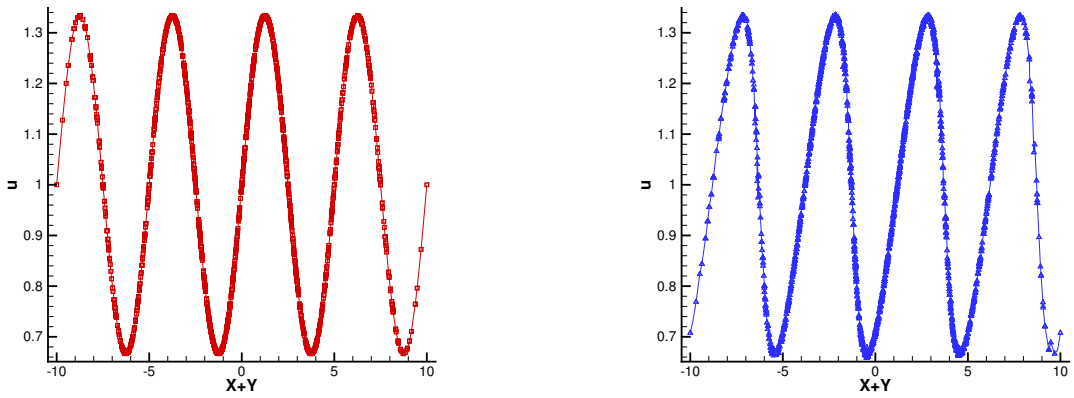
Periodic boundary conditions are imposed on both directions. The initial condition is advanced to $t = 0.6$ before shocks develop. We verify the numerical accuracy using all grid configurations in Fig. 3.20.

Figs. 4.8 and 4.9 show initial and final time solutions for two grid configurations. The wave steepening effect can be seen at the final time. Cell average solutions on unstructured grids in Fig. 4.9 are spread out randomly due to the unstructured nature of the grid.



(a) $t = 0$ (b) $t = 0.6$

Figure 4.8: Sine wave problem, Eq. (4.44), on triangular grid B on $x + y$ coordinate at $\nu = 0.7$. Symbols indicate cell averages. Internal reconstructions are shown in solid lines.

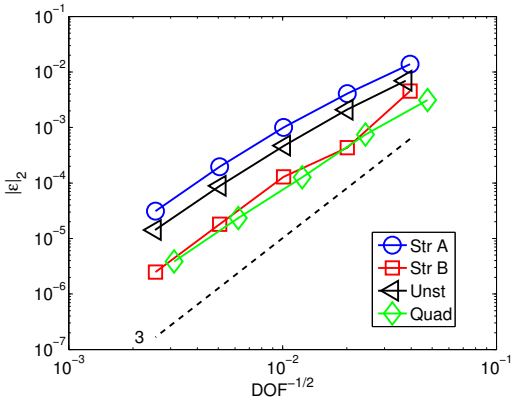


(a) $t = 0$ (b) $t = 0.6$

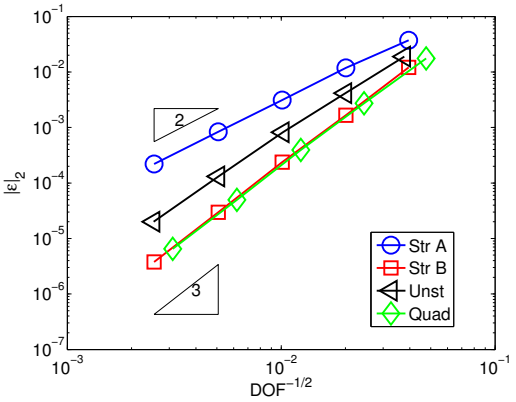
Figure 4.9: Sine wave problem, Eq. (4.44), on triangular unstructured grid on $x + y$ coordinate at $\nu = 0.7$. Symbols indicate cell averages. Internal reconstructions are shown in solid lines.

Results in Fig. 4.10 show the L_2 norm errors in conservative variables and primitive solution reconstruction variables for the sine wave problem. Errors in conservative variables converge approximately at $\mathcal{O}(h^3)$ with an exception of the triangular structured grid A, which converges at $\mathcal{O}(h^{2.6})$ see Fig. 4.10a. Errors in solution reconstruction variables indicate slight signs of mesh dependency for triangular structured grid A ($\mathcal{O}(h^2)$) and triangular unstructured grid ($\mathcal{O}(h^{2.7})$) cases, see Fig. 4.10b.

Convergence rates of the triangular structured grid A (Str A) vary depending on what error norm is chosen. The $L_2(\overline{u})$ norm, in general, fails to show the error incurred due to mesh alignment because conservative variables represent averaged quantities. On the other hand, the $L_2(u(\mathbf{x}))$ norm definitely captures the mesh alignment problem because it measures an error within the solution reconstruction.



(a) $L_2(\overline{u})$ norm error of conservative variable. Convergence rates of Str A: $\mathcal{O}(h^{2.6})$. The rest are $\mathcal{O}(h^3)$



(b) $L_2(u(\mathbf{x}))$ norm error of reconstruction. Convergence rates of Str A: $\mathcal{O}(h^2)$, Unst: $\mathcal{O}(h^{2.7})$. The rest are $\mathcal{O}(h^3)$

Figure 4.10: Sine wave problem, Eq. (4.44), error for various grids at $t = 10$ and $\nu = 0.7$. Triangular structured grid A (Str A), B (Str B), unstructured (Unst) and quadrilateral (Quad)

The mesh alignment problem was discussed in Sections 3.3.7 and 3.4 for linear advection, but the general idea applies to nonlinear advection as well. In nonlinear advection, the characteristics may be non-parallel due to non-constant wave speeds. Regardless of the solution initialization, characteristics may align with the mesh geometry and the mesh alignment problem can appear in time.

4.3.4 Numerical Comparison

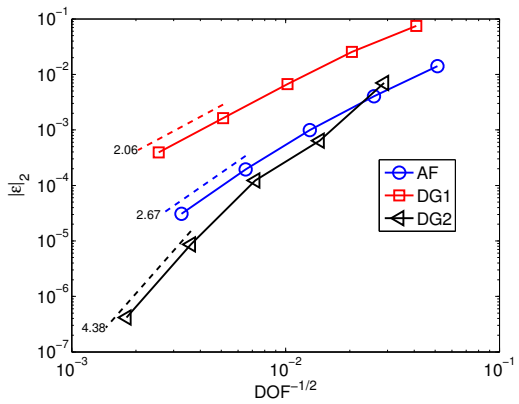
AF error results from the sine wave initial condition, Eq. (4.44), are compared to DG results. DG schemes considered in this comparison are the Runge-Kutta DG scheme with a linear ($p = 1$) and a quadratic ($p = 2$) spatial discretization orders, respectively referred to as DG1 and DG2. Two- and three-stage RK methods are used for DG1 and DG2, respectively. The computational grids for the DG results are based on the same grids shown in Fig. 3.20. For numerical comparisons made here, only the triangular structured grid A, B, and unstructured grid types are considered. Furthermore, the DG CFL stability is restricted by

$$\nu < \frac{1}{2p+1} \tag{4.45}$$

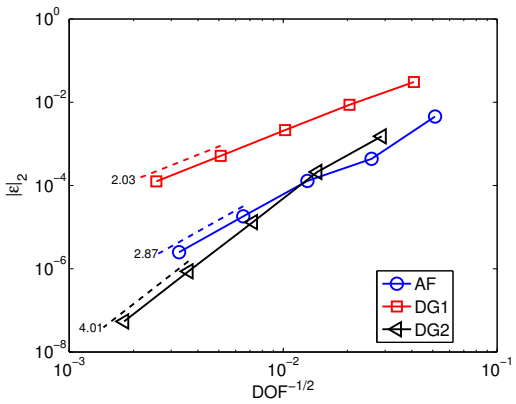
where p is the spatial discretization order [14].

We first compare the error in conservative variables using $L_2(\overline{u})$ norm in Fig. 4.11. The DG2 convergence rates are better than $\mathcal{O}(h^{p+1})$ for both grid types. DG schemes for the sine wave initial condition problem, Eq. (4.44), benefit from the regularity of the triangular structured grid B. The DG2 convergence rate is about one order less for the

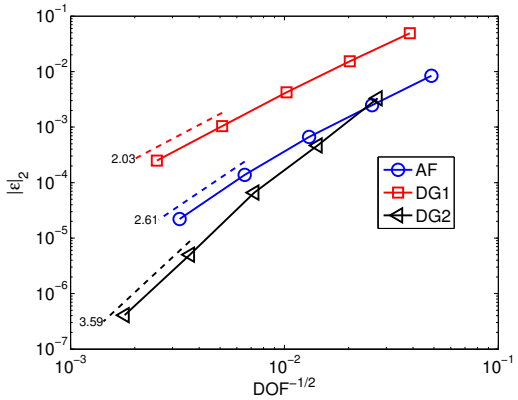
unstructured grid type. On the other hand, the convergence rates of AF do not exhibit a strong dependency on the grid type.



(a) Triangular structured grid A



(b) Triangular structured grid B

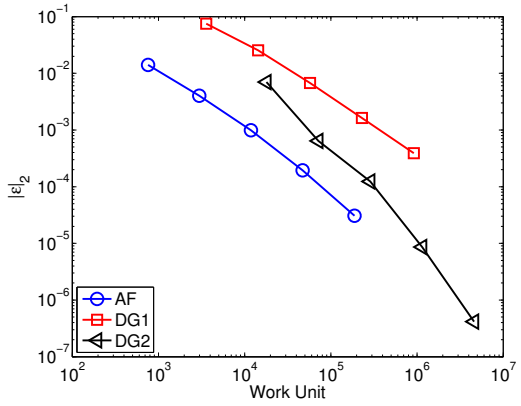


(c) Triangular unstructured grid

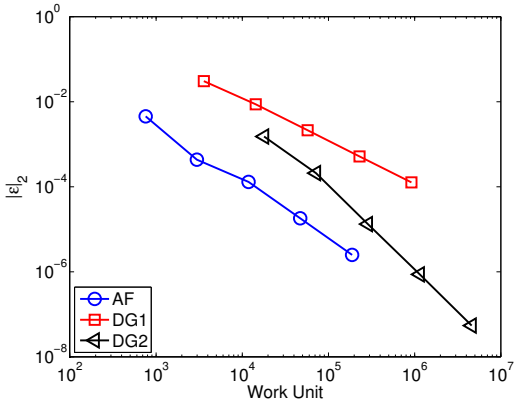
Figure 4.11: Sine wave problem, Eq. (4.44), conservative variable $L_2(\overline{u})$ error convergence results at $t = 0.6$. Courant numbers for AF, DG1, and DG2 are respectively $\nu = 0.7$, $\nu = 0.3$, and $\nu = 0.2$.

We have previously devised a common measure called the work unit in Eq. (3.86). The parameters for evaluating the WU are listed in Table 3.7 for both schemes. In Section 3.3.6, we have established that the WU, though not perfect, provides a *good* measure of efficiency with regard to the theoretical construction of each scheme.

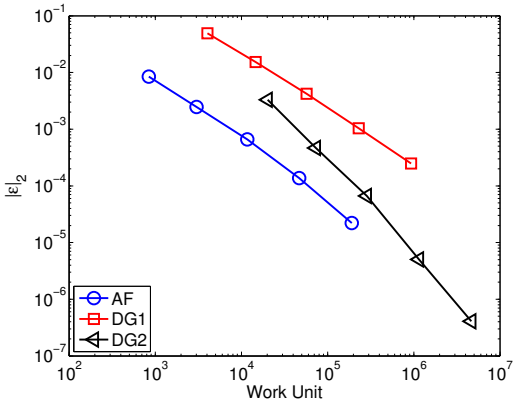
The $L_2(\overline{u})$ norm errors of conservative variables for all schemes are presented in Fig. 4.12. The results indicate that AF starts to show a theoretical advantage over all DG schemes. The WU for each scheme are listed in Tables 4.2, 4.3, and 4.4. To reach the same level of error, AF would take about one order of magnitude fewer WU than the DG2 counterpart. Indeed, AF starts seeing some lead over DG, compared to the linear advection case in Fig. 3.24.



(a) Triangular structured grid A



(b) Triangular structured grid B



(c) Triangular unstructured grid

Figure 4.12: Sine wave problem, Eq. (4.44), conservative variable $L_2(\overline{u})$ error convergence results measured as a function of WU at $t = 0.6$. Courant numbers for AF, DG1, and DG2 are respectively $\nu = 0.7$, $\nu = 0.3$, and $\nu = 0.2$.

Table 4.2: Sine wave problem, Eq. (4.44), $L_2(\overline{u})$ norm errors of conservative variables and WU for AF on unstructured grid at $t = 0.6$ and $\nu = 0.7$

DOF	Work Unit	$L_2(\overline{u})$	Order
4.21e+02	8.42e+02	8.42e-03	
1.51e+03	3.03e+03	2.48e-03	1.91
5.86e+03	1.17e+04	6.62e-04	1.95
2.34e+04	4.69e+04	1.37e-04	2.27
9.52e+04	1.90e+05	2.21e-05	2.61

Table 4.3: Sine wave problem, Eq. (4.44), $L_2(\overline{u})$ norm errors of conservative variables and WU for DG1 on unstructured grid at $t = 0.6$ and $\nu = 0.3$

DOF	Work Unit	$L_2(\overline{u})$	Order
6.72e+02	4.03e+03	4.91e-02	
2.44e+03	1.47e+04	1.54e-02	1.80
9.53e+03	5.72e+04	4.22e-03	1.90
3.82e+04	2.29e+05	1.04e-03	2.02
1.56e+05	9.33e+05	2.49e-04	2.03

Table 4.4: Sine wave problem, Eq. (4.44), $L_2(\overline{u})$ norm errors of conservative variables and WU for DG2 on unstructured grid at $t = 0.6$ and $\nu = 0.2$

DOF	Work Unit	$L_2(\overline{u})$	Order
1.34e+03	2.02e+04	3.30e-03	
4.88e+03	7.33e+04	4.67e-04	3.03
1.91e+04	2.86e+05	6.61e-05	2.87
7.64e+04	1.15e+06	5.04e-06	3.71
3.11e+05	4.67e+06	4.07e-07	3.59

CHAPTER 5

Limiting

When the solution contains large gradients, for higher than first-order numerical methods, non-physical oscillations may develop in the numerical solution. These oscillations, Gibbs phenomena, not only lead to aesthetically displeasing numerical solutions but also can be catastrophic when solving nonlinear systems. Nonlinear mechanisms, often called slope or flux limiters, are widely accepted procedures for higher than first-order numerical methods to curb spurious oscillations occurring in numerical solutions.

In the AF method, advective and acoustic problems are separately treated because they present different issues. We present distinct limiting principles for these problems. Advective problems have a very narrow domain of dependence and observe simple bounds. Acoustic problems have a large domain of dependence and obey no bounding principle. These observations are essential details which will be used for designing an accurate limiting procedure for each problem. And the limiting guidelines set the precedence for the future limiting directions, which do not depend on the old antiquated or ill-suited principles. The design principles of the AF limiter, as well as some examples in one and two dimensions, are presented.

5.1 Active Flux Method Limiting Basics

Godunov's theorem proved that any linear monotone schemes for solving hyperbolic partial differential equations can be at most first-order accurate [26]. This theorem does not, however, apply to nonlinear schemes and limiters are designed with this in mind. We lay out some requirements for AF limiters.

- Reduce spurious oscillations in solution
- Conservative
- Preserve high-order accuracy if possible

- Compact
- Intuitive

First and foremost, a limiter should either completely remove or attempt to reduce spurious oscillations. It is the main purpose of introducing a limiter to high-order numerical methods. The second point is to maintain the conservation of scheme by neither destroying nor introducing additional mass. It is particularly important for compressible flow numerical methods where conservation is imperative.

Early developments of limiters were based on the observation that oscillatory solutions are most often indicated by changing signs of solution gradients, as famously investigated by van Leer [83], and have employed to limit solutions based on some simple measure of these gradients. Many similar approaches soon became known as total variation diminishing (TVD) schemes where a high-order linear scheme is combined with some nonlinear mechanism involving ratios of solution spatial gradients [81]. TVD schemes have limitations for smooth solutions containing genuine extrema, and the clipping phenomena may occur, see Fig. 5.1. The third point of limiter design is to preserve a high-order accuracy by means other than TVD. For numerical schemes higher than second-order, it is evident that the TVD property would be a hindrance rather than helpful since essential high-order accurate solution features would be removed.

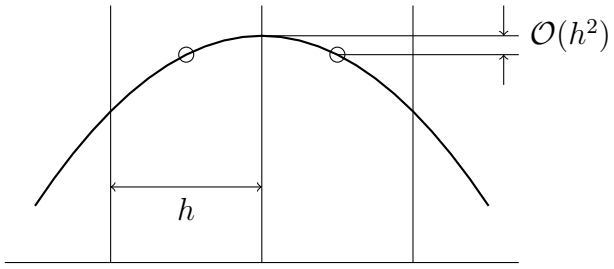


Figure 5.1: TVD schemes introduce an $\mathcal{O}(h^2)$ clipping of true extremum in higher than first-order solutions

The compactness property of limiter becomes necessary in truly multidimensional numerical schemes. Some high-order numerical methods rely heavily on an extended numerical stencil to determine the smoothest available solution, for example, ENO and WENO methods. This not only goes against on the principles of local nature of hyperbolic problems but also seems rather artificial. In more than one dimension, especially with arbitrarily oriented meshes, it presents additional computational expenses and efforts to overcome the complex geometrical challenges involving interpolations [31]. Therefore, a compact numerical footprint in limiting should benefit a multidimensional numerical scheme.

It is our belief that the current take on limiting emphasizes heavily on rather unnecessary and unnatural notion of discontinuity. The emphasis of discontinuity in limiting has perhaps been aggravated by the persistent belief in Riemann solvers from the CFD community. Regardless of the success of Riemann problems in numerical methods, it is important to shift our focus onto a new interpretation of limiting, one that is not based on addressing the discontinuous interface property. In the AF limiters, the conventional wisdom of limiting by revising the solution slope will be abandoned in lieu of more geometric interpretation. Chiefly, this shift is due to the design philosophy of AF which strives to move away from Riemann solvers. A correct conservative flux is a key to accomplishing a conservative numerical scheme, and it is the means by which one can define the limiting process. A geometric interpretation gives a lot of freedom in terms of deciding what kind of limiting strategy is best suited for a certain model problem.

Additionally, the limiting procedures must be intuitive with respect to the type of physics one tries to numerically replicate. The AF method considers physical origins of problems when designing numerical methods. This property is very relevant for designing limiters as well. The maximum principle is most commonly used in advection type problems where *a priori* bounds of the problem are known. However, acoustics problems are not subject to the same principle. Therefore, it is natural to consider different motivations for advection and acoustics limiters.

The AF method is a third-order scheme and odd-order accurate schemes are known to have good dissipation and dispersion relations [10]. Spurious oscillations generated from the AF advection scheme is already of much smaller amplitude and more limited extent than those of second-order schemes, and this makes it a good basis for limiting. The comparison of the third-order AF and second-order Lax-Wendroff schemes for linear advection of a square wave is shown in Fig. 5.2.

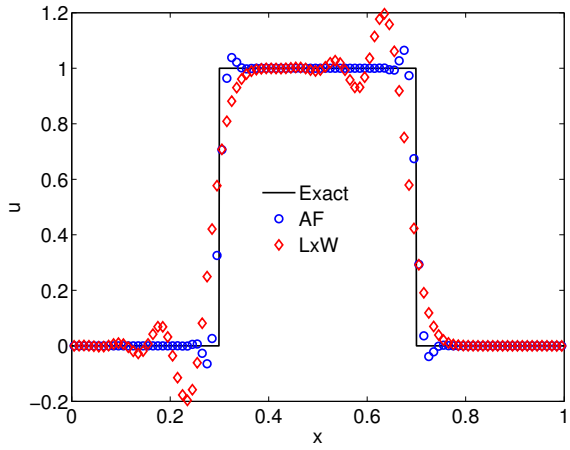


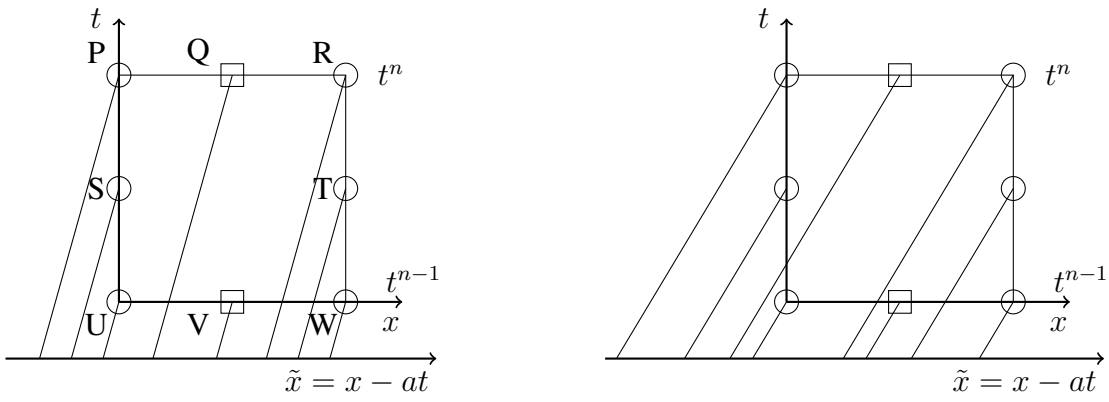
Figure 5.2: Third-order (AF) scheme compared to second-order (Lax-Wendroff) scheme for linear convection of square wave after one cycle with $N = 100$ and $\nu = 0.7$

5.2 One-Dimensional Limiter

Some one-dimensional limiter ideas are presented. The purpose of one-dimensional limiters is to closely understand the impact of new limiting design criteria laid out in the previous section, for AF and perhaps other high-order schemes. Also, having a clear understanding of the requirements in one-dimensional problems, a simpler numerical setting, will provide a better basis for multidimensional limiting efforts to come.

5.2.1 Limiting from the Past

One of the crucial requirements for a genuinely multidimensional limiter is that of compactness. It has traditionally been taken as the accepted practice to seek the validity of non-smooth data from extended spatial neighbors. With modest effort, the previous time data can be used to augment, if not replace, the data extracted from spatially extended neighbors. We refer to Fig. 5.3 to select information that may be of physical and numerical importance to the default data (PQR) at t^n. Characteristic lines carry relevant information and might be useful in the limiting procedure based on the past time history. In the characteristic coordinate, we can form relations between the default data (PQR) to the past time data.



(a) $\nu < 0.5$ (b) $\nu > 0.5$

Figure 5.3: One-dimensional characteristic coordinate, $\tilde{x}$, for linear problems. The order of information shifts from Courant numbers less than 0.5 to greater than that.

For hyperbolic partial differential equations, the notion of characteristics is a familiar one. Numerical solutions are designed, in some sense, to follow the physical axiom which is rooted in the mathematical theorem. Characteristics serve a dual purpose in that sense. Firstly, we have extensively used it throughout the development of numerical schemes for advection. Secondly, we shall now use it to serve the need for extra information in the limiting process for advection.

5.2.2 Maximum Principle Satisfying Limiter

The first investigation of limiting begins with the well-known maximum principle that exists in problems with known *a priori* bounds. The maximum principle states

$$m = \min_j u_j^n \leq u_j^{n+1} \leq \max_j u_j^n = M \tag{5.1}$$

where M and m stand for maximum and minimum bounds. A well-known application of this principle was proposed by Zhang and Shu [102, 103] who have proposed and shown maximum principle satisfying high-order schemes for scalar conservation laws in one and multidimensions.

Typically, the maximum principle is used to find the numerical solution bounds, a maximum M, and a minimum, m, from some extent of space. In our approach, we extract the bounds from a pool of data ranging from t^{n-1} to t^n, see Fig. 5.3. They must not be taken only from the past time cell averages because that would correspond to the TVD criterion that limits the scheme to second-order accuracy in smooth solution cases. They are estimated instead from the reconstruction of past time data. We replace the piecewise natural

quadratic reconstruction $u(\xi)$ with a limited reconstruction $\tilde{u}(\xi)$ that satisfies known local *a priori* bounds.

$$\tilde{u}(\xi) = \theta \left(u(\xi) - \bar{u}_j^n \right) + \bar{u}_j^n \tag{5.2}$$

where the scaling coefficient is

$$\theta = \min \left(\left| \frac{M - \bar{u}^n}{M_p - \bar{u}^n} \right|, \left| \frac{m - \bar{u}^n}{m_p - \bar{u}^n} \right|, 1 \right) \tag{5.3}$$

One set of bounds is determined from the available collection of numerical solution data $\{u_{j,\text{numerical}}\}$, which are discrete values.

$$m = \min(\{u_{j,\text{numerical}}\}), \quad M = \max(\{u_{j,\text{numerical}}\}) \tag{5.4}$$

Another set of bounds is determined from reconstructions of the numerical solution data. Because the reconstructions are quadratic polynomials, two set of bounds are not the same.

$$m_p = \min(u_j^n(\xi)), \quad M_p = \max(u_j^n(\xi)) \quad \xi \in [0, 1] \tag{5.5}$$

The bounds from reconstructions are sampled at Gaussian quadrature locations. But any other sample locations would be choices as long as the internal structure of a reconstruction can be represented.

So far, we have not yet discussed in what way the past time data is selected and utilized. The following set of guidelines is introduced for a possible way to utilize the past time data.

5.2.2.1 Genuine extremum

It is well known that quadratic polynomials permit internal extrema. This feature of quadratic polynomial necessitates the proper treatment of genuine extrema. Simply put, if we find a genuine extremum within the reconstruction, we would like to reflect that in the limiting procedure. In order to distinguish the genuine extrema, we use additional information from available data. In the framework of maximum principle satisfying limiters, the linear scaling coefficient, θ, is modified based on some high-order criteria.

We define a semi-empirical function based on high-order data, curvature, to indicate the validity of an extremum.

$$\psi = \text{minmod} \left(\text{minmod} \left[\frac{\bar{c}_j}{\bar{c}_n} \right], \text{minmod} \left[\frac{\bar{c}_n}{\bar{c}_j} \right] \right) \quad \forall j = 1, 3 \tag{5.6}$$

where $\bar{c}_n$ is the average curvature at t^n and $\bar{c}_j$ refers to three approximations of curvatures

from a combination of data available. Possible numerical stencil candidates from which curvatures are evaluated are suggested in Table 5.1 and inspired by Fig. 5.3. Note that the combination depends on the Courant number, ν. The genuine extremum indicator, ψ, would output a value smaller than or equal to 1 based on the comparison of the approximated curvatures. ψ near 1 indicates that the extremum is true while near 0 indicates the extremum is false. The final linear scaling coefficient becomes,

$$\Theta = \max(\theta, \psi) \tag{5.7}$$

Finally, the expression for limited reconstruction becomes the following.

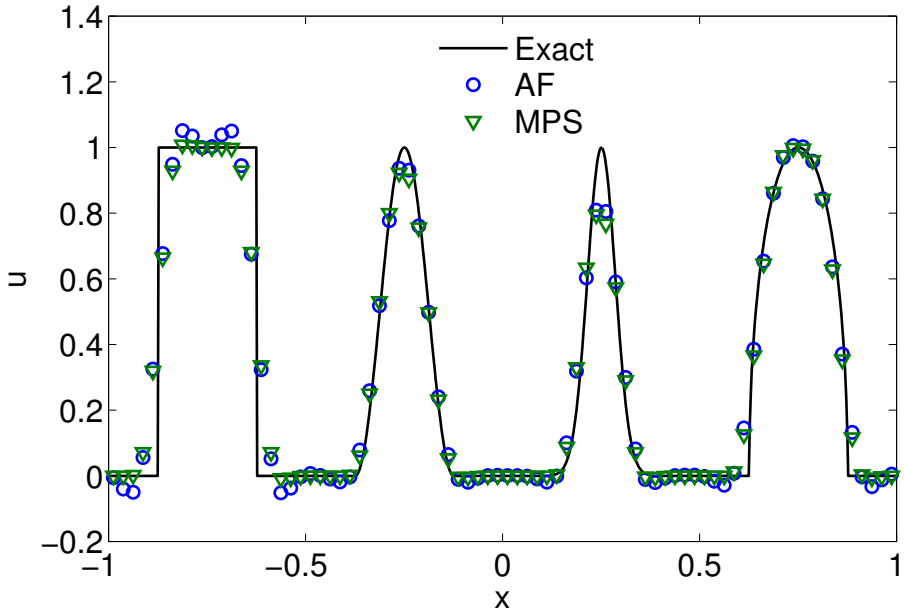
$$\tilde{u}(\xi) = \Theta(u(\xi) - \bar{u}^n) + \bar{u}^n \tag{5.8}$$

One untold rule of limiters is that they depend continuously on data. The maximum principle satisfying limiter certainly reproduces that rule. The linear advection results of this limiter are shown in Fig. 5.4.

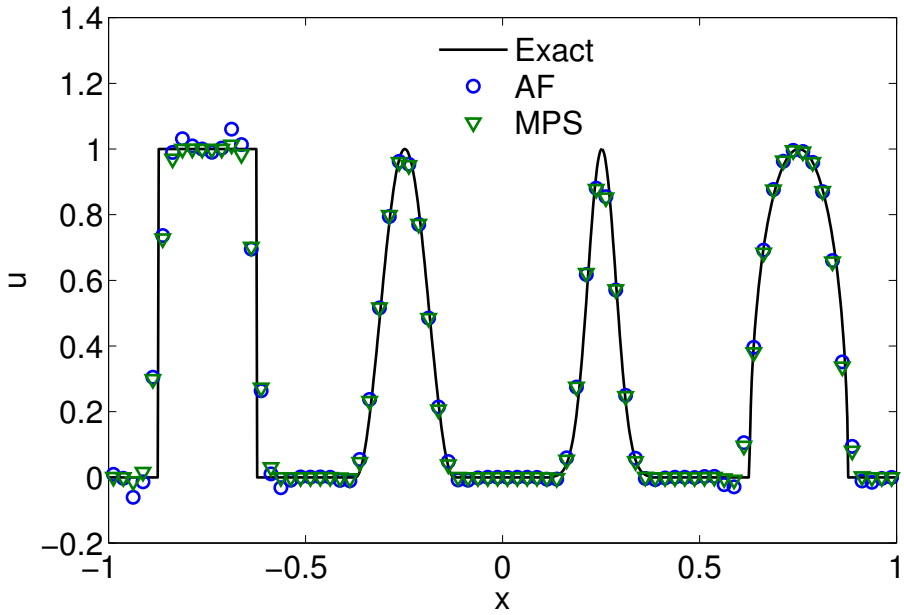
This limiter can retain the original shape of the reconstruction when the solution is smooth, as much as it can. In fact, the performance of the limiter depends on how one constructs the extremum indicator, which takes some careful empiricism. However, non-smooth solutions are damped according to the maximum principle, and spurious oscillations are minimized.

Table 5.1: Possible alternative stencil choices. See Fig. 5.3 for indices

Stencil	$\nu < 0.5$	$\nu > 0.5$
n, Default	PQR	PQR
$n-1$	UVW	UVW
Alternate 1	SUQ	SQU
Alternate 2	VRT	RVT



(a) $\nu = 0.4$



(b) $\nu = 0.8$

Figure 5.4: Zalesak test cases after one cycle with $N = 80$. Comparison between unlimited AF and limited by maximum principle (MPS). Notice the marked improvement of the limited solution, especially near sharp gradients. Smooth extrema solutions are well resolved.

5.2.3 Reconstruction Modifying Limiter

The reconstruction modifying limiter modifies the geometry of a reconstruction to control interface flux. The modifications to the geometry of a reconstruction are done without introducing any changes to the interface values or the cell average.

We introduce a parameter R which can be used to determine the presence of an internal extremum.

$$R = \frac{u_R - \bar{u}}{\bar{u} - u_L} \tag{5.9}$$

where u_L and u_R indicate left and right interface values. After some easy algebra, we observe that the default reconstruction (the parabola that exactly matches the given cell average and interface values) will have an internal extremum if the parameter lies outside the range $1/2 < R < 2$. Based on R, three regions are defined.

1. $1/2 < R < 2$: monotone region. No limiting is necessary.

2. $1/5 < R < 1/2$ or $2 < R < 5$: non-monotone region. But a monotone reconstruction *can* be constructed by "flattening" one half of the reconstruction.

3. All other R: non-monotone region. Some smoothing is possible by replacing the original reconstruction by two different halves of parabolas meeting at the center.

In the second region, the original reconstruction is non-monotone, but a monotone reconstruction *can* be constructed. The original reconstruction is replaced by a combination of a piecewise constant and a parabola in order to introduce a monotone reconstruction, see $R = 5$ case in Fig. 5.5. The conservation property provides an implicit constraint for constructing a modified reconstruction. As R varies, the internal point conjoining the piecewise constant and the parabola moves within the reconstruction in order to make sure the area under the modified reconstruction stays the same as before. Also in this region, the empirical criterion discussed in Section 5.2.2.1 is used to discourage unnecessary "flattening" of a genuine extremum.

In the third region, all limited reconstructions are non-monotone by default, but some smoothing can be introduced. A less extreme reconstruction is obtained by introducing two different parabolas. Because the original reconstruction is replaced by two different halves of parabolas, the internal extremum is reduced, see $R = 10, R = -10$ cases in Fig. 5.5. In the special circumstance, $R = -1$, the original reconstruction is accepted as genuine.

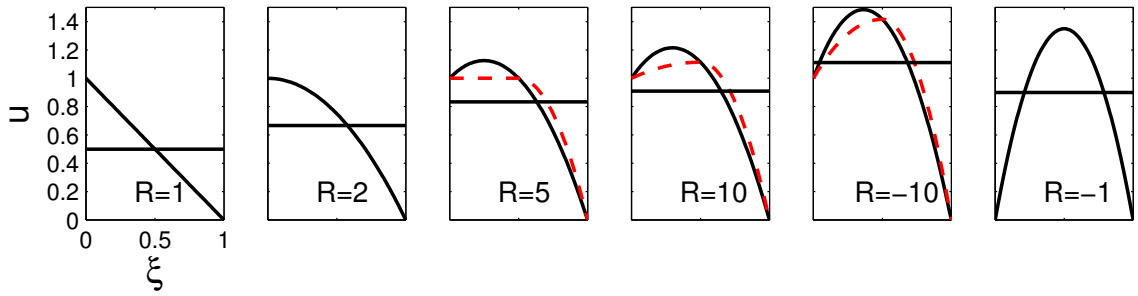
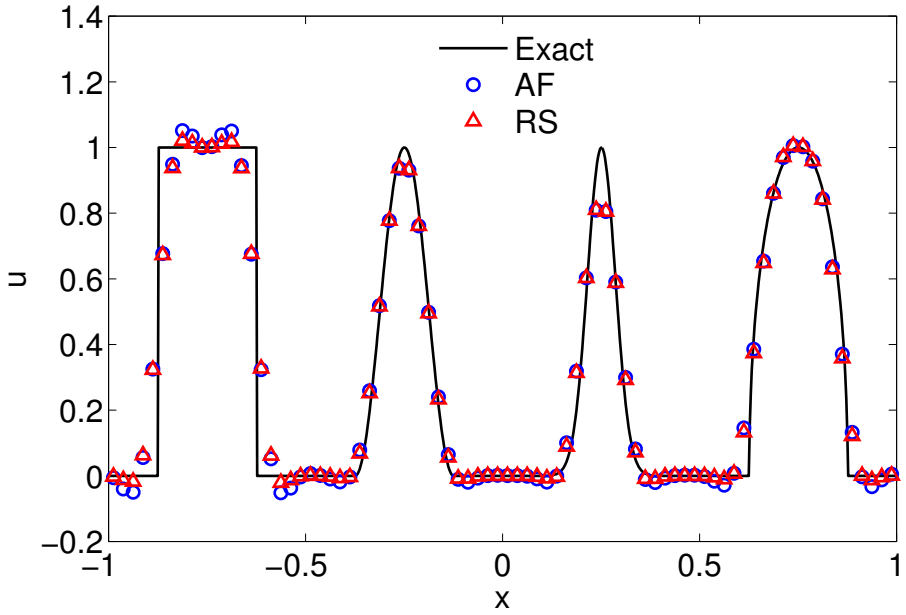
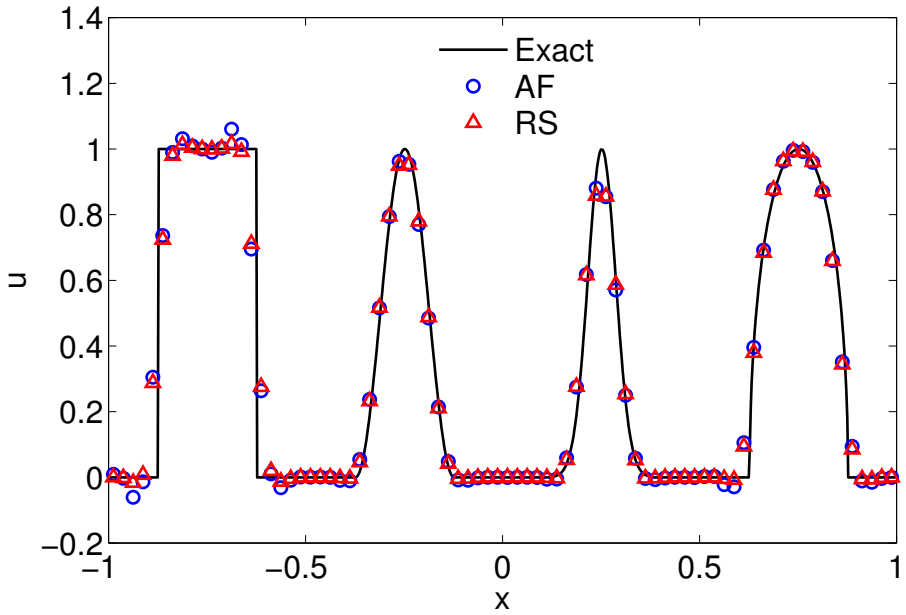


Figure 5.5: Modified/limited reconstructions for various R. Cell averages are indicated by piecewise constant lines. Limited reconstructions are red dashed lines. For reconstructions $R < 2$, no limiting is necessary. When $2 < R < 5$, the reconstruction is "flattened" to introduce a monotonic reconstruction. When $R > 5$, two different parabolas are introduced to reduce internal extremum. If $R = -1$, we accept the reconstruction.

The modified reconstruction preserves both the cell average and interface values. However, the difficulty of this limiting method is on the interface flux evaluation for $R > 5$ or $R < 1/5$. Since the reconstruction is made of two monotonicity preserving parabolas, the interface flux evaluation must take place in two integrations to accurately evaluate it. Although it is not conceptually difficult in one dimension, it may pose some difficulties in two-dimensional applications. The limiter performance for linear advection is shown in Fig. 5.6.



(a) $\nu = 0.4$



(b) $\nu = 0.8$

Figure 5.6: Zalesak test cases after one cycle with $N = 80$. Comparison between unlimited AF and limited by reconstruction modifying (RS). Notice the marked improvement of extrema solutions. Sharp discontinuities are not very well captured but better than the unlimited solution.

5.2.4 Accuracy Preserving Limiter of H. T. Huynh

The accuracy preserving limiter concept was introduced by H. T. Huynh [34] for the piecewise-parabolic dual-mesh scheme, a staggered-grid variant of Scheme V. He takes a one of a kind approach, not seen before in limiting. For these reasons, the accuracy preserving limiter of H. T. Huynh is investigated and put to trial in the AF advection schemes.

The limiter idea was based on a set of constraints ranging from a first-order to a third-order constraint that would select genuine extrema and suppress spurious oscillations. In the following sections, the constraints for the accuracy preserving limiter are presented.

5.2.4.1 First-Order Constraint

We introduce the geometric definition of monotonic reconstructions. It refers to reconstructions where no internal extremum is present within the cell extent.

(a) Type A (b) Type B

Figure 5.7: Monotonic reconstructions. Only concave down cases are shown.

There are two types of monotonic reconstructions, only concave down cases are shown in Fig. 5.7 and concave up cases can be imagined similarly. The differentiator is the location of reconstruction extremum. Type A attains an extremum at the right interface and the opposite for the Type B reconstruction.

We suppose that a first-order accurate scheme is obtained when the above monotonic reconstruction constraint is enforced. This will necessarily make the scheme monotone and will introduce some clipping at genuine extrema. For any quadratic reconstructions,

$$u(\xi) = a\xi^2 + b\xi + c \qquad \xi \in [0, 1] \tag{5.10}$$

we can define three constraints that would result in a monotonic reconstruction. Only the Type A reconstruction is considered as Type B is achieved by reflection. Three constraints to define a quadratic reconstruction are as follows.

1. $u(\xi = 1) = u_{j+1/2}$

2. $\frac{du}{d\xi}\Big|_{\xi=1} = 0$

3. $\bar{u}_j = \frac{1}{\Delta\xi}\int_0^1 u(\xi)d\xi$

While leaving the right interface value and the cell average unchanged, we force the derivative of the reconstruction at the right interface to be zero. This has the effect of "pushing" the internal extremum to the right interface. With these constraints, a monotonic reconstruction is found.

$$u^{\text{mon}}(\xi) = 3(\bar{u}_j - u_R)\xi^2 + 6(u_R - \bar{u}_j)\xi + (3\bar{u}_j - 2u_R) \tag{5.11}$$

From above, we obtain the limited left interface value for which the monotonic reconstruction is valid.

$$u_L^* = u^{\text{mon}}(0) = \bar{u}_j + 2(\bar{u}_j - u_R) \tag{5.12}$$

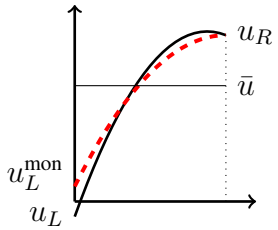


Figure 5.8: Comparison of reconstructions. The original non-monotonic reconstruction is shown in solid line. The monotonized reconstruction is shown in dashed red line.

Fig. 5.8 illustrates an example of an original reconstruction in solid line compared to a monotonized reconstruction in dashed line. Note that conservation is not violated. This limiter modifies the conservative flux through the right interface by modifying the reconstruction, rather than limiting the flux itself.

In its original debut [34], the first-order monotonic constraint was introduced using a median function. The median function returns the middle one of three numbers. This idea comes from the observation that monotonic solution is bounded by two extreme cases: piecewise constant and non-monotonic parabola. The interface value obtained from imposing the first-order constraint is

$$u_L^{1st} = \text{median}(\bar{u}_j, u_L, u_L^*) \tag{5.13}$$

where median and minmod functions are defined as follows.

$$\text{median}(a, b, c) = a + \text{minmod}(b - a, c - a) \tag{5.14}$$

$$\text{minmod}(a,b) = \frac{1}{2}(\text{sign}(a) + \text{sign}(b))\text{min}(|a|,|b|) \tag{5.15}$$

The median function continuously depends on data.

5.2.4.2 Second-Order Constraint

The second-order constraint is achieved by extending the stencil to include cell averages of either immediate neighbors. In order to determine the legitimacy of the extremum in the original reconstruction in Fig. 5.8, the cell average from right neighbor $\bar{u}_{j+1}$ is included. It, along with $\bar{u}_j$ and u_R, is used to approximate the second derivative, or the curvature, of the parabola extending to the right. The second derivative is

$$u'' = 3\left(\bar{u}_{j+1} + \bar{u}_j - 2u_R\right). \tag{5.16}$$

Now, a set of data, $\bar{u}_j$, u_R, and u'', is used to determine a second-order accurate approximation of the left interface value. Using Huynh's notation, the subscript XFR stands for extrapolating from the right.

$$u_{XFR} = 2\bar{u}_j - u_R + \frac{1}{6}u'' \tag{5.17}$$

This limit provides a second-order limit for the reconstruction, roughly speaking. Then, a new second-order accurate satisfying bound is defined.

$$u_L^{**} = \text{median}(\bar{u}_j, u_L, u_{XFR}) \tag{5.18}$$

Finally, the second-order accurate interface value is obtained from the median function.

$$u_L^{2nd} = \text{median}(\bar{u}_j, u_L^*, u_L^{**}) \tag{5.19}$$

5.2.4.3 Third-Order Constraint

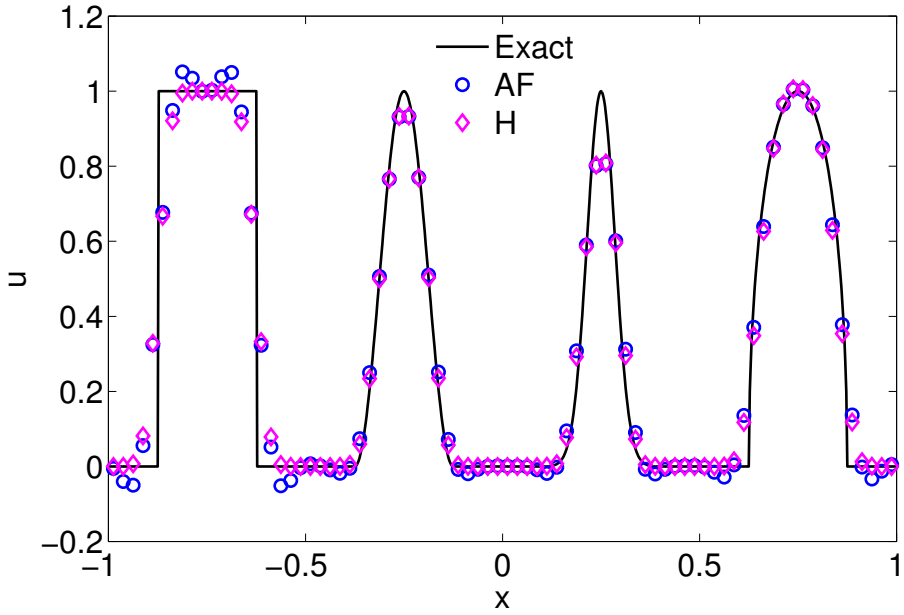
Lastly, the third-order constraint is considered by extending the curvature definition to its extreme, the largest allowable curvature. The interface value extrapolated from the right that has the largest allowable curvature, before violating the monotonicity condition, is $2u''$. The subscript LAC stands for the largest allowable curvature.

$$u_{LAC} = 2\bar{u}_j - u_R + \frac{1}{6}(2u'') \tag{5.20}$$

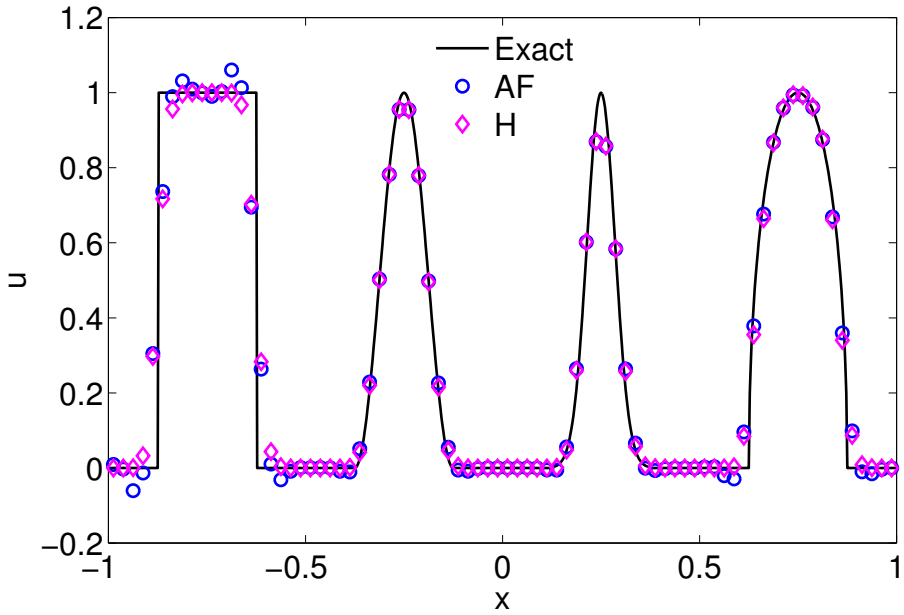
Finally, the third-order accurate interface value is obtained from the median function.

$$u_L^{3rd} = \text{median}(\bar{u}_j, u_L^*, u_{LAC}) \tag{5.21}$$

Because this constraint accurately approximates the limit of a bound to second-order, the interface value is evaluated to third-order. In Fig. 5.9, the limiter performance is presented for linear advection.



(a) $\nu = 0.4$



(b) $\nu = 0.8$

Figure 5.9: Zalesak test cases after one cycle with $N = 80$. Comparison between unlimited AF and limited by Huynh limiter (H). The overall performance of limiter is spectacular. The second- and third-order constraints are based on the spatial extension of data, not the point the new limiting paradigm looks for.

5.2.5 Accuracy Preserving Limiter Using the Past Time History

One variation of the accuracy preserving limiter of H. T. Huynh is presented. The high-order accuracy preserving constraints set in the last section require information from immediate neighbors. In other words, the numerical stencil is extended in spatial directions. While this is not algorithmically and mathematically challenging to understand, it violates the goal to remain compact. The compactness feature is appreciated in two or higher dimensions due to the logistics of data transfer. So, instead of expanding the numerical stencil in space, we choose to look into the past time history [75].

Now, to understand how we could utilize the past time solutions in the limiter design, we must revisit the second- and third-order constraints of the accuracy preserving limiter. The essential point of including the immediate neighbor cell average values is to approximate second derivatives of the solution. If the current cell reconstruction produced a solution that lies out of approximate bounds set by high-order constraints, then the solution must be limited.

A similar effect can be obtained by extending the numerical domain of dependence to include the past time solutions. Because solutions of hyperbolic partial differential equations are subject to a translational property as shown in Fig. 5.3, some relevant information can be obtained from the characteristic coordinate. Table 5.2 lists two sets of stencils, in addition to the current stencil at n. From these stencils, we construct the second- and third-order constraints, similar to the steps shown in Section 5.2.4.

Table 5.2: Temporal numerical stencils for accuracy preserving limiter using the past time history

Stencil	
n, Default	PQR
Like $j-1$	PSU
Like $j+1$	RTW

The first-order constraint from Huynh's limiter is unchanged. However, the second- and third-order constraints must be changed because we look at the past time history of the solution. Again, we only consider the Type A configuration in Fig. 5.7. The stencil with data points (RTW) is used to approximate the curvature, u''.

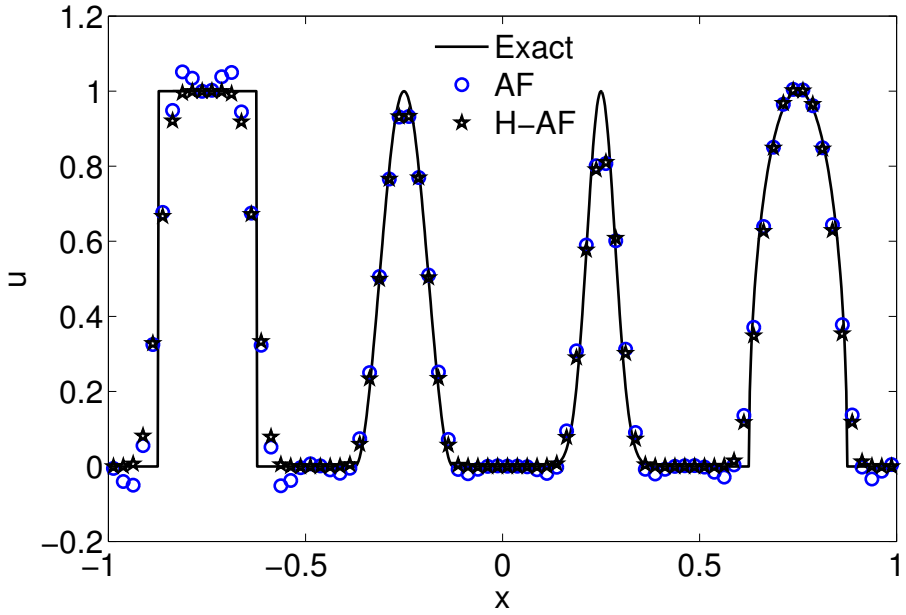
$$u'' = \frac{2}{\nu^2}\left(u^R - 2u^T + u^W\right). \tag{5.22}$$

where ν is the Courant number, u^R, u^T and u^W indicate values at R, T and W, respectively.

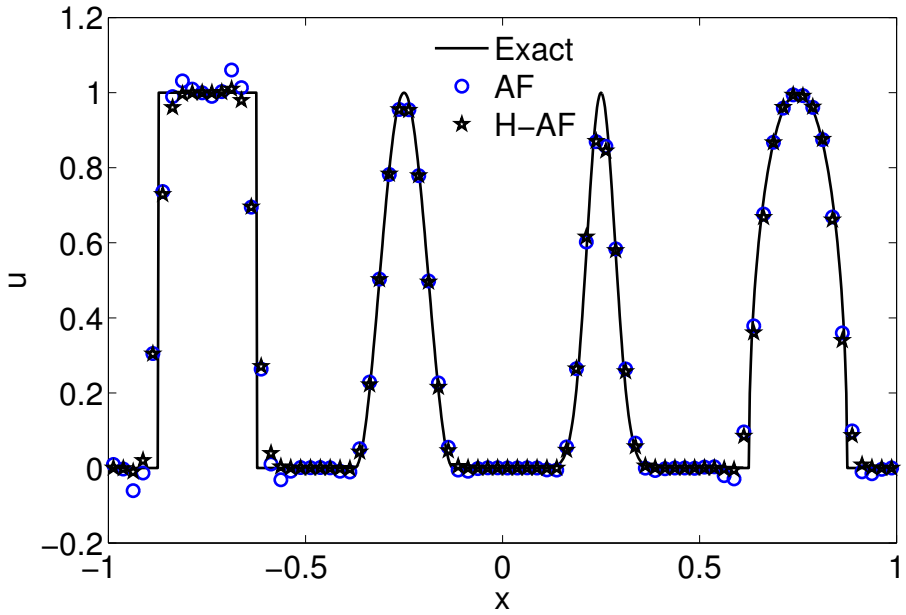
It is apparent that the curvature is dependent on the Courant number because characteristics depend on the Courant number.

The performance of the limiter for linear advection is shown in Fig. 5.10. However, the new adaptation of the accuracy preserving limiter develops small but noticeable numerical oscillations in the upwind side of discontinuities. From the practical standpoint, most smooth solutions are well resolved.

The past time history based accuracy preserving limiter is inherently upwind biased because the additional information used for high-order constraints are based on the characteristic information, see Fig. 5.3. This inherent deficiency will cause slight oscillations on the upwind side of sharp solution gradient changes. This was experimentally observed for Courant numbers $\nu > 0.7$, compare Figs. 5.9b and 5.10b.



(a) $\nu = 0.4$



(b) $\nu = 0.8$

Figure 5.10: Zalesak test cases after one cycle with $N = 80$. Comparison between unlimited AF and limited by the accuracy preserving limiter with temporal extension (H-AF).

5.3 Multidimensional Limiter

The design principles formerly mentioned in Section 5.1 is reminded again for the multidimensional limiter design. The intuitive requirement comes down to assigning physically

intuitive numerical processes to accurately represent the problem. In AF, we propose a new kind of operator splitting such that the numerical processes represented by each operator can be designed for maximal stability and accuracy, see Chapter 7. This decision impacts the limiter design accordingly. We believe that multidimensional limiters for advection and acoustics processes must differ. In following sections, we lay out some background as well as proposed approaches for advection and acoustics limiting.

It should be mentioned that there are other approaches for removing the unwanted oscillations in solutions, such as non-oscillatory reconstructions and artificial viscosity. However, as these approaches are a vast field of their own and do not represent the core belief of the AF method, we won't discuss them in this thesis.

5.3.1 Multidimensional Advection Limiting

For a large range of convection dominated problems, the slope limiting has been the *de facto* strategy. The widespread use of slope limiters is perhaps attributed to its conceptually simple mechanism. When the slope limiting ideas was introduced by van Leer [83, 85], the idea had not presented the contention on the subject that it has now, especially for multidimensional problems. It was based on one-dimensional flow problems with the use of discontinuous solution types. As we have stressed throughout the dissertation, these two observations are precisely the points of objection for the AF method. The difficulty in multidimensional limiting is still felt across high-order numerical methods communities as a good basis for a high-order limiting is scarce. Nevertheless, there has been a steady progress in limiting for the DG methods: based on high-order limiting [44, 43], artificial viscosity [62], non-oscillatory reconstruction [49, 93, 54, 63], and flux reconstruction [35]. In addition, a new class of limiting has been introduced called *a posteriori* limiting [20]. It introduces a posteriori candidate solution computed by the unlimited Arbitrary high-order DERivatives discontinuous Galerkin (ADER-DG) scheme and compare to a high-order DG solution to check if certain design criteria are satisfied. It shares some of our limiting criteria we set forth, and definitely should be studied further.

The RD schemes have gained respect for computation of steady solutions. Even though it's acclaimed fame resides in the steady solution fields, there has been continued research endeavor in achieving unsteady RD schemes and corresponding multidimensional limiter [33, 32, 4]. For unsteady flows, the RD schemes are either implicit in nature or require the use of high-order time integrator. But the aspect that RD schemes are naturally favored compared to other schemes perhaps is the fact that they are genuinely multidimensional and nonlinear at the same time.

One type of limiting based on the general slope limiting idea is considered as a preliminary candidate for deciding the viability of using the temporal extension of data to the past. The reason for choosing the slope limiting idea lies on the simplicity aspect. Because spurious oscillations are directly controllable by modifying the slopes, it makes the process of limiting quite straightforward. The total variation bounded (TVB) limiters have been available for multidimensional problems [77]. And more recently it was applied in a high-order compact SV scheme [89]. The inherent shortcoming of TVB limiters is its need for customization of the bound. The SV scheme requires a tunable parameter to be set for a particular type of problem.

In our first attempt to multidimensional advection limiting, we make the following suggestion. Let us not worry about the accuracy aspect of the limited solution for the moment. But some essential requirements for the multidimensional limiting must be addressed.

- Conservative
- Compact

The compactness requirement is introduced by seeking the past time solution data belonging to the same element for each limiting criterion. This is quite different from the generalized slope limiters which seek some information from neighboring cells. The limiting criterion we use here is a simple on-off switch defined by a ratio of curvatures at two time steps. This makes the scheme compact.

$$\text{Limiter} = \begin{cases} \text{on} & \text{if } \frac{1}{2} \leq \frac{\bar{c}_j^n}{\bar{c}_j^{n-1}} \leq 2 \\ \text{off} & \text{otherwise} \end{cases} \tag{5.23}$$

where $\bar{c}_j^n$ and $\bar{c}_j^{n-1}$ are average curvatures at time t^n and t^{n-1}. We propose a simple reconstruction for limiting the solution. When the limiter is "on", the following slope limiter is invoked.

$$q(\boldsymbol{\xi}) = \overline{q}_j + \nabla q_c^{\text{lim}}(\boldsymbol{\xi} - \boldsymbol{\xi}_c) \tag{5.24}$$

where ∇q_c^{lim} is the limited slope at the centroid. The easiest choice is $\nabla q_c^{\text{lim}} = \nabla q_c^n$ the slope available at current time step t^n. This makes the scheme conservative. Strictly speaking, this particular choice of the slope does not result in a slope limiter, because it still admits solutions that exceed minima and maxima. But it should help with spurious oscillations that arise from the curvature of the reconstruction.

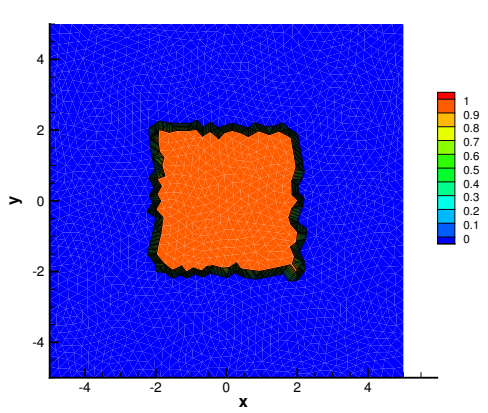
5.3.1.1 Preliminary results

A square initial wave is considered for two-dimensional linear advection.

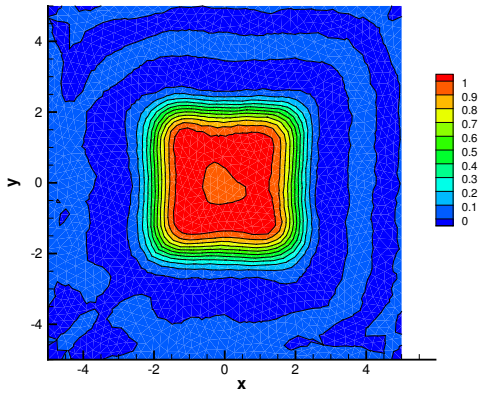
$$u_0(x,y) = \begin{cases} 1 & \text{if } |x|, |y| \leq 2.0 \\ 0 & \text{otherwise} \end{cases} \tag{5.25}$$

They are shown to provide the readers that some acceptable limited results can be achieved with the compact requirement that we extend our stencil in the spatial direction. More in-depth and specific analysis made towards limiting would definitely produce a better multidimensional limiter. Fig. 5.11 shows the solution contour plot after one cycle on the unstructured grid. The grid used for the solutions here are shown in Fig. 3.20. The limiter does not introduce excessive amounts of dissipation. In fact, the contour plots are qualitatively similar. But improvements would definitely help.

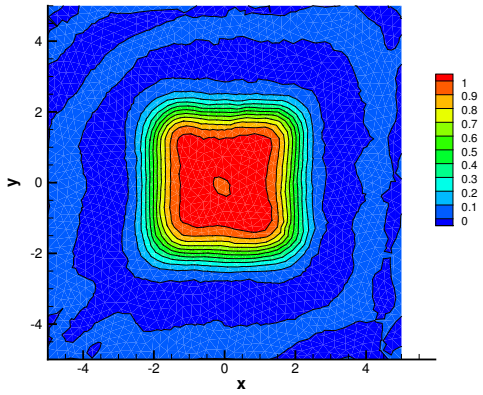
The next Fig. 5.12 shows the solution contour plot after one cycle on the triangular structured grid A. This grid configuration was notorious for producing mesh alignment issue. The limiter removes the mesh alignment issue while preserving the initial condition. This empirical result shows that additional efforts put towards developing a multidimensional limiter may remove some mesh alignment issues as well.



(a) Initial solution

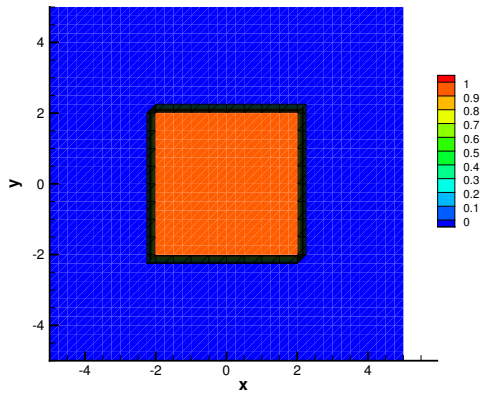


(b) Unlimited solution after one cycle at $t = 10$

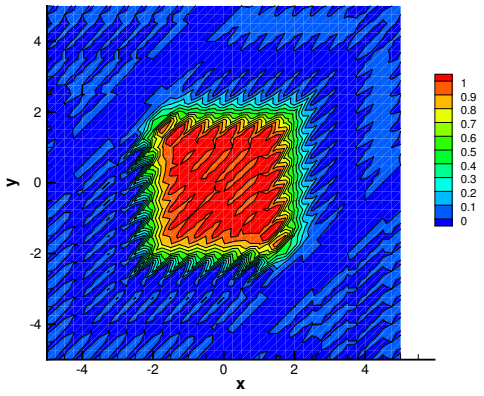


(c) Limited solution after one cycle at $t = 10$

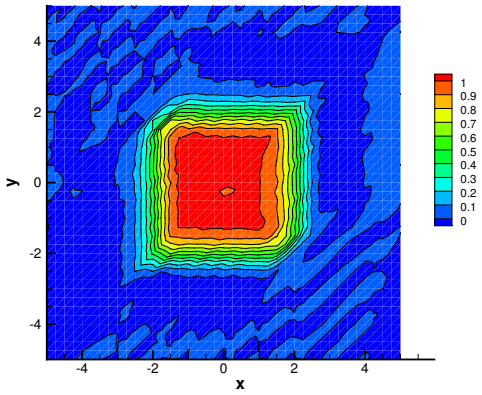
Figure 5.11: Square wave solution on triangular unstructured grid with $N = 814$, $\nu = 0.7$, and $\mathbf{a} = (1, 1)^T$



(a) Initial solution



(b) Unlimited solution after one cycle at $t = 10$



(c) Limited solution after one cycle at $t = 10$

Figure 5.12: Square wave solution on triangular structured grid A with $N = 800$, $\nu = 0.7$, and $\mathbf{a} = (1, 1)^T$

5.3.2 Multidimensional Acoustics Limiting

The maximum principle has been extensively used in the one-dimensional linear problems with the explicit knowledge of *a priori* bound on the solution. The Flux Corrected Transport (FCT) algorithm proposed by Boris and Book [9] when first introduced for one-dimensional transport equations, also shared the same basic principle. However, there is a *qualitative* aspect of FCT and the multidimensional extension of FCT introduced by Zalesak [98] that makes this particular approach more relevant as a method of choice for the multidimensional acoustics limiting.

FCT operates on the following principle. A low-order accurate but an oscillation-free flux, and a high-order accurate flux that is not oscillation free are introduced. Based on

some criterion, the anti-diffusive flux, which is the difference between high- and low-order accurate fluxes, is applied to the base scheme to achieve a crisp resolution. The process is based on a qualitative argument of deciding the "best" available result given two sets of data that differ in character. This draws a contrast to the single step slope limiting paradigm introduced earlier. Now there is no firm guideline for choosing the criterion used for finding the anti-diffusive flux. It happened to include solutions of the extended spatial stencil, but one could just as easily choose to include temporal spatial extensions without undue repercussion. The following suggestion is particularly a useful concept to utilize in the limiting process in general.

As it was noted before, multidimensional physics is inherently different from one-dimensional physics on which many numerical methods are based. Instead of a maximum principle, one must consider more natural criteria for deciding bounds of a problem. A more natural bound is that the integral of "energy" is conserved [25]. Note that this does not set *a priori* bound on individual scalar quantities. In particular, the limiting process can be loosely based on the temporal changes of the physically intuitive quantities. Notions of "driver" and "driven" quantities in the context of Lagrangian hydrodynamics equations were introduced by Lung [52], and Lung and Roe [53, 75]. Temporal changes of these quantities can be measured in terms of time derivatives, $\frac{\partial}{\partial t}$, $\frac{\partial^2}{\partial t^2}$, which coincidentally are parts of the Lax-Wendroff expansion in time. Clever use of this fact may result in a highly optimized high-order scheme with a highly optimized limiting scheme. I believe that the acoustics limiting must take place on the same values that Lung and Roe tried to achieve.

In no ways, the discussion on multidimensional acoustics limiting provided here is close to complete. There are a lot of challenges that lie ahead of us for lack of a good understanding of the matter.

CHAPTER 6

Nonlinear Hyperbolic System

The AF method for nonlinear hyperbolic systems is first studied with the pressureless Euler system. The pressureless Euler system is a subset of the Euler system in the limit of vanishing pressure, so only the advective part of the Euler system remains. It is also the system that we need to solve in an operator split approach of the Euler system.

In conservation form, the pressureless Euler system is written as follows.

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{v} = 0 \tag{6.1}$$

$$\frac{\partial \rho \mathbf{v}}{\partial t} + \nabla \cdot (\rho \mathbf{v} \otimes \mathbf{v}) = 0 \tag{6.2}$$

The pressureless Euler system has been used as simple models in cosmological dust cloud dynamical systems [99] or low-temperature plasma physics systems where hydrodynamical forces are predominant [11]. The pressureless Euler system has been intensively analyzed in one dimension, but only a few numerical approaches have been made in more than one dimension [95]. It is a hyperbolic conservation system despite not being strictly hyperbolic due to degenerate eigenvalues and non-distinct eigenvectors in more than one dimension. However, its true significance lies in that it allows us to study the nonlinear advective conservation system.

For the remainder of the chapter, numerical methods are established to solve the pressureless Euler system in both one and two dimensions. The AF method for nonlinear hyperbolic systems is accomplished in three stages. The first two stages, the nonconservative and conservative update stage, remain unchanged from nonlinear scalar conservation laws. However, an additional stage, called the reconciliation stage, must be considered in nonlinear systems. Algorithm 2.1 is consulted for the flow of the AF method in system applications.

6.1 Nonconservative System

It is instructive to consider nonconservative systems because when a system is written in terms of physical variables, it can reveal physical details associated with the system.

We first discuss the one-dimensional pressureless Euler system.

$$\frac{\partial \rho}{\partial t} + u\frac{\partial \rho}{\partial x} + \rho\frac{\partial u}{\partial x} = 0 \tag{6.3}$$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0 \tag{6.4}$$

It is evident that the velocity equation Eq. (6.4) is identical to the one-dimensional Burgers' equation. The numerical method developed for Burgers' equation, which is the characteristic tracing method, directly applies here, see Section 4.2. The density equation Eq. (6.3), however, consists of two terms with distinct properties: the advection and divergence term.

The two-dimensional pressureless Euler system is shown.

$$\frac{\partial \rho}{\partial t} + u\frac{\partial \rho}{\partial x} + v\frac{\partial \rho}{\partial y} + \rho\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0 \tag{6.5}$$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = 0 \tag{6.6}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} = 0 \tag{6.7}$$

A very similar structure is found. The velocity equations, Eqs. (6.6) and (6.7), describe a version of two-dimensional Burgers' system. This system is different from the two-dimensional Burgers' equation discussed in Section 4.3 because the velocity components are coupled. Now, a semi-Lagrangian method for advancing the multidimensional nonlinear advection system in time is presented.

6.1.1 Streamline Tracing Method for Multidimensional Systems

The characteristic tracing method for the multidimensional Burgers' system is now called a *streamline tracing* method as there are no generalized characteristic solutions for the multidimensional Burgers' system [76]. The streamline tracing method is a semi-Lagrangian method where numerical solutions in the Eulerian frame are obtained by advancing the solutions along streamlines. The streamlines are straight and composed of velocity vectors $\mathbf{v}$.

We discuss the two-dimensional Burgers' equation to convey the idea. Streamlines, like nonlinear characteristics, are nonlinear in nature. First-order corrections to the streamlines

are devised to obtain at least second-order accurate interface fluxes. For an arbitrary velocity vector, we introduce a simplified vector at a particular location $\mathbf{v}(\mathbf{x}, t) = \tilde{\mathbf{v}}(t)$. A Taylor expansion of the velocity vector $\tilde{\mathbf{v}}(t)$ is considered.

$$\tilde{\mathbf{v}}(t) = \tilde{\mathbf{v}}(0) + t\frac{\partial \tilde{\mathbf{v}}}{\partial t} \tag{6.8}$$

A Lax-Wendroff substitution is used to replace the time derivative, $\frac{\partial \tilde{\mathbf{v}}}{\partial t} = -\tilde{\mathbf{v}}(t) \cdot \nabla \tilde{\mathbf{v}}$.

$$\tilde{\mathbf{v}}(t) = \tilde{\mathbf{v}}(0) - t\tilde{\mathbf{v}}(t) \cdot \nabla \tilde{\mathbf{v}} \tag{6.9}$$

$$\Rightarrow \tilde{\mathbf{v}}(t) = \frac{\tilde{\mathbf{v}}(0)}{\mathbf{I} + t\nabla \tilde{\mathbf{v}}} \tag{6.10}$$

$$= \boldsymbol{\mathcal{J}}^{-1}\tilde{\mathbf{v}}(0) + \mathcal{O}(t^2) \tag{6.11}$$

where

$$\boldsymbol{\mathcal{J}} = (\mathbf{I} + t\nabla \tilde{\mathbf{v}}) = \begin{pmatrix} 1 + t\frac{\partial \tilde{u}}{\partial x} & t\frac{\partial \tilde{u}}{\partial y} \\ t\frac{\partial \tilde{v}}{\partial x} & 1 + t\frac{\partial \tilde{v}}{\partial y} \end{pmatrix} \tag{6.12}$$

and $\mathbf{I}$ is an identity matrix. Alternatively, Eq. (6.11) is equivalently written as follows.

$$\tilde{\mathbf{v}}(t) = \begin{pmatrix} \tilde{u}(t) \\ \tilde{v}(t) \end{pmatrix} = \begin{pmatrix} \tilde{u}(0) - t\left(\tilde{u}(0)\frac{\partial \tilde{u}}{\partial x} + \tilde{v}(0)\frac{\partial \tilde{u}}{\partial y}\right) \\ \tilde{v}(0) - t\left(\tilde{u}(0)\frac{\partial \tilde{v}}{\partial x} + \tilde{v}(0)\frac{\partial \tilde{v}}{\partial y}\right) \end{pmatrix} + \mathcal{O}(t^2) \tag{6.13}$$

This numerical treatment evidently illustrates the multidimensional nature of the scheme. There are small, $\mathcal{O}(t)$, but definite presence of cross-wind effects in both velocity components. The extension to three dimensions can be made easily.

Once the first-order corrected velocity $\tilde{\mathbf{v}}(t)$ is known, the advective portion of the solution is advanced in time by a semi-Lagrangian method. For nonconservative variable reconstructions at time t^n,

$$\mathbf{q}^n(\boldsymbol{\xi}) = \sum_{i=1}^{7} c_i(\mathbf{q}) \cdot \phi_i(\boldsymbol{\xi}) \tag{6.14}$$

where

$$\mathbf{q} = \begin{pmatrix} \rho \\ u \\ v \end{pmatrix} \tag{6.15}$$

and $c_i(\mathbf{q})$ indicates the Lagrange basis coefficients of $\mathbf{q}$. The coefficient and basis function definitions are shown in Table 3.5. The nonconservative variables $\mathbf{q}^{n+1}$ are obtained by

tracing the streamlines to origins $\boldsymbol{\xi}_F$.

$$\mathbf{q}^{n+1}(\boldsymbol{\xi}_I) = \mathbf{q}^n(\boldsymbol{\xi}_F) \tag{6.16}$$

$$\boldsymbol{\xi}_F = \boldsymbol{\xi}_I - \Delta t \mathbf{J}^{-1}\tilde{\mathbf{v}}^n(\Delta t; \boldsymbol{\xi}_I) \tag{6.17}$$

where $\tilde{\mathbf{v}}^n(\Delta t; \boldsymbol{\xi}_I)$ indicates the nonlinear corrected streamline velocity at t^{n+1} and $\boldsymbol{\xi}_I$. Half-time step values are obtained similarly.

$$\mathbf{q}^{n+1/2}(\boldsymbol{\xi}_I) = \mathbf{q}^n(\boldsymbol{\xi}_H) \tag{6.18}$$

$$\boldsymbol{\xi}_H = \boldsymbol{\xi}_I - \frac{\Delta t}{2}\mathbf{J}^{-1}\tilde{\mathbf{v}}^n\left(\frac{\Delta t}{2}; \boldsymbol{\xi}_I\right) \tag{6.19}$$

Note that the first-order corrected velocity $\tilde{\mathbf{v}}(t)$ directly depends on time, and would result in different values for different time steps Δt and $\frac{\Delta t}{2}$, see Eq. (6.13).

The density equation in Eq. (6.5) is comprised of two distinct terms. The numerical method for the advective term has been developed in this section. The other term is the velocity divergence term. It couples the velocity to density, or vice versa. In other words, by introducing the velocity divergence, the density equation can be formed into a divergence form for the conservation of system. The numerical method for treating the velocity divergence term is discussed in the following section.

6.1.2 Streamtube Area

The particle path in the pressureless Euler system is described in accordance with the mass conservation.

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{v} = 0 \tag{6.20}$$

Mass in a streamtube, ρA, stays constant in time. The streamlines are straight without pressure. For each particle path with some velocity vector $\mathbf{v}$, a new position for the particle is found in time. The expression for the new position is

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{v}t. \tag{6.21}$$

For small vectors, it transforms as

$$d\mathbf{x} = d\mathbf{x}_0 + \nabla \mathbf{x}_0 = (\mathbf{I} + t\nabla \mathbf{v}) d\mathbf{x}_0. \tag{6.22}$$

And the change in a streamtube area is found as follows.

$$dA = \det\left(\mathbf{I} + t\nabla\mathbf{v}\right) dA_0 \tag{6.23}$$

So, the final density taking into account of the streamtube area change becomes

$$\rho(t) = \frac{\rho_0}{\det\left(\mathbf{I} + t\nabla\mathbf{v}\right)}. \tag{6.24}$$

This formula describing the conservation of mass is exact for all t until streamlines collide. Now, streamtube area changes in one- and two-dimensional pressureless Euler system are discussed in detail.

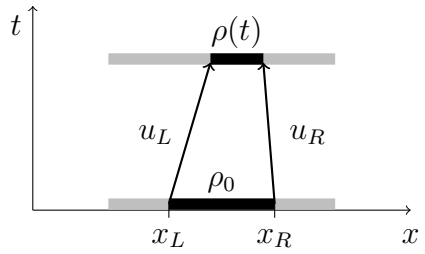


Figure 6.1: Streamtube area change due to the velocity divergence in 1D

In one dimension, the final density at time t is

$$\rho(t) = \frac{\rho_0}{\det\left(\mathbf{I} + t\nabla\mathbf{v}\right)} = \frac{\rho^n(\xi_F)}{1 + t\frac{\partial u}{\partial x}\Big|_{\xi_F}} \tag{6.25}$$

where $\rho_0 = \rho^n(\xi_F)$ is the density value obtained from the streamline tracing method. In Fig. 6.1, a one-dimensional example illustrates the streamtube area change, therefore the density change, due to the velocity divergence. The velocity divergence term is evaluated from the gradient of Lagrange basis functions, see Table 3.2 for definitions of the basis functions.

$$\frac{\partial u}{\partial x}\Big|_{\xi_F} = \sum_{i=1}^{3} c_i(u^n(\xi_F)) \cdot \frac{\partial \xi}{\partial x}\frac{\partial \phi_i}{\partial \xi} \tag{6.26}$$

$$= \sum_{i=1}^{3} c_i(u^n(\xi_F)) \cdot J^{-1}\frac{\partial \phi_i}{\partial \xi} \tag{6.27}$$

where $c_i(u^n(\xi_F))$ is the velocity coefficient interpolated at the streamline origin location ξ_F. The velocity gradients at the half time step are defined in an analogous manner. Note that the Lagrange basis functions are defined in the reference coordinate. The transformation

from the reference to the physical coordinate is accomplished by the mapping Jacobian $J^{-1} = \frac{1}{\Delta x}$.

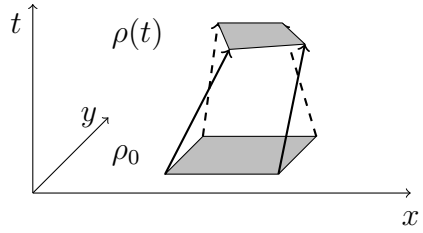


Figure 6.2: Streamtube area change and twisting from the velocity divergence in 2D

In two dimensions, the final density at t is evaluated as follows.

$$\rho(t) = \frac{\rho_0}{\det\left(\mathbf{I} + t\nabla\mathbf{v}\right)} = \frac{\rho^n(\boldsymbol{\xi}_F)}{\det\left(\boldsymbol{\mathcal{J}}|_{\boldsymbol{\xi}_F}\right)} = \frac{\rho^n(\boldsymbol{\xi}_F)}{1 + t\nabla\cdot\mathbf{v}|_{\boldsymbol{\xi}_F} + t^2\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial v}{\partial x}\frac{\partial u}{\partial y}\right)\Big|_{\boldsymbol{\xi}_F}} \tag{6.28}$$

where both the numerator and denominator terms are evaluated from the quantities at the streamline origin $\boldsymbol{\xi}_F$. The denominator is the determinant of matrix $\boldsymbol{\mathcal{J}}$, Eq. (6.12), evaluated at the origin.

$$\det\left(\boldsymbol{\mathcal{J}}|_{\boldsymbol{\xi}_F}\right) = 1 + t\nabla\cdot\mathbf{v}|_{\boldsymbol{\xi}_F} + t^2\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial v}{\partial x}\frac{\partial u}{\partial y}\right)\Big|_{\boldsymbol{\xi}_F} \tag{6.29}$$

where the first-order term governs the change of the streamtube area and the second-order terms define the effect of twisting due to non-uniform velocity vectors. In Fig. 6.2, an example of the two-dimensional velocity divergence is illustrated. The velocity gradients used in the first- and second-order terms are evaluated from the gradients of Lagrange basis functions, see Table 3.5 for definitions of basis functions. For example, u velocity gradients are obtained as

$$\frac{\partial u}{\partial \mathbf{x}}\Big|_{\boldsymbol{\xi}_F} = \sum_{i=1}^{7} c_i(u^n(\boldsymbol{\xi}_F))\cdot\frac{\partial\boldsymbol{\xi}}{\partial\mathbf{x}}\frac{\partial\phi_i}{\partial\boldsymbol{\xi}} \tag{6.30}$$

$$= \sum_{i=1}^{7} c_i(u^n(\boldsymbol{\xi}_F))\cdot\mathbf{J}^{-1}\frac{\partial\phi_i}{\partial\boldsymbol{\xi}} \tag{6.31}$$

where $c_i(u^n(\xi_F))$ is the velocity coefficient interpolated at the streamline origin location ξ_F. The velocity gradients at the half time step are defined in an analogous manner. Note that the Lagrange basis functions are defined in the reference coordinate. The transformation from the reference to physical coordinate is accomplished by the mapping Jacobian $\mathbf{J}^{-1}$

matrix defined in Eq. (3.59). The numerical consistency of the two-dimensional streamtube area change term, Eq. (6.29), is verified in Appendix A.

6.2 System of Conservation Laws

The integral form of the conservation system is presented as follows.

$$\int_\Omega \frac{\partial \mathbf{u}}{\partial t} dV + \int_\Omega \nabla \cdot \mathbf{F} dV = 0 \tag{6.32}$$

$$\frac{d}{dt}\int_\Omega \mathbf{u} dV + \oint_{\partial\Omega} \mathbf{F} \cdot \mathbf{n} dS = 0 \tag{6.33}$$

where the conservative variables and fluxes for the pressureless Euler system are

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \rho u & \rho v \\ \rho uu & \rho uv \\ \rho uv & \rho vv \end{pmatrix}. \tag{6.34}$$

The general form of discrete conservation law for systems is presented.

$$\overline{\mathbf{u}}_j^{n+1} = \overline{\mathbf{u}}_j^n - \frac{\Delta t}{\Omega_j} \sum_{l=1}^{\text{sides}} \overline{\mathbf{F}}_l \cdot \mathbf{n}_l \tag{6.35}$$

where $\mathbf{n}_l$ is a scaled outward normal vector of side l. The flux tensor on each interface $\overline{\mathbf{F}}_l$ is evaluated numerically according to the composite Simpson's formula described in Eq. (3.73). The conservative flux is second-order accurate, and the conservation stage produces a third-order scheme.

In the AF method, dealing with the system of conservation laws introduces some nuance that scalar conservation laws don't pose. One may notice that conservative variables $\overline{\mathbf{u}}$ obtained from the discrete conservation law for systems, Eq. (6.35), have different a form from the individual components of the interface flux. Most notably that the flux tensor elements are comprised of nonconservative variables $\mathbf{q}$. Evolution stages of AF involve updating individual nonconservative variables. This feature, in general, is very different from conventional compressible flow solvers where fluxes are *passive* functions of conservative variables. As a result, the AF method requires an additional step called a "reconciliation" after the conservation step is completed.

6.3 Reconciliation for System

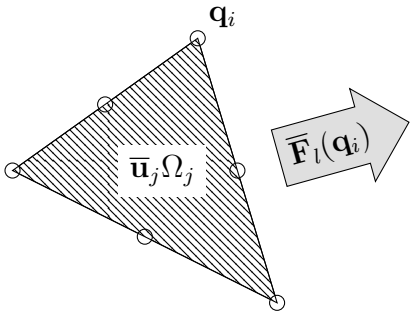


Figure 6.3: Finite volume element with AF nomenclature. The element consists of the cell average $\overline{\mathbf{u}}_j$, primitive variables defined at Lagrange basis locations $\mathbf{q}_i$.

The reconciliation process is perhaps unique to the AF scheme. Let us recall the two distinct update procedures in AF: the evolution and conservation update stages in Algorithm 2.1.

First, Fig. 6.3 is introduced for a proper nomenclature system. In the evolution update stage, primitive or nonconservative variables $\mathbf{q}$ are updated explicitly in time by an appropriate numerical method. This stage is an efficient way to embody different aspects of physics.

In the conservation update stage, discrete conservation laws are solved by an FV scheme. But it is important to stress that conservative fluxes are functions of nonconservative variables defined at Lagrange basis locations.

$$\overline{\mathbf{F}}_l = f(\mathbf{q}_i) \qquad \forall i \in l \tag{6.36}$$

where l is the triangle face/side index. Conservative fluxes are *not* based on cell average quantities $\overline{\mathbf{u}}_j$, unlike it is often the case in FV methods. Therefore, we see that for each iteration of the AF method, conservative variables are influenced by the primitive variables, but not the other way around. This shows the need for the reconciliation. The reconciliation procedure provides a feedback from conservative variables to primitive reconstructions by assigning appropriate bubble functions.

We have formerly introduced the bubble function, see Section 3.3.4, as a way to enforce conservation and maintain the compactness of the scheme. For nonlinear systems, approaches for obtaining bubble functions require some subtle differences. The subtlety is due to the highly coupled nature of the primitive variables and conservative variables. Now reconciliation procedures for one- and two-dimensional nonlinear systems are discussed.

6.3.1 Reconciliation in One-Dimensional System

In the one-dimensional pressureless Euler system, the conservative variables and primitive variables are shown as follows.

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho u \end{pmatrix}, \qquad \mathbf{q} = \begin{pmatrix} \rho \\ u \end{pmatrix} \tag{6.37}$$

In discrete conservation systems, the area averaged quantities of conservative variables are more relevant. Therefore, the following variables are used for the reconciliation instead.

$$\overline{\mathbf{u}}_j = \begin{pmatrix} \overline{\rho}_j \\ \overline{\rho u}_j \end{pmatrix}, \qquad \mathbf{q}_i = \begin{pmatrix} \rho_i \\ u_i \end{pmatrix} \tag{6.38}$$

where j is a cell index and i is a local Lagrange function index of the cell, see Fig. 6.4.

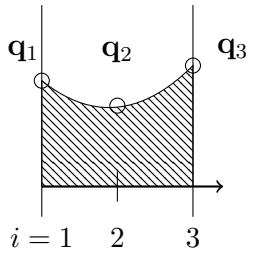


Figure 6.4: Element j reference coordinate nomenclature for the reconciliation procedure in 1D

At the end of the conservation update procedure, all conservative variables have their newest values $\overline{\mathbf{u}}_j^{n+1}$ and nonconservative variables located on interfaces have their newest values $\mathbf{q}_1^{n+1}$ and $\mathbf{q}_3^{n+1}$. As of now, the interior nonconservative variable $\mathbf{q}_2^{n+1}$ is undefined. For brevity, the superscripts are dropped if no confusion arises. We know the primitive variables are located at Lagrange basis locations and form a quadratic reconstruction, see Table 3.2 for definitions of coefficients and basis functions.

$$\mathbf{q}(\xi) = \sum_{i=1}^{3} c_i(\mathbf{q}) \cdot \phi_i(\xi) \tag{6.39}$$

From this, relations between conservative variables and primitive reconstructions can be

established.

$$\overline{\rho}_j = \frac{1}{\Delta\xi}\int_{\Omega_j} \rho(\xi) dV \tag{6.40}$$

$$\overline{\rho u}_j = \frac{1}{\Delta\xi}\int_{\Omega_j} \rho(\xi)\cdot u(\xi) dV \tag{6.41}$$

The density relation, Eq. (6.40), is approximately established by the use of the Simpson's rule. The Simpson's rule can produce a third-order accurate integration, which happens to be the accuracy of the scheme.

$$\overline{\rho}_j = \frac{1}{\Delta\xi}\int_{\Omega_j} \rho(\xi) dV \approx \frac{1}{6}\left(\rho_1 + 4\rho_2 + \rho_3\right) + \mathcal{O}(\Delta\xi^4) \tag{6.42}$$

where ρ_2 is the only unknown. Solving for it defines the density reconstruction completely. However, the momentum relation Eq. (6.41) requires some clarification due to the fact that it is a product of two quadratic reconstructions. We make an approximation of Eq. (6.41) to simplify the process.

$$\overline{\rho u}_j = \frac{1}{\Delta\xi}\int_{\Omega_j} \rho(\xi)\cdot u(\xi) dV \approx \frac{1}{\Delta\xi}\int_{\Omega_j} \sum_{i=1}^{3} \left[(\rho_i u_i)\cdot \phi_i(\xi)\right] dV \tag{6.43}$$

$$\approx \frac{1}{6}\left(\rho_1 u_1 + 4\rho_2 u_2 + \rho_3 u_3\right) + \mathcal{O}(\Delta\xi^4) \tag{6.44}$$

Since we already know u_1 and u_3 from the evolution of velocity variables on interfaces, the only unknown is u_2. The interior density value ρ_2 is obtained from solving Eq. (6.42) so u_2 is the only unknown in Eq. (6.44). And the velocity reconstruction $u(\xi)$ is defined.

6.3.2 Reconciliation in Two-Dimensional System

The reconciliation in the two-dimensional pressureless Euler system is presented. But the idea can be extended to all other systems.

$$\overline{\mathbf{u}}_j = \begin{pmatrix} \overline{\rho}_j \\ \overline{\rho u}_j \\ \overline{\rho v}_j \end{pmatrix}, \quad \mathbf{q}_i = \begin{pmatrix} \rho_i \\ u_i \\ v_i \end{pmatrix} \tag{6.45}$$

where j is a cell index, and i is a nodal index belonging to the same cell, see Fig. 6.5.

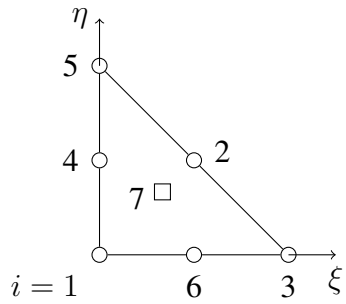


Figure 6.5: Element j reference coordinate nomenclature for the reconciliation procedure in 2D. In a reference element, $\Delta\xi = \Delta\eta$.

At the end of the conservation update procedure, all conservative variables $\overline{\mathbf{u}}_j^{n+1}$ and primitive variables $\mathbf{q}_i^{n+1}$ have their newest values. For brevity, the superscripts are dropped if no confusion arises. We know the primitive variables are located on the Lagrange basis locations and form a quadratic reconstruction. The coefficients and basis function definitions are shown in Table 3.5.

$$\mathbf{q}(\boldsymbol{\xi}) = \sum_{i=1}^{6} c_i(\mathbf{q}) \cdot \phi_i(\boldsymbol{\xi}) \tag{6.46}$$

The relation between conservative variables and primitive reconstruction can be established as follows.

$$\overline{\rho}_j = \frac{1}{\Omega_j} \int_{\Omega_j} \rho(\boldsymbol{\xi}) dV \tag{6.47}$$

$$\overline{\rho u}_j = \frac{1}{\Omega_j} \int_{\Omega_j} \rho(\boldsymbol{\xi}) \cdot u(\boldsymbol{\xi}) dV \tag{6.48}$$

$$\overline{\rho v}_j = \frac{1}{\Omega_j} \int_{\Omega_j} \rho(\boldsymbol{\xi}) \cdot v(\boldsymbol{\xi}) dV \tag{6.49}$$

In two dimensions, we suggest a slightly different approach for approximation. The reason for a sudden shift is not apparent and needs more elaboration. In one-dimensional problems, the quadratic reconstruction is fully defined by three DOF. However, in two-dimensional problems, we introduced a bubble function to make a cubic reconstruction polynomial. This increased the DOF to seven per element. For the moment, we leave bubble functions out of the reconstructions.

The density reconstruction is defined with ease.

$$\overline{\rho}_j = \frac{1}{\Omega_j} \int_{\Omega_j} \rho(\boldsymbol{\xi}) dV = \frac{1}{3} \left(\rho_2 + \rho_4 + \rho_6\right) + \frac{9}{20} \rho_7 \tag{6.50}$$

where ρ_7 is the only unknown. However, the following approximation for the momentum equations is suggested. We accept the conservative variables are obtained by a combination of primitive reconstructions. At the same time, we also make sure that the integral of individual primitive reconstruction is consistent with the quadratic quadrature rule provided in Eq. (3.77).

$$\overline{\rho u}_j = \frac{1}{\Omega_j}\int_{\Omega_j} \rho(\boldsymbol{\xi}) \cdot u(\boldsymbol{\xi}) dV \approx \overline{\rho}_j \overline{u}_j + \Delta\rho u_j \tag{6.51}$$

$$\overline{\rho v}_j = \frac{1}{\Omega_j}\int_{\Omega_j} \rho(\boldsymbol{\xi}) \cdot v(\boldsymbol{\xi}) dV \approx \overline{\rho}_j \overline{v}_j + \Delta\rho v_j \tag{6.52}$$

where $\Delta\rho u_j$ and $\Delta\rho v_j$ terms are truncation errors resulting from the approximation, and approximating these truncation terms will be the approach by which we reach an accurate reconciliation.

The truncation terms are approximated so that Eqs. (6.51) and (6.52) are approximated to $\mathcal{O}(\Delta\xi^4)$. The truncation terms are obtained by a Taylor expansion analysis. Taylor expansions of primitive reconstructions $\mathbf{q}(\boldsymbol{\xi})$ about $i = 1$ are substituted into momentum equations, Eqs. (6.51) and (6.52). This reveals the following truncation terms are sufficient for $\mathcal{O}(\Delta\xi^4)$ accurate approximation of Eqs. (6.51) and (6.52).

$$\Delta\rho u_j = \frac{1}{36}\left(\frac{d\rho}{dx}\left(2\frac{du}{dx} - \frac{du}{dy}\right) + \frac{d\rho}{dy}\left(\frac{du}{dy} - \frac{du}{dx}\right)\right)\Delta\xi^2 \tag{6.53}$$

$$\Delta\rho v_j = \frac{1}{36}\left(\frac{d\rho}{dx}\left(2\frac{dv}{dx} - \frac{dv}{dy}\right) + \frac{d\rho}{dy}\left(\frac{dv}{dy} - \frac{dv}{dx}\right)\right)\Delta\xi^2 \tag{6.54}$$

with

$$\begin{aligned}
\frac{d\rho}{dx} &= \frac{1}{\Delta\xi}(\rho_3 - \rho_1), \quad & \frac{d\rho}{dy} &= \frac{1}{\Delta\xi}(\rho_5 - \rho_1) \\
\frac{du}{dx} &= \frac{1}{\Delta\xi}(u_3 - u_1), \quad & \frac{du}{dy} &= \frac{1}{\Delta\xi}(u_5 - u_1) \\
\frac{dv}{dx} &= \frac{1}{\Delta\xi}(v_3 - v_1), \quad & \frac{dv}{dy} &= \frac{1}{\Delta\xi}(v_5 - v_1)
\end{aligned} \tag{6.55}$$

where the indices are found in Fig. 6.5. Once the truncation terms are defined fully, the only unknowns in the reconciliation procedure are $\overline{u}_j$ and $\overline{v}_j$ in Eqs. (6.51) and (6.52), which are

then solved for.

$$\overline{u}_j = \frac{1}{\overline{\rho}_j}\left(\overline{\rho u}_j - \Delta\rho u_j\right) \tag{6.56}$$

$$\overline{v}_j = \frac{1}{\overline{\rho}_j}\left(\overline{\rho v}_j - \Delta\rho v_j\right) \tag{6.57}$$

Lastly, the reconciliation procedure concludes with defining the bubble functions.

$$\rho_7 = \frac{20}{9}\left(\overline{\rho}_j - \frac{1}{3}\left(\rho_2 + \rho_4 + \rho_6\right)\right) \tag{6.58}$$

$$u_7 = \frac{20}{9}\left(\overline{u}_j - \frac{1}{3}\left(u_2 + u_4 + u_6\right)\right) \tag{6.59}$$

$$v_7 = \frac{20}{9}\left(\overline{v}_j - \frac{1}{3}\left(v_2 + v_4 + v_6\right)\right) \tag{6.60}$$

In the next iteration of solution, all primitive reconstructions are consistently defined including the bubble coefficients.

$$\mathbf{q}(\boldsymbol{\xi}) = \sum_{i=1}^{7} c_i(\mathbf{q}) \cdot \phi_i(\boldsymbol{\xi}) \tag{6.61}$$

where $\mathbf{q} = (\rho, u, v)^T$.

6.3.2.1 Two-Dimensional Euler System Reconciliation

For the sake of completeness, the reconciliation process for the Euler system is discussed. The Euler system is introduced by adding the conservation of energy equation to the pressureless Euler system.

$$\frac{\partial \rho E}{\partial t} + \nabla \cdot \rho \mathbf{v} H = 0 \tag{6.62}$$

where $E = e + \frac{1}{2}\mathbf{v}\cdot\mathbf{v}$ is the specific total energy and $H = E + \frac{p}{\rho}$ is the specific total enthalpy. Additionally, the equation of state for an ideal gas is $p = (\gamma - 1)\rho e$. Disregarding some details of the Euler system, we present the reconciliation for total energy and pressure.

Conservation of energy is related to pressure variable via the equation of state. Unlike the reconstructions for velocity, the relation between total energy and pressure is not very obvious. Separating the total energy into the internal $\mathcal{I}$ and the kinetic $\mathcal{K}$ part, we can see

that the pressure is related to the total energy as follows.

$$\begin{aligned}\overline{\rho E}_j &= \frac{1}{\Omega_j}\left(\int_{\Omega_j}\mathcal{I}(\boldsymbol{\xi})dV + \int_{\Omega_j}\mathcal{K}(\boldsymbol{\xi})dV\right) \\ &= \frac{1}{\Omega_j}\int_{\Omega_j}\frac{p(\boldsymbol{\xi})}{\gamma-1}dV + \frac{1}{2}\frac{\frac{1}{\Omega_j}\int_{\Omega_j}\left[\left(\rho(\boldsymbol{\xi})u(\boldsymbol{\xi})\right)^2 + \left(\rho(\boldsymbol{\xi})v(\boldsymbol{\xi})\right)^2\right]dV}{\frac{1}{\Omega_j}\int_{\Omega_j}\rho(\boldsymbol{\xi})dV} + \Delta\rho E_j \\ &= \frac{\overline{p}_j}{\gamma-1} + \frac{1}{2}\frac{\left(\overline{\rho u}_j\right)^2 + \left(\overline{\rho v}_j\right)^2}{\overline{\rho}_j} + \Delta\rho E_j \end{aligned} \tag{6.63}$$

where $\Delta\rho E_j$ is the truncation error in total energy. The same analysis procedure as pressureless Euler system is applied to find the truncation error in energy. The total energy truncation error is

$$\Delta\rho E_j = \frac{1}{36}\rho_1\left(\left(\frac{du}{dx}-\frac{du}{dy}\right)^2 + \frac{du}{dx}\frac{du}{dy} + \left(\frac{dv}{dx}-\frac{dv}{dy}\right)^2 + \frac{dv}{dx}\frac{dv}{dy}\right)\Delta\xi^2. \tag{6.64}$$

With derivative approximations from Eq. (6.55), the discrete relation is approximated to $\mathcal{O}(\Delta\xi^4)$. The pressure reconstructed average is found by rearranging the relation Eq. (6.63).

$$\overline{p}_j = (\gamma-1)\left(\overline{\rho E}_j - \frac{1}{2}\frac{(\overline{\rho u}_j)^2 + (\overline{\rho v}_j)^2}{\overline{\rho}_j} - \Delta\rho E_j\right) \tag{6.65}$$

The bubble function for pressure p_7 is defined finally as follows.

$$p_7 = \frac{20}{9}\left(\overline{p}_j - \frac{1}{3}(p_2+p_4+p_6)\right) \tag{6.66}$$

In the next iteration of solution, all primitive reconstructions are consistently defined, including the bubble coefficients. Bubble functions for other primitive reconstructions are shown in Eqs. (6.58), (6.59), and (6.60).

$$\mathbf{q}(\boldsymbol{\xi}) = \sum_{i=1}^{7} c_i(\mathbf{q})\cdot\phi_i(\boldsymbol{\xi}) \tag{6.67}$$

where $\mathbf{q} = (\rho, u, v, p)^T$.

6.4 Numerical Accuracy of One-Dimensional Pressureless Euler

Errors in conservative variables are measured by the $L_2(\overline{\mathbf{u}})$ norm definition. It is formally used for error measurements in finite volume methods.

$$L_2(\overline{\mathbf{u}}) = \left\{ \frac{1}{\Omega} \sum_{j=1}^{N} \left| \overline{\mathbf{u}}_j \Omega_j - \overline{\mathbf{u}}_j^{\text{exact}} \Omega_j \right|^2 \right\}^{1/2} \tag{6.68}$$

where $\overline{\mathbf{u}} = (\overline{\rho}, \overline{\rho u}, \overline{\rho v})^T$.

Errors in solution reconstruction variables are measured by the $L_2(\mathbf{q}(\mathbf{x}))$ norm definition. This error norm measure provides an additional perspective in quantifying the solution quality such as the property of continuous reconstruction.

$$L_2(\mathbf{q}(\mathbf{x})) = \left\{ \frac{1}{\Omega} \sum_{j=1}^{N} \int_{\Omega_j} \left| \mathbf{q}_j(\mathbf{x}) - \mathbf{q}_j^{\text{exact}}(\mathbf{x}) \right|^2 dV \right\}^{1/2} \tag{6.69}$$

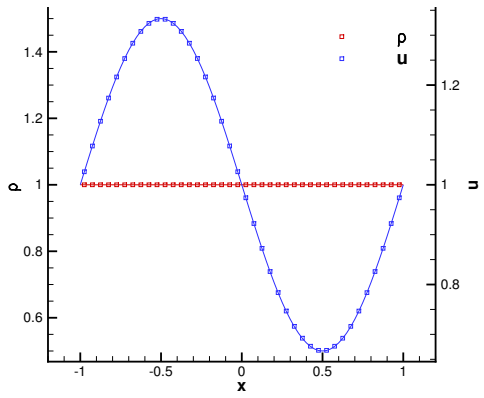
where $\mathbf{q} = (\rho, u, v)^T$.

A sine wave initial velocity condition is considered.

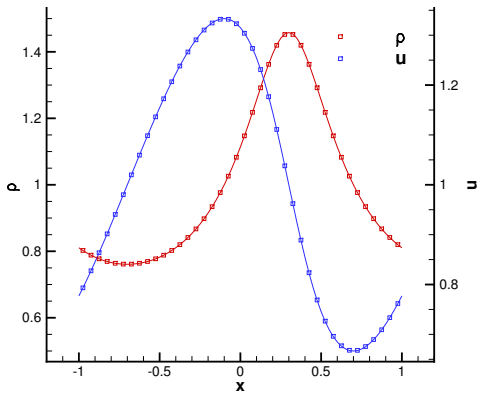
$$\rho_0(x) = 1 \tag{6.70}$$

$$u_0(x) = 1 + \frac{1}{3} \sin\left(\pi(x+1)\right) \tag{6.71}$$

where $x \in [-1, 1]$. Periodic boundary conditions are applied on boundaries. Initially constant density develops into concentrations of high and low regions, Fig. 6.6. This illustrates nonlinear system coupling of density and velocity.



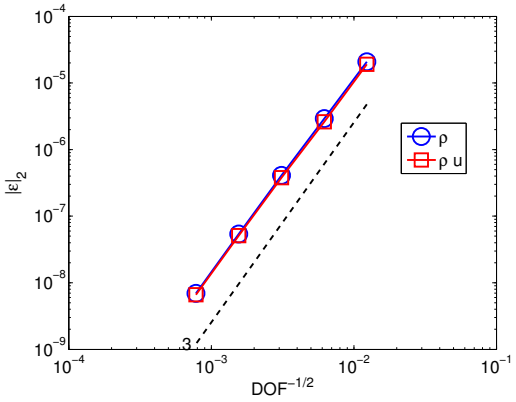
(a) Initial conditions



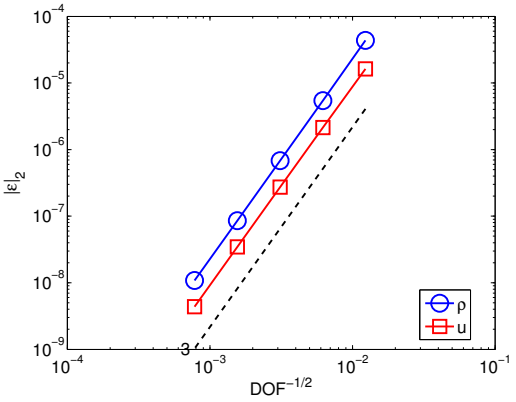
(b) Solutions at $t = 0.3$

Figure 6.6: One-dimensional pressureless Euler problem, Eqs. (6.70) and (6.71), with $N = 40$ and $\nu = 0.7$

In order to examine the numerical accuracy as well as the consistency of reconciliation process, numerical errors in both conservative variables and nonconservative reconstructions are analyzed. In Fig. 6.7, the numerical errors of conservative and primitive variables are shown. Errors in all variables converge at the designed rate $\mathcal{O}(h^3)$.



(a) $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables



(b) $L_2(\mathbf{q})$ norm errors of solution reconstructions

Figure 6.7: One-dimensional pressureless Euler problem, Eqs. (6.70) and (6.71), error convergence results

6.5 Numerical Accuracy of Two-Dimensional Pressureless Euler

For the two-dimensional pressureless Euler equations, we consider two numerical accuracy tests. These test problems are not representative of any physical system. In fact, these are purely designed for numerical error convergence tests. There are no known analytic tests with significant complexity for the pressureless Euler system of two or higher dimensions. There are some simple analytic tests that are quadratic in nature [97], but these serve no purpose for us because the truncation errors of AF are of $\mathcal{O}(h^3)$. For this reason, an iterative method is sought to find the solution at the final time.

6.5.1 Test 1

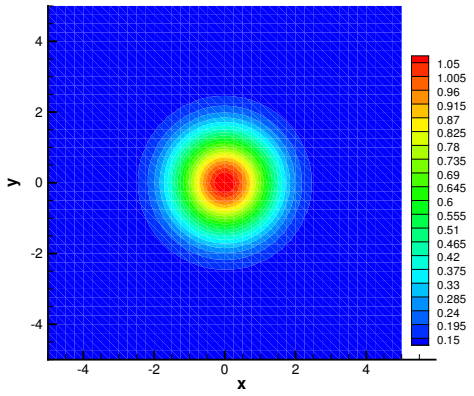
Initial conditions are provided for test problem 1.

$$\rho_0(x,y) = 0.1 + \exp\left(-0.5\left(x^2+y^2\right)\right) \tag{6.72}$$

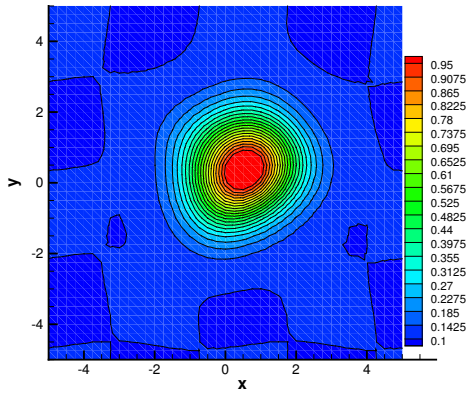
$$u_0(x,y) = \frac{1}{3}\cos\left(\frac{2\pi}{5}y\right) \tag{6.73}$$

$$v_0(x,y) = \frac{1}{3}\sin\left(\frac{2\pi}{5}x\right) \tag{6.74}$$

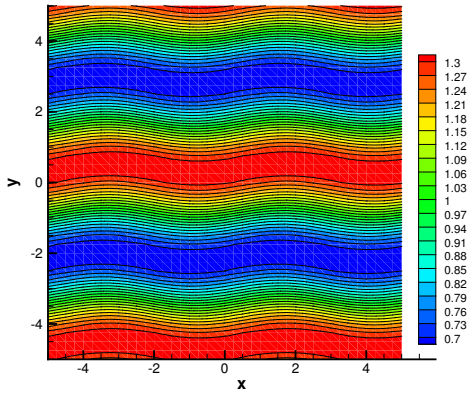
An initially radially symmetric Gaussian density profile is distorted due to a sinusoidally varying velocity in both directions, see Fig. 6.8.



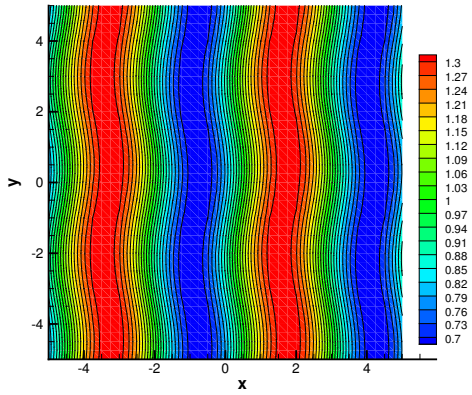
(a) Density at $t = 0$



(b) Density at $t = 0.4$



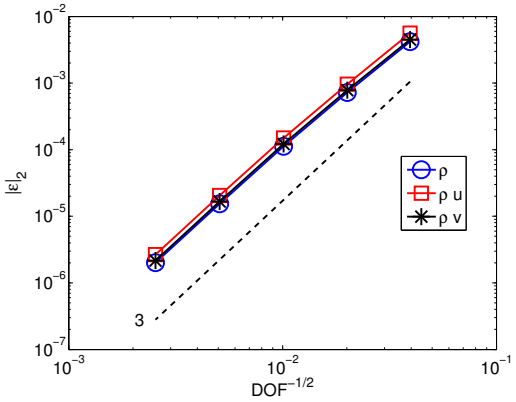
(c) u velocity at $t = 0.4$



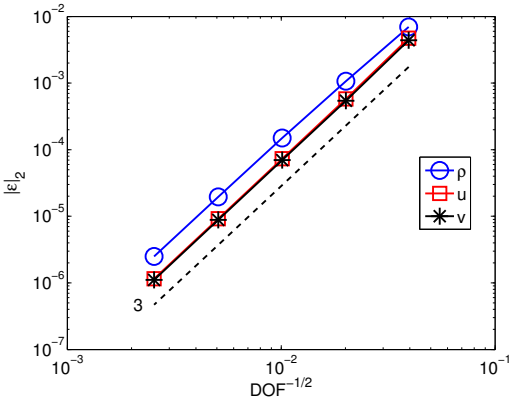
(d) v velocity at $t = 0.4$

Figure 6.8: Test problem 1 solutions on triangular grid B at $\nu = 0.7$

Numerical errors in conservative variables and primitive solution reconstructions are shown in Fig. 6.9 for the triangular grid B type. Errors of all variables converge at the designed rate $\mathcal{O}(h^3)$.



(a) $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables

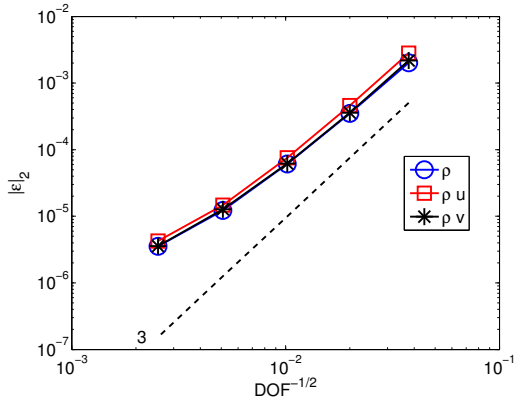


(b) $L_2(\mathbf{q}(\mathbf{x}))$ norm errors of solution reconstructions

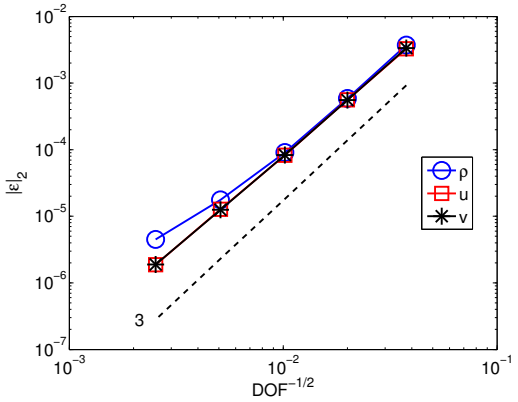
Figure 6.9: Test problem 1 error convergence results on triangular grid B at $t = 0.4$ and $\nu = 0.7$

Now, numerical errors of conservative variables and primitive solution reconstructions are shown in Fig. 6.10 for the unstructured grid type. However, unlike for triangular grid B, the errors now converge at a rate lower than $\mathcal{O}(h^3)$. The $L_2(\mathbf{q}(\mathbf{x}))$ norm error measure in primitive reconstructions, Fig. 6.10b, display individual error convergence rates. The two velocity reconstruction error norms show convergence rates of $\mathcal{O}(h^{2.7})$ whereas the density reconstruction error convergence shows the rate of $\mathcal{O}(h^2)$. And the conservative variable error convergence rates are $\mathcal{O}(h^2)$ due to the density rate shown in Fig. 6.10a.

There are two possible sources for this behavior. The first is related to the mesh alignment problem discussed in Sections 3.3.7 and 3.4. In general, nonlinear advection problems show varying velocity directions which make alignment with mesh geometry much more likely. The second is more numerically oriented. In pressureless Euler system, density incurs more error than velocity because density is modified to seek the conservation of mass via streamtube area change correction discussed in Section 6.1.2. The difference between density and velocity error levels is not big. However, as the mesh size reduces close to $h \to 0$, the truncation errors from two different sources start to show noticeable effects on the density convergence.



(a) $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables



(b) $L_2(\mathbf{q}(\mathbf{x}))$ norm errors of solution reconstructions

Figure 6.10: Test problem 1 error convergence results on unstructured grid at $t = 0.4$ and $\nu = 0.7$

6.5.2 Test 2

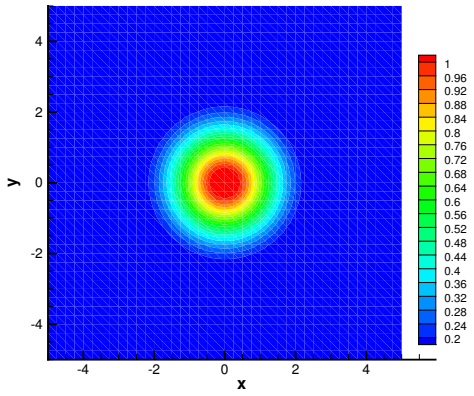
The initial conditions are provided for test problem 2.

$$\rho_0(x, y) = 0.1 + \exp\left(-0.5\left(x^2 + y^2\right)\right) \tag{6.75}$$

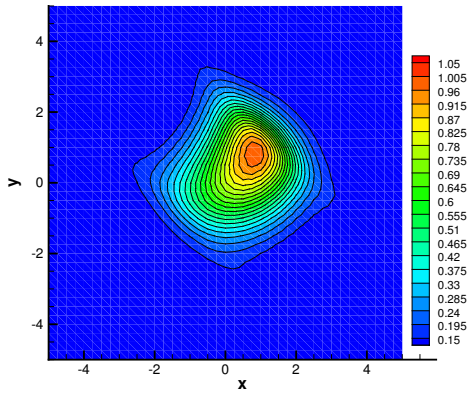
$$u_0(x, y) = \frac{1}{3}\cos\left(\frac{2\pi}{5}(x + y)\right) \tag{6.76}$$

$$v_0(x, y) = \frac{1}{3}\sin\left(\frac{2\pi}{5}(x + y)\right) \tag{6.77}$$

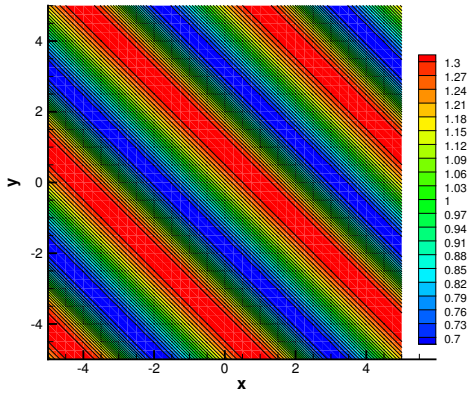
The initially radially symmetric Gaussian density profile is distorted due to the sinusoidally varying velocity in the $x + y$ direction, see Fig. 6.11.



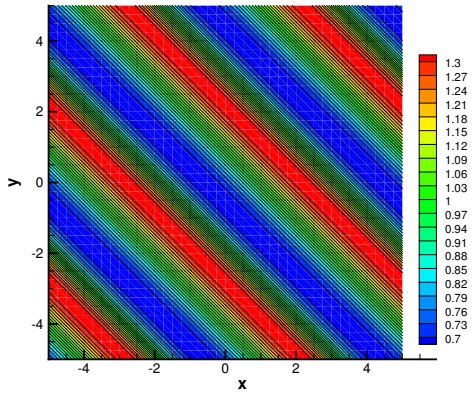
(a) Density at $t = 0$



(b) Density at $t = 0.4$



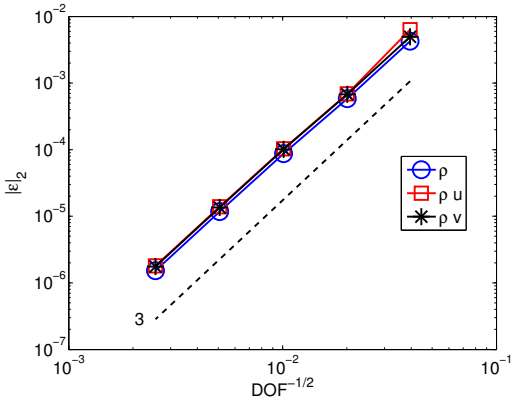
(c) u velocity at $t = 0.4$



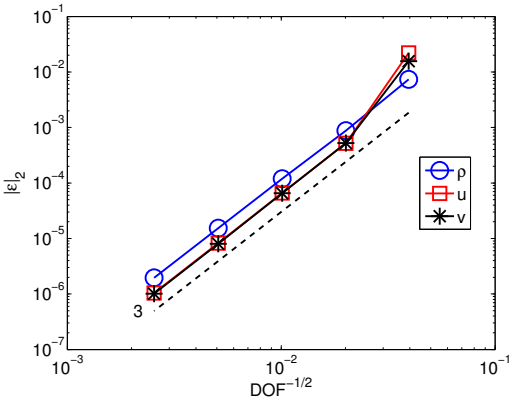
(d) v velocity at $t = 0.4$

Figure 6.11: Test problem 2 solutions on triangular grid B at $\nu = 0.7$

The numerical errors in the conservative variables and primitive solution reconstructions are shown in Fig. 6.12 for triangular grid B type. All errors converge at the designed rate $\mathcal{O}(h^3)$.



(a) $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables

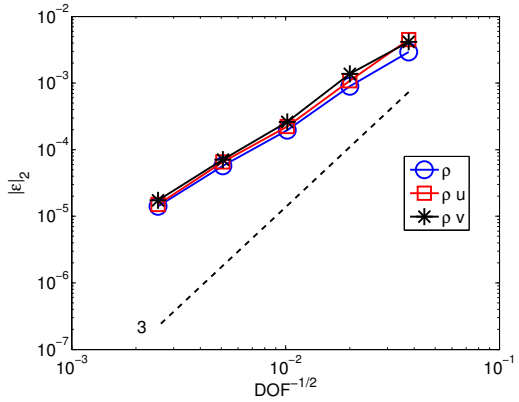


(b) $L_2(\mathbf{q}(\mathbf{x}))$ norm errors of solution reconstructions

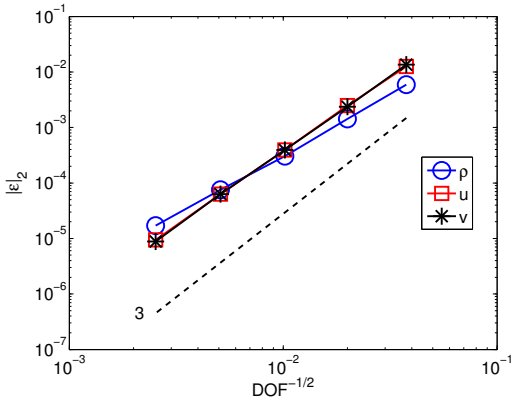
Figure 6.12: Test problem 2 error convergence results on triangular grid B at $t = 0.4$ and $\nu = 0.7$

The numerical errors of conservative and primitive variables are shown in Fig. 6.13 for unstructured grid type. However, unlike for triangular grid B, the errors now converge at a rate lower than $\mathcal{O}(h^3)$. The two velocity reconstruction error convergence shows the rate of $\mathcal{O}(h^{2.7})$ whereas the density reconstruction error convergence shows the rate of $\mathcal{O}(h^2)$. And the error norms of conservative variables shown in Fig. 6.10a all converge at $\mathcal{O}(h^2)$ due to density.

There are two possible sources for this behavior. The first is related to the mesh alignment problem discussed in Sections 3.3.7 and 3.4. In general, nonlinear advection problems show varying velocity directions which make alignment with mesh geometry much more likely. The second is more numerically oriented. In pressureless Euler system, density incurs more error than velocity because density is modified to seek the conservation of mass via streamtube area change correction discussed in Section 6.1.2. The difference between density and velocity error levels is not big. However, as the mesh size reduces close to $h \to 0$, the truncation errors from two different sources start to show noticeable effects on the density convergence.



(a) $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables



(b) $L_2(\mathbf{q}(\mathbf{x}))$ norm errors of solution reconstructions

Figure 6.13: Test problem 2 error convergence results on triangular unstructured grid at $t = 0.4$ and $\nu = 0.7$.

6.6 Numerical Comparison of Two-Dimensional Pressureless Euler

Previously studied test problems in Section 6.5 are compared to two DG schemes. DG schemes considered in this comparison are time explicit Runge-Kutta DG schemes. For the spatial discretization, a linear ($p = 1$) and a quadratic ($p = 2$) basis functions are used, and are respectively referred to as DG1 and DG2. Two- and three-stage RK methods are used for DG1 and DG2, respectively. For numerical comparisons made here, only the triangular structured grid B and the unstructured grid types are considered. For DG schemes, the CFL stability is restricted by

$$\nu < \frac{1}{2p+1} \tag{6.78}$$

where p is the spatial discretization order [14].

We have previously devised a common measure called the work unit (WU) in Eq. (3.86). The parameters for evaluating the WU are listed in Table 3.7 for both schemes. Even though the WU provides a theoretical standard to compare the efficiency of different schemes, it makes no allowance for the arithmetical complexity of schemes. For a system of equations, the mathematical complexity of DG is largely unchanged from that of scalar conservation laws. Conservation systems are formulated in the weak form, leading to integrations along element boundaries and inside elements. Conservative fluxes across boundaries are evaluated from upwind-biased data. Many numerical upwind flux formulations are available. In the pressureless Euler equations, the exact Godunov's flux is used. Since the way in which

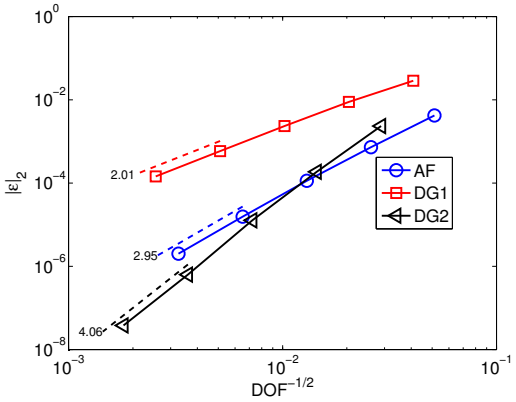
conservation systems are solved is largely unchanged from applications to scalar conservation laws, the DG method creates a mathematically streamlined framework for systems.

Unlike DG, the AF method takes a two stage approach. The first stage, the nonconservative stage, a semi-Lagrangian method is used to update the variables in advection systems but other equivalent methods can be used. And the conservation form of the system is solved in the second stage. Because two stages solve for different variables, the reconciliation between the conservative and reconstruction variables takes place at the end of the conservation stage. The flexible design aspect of the AF nonconservative stage provides many choices for improvements. Given the perspectives of each method, we understand the WU is neither the best nor the universal standard, but it can be understood as a *good* measure of efficiency with regard to the theoretical construction of each scheme.

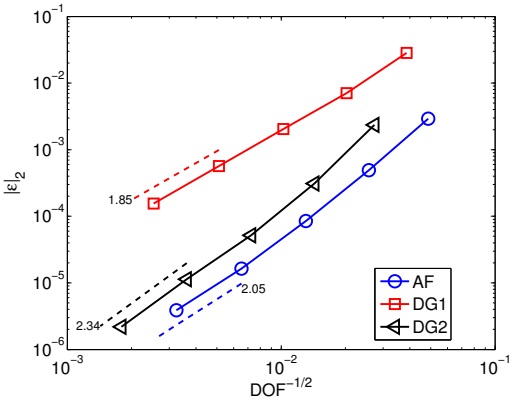
6.6.1 Test 1

We first compare the $L_2(\overline{\mathbf{u}})$ norm errors in Fig. 6.14. Since x and y momentum errors are quite similar, only the x momentum errors are shown. As previously discussed, the two-dimensional pressureless Euler results for test problem 1 exhibit convergence rates of $\mathcal{O}(h^3)$ and $\mathcal{O}(h^2)$ in conservative variable errors for the triangular structured grid B and the unstructured grid type, respectively. The unstructured grid type shows less than normal convergence rates due to the mesh alignment problem associated with the streamline tracing method.

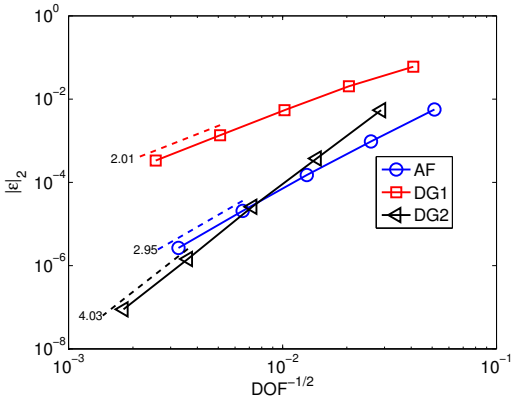
DG2 results are comparable to AF in this nonlinear problem. These DG2 results on the triangular structured grid B, Figs. 6.14a and 6.14c, achieve a far better error convergence rate, $\mathcal{O}(h^4)$, than for the unstructured grid counterpart, $\mathcal{O}(h^{2.3})$. This behavior is similar to the AF scheme, however, seems to stem from a different reason. The triangular structured grid B provides a nearly one-dimensional problem for some interface edges, as a result, the error convergence becomes superior to the unstructured grid case.



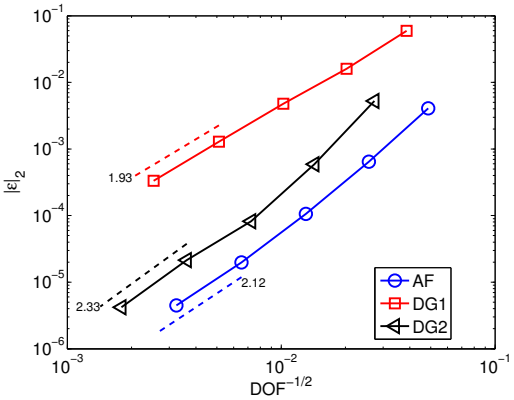
(a) Density error $L_2(\overline{\rho})$ convergence on triangular structured grid B



(b) Density error $L_2(\overline{\rho})$ convergence on triangular unstructured grid



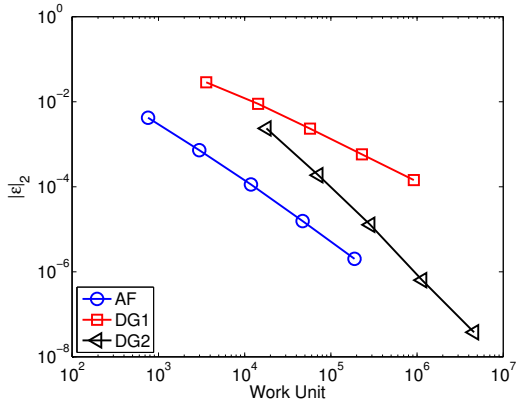
(c) x momentum error $L_2(\overline{\rho u})$ convergence on triangular structured grid B



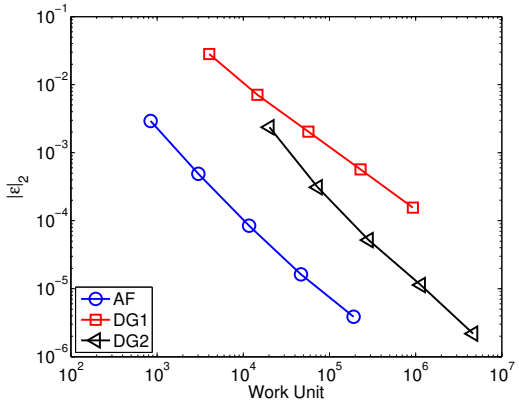
(d) x momentum error $L_2(\overline{\rho u})$ convergence on triangular unstructured grid

Figure 6.14: Test problem 1 error convergence results at $t = 0.4$. Courant numbers for AF, DG1, and DG2 are respectively $\nu = 0.7$, $\nu = 0.3$ and, $\nu = 0.2$.

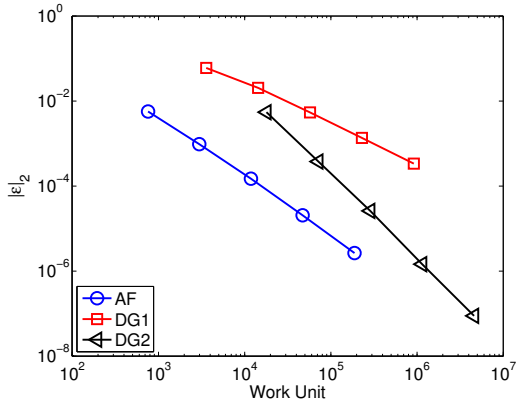
The L_2 norm errors in conservative variables for all schemes are compared against WU in Fig. 6.15. The results indicate that AF starts to see some advantage over all DG schemes. The AF scheme shows a superior efficiency property in both the density and x momentum errors. The error in y momentum is very similar to the error in x momentum. To reach the same level of error on the unstructured grid, the AF scheme would take around one order of magnitude fewer work units than the DG2 counterpart, see Tables 6.1 and 6.3.



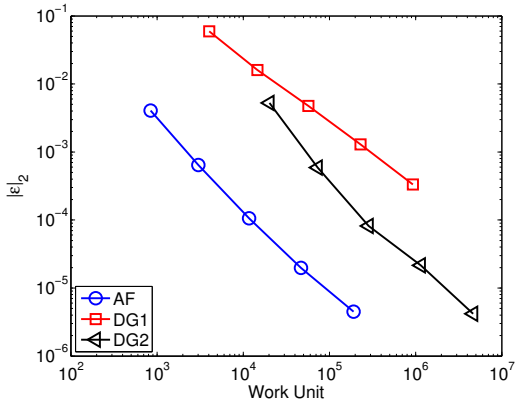
(a) Density error $L_2(\overline{\rho})$ convergence on triangular structured grid B



(b) Density error $L_2(\overline{\rho})$ convergence on triangular unstructured grid



(c) x momentum error $L_2(\overline{\rho u})$ convergence on triangular structured grid B



(d) x momentum error $L_2(\overline{\rho u})$ convergence on triangular unstructured grid

Figure 6.15: Test problem 1 error convergence results measured as a function of WU at $t = 0.4$. Courant numbers for AF, DG1, and DG2 are respectively $\nu = 0.7$, $\nu = 0.3$, and $\nu = 0.2$.

Table 6.1: Test problem 1 $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables and WU for AF on unstructured grid at $t = 0.4$ and $\nu = 0.7$

DOF	Work Unit	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order
4.21e+02	8.42e+02	2.92e-03		4.06e-03	
1.51e+03	3.03e+03	4.90e-04	2.79	6.44e-04	2.88
5.86e+03	1.17e+04	8.46e-05	2.59	1.06e-04	2.67
2.34e+04	4.69e+04	1.63e-05	2.38	1.98e-05	2.42
9.52e+04	1.90e+05	3.88e-06	2.05	4.49e-06	2.12

Table 6.2: Test problem 1 $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables and WU for DG1 on unstructured grid at $t = 0.4$ and $\nu = 0.3$

DOF	Work Unit	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order
6.72e+02	4.03e+03	2.83e-02		4.05e-02	
2.44e+03	1.47e+04	7.08e-03	2.15	1.06e-02	2.08
9.53e+03	5.72e+04	2.04e-03	1.83	2.99e-03	1.86
3.82e+04	2.29e+05	5.68e-04	1.84	8.05e-04	1.89
1.56e+05	9.33e+05	1.55e-04	1.85	2.10e-04	1.92

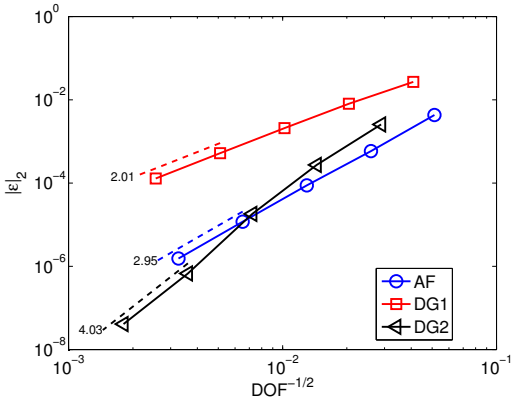
Table 6.3: Test problem 1 $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables and WU for DG2 on unstructured grid at $t = 0.4$ and $\nu = 0.2$

DOF	Work Unit	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order
1.34e+03	2.02e+04	2.36e-03		3.39e-03	
4.88e+03	7.33e+04	3.09e-04	3.15	4.46e-04	3.14
1.91e+04	2.86e+05	5.18e-05	2.63	6.18e-05	2.90
7.64e+04	1.15e+06	1.13e-05	2.19	1.36e-05	2.18
3.11e+05	4.67e+06	2.19e-06	2.34	2.61e-06	2.35

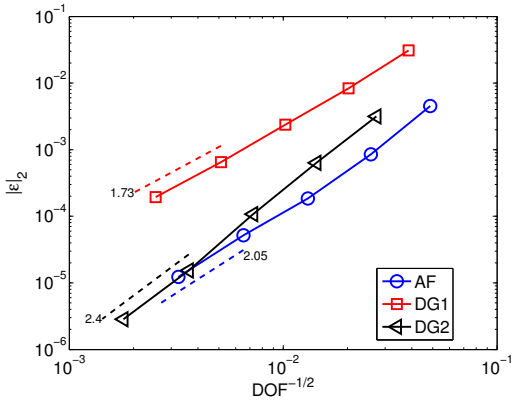
6.6.2 Test 2

The $L_2(\overline{\mathbf{u}})$ norm errors in conservative variables are shown in Fig. 6.16 for all schemes. Since the x and y momentum errors are quite similar, only the x momentum errors are shown. As previously discussed, the AF method convergence rates are $\mathcal{O}(h^3)$ and $\mathcal{O}(h^2)$ for the triangular structured grid B and unstructured grid types, respectively. The unstructured grid results in a less than normal convergence rates due to the mesh alignment problem associated with the streamline tracing method.

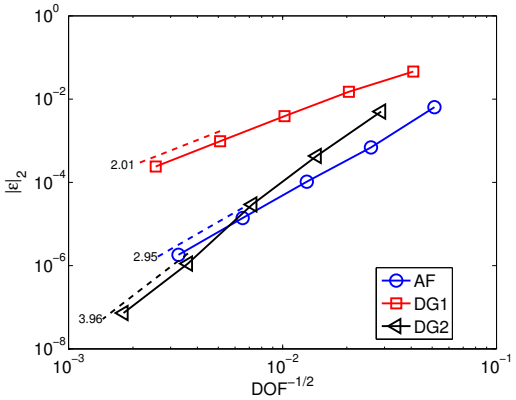
DG2 results are comparable to the AF results in this nonlinear problem. The DG2 results on the triangular structured grid B, Figs. 6.16a and 6.16c, achieve a far better error convergence rate, $\mathcal{O}(h^{4.3})$, than the unstructured grid counterpart, $\mathcal{O}(h^{2.4})$. This behavior is similar to the AF scheme, however, seems to stem from a different reason. The triangular structured grid B provides a nearly one-dimensional problem for some interface edges, as a result, the error convergence becomes superior to the result on the unstructured grid.



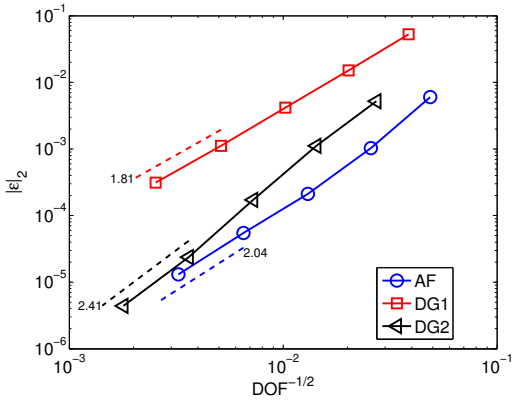
(a) Density error $L_2(\overline{\rho})$ convergence on triangular structured grid B



(b) Density error $L_2(\overline{\rho})$ convergence on triangular unstructured grid



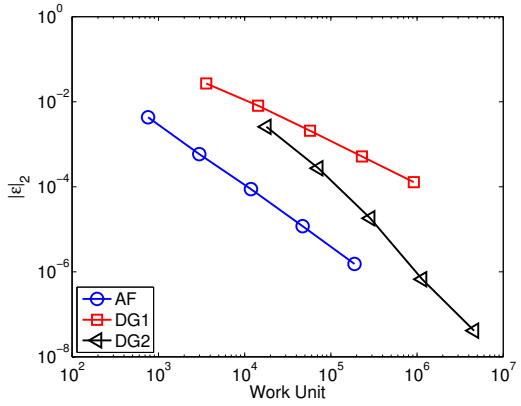
(c) x momentum error $L_2(\overline{\rho u})$ convergence on triangular structured grid B



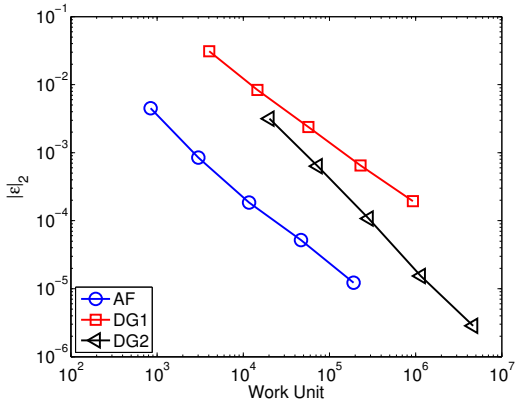
(d) x momentum error $L_2(\overline{\rho u})$ convergence on triangular unstructured grid

Figure 6.16: Test problem 2 error convergence results at $t = 0.4$. Courant numbers for AF, DG1, and DG2 are respectively $\nu = 0.7$, $\nu = 0.3$ and, $\nu = 0.2$.

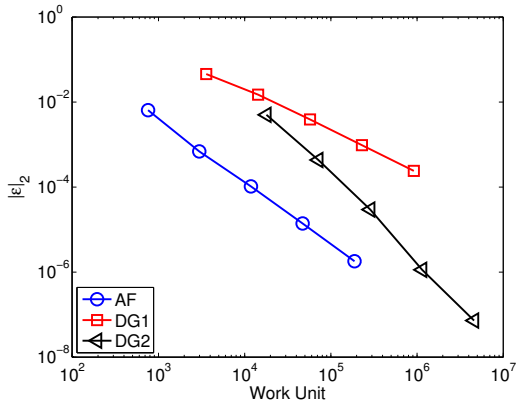
The L_2 norm of conservative variable errors for all schemes are compared against WU in Fig. 6.17. The results indicate that AF starts to see some advantage over all DG schemes. The AF scheme show superior efficiency property in both the density and x momentum errors. The error in y momentum is very similar to the error in x momentum. To reach the same level of error on the unstructured grid, the AF scheme would take a bit shy of one order of magnitude fewer WU than the DG2 counterpart, see Tables 6.4 and 6.6.



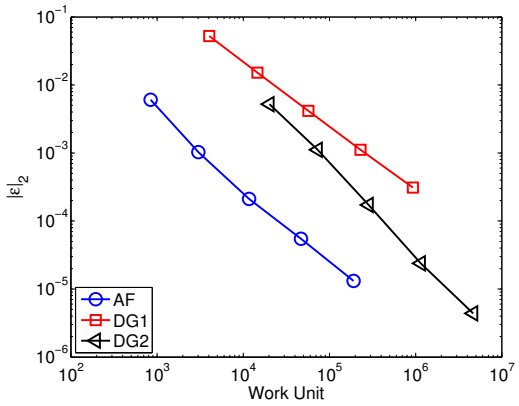
(a) Density error $L_2(\overline{\rho})$ convergence on triangular structured grid B



(b) Density error $L_2(\overline{\rho})$ convergence on triangular unstructured grid



(c) x momentum error $L_2(\overline{\rho u})$ convergence on triangular structured grid B



(d) x momentum error $L_2(\overline{\rho u})$ convergence on triangular unstructured grid

Figure 6.17: Test problem 2 error convergence results measured as a function of WU at $t = 0.4$. Courant numbers for AF, DG1, and DG2 are respectively $\nu = 0.7$, $\nu = 0.3$, and $\nu = 0.2$.

Table 6.4: Test problem 2 $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables and WU for AF on unstructured grid at $t = 0.4$ and $\nu = 0.7$

DOF	Work Unit	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order
4.21e+02	8.42e+02	4.51e-03		6.04e-03	
1.51e+03	3.03e+03	8.48e-04	2.61	1.03e-03	2.77
5.86e+03	1.17e+04	1.85e-04	2.25	2.11e-04	2.34
2.34e+04	4.69e+04	5.17e-05	1.84	5.49e-05	1.95
9.52e+04	1.90e+05	1.23e-05	2.05	1.31e-05	2.04

Table 6.5: Test problem 2 $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables and WU for DG1 on unstructured grid at $t = 0.4$ and $\nu = 0.3$

DOF	Work Unit	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order
6.72e+02	4.03e+03	3.10e-02		4.26e-02	
2.44e+03	1.47e+04	8.36e-03	2.03	1.09e-02	2.12
9.53e+03	5.72e+04	2.38e-03	1.85	2.97e-03	1.90
3.82e+04	2.29e+05	6.50e-04	1.87	7.92e-04	1.90
1.56e+05	9.33e+05	1.94e-04	1.73	2.24e-04	1.80

Table 6.6: Test problem 2 $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables and WU for DG2 on unstructured grid at $t = 0.4$ and $\nu = 0.2$

DOF	Work Unit	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order
1.34e+03	2.02e+04	3.16e-03		4.55e-03	
4.88e+03	7.33e+04	6.35e-04	2.49	8.13e-04	2.67
1.91e+04	2.86e+05	1.08e-04	2.60	1.25e-04	2.75
7.64e+04	1.15e+06	1.54e-05	2.81	1.74e-05	2.84
3.11e+05	4.67e+06	2.84e-06	2.40	3.15e-06	2.44

6.7 Complex Problem

To demonstrate a dynamic effect of the pressureless Euler system, we consider a test problem with the complex rotating velocity field. In fact, this can be regarded as a vortex problem without the presence of pressure.

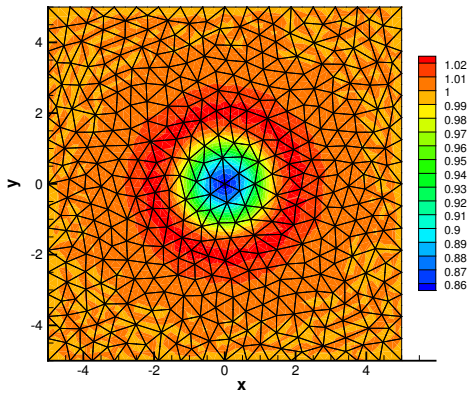
$$\rho_0(x, y) = 0.1 \tag{6.79}$$

$$u_0(x, y) = -\frac{10y}{2\pi} \exp\left(0.25(1 - r^2)\right) \tag{6.80}$$

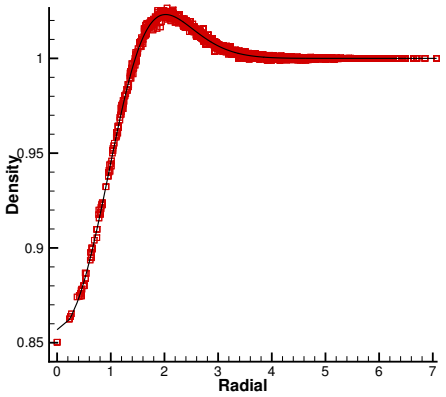
$$v_0(x, y) = \frac{10x}{2\pi} \exp\left(0.25(1 - r^2)\right) \tag{6.81}$$

where $r = \sqrt{x^2 + y^2}$.

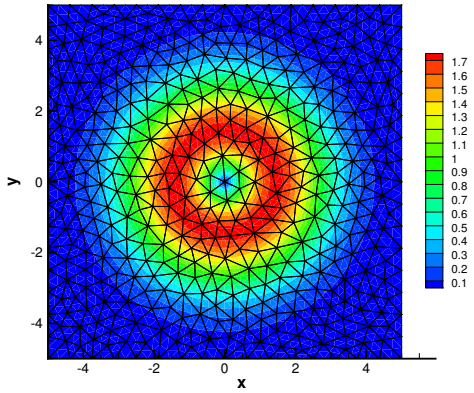
The complex rotating nature of the velocity field creates a radially symmetric accumulating density front at final time, Fig. 6.18. A shock created by the pressureless Euler system develops into what is known as a 'delta shock' due to the lack of diffusive mechanisms and pressure. As the focus of considering the pressureless Euler system is to understand a system of conservation laws, we do not consider problems containing shocks here.



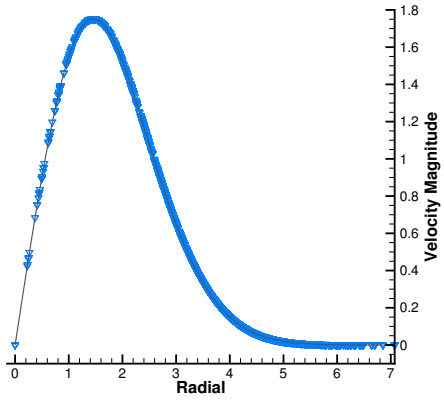
(a) Density contour solution. Unstructured grid is used.



(b) Density solution in radial coordinate. Black solid line indicates iteratively obtained 'exact' solution.



(c) Velocity magnitude contour solution. Unstructured grid is used.



(d) Velocity magnitude solution in radial coordinate. Black solid line indicates iteratively obtained 'exact' solution.

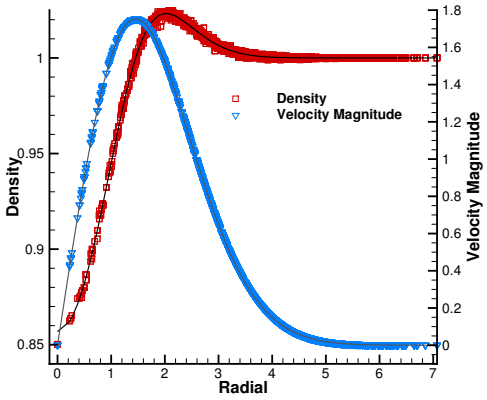
Figure 6.18: Rotating velocity test problem on triangular unstructured grid at $t = 0.2$ and $\nu = 0.7$

The velocity and density relation are well illustrated in Fig. 6.19a. As an initially constant density is exposed to a radially outward moving velocity field, a trough and a peak in density are produced at locations corresponding to low and high velocity magnitudes. It is evident from the relative error measurements in Fig. 6.19b that the density solution exhibits a behavior that is more oscillatory than the velocity solution. A relative error is defined as $|q_j - q_j^{\text{exact}}|$ where q_j^{exact} is an iteratively obtained exact solution.

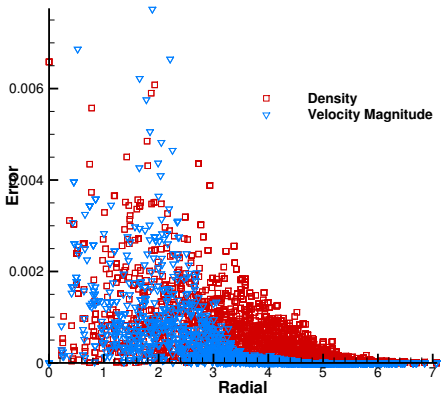
Now, possible sources of density solution errors are discussed. They have been previously discussed, but are reiterated for clarity. The first is related to the mesh alignment

problem discussed in Sections 3.3.7 and 3.4. In general, nonlinear advection problems show varying velocity directions which make alignment with mesh geometry much more likely. The second source of error, and the most obvious one, is from the numerical implementation aspect. The density solution obtains the correction from the velocity field. This suggests that the error in density will inherit some of the truncation error of velocity. This has been rather empirically observed for the pressureless Euler test problems and in Appendix A that the error norm level of velocity components is lower than the error norm level of density after density has been numerically treated with the streamtube area changing effect.

The advective problems shown in Chapter 3 for linear advection, Chapter 4 for nonlinear scalar advection, and in the current chapter for nonlinear advection systems display some signs of mesh alignment problems, as manifested in error convergence results. We have established that these are due to the lack of diffusive mechanisms in advective numerical schemes. In the next Chapter 7, we will be able to check out whether adding acoustic disturbances, which provide a process similar to an averaging process, makes any difference on the outcome.



(a) Density and velocity magnitude solution. Solid lines are numerically iterated exact solutions.



(b) Density and velocity magnitude relative error

Figure 6.19: Density and velocity magnitude solutions plotted in radial coordinate at $t = 0.2$ and $\nu = 0.7$

CHAPTER 7

Towards Multidimensional Nonlinear System

The AF method for multidimensional nonlinear systems is introduced. In particular, the method for the Euler equations is developed but can also be extended to solve other conservation laws. We will decompose the Euler system into advective and acoustic processes and develop different numerical methods. Although there is no computationally significant difference between these two processes in a one-dimensional system, they exhibit vastly different physical properties in multidimensions. We employ an individually appropriate nonlinear scheme for each process. The advective process is numerically modeled by the streamline tracing method developed in Chapter 6. The acoustic process is accomplished by a numerical method that is an extension of the Poisson's integral solution to the initial value problem for scalar wave equations, for which some discussion is presented in this chapter. Physical variables that are individually evolved by two distinct processes are coupled using a new nonlinear operator splitting method. Since each numerical process is designed specifically to satisfy the respective stability associated with the type of physics, the nonlinear operator splitting method is maximally stable as the CFL criterion allows it to be. In addition, the operator splitting method returns methods for linear problems when linear problems are considered.

7.1 Multidimensional Euler Equations

The unsteady multidimensional Euler equations for ideal compressible flows are presented in a conservative vector form.

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{v} = 0 \tag{7.1}$$

$$\frac{\partial \rho \mathbf{v}}{\partial t} + \nabla \cdot (\rho \mathbf{v} \otimes \mathbf{v} + \mathbf{I}p) = 0 \tag{7.2}$$

$$\frac{\partial \rho E}{\partial t} + \nabla \cdot \rho \mathbf{v} H = 0 \tag{7.3}$$

where $E = e + \frac{1}{2}\mathbf{v} \cdot \mathbf{v}$ and $H = E + \frac{p}{\rho}$ are the specific total energy and specific total enthalpy. The equation of state is

$$p = \rho e(\gamma - 1) \tag{7.4}$$

where γ is the specific heat ratio and e is the specific internal energy.

7.1.1 Nonconservative System

A nonconservative system bears a larger importance in AF for a sensible numerical discretization. A Lax-Wendroff type discretization is applied to the physical disturbance systems that are easily distinguished in the nonconservative system form. Furthermore, this provides an intuitive understanding of how each physical disturbance evolves in time.

The nonconservative Euler system is shown below.

$$\frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho + \rho \nabla \cdot \mathbf{v} = 0 \tag{7.5}$$

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} + \frac{1}{\rho} \nabla p = 0 \tag{7.6}$$

$$\frac{\partial p}{\partial t} + \mathbf{v} \cdot \nabla p + \gamma p \nabla \cdot \mathbf{v} = 0 \tag{7.7}$$

where $\mathbf{v} = (u, v)^T$ is a velocity vector. Partial differential equations of density, velocity, and pressure exhibit two distinct processes: advection and acoustics. A compact notation of the nonconservative Euler system is introduced.

$$\frac{\partial \mathbf{q}}{\partial t} + \mathcal{K}\mathbf{q} + \mathcal{H}\mathbf{q} = 0 \tag{7.8}$$

where the nonconservative vector is defined as

$$\mathbf{q} = \begin{pmatrix} \rho \\ u \\ v \\ p \end{pmatrix} \tag{7.9}$$

and the advection and acoustics operator matrices are respectively defined as

$$\mathcal{K} = k \begin{pmatrix} \mathbf{v}\cdot\nabla() & 0 & 0 & 0 \\ 0 & \mathbf{v}\cdot\nabla() & 0 & 0 \\ 0 & 0 & \mathbf{v}\cdot\nabla() & 0 \\ 0 & 0 & 0 & \mathbf{v}\cdot\nabla() \end{pmatrix} \tag{7.10}$$

and

$$\mathcal{H} = h \begin{pmatrix} 0 & \rho\frac{\partial()}{\partial x} & \rho\frac{\partial()}{\partial y} & 0 \\ 0 & 0 & 0 & \frac{1}{\rho}\frac{\partial()}{\partial x} \\ 0 & 0 & 0 & \frac{1}{\rho}\frac{\partial()}{\partial y} \\ 0 & \gamma p\frac{\partial()}{\partial x} & \gamma p\frac{\partial()}{\partial y} & 0 \end{pmatrix}. \tag{7.11}$$

The coefficients k and h indicate the terms associated with the advective and acoustics processes, respectively. In case $k = 1$ and $h = 0$, the nonlinear advection system is recovered, and vice versa. These coefficients also serve as a mark to denote which terms belong to the advective or acoustic process.

In addition, we introduce entropy which is a measure of disorder in a system.

$$s = c_v \log(p/\rho^\gamma) + C \tag{7.12}$$

Under the isentropic condition, the speed of sound is defined as follows.

$$\left(\frac{\partial p}{\partial \rho}\right)\bigg|_s \equiv a^2 = \frac{\gamma p}{\rho} \tag{7.13}$$

The entropy equation is derived from Eqs. (7.12) and (7.13).

$$\frac{\partial s}{\partial t} + \mathbf{v}\cdot\nabla s = 0 \tag{7.14}$$

From this, we conclude that the entropy equation is a passive scalar equation, and so long as there aren't any entropy generating mechanisms entropy should remain constant. It also serves as a way to measure an imbalance between two numerical methods that make up the Euler system.

7.2 Constructing Nonlinear Advection & Acoustics Systems

We have established that the nonconservative system reveals two physical operators that are only linearly coupled. We introduce numerically accurate nonlinear systems that reproduce individual physical processes. Each nonlinear system is a complete procedure on its own, and a second-order accurate procedure. Because of the differences in mechanics, two numerical methods are completely different in approach. Along with the description of systems, brief introductions of the numerical methods are provided.

The discretization method employed is a Lax-Wendroff procedure [46]. Both advection and acoustic operators are discretized to second-order accuracy in time and space. When there is no confusion, partial derivatives will be denoted by subscripts.

$$\mathbf{q}(t) = \mathbf{q}_0 + t\mathbf{q}_t + \frac{t^2}{2}\mathbf{q}_{tt} + \mathcal{O}(t^3) \tag{7.15}$$

where $\mathbf{q}_0 = \mathbf{q}(t = 0)$.

7.2.1 Nonlinear Advection System

A second-order accurate discretization of the advection system is discussed. We recall the advection operator.

$$\mathcal{K}\mathbf{q} = k\mathbf{v} \cdot \nabla\mathbf{q} \tag{7.16}$$

Linear advective system components of the Euler system are

$$\rho_t + k\mathbf{v} \cdot \nabla\rho = \rho_t + \mathcal{K}\rho = 0 \tag{7.17}$$

$$\mathbf{v}_t + k\mathbf{v} \cdot \nabla\mathbf{v} = \mathbf{v}_t + \mathcal{K}\mathbf{v} = 0 \tag{7.18}$$

$$p_t + k\mathbf{v} \cdot \nabla p = p_t + \mathcal{K}p = 0 \tag{7.19}$$

These only represent first-order advective disturbances in the system.

For a second-order nonlinear advection scheme, so that nonlinear advection disturbances can be modeled, we must introduce both second-order and nonlinear advection terms. A second-order term is what makes the advection system discretization second-order accurate, but is present even for linear problems. A nonlinear second-order term is what makes the advection system discretization nonlinear, and is not present in linear problems. These terms are obtained by taking time derivatives of the linear operators, Eqs. (7.17)–(7.19), and replacing them by the corresponding spatial derivatives, or equivalently known

as a Lax-Wendroff procedure. Because of the systematic nature of the advection system, only one derivation is sufficient. All the rest is easily found by following the same steps. We take the density equation for instance. By repeated use of the Lax-Wendroff procedure, we obtain the following second-order derivative term for density.

$$\begin{aligned}\rho_{tt} &= -(\mathcal{K}\rho)_t = -k(\mathbf{v}_t \cdot \nabla\rho + \mathbf{v} \cdot (\nabla\rho)_t) \\ &= k^2\left((\mathbf{v}\cdot\nabla\mathbf{v})\cdot\nabla\rho + \mathbf{v}\cdot\nabla(\mathbf{v}\cdot\nabla\rho)\right) \\ &= \left(\underbrace{\mathcal{K}_t}_{\text{nonlinear}} + \underbrace{\mathcal{K}^2}_{\text{second-order}}\right)\rho \end{aligned} \tag{7.20}$$

The complete derivations for all variables are given in Appendix C.

The complete second-order nonlinear advection system discretization is described by the following form using advection operators, provided all variables are similarly discretized to second-order.

$$\mathbf{q}(t) = \mathbf{q}_0 - t\mathcal{K}\mathbf{q} + \frac{t^2}{2}\left(\mathcal{K}_t + \mathcal{K}^2\right)\mathbf{q} \tag{7.21}$$

This nonlinear advection system is completely equivalent to the nonlinear advection scheme studied in Section 6.1. Instead of being discretized in space, this form of nonlinear advection operator is discretized in time. For fully explicit schemes like the AF advection scheme, they are equivalent. The equivalence can be relegated to the readers but the general spirit is understood by making a substitution of $\mathbf{x} = -\mathbf{v}t$ in Eq. (7.21).

We revisit the characteristic tracing method that was developed in Section. 6.1 for the multidimensional Burgers' equation. All characteristic tracings take place in the reference space. Consequently, it is convenient to discuss the updates in terms of interpolations. For nonconservative variable reconstructions at time t^n,

$$\mathbf{q}^n(\boldsymbol{\xi}) = \sum_{i=1}^{7} c_i(\mathbf{q}) \cdot \phi_i(\boldsymbol{\xi}) \tag{7.22}$$

where $c_i(\mathbf{q})$ indicates Lagrange basis coefficients of variable $\mathbf{q}$. Coefficients and basis function definitions are summarized in Table 3.5. A new nonconservative variable $\mathbf{q}^{n+1}$ is simply found by tracing the streamline to its origin $\boldsymbol{\xi}_F$.

$$\mathbf{q}^{n+1}(\boldsymbol{\xi}_I) = \mathbf{q}^n(\boldsymbol{\xi}_F) \tag{7.23}$$

$$\boldsymbol{\xi}_F = \boldsymbol{\xi}_I - \Delta t \mathbf{J}^{-1}\tilde{\mathbf{v}}^n(\Delta t; \boldsymbol{\xi}_I) \tag{7.24}$$

where $\tilde{\mathbf{v}}^n(\Delta t; \boldsymbol{\xi_I})$ indicates the nonlinear corrected streamline velocity at t^{n+1} and at the interface $\boldsymbol{\xi_I}$. The detailed derivation for the nonlinear corrected streamline velocity is shown in Eq. (6.13). A half-time step solution can be found using the same formula with the modified time step $\frac{\Delta t}{2}$.

7.2.2 Nonlinear Acoustics System

A second-order accurate discretization of the acoustic disturbance system is established. The acoustic operators are recalled.

$$\mathcal{H}\rho = h\rho\nabla\cdot\mathbf{v}, \quad \mathcal{H}\mathbf{v} = h\frac{1}{\rho}\nabla p, \quad \mathcal{H}p = h\gamma p\nabla\cdot\mathbf{v} \tag{7.25}$$

The linear acoustic terms of the Euler system are

$$\rho_t + h\rho\nabla\cdot\mathbf{v} = \rho_t + \mathcal{H}\rho = 0 \tag{7.26}$$

$$\mathbf{v}_t + h\frac{1}{\rho}\nabla p = \mathbf{v}_t + \mathcal{H}\mathbf{v} = 0 \tag{7.27}$$

$$p_t + h\gamma p\nabla\cdot\mathbf{v} = p_t + \mathcal{H}p = 0 \tag{7.28}$$

These only represent first-order acoustic disturbances in the system.

For a second-order nonlinear acoustic scheme, we must introduce high-order (nonlinear and second-order) terms. The discretization process, the Lax-Wendroff procedure, is identical to the advection system. However, since all linear acoustic terms are different,

they would result in different high-order terms.

$$\begin{aligned}
\rho_{tt} &= -(\mathcal{H}\rho)_t = -h\left(\rho_t \nabla\cdot\mathbf{v} + \rho(\nabla\cdot\mathbf{v})_t\right) \\
&= h^2\left(\nabla^2 p - \frac{1}{\rho}\nabla p\cdot\nabla\rho + \rho(\nabla\cdot\mathbf{v})^2\right) \\
&= (\mathcal{H}_t + \mathcal{H}^2)\rho \qquad (7.29)\\
\mathbf{v}_{tt} &= -(\mathcal{H}\mathbf{v})_t = -h\left(\left(\frac{1}{\rho}\right)_t \nabla p + \frac{1}{\rho}\nabla p_t\right) \\
&= h^2\left(\frac{\gamma p}{\rho}\nabla\nabla\cdot\mathbf{v} + \frac{(\gamma-1)}{\rho}\nabla p\nabla\cdot\mathbf{v}\right) \\
&= (\mathcal{H}_t + \mathcal{H}^2)\mathbf{v} \qquad (7.30)\\
p_{tt} &= -(\mathcal{H}p)_t = -h\left(\gamma p_t\nabla\cdot\mathbf{v} + \gamma p\left(\nabla\cdot\mathbf{v}\right)_t\right) \\
&= h^2\left(\frac{\gamma p}{\rho}\nabla^2 p - \frac{\gamma p}{\rho^2}\nabla p\cdot\nabla\rho + \gamma^2 p(\nabla\cdot\mathbf{v})^2\right) \\
&= \left(\underbrace{\mathcal{H}_t}_{\text{nonlinear}} + \underbrace{\mathcal{H}^2}_{\text{second-order}}\right) p \qquad (7.31)
\end{aligned}$$

Now, the first-order and second-order derivatives are completely defined. The second-order nonlinear acoustic system discretization is shown in Eq. (7.32).

$$\mathbf{q}(t) = \mathbf{q}_0 - t\mathcal{H}\mathbf{q} + \frac{t^2}{2}\left(\mathcal{H}_t + \mathcal{H}^2\right)\mathbf{q} \qquad (7.32)$$

It has been stressed over that we cannot use the same numerical procedures for both the advection and acoustics processes. Wave systems pose different propagation principles in more than one dimension [17]. The omnidirectional property of two-dimensional acoustic processes is well modeled by the method of spherical means which has been previously[1] investigated by [23, 22].

Here we only try to lay out a fundamental version of the numerical procedure for the nonlinear acoustics system. For a given function $f(\mathbf{x}, t)$, we define the spherical mean of the function as $M_R\left\{f(\mathbf{x}, t)\right\}$. This is the mean value of $f(\mathbf{x}, t)$ over a sphere of radius $R = at$ centered on $\mathbf{x}$. An initial value problem for a scalar wave equation is considered.

$$\frac{\partial^2\phi}{\partial t^2} = a^2\nabla^2\phi \qquad (7.33)$$

[1] and in the future by D. Fan in her upcoming dissertation which discusses the nonlinear acoustics system in detail.

where a is the local speed of sound. A solution to this problem is found.

$$\phi(\mathbf{x},t) = \frac{\partial}{\partial t}\left(tM_R\left\{\phi(\mathbf{x},0)\right\}\right) + atM_R\left\{\frac{\partial}{\partial t}\phi(\mathbf{x},0)\right\} \tag{7.34}$$

Now, an alternative form of the solution is found as follows.

$$\phi(\mathbf{x},t) = \phi(\mathbf{x},0) + tM_R\left\{\frac{\partial}{\partial t}\phi(\mathbf{x},0)\right\} + \int_0^R tM_t\left\{\nabla^2\phi(\mathbf{x},0)\right\}dt \tag{7.35}$$

This form of the solution has the second and third terms that are proportional to Δt and Δt^2. This resembles the nonlinear discretization we have developed using the Taylor expansion. For that reason, these terms are exploited in completing the first- and second-order terms shown in Eq. (7.32).

$$\begin{aligned}\rho(t) = \rho_0 - tM_R\left\{\rho\nabla\cdot\mathbf{v}\right\} + \int_0^R tM_t\left\{\nabla^2 p\right\}dt+ \\ \frac{t^2}{2}\left(-\frac{1}{\rho}\left(M_R\left\{\nabla p\right\}\cdot M_R\left\{\nabla\rho\right\}\right) + \rho\left(M_R\left\{\rho\nabla\cdot\mathbf{v}\right\}\right)^2\right)\end{aligned} \tag{7.36}$$

$$\begin{aligned}\mathbf{v}(t) = \mathbf{v}_0 - tM_R\left\{\frac{1}{\rho}\nabla p\right\} + \int_0^R tM_t\left\{\nabla^2\mathbf{v}\right\}dt+ \\ \frac{t^2}{2}\left(-\frac{(\gamma-1)}{\rho}\left(M_R\left\{\nabla p\right\}\cdot M_R\left\{\nabla\cdot\mathbf{v}\right\}\right)\right)\end{aligned} \tag{7.37}$$

$$\begin{aligned}p(t) = p_0 - tM_R\left\{\gamma p\nabla\cdot\mathbf{v}\right\} + \int_0^R tM_t\left\{a^2\nabla^2 p\right\}dt+ \\ \frac{t^2}{2}\left(-\frac{a^2}{\rho}\left(M_R\left\{\nabla p\right\}\cdot M_R\left\{\nabla\rho\right\}\right) + \gamma^2 p\left(M_R\left\{\rho\nabla\cdot\mathbf{v}\right\}\right)^2\right)\end{aligned} \tag{7.38}$$

where spherical means of nonconservative variables $M_R\left\{\mathbf{q}(\mathbf{x})\right\}$ are exactly retrieved by the following formula.

$$M_R\left\{\mathbf{q}(\mathbf{x})\right\} = \frac{1}{2\pi R}\int_0^{2\pi}\int_0^R \frac{r\mathbf{q}(x+r\cos\theta, y+r\sin\theta)}{\sqrt{R^2-r^2}}drd\theta \tag{7.39}$$

Also note $\nabla^2\mathbf{v} = (\nabla^2 u, \nabla^2 v)^T$ in Eq. (7.37), not the usual Hessian.

7.3 Linear Operator Splitting

Two nonlinear operators have been established in the previous section. A straightforward way to approximate the Euler equations is to apply a linear combination of two independent

but nonlinear operators simultaneously.

$$\mathbf{q}(t) = \mathbf{q}_0 - t(\mathcal{K} + \mathcal{H})\mathbf{q} + \frac{t^2}{2}\left((\mathcal{K}\mathbf{q})_t + (\mathcal{H}\mathbf{q})_t\right) \tag{7.40}$$

It does not, however, result in a complete nonlinear scheme that models the *nonlinear* and *multidimensional* nature of the Euler system. The deficiency of this approach is revealed upon a close inspection of the second-order terms in Eq. (7.40) compared to the complete second-order expansion of the multidimensional Euler system detailed in Appendix B. The following *nonlinear interaction* terms are not accounted for in the linear operator splitting approach.

$$\delta s = \frac{t^2}{2} kh \left(\frac{\nabla p \cdot \nabla s}{\rho} \right) \tag{7.41}$$

$$\delta \rho = \frac{t^2}{2} kh \left(2\mathbf{v} \cdot \nabla \rho (\nabla \cdot \mathbf{v}) + 2\rho(\mathbf{v} \cdot (\nabla \nabla \cdot \mathbf{v})) + \frac{\nabla p \cdot \nabla \rho}{\rho} + \rho(u_x^2 + 2u_y v_x + v_y^2) \right) \tag{7.42}$$

$$\delta u = \frac{t^2}{2} kh \left(-2\frac{p_x(u\rho_x + v\rho_y)}{\rho^2} + 2\frac{up_{xx} + vp_{xy}}{\rho} + \frac{\nabla p \cdot \nabla u}{\rho} + \frac{p_x u_x + p_y v_x}{\rho} \right) \tag{7.43}$$

$$\delta v = \frac{t^2}{2} kh \left(-2\frac{p_y(u\rho_x + v\rho_y)}{\rho^2} + 2\frac{up_{xy} + vp_{yy}}{\rho} + \frac{\nabla p \cdot \nabla v}{\rho} + \frac{p_x u_y + p_y v_y}{\rho} \right) \tag{7.44}$$

$$\delta p = \frac{t^2}{2} kh \left(2\gamma \mathbf{v} \cdot \nabla p (\nabla \cdot \mathbf{v}) + 2\gamma p(\mathbf{v} \cdot (\nabla \nabla \cdot \mathbf{v})) + \frac{\nabla p \cdot \nabla p}{\rho} + \gamma p(u_x^2 + 2u_y v_x + v_y^2) \right) \tag{7.45}$$

The prevailing thought is that these terms are *nonlinear* and *interactive* by nature and that the treatment in numerical schemes must represent that. In the following section, we will discuss how to account for the nonlinear interaction terms by introducing nonlinear numerical operator splitting procedures.

7.4 Nonlinear Operator Splitting - Evolution of Nonconservative Variables by Different Nonlinear Operators

We first present the complete second-order nonlinear Euler system using the operator notations.

$$\mathbf{q}(t) = \mathbf{q}_0 - t(\mathcal{K}+\mathcal{H})\mathbf{q} + \frac{t^2}{2}\Big((\mathcal{K}+\mathcal{H})\mathbf{q}\Big)_t \tag{7.46}$$

$$= \mathbf{q}_0 - t(\mathcal{K}+\mathcal{H})\mathbf{q} + \frac{t^2}{2}\Big(\mathcal{K}_t\mathbf{q} + \mathcal{K}\mathbf{q}_t + \mathcal{H}_t\mathbf{q} + \mathcal{H}\mathbf{q}_t\Big) \tag{7.47}$$

$$= \mathbf{q}_0 - t(\mathcal{K}+\mathcal{H})\mathbf{q} + \frac{t^2}{2}\Big(\underbrace{(\mathcal{K}_t+\mathcal{K}^2)\mathbf{q}}_{\text{nonlinear advection}} + \underbrace{(\mathcal{H}_t+\mathcal{H}^2)\mathbf{q}}_{\text{nonlinear acoustics}} + \mathcal{K}\mathcal{H}\mathbf{q} + \mathcal{H}\mathcal{K}\mathbf{q}\Big) \tag{7.48}$$

This reveals the structure of the exact system discretization and the reason why a straightforward linear operator splitting, Eq. (7.40), cannot provide a complete second-order accurate discretization. The terms $\mathcal{K}\mathcal{H}\mathbf{q}$ and $\mathcal{H}\mathcal{K}\mathbf{q}$ are described as the nonlinear interaction terms that arise when two operators act on one another in a particular order. They must be evaluated nonlinearly. These terms are listed in Appendix C.

It is not, however, necessary to include all these interaction terms explicitly. Most of them can be incorporated into some modification of the first-order coefficients. In the following subsections, we investigate how this works out for various orderings of the advective and acoustic operators. For this purpose, we introduce the following notations to emphasize that operators may be originated from different time steps from the coefficients.

$$\mathcal{K}^a\rho^b = k\mathbf{v}^a\cdot\nabla\rho^b, \qquad \mathcal{H}^a p^b = h\gamma p^b\nabla\cdot\mathbf{v}^a \tag{7.49}$$

where superscripts a and b indicate different time steps.

7.4.1 Nonsymmetric Operator Splitting A: Advection $\rightarrow$ Acoustics

First, the nonsymmetric operator splitting A in which the advection operator is applied prior to applying the acoustic operator to data is described.

Quadratic nonconservative variables $\mathbf{q}^n$ are available at the time t^n in Eq. (7.22). Then, we proceed to advance the nonconservative variables in time by applying the advection

operator followed by the acoustic operator.

$$\mathbf{q}^* = \mathbf{q}^n - t\mathcal{K}^n\mathbf{q}^n + \frac{t^2}{2}(\mathcal{K}^n\mathbf{q}^n)_t \tag{7.50}$$

$$\mathbf{q}^{n+1} = \mathbf{q}^* - t\mathcal{H}^*\mathbf{q}^* + \frac{t^2}{2}(\mathcal{H}^*\mathbf{q}^*)_t \tag{7.51}$$

The star states, $\mathbf{q}^*$, are partially updated states by the advective operator. It is important to realize that the acoustic operator is acting on this partially updated data. The advective update method, Eq. (7.50), is equivalent to the streamline tracing method as mentioned before in Section 7.2. So, the advective solution is obtained by interpolating in a solution reconstruction polynomial at the point $\mathbf{x}^n$, the streamline origin.

$$\mathbf{x}^n = \begin{bmatrix} -tu^n + t^2(u^n u_x^n + v^n u_y^n) + \frac{t^2}{2}\frac{p_x^n}{\rho^n} \\ -tv^n + t^2(u^n v_x^n + v^n v_y^n) + \frac{t^2}{2}\frac{p_y^n}{\rho^n} \end{bmatrix} \tag{7.52}$$

$$= -t\left[\begin{pmatrix} 1 + t\frac{\partial u^n}{\partial x} & t\frac{\partial u^n}{\partial y} \\ t\frac{\partial v^n}{\partial x} & 1 + t\frac{\partial v^n}{\partial y} \end{pmatrix}^{-1} \mathbf{v}^n - \frac{t}{2}\frac{\nabla p^n}{\rho^n}\right] \tag{7.53}$$

$$= -t\left[(\boldsymbol{\mathcal{J}}^{\boldsymbol{n}})^{-1}\,\tilde{\mathbf{v}}^n - \frac{t}{2}\frac{\nabla p^n}{\rho^n}\right] \tag{7.54}$$

Note that we are using the physical coordinate, $\mathbf{x}^n$, but it is clear the streamline origin in the reference coordinate is found by the following relation.

$$\boldsymbol{\xi}_F = \boldsymbol{\xi}_I + \Delta t\mathbf{J}^{-1}\mathbf{x}^n \tag{7.55}$$

It is important to discuss the streamline origin expression in Eq. (7.54). It consists of the nonlinear streamline velocity used in nonlinear advection, Eq. (6.11), and a pressure gradient term. The pressure gradient term can be physically interpreted as the acceleration of a fluid particle in the presence of pressure. Streamlines can be curved due to the pressure in the Euler system. Formerly in Chapter 6, the pressureless Euler system didn't include pressure gradients in streamline origins because there is no pressure in that system.

The numerical method discussed for nonlinear acoustics system, in Section 7.2, is applied to update the nonconservative quantities to the final time $\mathbf{q}^{n+1}$. Because the acoustic operator is based on the partially updated data $\mathbf{q}^*$, the resulting operation would couple these two physical processes. This is one way to numerically achieve a coupling of two distinct phenomena. To see this, the final operator stage, Eq. (7.51), can be expressed in

one step.

$$\mathbf{q}^{n+1} = \mathbf{q}^n - t(\mathcal{K}^n + \mathcal{H}^*)\mathbf{q}^n + \frac{t^2}{2}\left(((\mathcal{K}^n)^2 + \mathcal{K}_t^n)\mathbf{q}^n + 2\mathcal{H}^*\mathcal{K}^n\mathbf{q}^n + ((\mathcal{H}^*)^2 + \mathcal{H}_t^*)\mathbf{q}^n\right) \quad (7.56)$$

The structure of second-order time terms in Eq. (7.56) shows clearly two nonlinear system specific terms and one interaction term. The coupling between two processes is evident from the mixed superscript notations in the second-order time terms. Because only the nonlinear interaction terms of $\mathcal{H}\mathcal{K}\mathbf{q}$ are partially taken care of, there are remaining second-order nonlinear interaction terms that have not been accounted for in Eq. (7.48). They correspond to $\mathcal{K}\mathcal{H}\mathbf{q}$ and are explicitly shown as follows.

$$\delta s = 0 \quad (7.57)$$

$$\delta\rho = -\frac{t^2}{2}kh\left(\rho(u_x^2 + 2u_y v_x + v_y^2)\right) \quad (7.58)$$

$$\delta u = -\frac{t^2}{2}kh\left(\frac{p_x u_x + p_y v_x}{\rho}\right) \quad (7.59)$$

$$\delta v = -\frac{t^2}{2}kh\left(\frac{p_x u_y + p_y v_y}{\rho}\right) \quad (7.60)$$

$$\delta p = -\frac{t^2}{2}kh\left(\gamma p(u_x^2 + 2u_y v_x + v_y^2)\right) \quad (7.61)$$

Some of these terms can be written in vector forms.

$$\delta\rho = -\frac{t^2}{2}kh\left(\rho[\text{tr}(\nabla\mathbf{v})^2 - 2\text{det}(\nabla\mathbf{v})]\right) \quad (7.62)$$

$$\delta\mathbf{v} = -\frac{t^2}{2}kh\left(\frac{\nabla p\cdot\nabla\mathbf{v} + \nabla p\times(\nabla\times\mathbf{v})}{\rho}\right) \quad (7.63)$$

$$\delta p = -\frac{t^2}{2}kh\left(\gamma p[\text{tr}(\nabla\mathbf{v})^2 - 2\text{det}(\nabla\mathbf{v})]\right) \quad (7.64)$$

The nonlinear interaction terms in this form suggest there may be of some physical significance in these interaction terms. The density and pressure interaction terms describe the dilation and rotation of streamtube due to the velocity field. The velocity interaction terms are due to the pressure velocity interaction involving some gyroscopic effect. They are reminiscent of some form of physical phenomena, but it is difficult to say what interpretation is correct.

These nonlinear terms are second-order terms, but they must be properly implemented in order for the numerical scheme to be stable and accurate. One noteworthy outcome is that the entropy equation is completely accounted for, with $\delta s = 0$.

One way to implement these nonlinear terms is by directly replacing them with numerically accurate expressions. The idea behind it is that these are high-order terms, therefore, some approximation would suffice. However, some circumstantial evidence shows the direct implementation of interaction terms may not result in the desired outcome, see Appendix D for numerical results of the Euler equations using the nonsymmetric operator splitting method.

7.4.2 Nonsymmetric Operator Splitting B: Acoustics → Advection

This operator splitting approach is similar to the nonsymmetric operator splitting A. However, we proceed to apply the nonlinear acoustics system operator first then apply the nonlinear advection system on the partially updated data.

$$\mathbf{q}^* = \mathbf{q}^n - t\mathcal{H}^n\mathbf{q}^n + \frac{t^2}{2}(\mathcal{H}^n\mathbf{q}^n)_t \tag{7.65}$$

$$\mathbf{q}^{n+1} = \mathbf{q}^* - t\mathcal{K}^*\mathbf{q}^* + \frac{t^2}{2}(\mathcal{K}^*\mathbf{q}^*)_t \tag{7.66}$$

The nonlinear acoustics system numerical method, discussed in Section 7.2, is applied to update the nonconservative quantities to the intermediate time $\mathbf{q}^*$.

The final solution $\mathbf{q}^{n+1}$ is found by applying the advection operator in Eq. (7.66). The streamline origin $\mathbf{x}^*$ in a quadratic reconstruction is defined as follows.

$$\mathbf{x}^* = \begin{pmatrix} -tu^* + t^2(u^*u_x^* + v^*u_y^*) - \frac{t^2}{2}\frac{p_x^*}{\rho^*} \\ -tv^* + t^2(u^*v_x^* + v^*v_y^*) - \frac{t^2}{2}\frac{p_y^*}{\rho^*} \end{pmatrix} \tag{7.67}$$

$$= -t\left[\begin{pmatrix} 1 + t\frac{\partial u^*}{\partial x} & t\frac{\partial u^*}{\partial y} \\ t\frac{\partial v^*}{\partial x} & 1 + t\frac{\partial v^*}{\partial y} \end{pmatrix}^{-1}\mathbf{v}^* + h\frac{t}{2}\frac{\nabla p^*}{\rho^*}\right] \tag{7.68}$$

$$= -t\left[(\boldsymbol{\mathcal{J}}^*)^{-1}\,\tilde{\mathbf{v}}^* - \frac{t}{2}\frac{\nabla p^*}{\rho^*}\right] \tag{7.69}$$

The reference coordinate of a streamline origin is found as follows.

$$\boldsymbol{\xi}_F = \boldsymbol{\xi}_I + \Delta t\mathbf{J}^{-1}\mathbf{x}^* \tag{7.70}$$

There is a minor difference in the definition of the interpolation location, $\mathbf{x}^*$ compared to the nonsymmetric operator splitting A. This interpolation location has an opposite sign of the pressure gradient correction term in Eq. (7.54).

Equivalently, the final operator stage, Eq. (7.66), can be expressed in one step.

$$\mathbf{q}^{n+1} = \mathbf{q}^n - t(\mathcal{H}^n + \mathcal{K}^*)\mathbf{q}^n + \frac{t^2}{2}\left(((\mathcal{H}^n)^2 + \mathcal{H}_t^n)\mathbf{q}^n + 2\mathcal{K}^*\mathcal{H}^n\mathbf{q}^n + ((\mathcal{K}^*)^2 + \mathcal{K}_t^*)\mathbf{q}^n\right) \quad (7.71)$$

The structure of second-order terms in Eq. (7.71) is now clearly distinguishable as two operator specific nonlinear terms and one interaction term. Because only the nonlinear interaction terms of $\mathcal{K}\mathcal{H}\mathbf{q}$ are partially taken care of, there are remaining second-order nonlinear interaction terms. They correspond to $\mathcal{H}\mathcal{K}\mathbf{q}$ and are explicitly written.

$$\delta s = 0 \quad (7.72)$$

$$\delta\rho = \frac{t^2}{2}kh\left(\rho(u_x^2 + 2u_y v_x + v_y^2)\right) \quad (7.73)$$

$$\delta u = \frac{t^2}{2}kh\left(\frac{p_x u_x + p_y v_x}{\rho}\right) \quad (7.74)$$

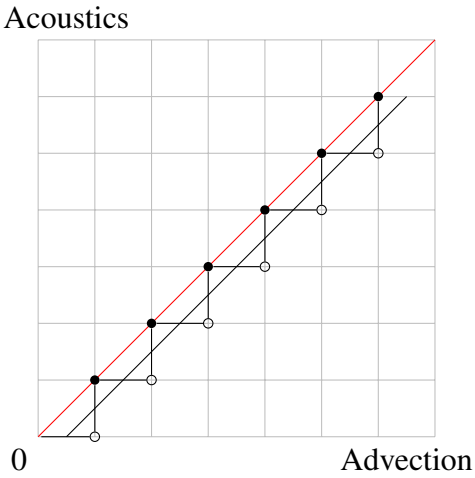
$$\delta v = \frac{t^2}{2}kh\left(\frac{p_x u_y + p_y v_y}{\rho}\right) \quad (7.75)$$

$$\delta p = \frac{t^2}{2}kh\left(\gamma p(u_x^2 + 2u_y v_x + v_y^2)\right) \quad (7.76)$$

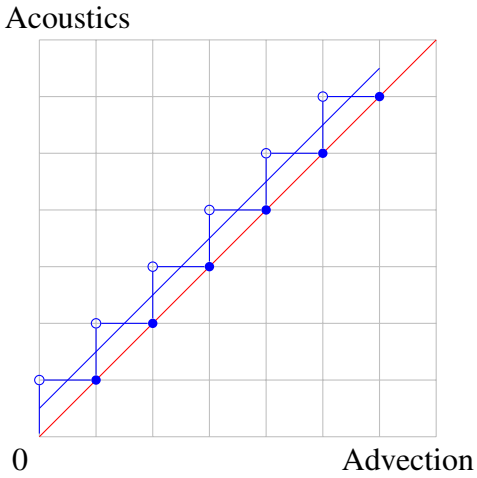
These nonlinear terms are implemented numerically accurately, by replacing the corresponding terms with numerically accurate procedures. The idea behind this choice is that these terms are small and high-order terms, therefore, an approximation would suffice.

Interestingly, the order in which operations are applied leads to the same magnitude of nonlinear interaction terms, but with opposite signs. Compare Eqs. (7.58)–(7.61) to (7.73)–(7.76). This peculiar symmetry opens up a new avenue for an operator splitting method. By averaging or alternating the two nonsymmetric operator splitting methods, we expect to eliminate these terms. This will lead us to reinvent Strang splitting [79], with the difference that Strang splitting was motivated by cases in which the linear operators do not commute. In our case, the linear operators do commute but their nonlinear versions do not.

Two schematics illustrating each nonsymmetric operator splitting procedures are shown in Fig. 7.1. For every complete time step in each nonsymmetric operator splitting procedures, the solution returns to the true solution along the diagonal. But the time averaged effect would always be under- or over-approximations of the true solution. It is easy to convince ourselves that the Euler equations in their natural form are comprised of *balanced* operators. In that sense, neither of the nonsymmetric operator splitting procedures presented so far is adequate to represent the system well.



(a) Nonsymmetric operator splitting A



(b) Nonsymmetric operator splitting B

Figure 7.1: Nonsymmetric operator splittings. Solid markers indicate full-time iterations. Empty markers indicate intermediate stages. The true solution is represented in the red line.

7.4.3 Symmetric Operator Splitting

In the previous section, two nonsymmetric operator splitting procedures resulted in nonlinear interaction terms with the same magnitude and opposite signs, Eqs. (7.58)–(7.61) and (7.73)–(7.76). This brings an interesting prospect for a new type of nonlinear Euler system operator splitting. The process of symmetrizing the operators can recover high-order approximations in their entirety. We consider two previously discussed nonsymmetric operator splittings denoted by $\mathbf{q}_A$ and $\mathbf{q}_B$. They are simultaneously evaluated.

$$\mathbf{q}_A^* = \mathbf{q}^n - t\mathcal{K}^n\mathbf{q}^n + \frac{t^2}{2}(\mathcal{K}^n\mathbf{q}^n)_t \tag{7.77}$$

$$\mathbf{q}_A^{n+1} = \mathbf{q}^* - t\mathcal{H}^*\mathbf{q}^* + \frac{t^2}{2}(\mathcal{H}^*\mathbf{q}^*)_t \tag{7.78}$$

$$\mathbf{q}_B^* = \mathbf{q}^n - t\mathcal{H}^n\mathbf{q}^n + \frac{t^2}{2}(\mathcal{H}^n\mathbf{q}^n)_t \tag{7.79}$$

$$\mathbf{q}_B^{n+1} = \mathbf{q}^* - t\mathcal{K}^*\mathbf{q}^* + \frac{t^2}{2}(\mathcal{K}^*\mathbf{q}^*)_t \tag{7.80}$$

Upon their complete evaluation, a new solution is evaluated by averaging the final solutions of $\mathbf{q}_A^{n+1}$ and $\mathbf{q}_B^{n+1}$.

$$\mathbf{q}^{n+1} = \frac{1}{2}\left(\mathbf{q}_A^{n+1} + \mathbf{q}_B^{n+1}\right) + \mathcal{O}(t^3) \tag{7.81}$$

It is also graphically understood in Fig. 7.2. The symmetric operator splitting involves both nonsymmetric operator splittings and finds the average of two procedures at the full-time step. Therefore, it approximates close to the true solution compared to either of the nonsymmetric operator splitting methods.

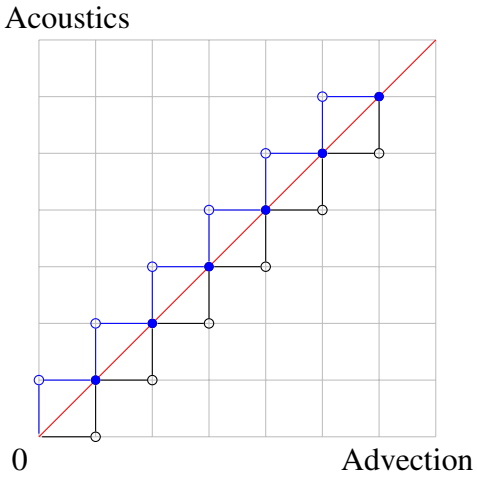


Figure 7.2: Symmetric operator splitting. Two nonsymmetric operator splitting processes are concurrently evaluated and averaged at the complete cycle. Solid markers indicate full-time iterations. Empty markers indicate intermediate stages. The true solution is represented in the red line.

7.4.4 Staggered Operator Splitting

An alternative to the symmetric operator splitting procedure in the previous section is the staggered operator splitting procedure. It is a second-order accurate operator splitting procedure, but with reduced computational cost. The staggered operator splitting is closely related to Strang splitting [79] even though the intents are very different.

Alternating between the nonsymmetric operator splitting A and B, the staggered operator splitting procedure always begins and finishes up with the same type of operator.

$$\mathbf{q}_A^* = \mathbf{q}^n - t\mathcal{H}^n\mathbf{q}^n + \frac{t^2}{2}(\mathcal{H}^n\mathbf{q}^n)_t \tag{7.82}$$

$$\mathbf{q}_A^{n+1} = \mathbf{q}_A^* - t\mathcal{K}^*\mathbf{q}_A^* + \frac{t^2}{2}(\mathcal{K}^*\mathbf{q}_A^*)_t \tag{7.83}$$

$$\mathbf{q}_B^{**} = \mathbf{q}_A^{n+1} - t\mathcal{K}^{n+1}\mathbf{q}_A^{n+1} + \frac{t^2}{2}(\mathcal{K}^{n+1}\mathbf{q}_A^{n+1})_t \tag{7.84}$$

$$\mathbf{q}_B^{n+2} = \mathbf{q}_B^{**} - t\mathcal{H}^{**}\mathbf{q}_B^{**} + \frac{t^2}{2}(\mathcal{H}^{**}\mathbf{q}_B^{**})_t \tag{7.85}$$

where superscripts $*$ and $**$ are provisional steps leading to a completion of a full-time step.

The feature of staggered operator splitting where the same operator is used for both starting and finishing up the procedure is reminiscent of Strang splitting. In Strang splitting, we consider the following linear operators of $\mathcal{C}$ and $\mathcal{D}$ where each operator is only first-order accurate.

$$e^{t(\mathcal{C}+\mathcal{D})} = e^{\frac{t}{2}\mathcal{C}} e^{t\mathcal{D}} e^{\frac{t}{2}\mathcal{C}} + \mathcal{O}(t^3). \tag{7.86}$$

Even though the application of operator splitting in AF and Strang splitting have different initial intent, the final outcomes produce the same order of accuracy.

Advective streamline origins can be interpreted in many ways, and are probably justifiable in many senses. Because of the symmetric nature of the staggered operator splitting, it is quite inconsequential as to what definitions of streamline origins are used. In this case, we choose one version of the streamline origin, Eq. (6.11), without the pressure gradient correction.

$$\mathbf{x} = \begin{pmatrix} -tu + t^2(uu_x + vu_y) \\ -tv + t^2(uv_x + vv_y) \end{pmatrix} \tag{7.87}$$

$$= -t\boldsymbol{\mathcal{J}}^{-1}\tilde{\mathbf{v}} \tag{7.88}$$

and in the reference coordinate

$$\boldsymbol{\xi}_F = \boldsymbol{\xi}_I + \Delta t \mathbf{J}^{-1}\mathbf{x} \tag{7.89}$$

In Fig. 7.3, we see the schematic of the staggered operator splitting. Each complete iteration produces the true solution.

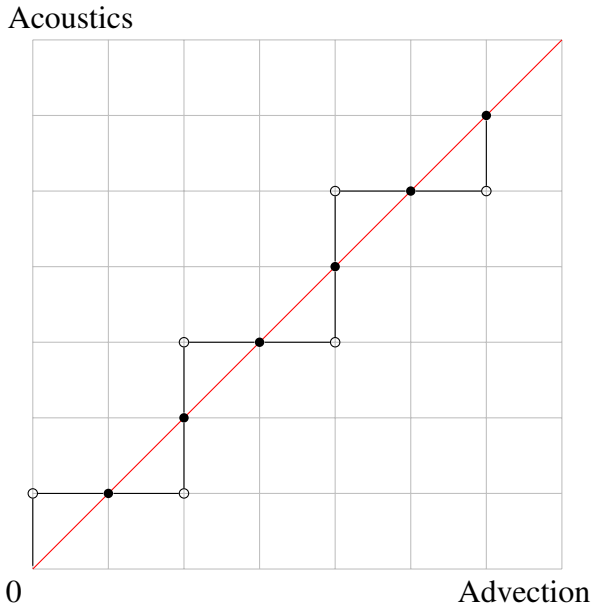


Figure 7.3: Staggered operator splitting. Each marker indicates an operator. Solid markers indicate a completion of one full-time iteration. Empty markers indicate intermediate stages. The true solution is represented in the red line.

7.4.5 Nonlinear Operator Splitting for Linear Problems

Aforementioned nonlinear operator splitting methods are developed for solving *nonlinear* Euler equations. It is important to emphasize that the nonlinear operator splitting methods must retrieve linear problems. Linear problems are characterized by constant operator coefficients. Therefore, all time dependent operator terms in linear problems are subject to cancellations. The exact nonlinear system is recalled Eq. (7.48). For linear problems, operators with the time dependent coefficients are omitted, and the expression simplifies as follows.

$$\mathbf{q}(t) = \mathbf{q}_0 - t(\mathcal{K}+\mathcal{H})\mathbf{q} + \frac{t^2}{2}\left((\cancel{\mathcal{K}_t} + \mathcal{K}^2)\mathbf{q} + (\cancel{\mathcal{H}_t} + \mathcal{H}^2)\mathbf{q} + \mathcal{K}\mathcal{H}\mathbf{q} + \mathcal{H}\mathcal{K}\mathbf{q}\right) \tag{7.48}$$

$$= \mathbf{q}_0 - t(\mathcal{K}+\mathcal{H})\mathbf{q} + \frac{t^2}{2}\left(\mathcal{K}^2\mathbf{q} + 2\mathcal{K}\mathcal{H}\mathbf{q} + \mathcal{H}^2\mathbf{q}\right) \tag{7.90}$$

where $\mathcal{K}\mathcal{H}\mathbf{q} = \mathcal{H}\mathcal{K}\mathbf{q}$ since coefficients are constant. The nonlinear operator splitting methods, Eqs. (7.56) and (7.71), will return the exact form for linear problems Eq. (7.90) given constant coefficients.

7.5 Pursuit of Complete Nonlinear Operator Splitting

Even though symmetric and staggered operator splitting methods produce a second-order accurate discretization in time and space for the Euler equations, the nonlinear operator splitting method poses a few challenges in terms of achieving a truly multidimensional numerical scheme. The difficulty stems from the observation that the means by which a second-order accurate discretization was obtained is due to the cancellation of nonlinear interaction terms, not due to the complete discretization. At the present time, it is difficult to say whether this would have an adverse effect towards the ultimate goal of solving multidimensional systems.

Here we suggest a simple modification to nonsymmetric operator splitting A, Section 7.4.1, in order to achieve a *complete* nonlinear operator splitting method. This means all nonlinear interaction terms are incorporated in the nonlinear operator splitting.

$$\mathbf{q}^* = \mathbf{q}^n - t\mathcal{K}^n\mathbf{q}^n + \frac{t^2}{2}(\mathcal{K}^n\mathbf{q}^n)_t \tag{7.91}$$

$$\mathbf{q}^{n+1} = \mathbf{q}^* - t\overline{\mathcal{H}}\mathbf{q}^* + \frac{t^2}{2}(\overline{\mathcal{H}}\mathbf{q}^*)_t \tag{7.92}$$

where $\overline{\mathcal{H}}$ is a time averaged acoustics operator.

$$\overline{\mathcal{H}} = \frac{1}{2}\left(\mathcal{H}^n + \mathcal{H}^*\right) \tag{7.93}$$

For example, the time averaged gradients are listed below. These are the gradients used in the time averaged acoustic operator, $\overline{\mathcal{H}}$.

$$\overline{\frac{\partial u}{\partial x}} = \frac{1}{2}\left(\left.\frac{\partial u^n}{\partial x}\right|_{\mathbf{x}^*} + \left.\frac{\partial u^*}{\partial x}\right|_{\mathbf{x}}\right) \tag{7.94}$$

$$\overline{\frac{\partial v}{\partial y}} = \frac{1}{2}\left(\left.\frac{\partial v^n}{\partial y}\right|_{\mathbf{x}^*} + \left.\frac{\partial v^*}{\partial y}\right|_{\mathbf{x}}\right) \tag{7.95}$$

$$\overline{\frac{\partial p}{\partial x}} = \frac{1}{2}\left(\left.\frac{\partial p^n}{\partial x}\right|_{\mathbf{x}^*} + \left.\frac{\partial p^*}{\partial x}\right|_{\mathbf{x}}\right) \tag{7.96}$$

$$\overline{\frac{\partial p}{\partial y}} = \frac{1}{2}\left(\left.\frac{\partial p^n}{\partial y}\right|_{\mathbf{x}^*} + \left.\frac{\partial p^*}{\partial y}\right|_{\mathbf{x}}\right) \tag{7.97}$$

The first terms are time data n gradients evaluated at a point $\mathbf{x}^*$, a streamline origin. The second terms are velocity gradients of partially updated data, after advection, evaluated at a point $\mathbf{x}$, an interface location. Since one gradient comes from time step n and another from $*$, this is a time averaged gradient. The individual velocity gradient components are

shown below.

$$\left.\frac{\partial u^n}{\partial x}\right|_{\mathbf{x}^*} = u_x - t(uu_{xx} + vu_{xy}) \tag{7.98}$$

$$\left.\frac{\partial u^*}{\partial x}\right|_{\mathbf{x}} = u_x - t(u_x u_x + u_y v_x + uu_{xx} + vu_{xy}) \tag{7.99}$$

$\left.\frac{\partial u^n}{\partial x}\right|_{\mathbf{x}^*}$ can be regarded as the x-derivative of a quadratic reconstruction $u^n(\mathbf{x})$ evaluated at $\mathbf{x}^*$, a streamline origin. And $\left.\frac{\partial u^*}{\partial x}\right|_{\mathbf{x}}$ is the x-derivative of a partially updated quadratic data $u^*(\mathbf{x})$.

There are a lot of analysis left to completely understand this form of a discretization. We simply present the possibility without any numerical implementations.

7.6 System of Conservation Laws

The conservation step completes the scheme as a whole. Conservative fluxes, individually second-order, produce a third-order conservative scheme by the virtue of the discrete Gauss theorem. The integral conservation systems are shown as follows.

$$\int_\Omega \frac{\partial \mathbf{u}}{\partial t} dV + \int_\Omega \nabla \cdot \mathbf{F} dV = 0 \tag{7.100}$$

$$\frac{d}{dt}\int_\Omega \mathbf{u} dV + \oint_{\partial\Omega} \mathbf{F} \cdot \mathbf{n} dS = 0 \tag{7.101}$$

where conservative variables and fluxes for the Euler system are

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho E \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \rho u & \rho v \\ \rho uu + p & \rho uv \\ \rho uv & \rho vv + p \\ \rho uH & \rho vH \end{pmatrix}. \tag{7.102}$$

The discrete conservation system form is presented.

$$\overline{\mathbf{u}}_j^{n+1} = \overline{\mathbf{u}}_j^n - \frac{\Delta t}{\Omega_j} \sum_{l=1}^{\text{sides}} \overline{\mathbf{F}}_l \cdot \mathbf{n}_l \tag{7.103}$$

and $\mathbf{n}_l$ is a scaled outward normal vector of an interface l. The flux tensor on each interface edge $\overline{\mathbf{F}}_l$ is evaluated numerically according to the composite Simpson's formula described in Eq. (3.73).

As discussed in Chapter 6 in great detail, the AF method for the system of conservation

laws require a reconciliation step, see Sections 6.3 and 6.3.2.1.

7.7 Numerical Test

The isentropic vortex test problem is used to verify the accuracy of the multidimensional Euler code [96]. Initial conditions are summarized below.

$$\rho = (T_\infty + \delta T)^{1/(\gamma-1)} \tag{7.104}$$

$$u = (u_\infty + \delta u) \tag{7.105}$$

$$v = (v_\infty + \delta v) \tag{7.106}$$

$$p = C\rho^\gamma \tag{7.107}$$

Changes in temperature and velocity components are defined under the isentropic assumption.

$$\delta T = -\frac{(\gamma-1)\beta^2}{8\gamma\pi^2} e^{(1-r^2)} \tag{7.108}$$

$$\delta u = -y\frac{\beta}{2\pi} e^{(1-r^2)/2} \tag{7.109}$$

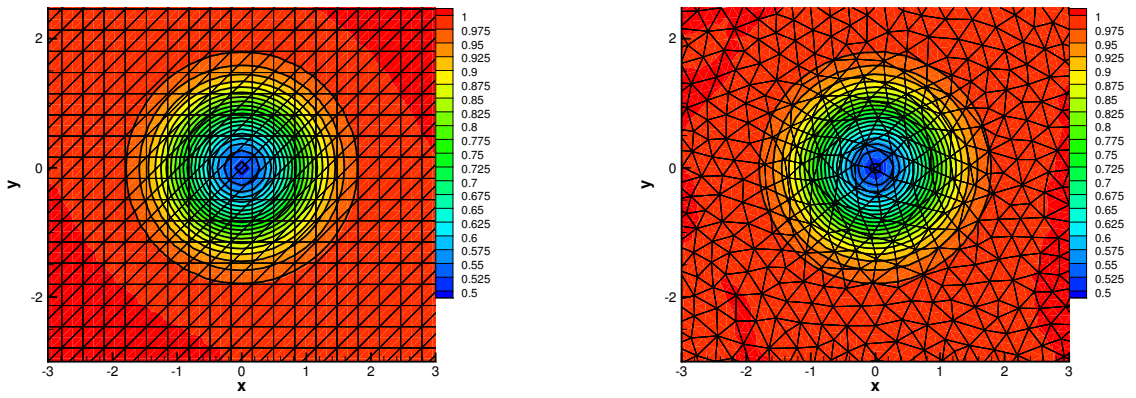
$$\delta v = x\frac{\beta}{2\pi} e^{(1-r^2)/2} \tag{7.110}$$

where $C = 1$, $\gamma = 1.4$, $r^2 = ((x - x_0)^2 + (y - y_0)^2)$, and $\beta = 5$ is the vortex strength.

We consider a horizontally moving isentropic vortex. Free parameters are summarized below.

- $(x_0, y_0) = (0, 0)$ and $(x, y) \in [-10, 10]$
- $\rho_\infty = 1$, $u_\infty = 1$, $v_\infty = 0$, $M_\infty = \sqrt{1.4}^{-1}$
- $\nu = 0.7$
- Periodic boundary conditions in x- and y-directions

The staggered operator splitting method, discussed in Section 7.4.4, is used for the nonconservative update stage.



(a) Triangular structured with $N = 3844$

(b) Triangular unstructured with $N = 3869$

Figure 7.4: Density solution after one cycle at $t = 20$. The final solution is colored according to the color map while the initial solution is indicated as black level lines.

Two types of computational grids are used as shown in Fig. 7.4. The grids used in the Euler system are different from the grids used in advective problems. The computational domain is enlarged to $[-10, 10]$ in order to reduce the interaction of vortex shedding due to the periodic boundary conditions.

Errors in both conservative variables and reconstructions are examined. As a reminder, the error norm definitions are recalled. Errors in conservative variables are measured in the $L_2(\overline{\mathbf{u}})$ norm. It is formally used for error measurements in finite volume methods where cell averages are explicitly known.

$$L_2(\overline{\mathbf{u}}) = \left\{ \frac{1}{\Omega} \sum_{j=1}^{N} \left| \overline{\mathbf{u}}_j \Omega_j - \overline{\mathbf{u}}_j^{\text{exact}} \Omega_j \right|^2 \right\}^{1/2} \tag{7.111}$$

where $\overline{\mathbf{u}} = \left(\overline{\rho}, \overline{\rho u}, \overline{\rho v}, \overline{\rho E}\right)^T$.

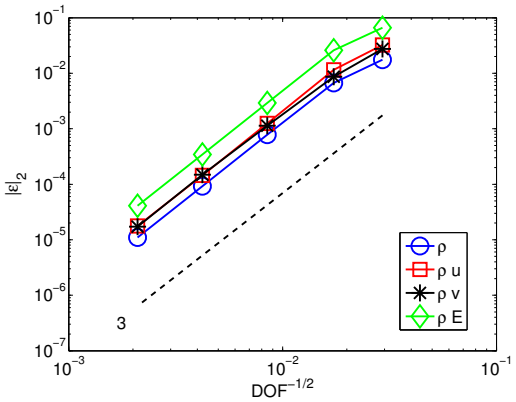
The nonconservative stage of the AF method involves interpolations in solution reconstructions. Since a solution reconstruction is continuously defined in an element, the error norm used for conservative variables is not adequate to represent the projection of the exact solution onto the numerical solution at interpolation points.

$$L_2(\mathbf{q}(\mathbf{x})) = \left\{ \frac{1}{\Omega} \sum_{j=1}^{N} \int_{\Omega_j} \left| \mathbf{q}_j(\mathbf{x}) - \mathbf{q}_j^{\text{exact}}(\mathbf{x}) \right|^2 dV \right\}^{1/2} \tag{7.112}$$

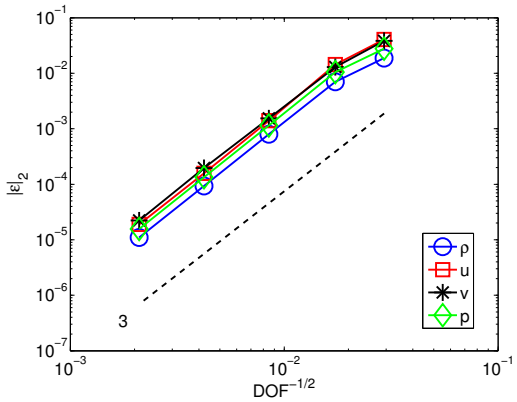
where $\mathbf{q} = (\rho, u, v, p)^T$. For consistently defined schemes, errors measured by these two definitions of error norms should be similar. Therefore, both of these definitions of error

measurements are used to check whether the solutions from the conservation and nonconservation stages correspond well.

We observe third-order convergence rates for the unstructured grid type in both error norm measures in Fig. 7.5. And the same rate of convergence is observed for the structured grid type in Fig. 7.6. In addition, conservative variables errors are summarized in Tables 7.1 and 7.2 for the unstructured and structured grid types, respectively. Note that only the x-component of the momentum error is tabulated because the y-component of the momentum error is very similar.

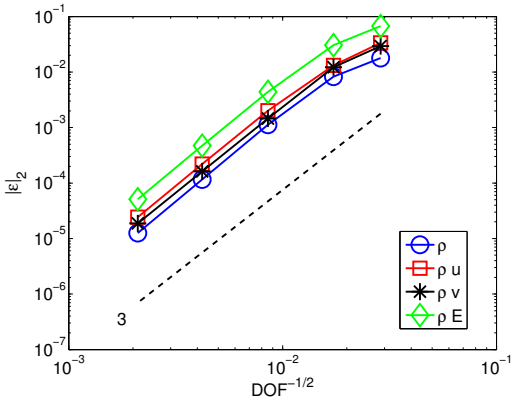


(a) $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables

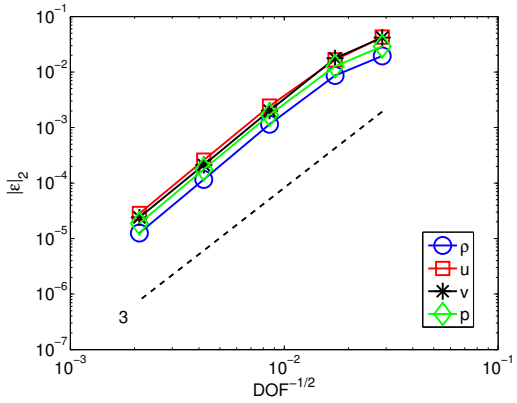


(b) $L_2(\mathbf{q}(\mathbf{x}))$ norm errors of solution reconstructions

Figure 7.5: Error convergence results on unstructured grid at $t = 20$ and $\nu = 0.7$



(a) $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables



(b) $L_2(\mathbf{q}(\mathbf{x}))$ norm errors of solution reconstructions

Figure 7.6: Error convergence results on structured grid at $t = 20$ and $\nu = 0.7$

Table 7.1: $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables on unstructured grid at $t = 20$ and $\nu = 0.7$

DOF	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order	$L_2(\overline{\rho E})$	Order
1.16e+03	1.75e-02		3.23e-02		6.59e-02	
3.32e+03	6.71e-03	1.82	1.15e-02	1.96	2.60e-02	1.77
1.39e+04	7.84e-04	3.00	1.24e-03	3.11	2.92e-03	3.06
5.62e+04	9.24e-05	3.06	1.45e-04	3.08	3.43e-04	3.07
2.27e+05	1.10e-05	3.05	1.77e-05	3.02	4.08e-05	3.05

Table 7.2: $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables on structured grid at $t = 20$ and $\nu = 0.7$

DOF	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order	$L_2(\overline{\rho E})$	Order
1.21e+03	1.78e-02		3.34e-02		6.72e-02	
3.34e+03	8.25e-03	1.52	1.32e-02	1.84	3.06e-02	1.55
1.37e+04	1.13e-03	2.82	2.00e-03	2.66	4.40e-03	2.74
5.65e+04	1.16e-04	3.22	2.19e-04	3.12	4.71e-04	3.16
2.26e+05	1.26e-05	3.21	2.43e-05	3.18	5.09e-05	3.21

Additionally, errors of some vortex scalar quantities are measured. The measure of entropy is defined as Eq. (7.113).

$$s = \ln\left(\frac{p}{\rho^{\gamma}}\right) \tag{7.113}$$

The numerical entropy generation indicates how dissipative a numerical scheme is. A velocity magnitude is defined as Eq. (7.114).

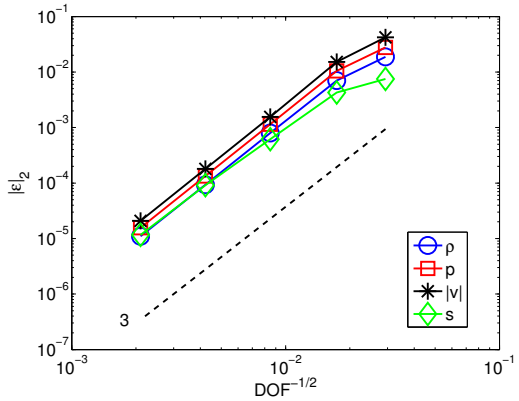
$$|\mathbf{v}| = \sqrt{\mathbf{v}\cdot\mathbf{v}} \tag{7.114}$$

Since both entropy and the velocity magnitude are defined by solution reconstruction variables that are continuously defined in the computational domain, the error norm definition Eq. (7.112) is used.

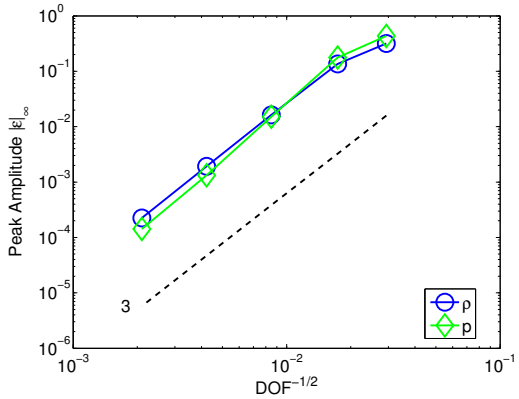
The peak amplitude of density and pressure in the vortex can serve as a measure of numerical dissipation. The location for minimum pressure is iteratively found, then used to evaluate the peak amplitude of density and pressure. The maximum norm is used to measure the peak amplitude error of density and pressure.

$$L_{\max}(\mathbf{q}(\mathbf{x})) = L_{\infty}(\mathbf{q}(\mathbf{x})) = \max|\mathbf{q}(\mathbf{x}) - \mathbf{q}^{\text{exact}}(\mathbf{x})| \tag{7.115}$$

where $\mathbf{q} = (\rho, p)^T$. The errors in vortex related quantities are presented in Figs. 7.7 and 7.8 for the unstructured and structured grid type. Tables 7.3 and 7.5 show the summary of errors for the unstructured and structured grid type. Tables 7.4 and 7.6 show the summary of peak vortex amplitude errors for the unstructured and structured grid type.

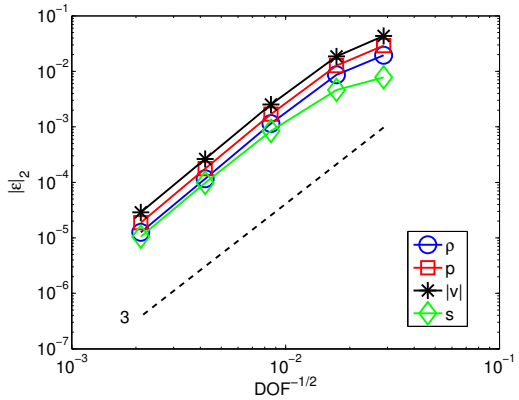


(a) Vortex scalar $L_2(\mathbf{q}(\mathbf{x}))$ norm error. $|\mathbf{v}|$ is a velocity magnitude.

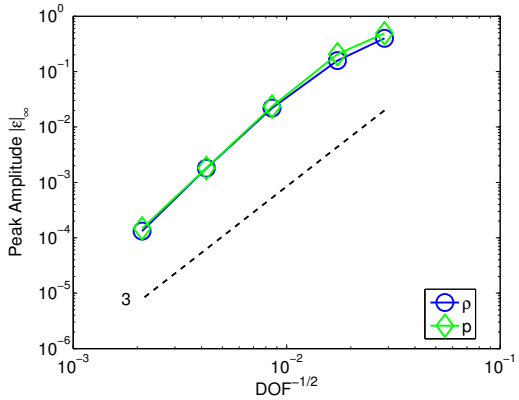


(b) Peak vortex amplitude error in density and pressure

Figure 7.7: Vortex scalar quantity error convergence results on unstructured grid at $t = 20$ and $\nu = 0.7$



(a) Vortex scalar $L_2(\mathbf{q}(\mathbf{x}))$ norm error. $|\mathbf{v}|$ is a velocity magnitude.



(b) Peak vortex amplitude error in density and pressure

Figure 7.8: Vortex scalar quantity error convergence results on structured grid at $t = 20$ and $\nu = 0.7$

Table 7.3: Errors of some vortex scalar quantities on unstructured grid at $t = 20$ and $\nu = 0.7$

DOF	$L_2(\rho)$	Order	$L_2(p)$	Order	$L_2(\|\mathbf{v}\|)$	Order	$L_2(s)$	Order
1.16e+03	1.88e-02		2.77e-02		4.22e-02		7.52e-03	
3.32e+03	7.01e-03	1.87	1.05e-02	1.84	1.52e-02	1.95	4.26e-03	1.08
1.39e+04	7.95e-04	3.05	1.14e-03	3.10	1.56e-03	3.19	6.18e-04	2.70
5.62e+04	9.30e-05	3.07	1.34e-04	3.07	1.79e-04	3.09	9.18e-05	2.73
2.27e+05	1.10e-05	3.06	1.57e-05	3.07	2.08e-05	3.09	1.17e-05	2.95

Table 7.4: Peak vortex amplitude errors on unstructured grid at $t = 20$ and $\nu = 0.7$

DOF	$x_{p_{\min}}$	$y_{p_{\min}}$	$L_\infty(\rho)$	Order	$L_\infty(p)$	Order
1.16e+03	-1.29e-01	-1.60e-01	3.15e-01		4.32e-01	
3.32e+03	2.51e-02	-1.00e-01	1.35e-01	1.61	1.77e-01	1.70
1.39e+04	-2.39e-03	3.47e-03	1.63e-02	2.96	1.51e-02	3.44
5.62e+04	1.64e-03	-2.66e-03	1.93e-03	3.06	1.33e-03	3.48
2.27e+05	-3.53e-04	-7.76e-04	2.25e-04	3.08	1.42e-04	3.21

Table 7.5: Errors of some vortex scalar quantities on structured grid at $t = 20$ and $\nu = 0.7$

DOF	$L_2(\rho)$	Order	$L_2(p)$	Order	$L_2(\|\mathbf{v}\|)$	Order	$L_2(s)$	Order
1.21e+03	1.94e-02		2.91e-02		4.34e-02		7.76e-03	
3.34e+03	8.61e-03	1.61	1.27e-02	1.64	1.86e-02	1.68	4.63e-03	1.02
1.37e+04	1.14e-03	2.86	1.69e-03	2.85	2.53e-03	2.82	8.42e-04	2.41
5.65e+04	1.16e-04	3.23	1.76e-04	3.20	2.63e-04	3.20	9.80e-05	3.04
2.26e+05	1.26e-05	3.21	1.87e-05	3.24	2.89e-05	3.19	1.04e-05	3.23

Table 7.6: Peak vortex amplitude errors on structured grid at $t = 20$ and $\nu = 0.7$

DOF	$x_{p_{\min}}$	$y_{p_{\min}}$	$L_\infty(\rho)$	Order	$L_\infty(p)$	Order
1.21e+03	-3.34e-01	-8.74e-01	3.99e-01		4.91e-01	
3.34e+03	-1.14e-02	1.05e-02	1.57e-01	1.84	2.07e-01	1.71
1.37e+04	4.76e-03	-1.29e-02	2.19e-02	2.79	2.33e-02	3.09
5.65e+04	1.52e-03	-7.79e-04	1.80e-03	3.53	1.82e-03	3.60
2.26e+05	1.24e-04	-7.98e-04	1.31e-04	3.78	1.48e-04	3.63

7.8 Numerical Comparison

For a numerical comparison between AF and DG schemes, all methods employ the same grid types and initial conditions to ensure the comparison is fair and meaningful. The initial condition is the moving vortex problem summarized in Section 7.7. For DG, spatial discretization orders of $p = 1$ and $p = 2$ are considered. A fourth-order time explicit Runge-Kutta integration method is used for both spatial discretization orders. For DG schemes, the stability is restricted by

$$\nu < \frac{1}{2p+1} \tag{7.116}$$

where p is the order of spatial discretization [14]. Furthermore, in the Euler equations, the time step is modified to include a factor to ensure the stability of the scheme.

$$\Delta t = \frac{\nu d_{\text{min}}}{|s|_{\text{max}}} \tag{7.117}$$

where s is a wave speed (eigenvalue) of the problem and d_{min} is a hydraulic diameter of a cell. It is defined as follows.

$$d_{\text{min}} = \min_{i=1}^{N} \left(\frac{2A_i}{P_i} \right) \tag{7.118}$$

where A_i is an area and P_i is a perimeter of cell i.

We have previously devised a common measure called the work unit (WU) in Eq. (3.86). Parameters for the WU calculation are summarized in Table 7.7. The WU incorporates both spatial and temporal discretization parameters in order to compare a fully discrete scheme to a semi-discrete one.

Table 7.7: Summary of parameters used for WU in AF and DG for the Euler equations

Scheme	DOF	Computation stages per iteration	ν_{max}
AF	$N_{\text{cell}} + \frac{1}{2}N_{\text{face}} + \frac{1}{6}N_{\text{vertex}}$	2	1
DGp	$(p+1)(p+2)/2 \times N_{\text{cell}}$	4 (RK 4)	$\frac{1}{2p+1}$

Even though the WU provides a theoretical standard to compare the efficiency of different schemes, it makes no allowance for the arithmetical complexity of schemes. For a system of equations, the mathematical complexity of DG is largely unchanged from that of scalar conservation laws. Conservation systems are formulated in the weak form, leading to performing integrations along element boundaries and inside elements. In order to ensure physical propagation of information, upwind numerical fluxes are used. Approxi-

mate Riemann solvers are usually used to find the upwind numerical flux. The Roe flux [68] is used. Since the way in which conservation systems are solved is largely unchanged from applications to scalar conservation laws, the DG method provides a mathematically streamlined framework for systems.

Unlike DG, the AF method takes a two stage approach with its second stage being the conservation stage. In the first stage, two nonconservative numerical schemes are solved. Each scheme try to address a proper numerical method for information propagation in multidimensional hyperbolic problems. The advective scheme is based on a semi-Lagrangian scheme and the acoustics scheme evaluates the multidimensional acoustic disturbances using spherical means. The advective and acoustic disturbance terms are combined in a nonlinear operator splitting method. Nonconservative methods can take various forms. Because there are many choices available for either methods, the AF method can be formulated to satisfy a variety of design goals; for instance to represent multidimensional problem aspects more faithfully. At the same time, this feature presents a major difference in approximating the computational cost from other conventional numerical methods for hyperbolic problems. Given the perspectives of each method, we understand the WU is neither the best nor the universal standard, but it can be understood as a *good* measure of efficiency with regard to the theoretical construction of each scheme.

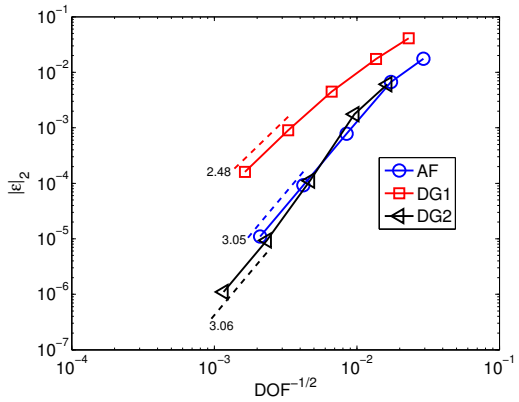
Errors for the three methods (AF, DG1, DG2) will be compared for unstructured grids in Section 7.8.1. The comparison will be made first of all for mesh size (DOF) and then allowance will be made for the number of updates required by each method to advance through a given time (WU). On this basis, the AF method is a very clear winner. Then, the comparison is made on the basis of the wall clock time, and this time, AF and DG2 are joint winners. However, no optimization has been performed on the codes, and we believe that AF will benefit more from this than DG2.

The comparison is repeated, with broadly similar results, for structured grids in Section 7.8.2, with final comparisons between structured and unstructured grids begin given in Section 7.8.3

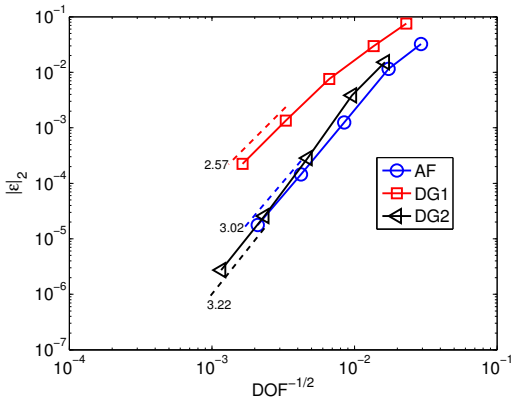
7.8.1 Unstructured Grid

In Fig. 7.9, error convergence rates of AF, DG1 and DG2 schemes are presented. The numerical errors are measured by the $L_2(\overline{\mathbf{u}})$ norm in conservative variables, defined in Eq. (7.111). For DG schemes, an error measured on the basis of the cell average solution may reveal better convergence properties [90]. This can be attributed to the cancellation of error based on the choice of a basis. DG1 scheme achieves $\mathcal{O}(h^{2.5})$ which is better than

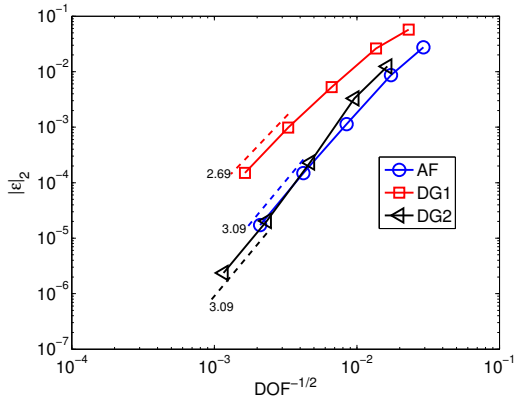
the rate $\mathcal{O}(h^{p+1})$ it normally achieves but less than the rate $\mathcal{O}(h^{2p+1})$ that might be hoped for with superconvergence. However, AF and DG2 schemes are not affected by this slight superconvergence effect, and the convergence rates are $\mathcal{O}(h^3)$ for all variables.



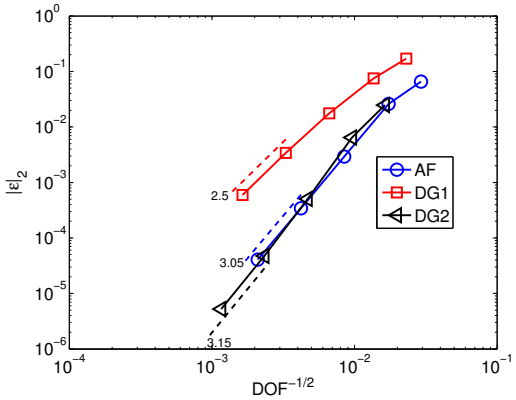
(a) Density error $L_2(\overline{\rho})$ convergence



(b) x-momentum error $L_2(\overline{\rho u})$ convergence



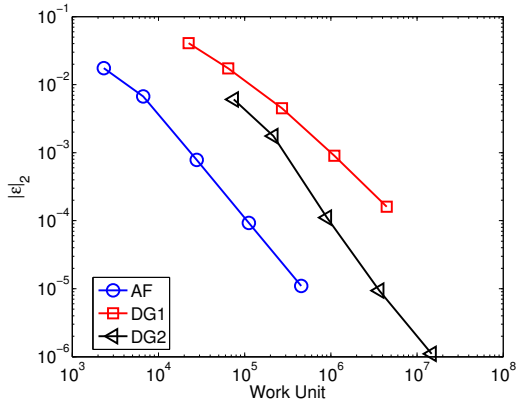
(c) y-momentum error $L_2(\overline{\rho v})$ convergence



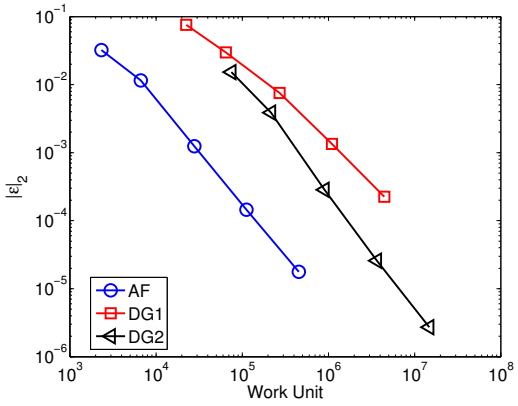
(d) Total energy error $L_2(\overline{\rho E})$ convergence

Figure 7.9: Error convergence comparison on unstructured grid at $t = 20$. Courant numbers are $\nu = 0.7$, $\nu = 0.3$, and $\nu = 0.2$ for AF, DG1, and DG2.

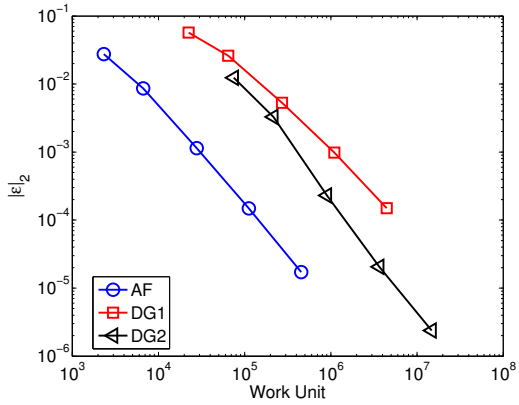
The $L_2(\overline{\mathbf{u}})$ norm errors of three schemes are measured against WU in Fig. 7.10. AF and DG2 exhibited similar convergence in Fig. 7.9. However, using the common WU reveals a marked reduction of WU needed to achieve the same level of error for AF compared to DG2. On average, one order of magnitude fewer WU is required to achieve the same level of error for AF in comparison to DG2.



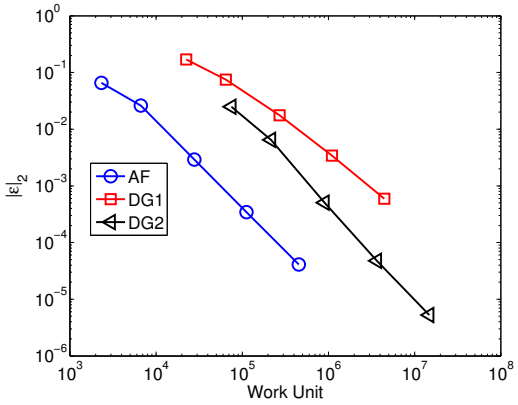
(a) Density error $L_2(\overline{\rho})$ convergence



(b) x-momentum error $L_2(\overline{\rho u})$ convergence



(c) y-momentum error $L_2(\overline{\rho v})$ convergence



(d) Total energy error $L_2(\overline{\rho E})$ convergence

Figure 7.10: Error convergence comparison as a function of WU on unstructured grid at $t = 20$. Courant numbers are $\nu = 0.7$, $\nu = 0.3$, and $\nu = 0.2$ for AF, DG1, and DG2.

Tables 7.8, 7.9 and 7.10 summarize WU and errors for the moving vortex problem for AF, DG1 and DG2.

Table 7.8: $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables for AF on unstructured grid at $t = 20$ and $\nu = 0.7$

DOF	WU	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order	$L_2(\overline{\rho E})$	Order
1.16e+03	2.32e+03	1.75e-02		3.23e-02		6.59e-02	
3.32e+03	6.65e+03	6.71e-03	1.82	1.15e-02	1.96	2.60e-02	1.77
1.39e+04	2.78e+04	7.84e-04	3.00	1.24e-03	3.11	2.92e-03	3.06
5.62e+04	1.12e+05	9.24e-05	3.06	1.45e-04	3.08	3.43e-04	3.07
2.27e+05	4.54e+05	1.10e-05	3.05	1.76e-05	3.02	4.08e-05	3.05

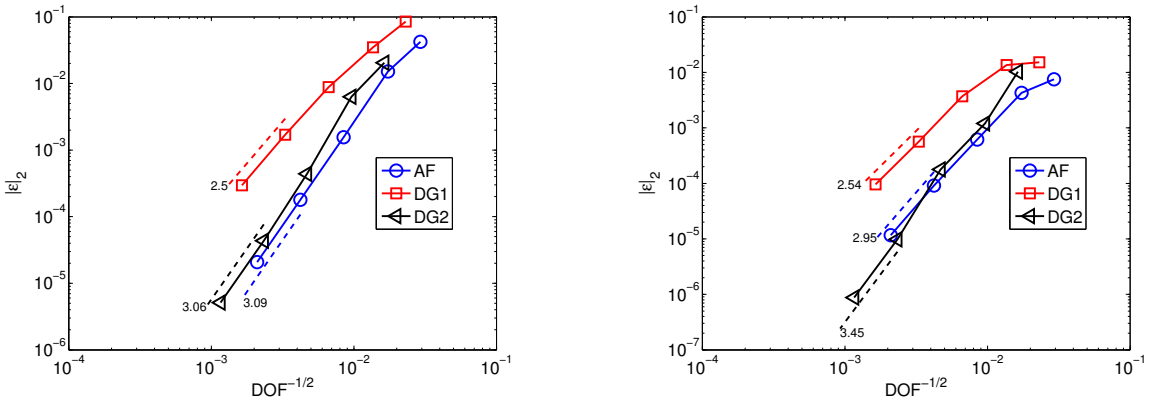
Table 7.9: $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables for DG1 on unstructured grid at $t = 20$, and $\nu = 0.3$

DOF	WU	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order	$L_2(\overline{\rho E})$	Order
1.87e+03	2.25e+04	4.09e-02		7.57e-02		1.70e-01	
5.39e+03	6.47e+04	1.73e-02	1.63	2.97e-02	1.77	7.47e-02	1.56
2.26e+04	2.71e+05	4.49e-03	1.88	7.55e-03	1.91	1.76e-02	2.02
9.17e+04	1.10e+06	9.04e-04	2.29	1.35e-03	2.46	3.40e-03	2.34
3.71e+05	4.45e+06	1.60e-04	2.48	2.25e-04	2.57	5.93e-04	2.50

Table 7.10: $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables for DG2 on unstructured grid at $t = 20$ and $\nu = 0.2$

DOF	WU	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order	$L_2(\overline{\rho E})$	Order
3.74e+03	7.49e+04	6.05e-03		1.52e-02		2.49e-02	
1.08e+04	2.16e+05	1.75e-03	2.34	3.86e-03	2.60	6.46e-03	2.55
4.52e+04	9.05e+05	1.11e-04	3.84	2.83e-04	3.64	5.05e-04	3.55
1.83e+05	3.67e+06	9.36e-06	3.54	2.58e-05	3.43	4.74e-05	3.38
7.41e+05	1.48e+07	1.10e-06	3.06	2.72e-06	3.22	5.26e-06	3.15

In addition, some errors related to a vortex are shown in Fig. 7.11. We examine the velocity magnitude as well as entropy error convergence. The entropy error of DG2 shows, Fig. 7.11b, the convergence rate of $\mathcal{O}(h^{3.5})$. On the other hand, the entropy error of AF converges at the designed rate of $\mathcal{O}(h^3)$. Entropy, which is a scalar output, may be subject to the superconvergence property of DG. As a result, the particular definition of L_2 norm in Eq. (7.111) and a choice of the basis affect the convergence rate in DG.

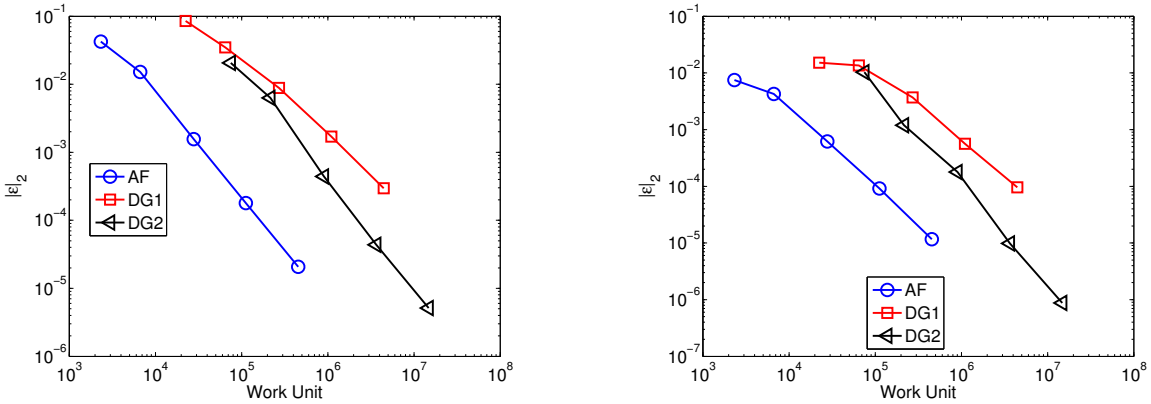


(a) Velocity magnitude error

(b) Entropy error

Figure 7.11: L_2 norm of velocity and entropy error on unstructured grids

The velocity and entropy errors are measured as a function of the WU in Fig. 7.12. As shown previously for the conservative variables, the same level of error is achieved with an order of magnitude fewer WU for AF in comparison to DG2. Tables 7.11, 7.12 and 7.13 summarize the isentropic vortex related error quantities for AF, DG1 and DG2.



(a) Velocity magnitude error

(b) Entropy error

Figure 7.12: L_2 norm of velocity and entropy errors as a function of WU on unstructured grid

Table 7.11: L_2 norm of velocity and entropy errors for AF on unstructured grid at $t = 20$ and $\nu = 0.7$

DOF	WU	$L_2(\|\mathbf{v}\|)$	Order	$L_2(s)$	Order
1.16e+03	2.32e+03	4.22e-02		7.52e-03	
3.32e+03	6.65e+03	1.52e-02	1.95	4.26e-03	1.08
1.39e+04	2.78e+04	1.56e-03	3.19	6.18e-04	2.70
5.62e+04	1.12e+05	1.79e-04	3.09	9.18e-05	2.73
2.27e+05	4.54e+05	2.08e-05	3.09	1.17e-05	2.95

Table 7.12: L_2 norm of velocity and entropy errors for DG1 on unstructured grid at $t = 20$ and $\nu = 0.3$

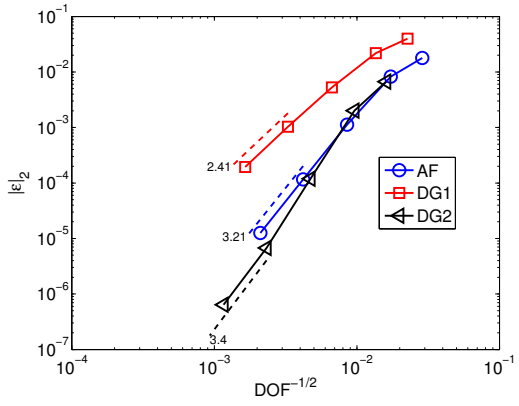
DOF	WU	$L_2(\|\mathbf{v}\|)$	Order	$L_2(s)$	Order
1.87e+03	2.25e+04	8.53e-02		1.51e-02	
5.39e+03	6.47e+04	3.48e-02	1.70	1.36e-02	0.20
2.26e+04	2.71e+05	8.81e-03	1.92	3.70e-03	1.81
9.17e+04	1.10e+06	1.70e-03	2.35	5.63e-04	2.69
3.71e+05	4.45e+06	2.95e-04	2.50	9.56e-05	2.54

Table 7.13: L_2 norm of velocity and entropy errors for DG2 on unstructured grid at $t = 20$ and $\nu = 0.2$

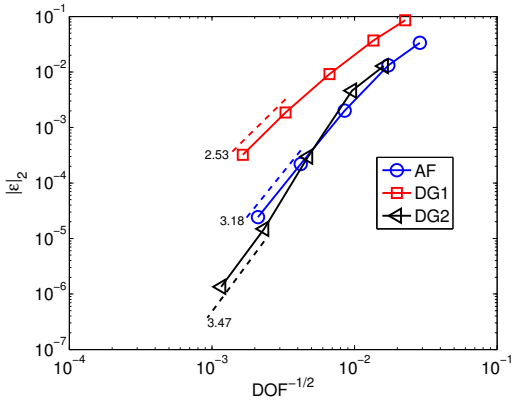
DOF	WU	$L_2(\|\mathbf{v}\|)$	Order	$L_2(s)$	Order
3.74e+03	7.49e+04	2.05e-02		1.03e-02	
1.08e+04	2.16e+05	6.29e-03	2.24	1.20e-03	4.07
4.52e+04	9.05e+05	4.40e-04	3.71	1.78e-04	2.66
1.83e+05	3.67e+06	4.36e-05	3.30	9.80e-06	4.14
7.41e+05	1.48e+07	5.14e-06	3.06	8.79e-07	3.45

7.8.2 Structured Grid

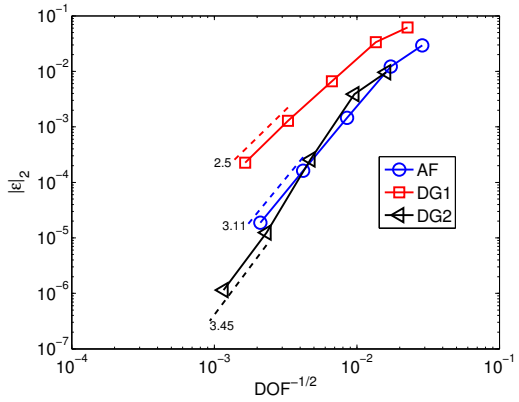
In Fig. 7.13, the convergence rates of AF, DG1 and DG2 schemes are presented. Numerical errore are measured by the $L_2(\overline{\mathbf{u}})$ norm, defined in Eq. (7.111). For DG schemes, an error measured on the basis of the cell average solution may reveal better convergence properties. This can be attributed to the cancellation of error based on the choice of basis. DG1 scheme achieves $\mathcal{O}(h^{2.5})$. DG2 error converges at about $\mathcal{O}(h^{3.5})$. However, it seems to be correlated the structured computational grids. The AF error convergence rate is also marginally better than $\mathcal{O}(h^3)$.



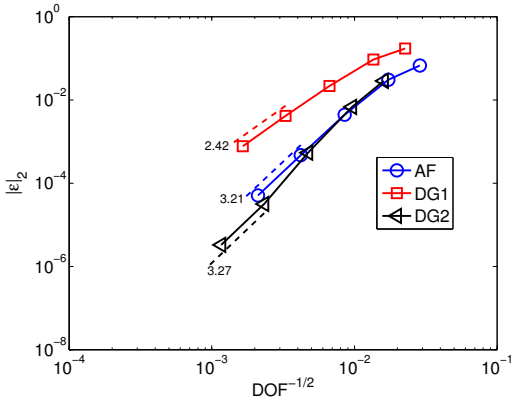
(a) Density error $L_2(\overline{\rho})$ convergence



(b) x-momentum error $L_2(\overline{\rho u})$ convergence



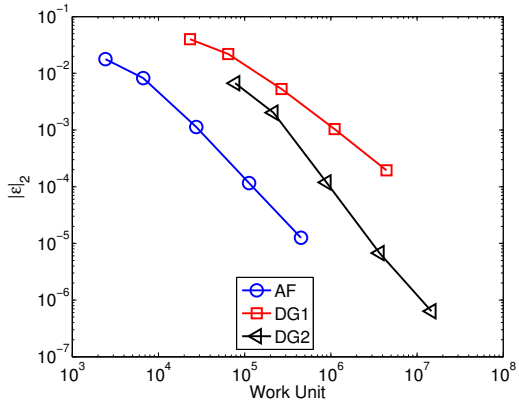
(c) y-momentum error $L_2(\overline{\rho v})$ convergence



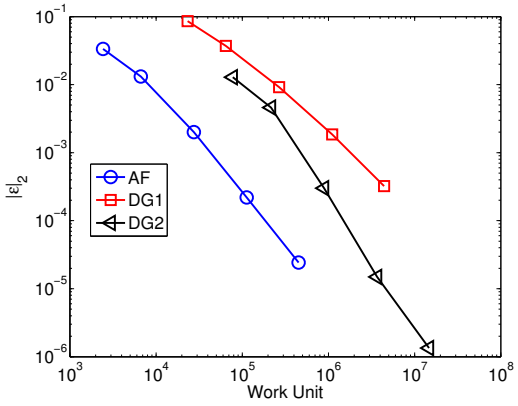
(d) Total energy error $L_2(\overline{\rho E})$ convergence

Figure 7.13: $L_2(\overline{\mathbf{u}})$ norm of conservative variable errors on structured grid at $t = 20$. Courant numbers are $\nu = 0.7$, $\nu = 0.3$, and $\nu = 0.2$ for AF, DG1, and DG2.

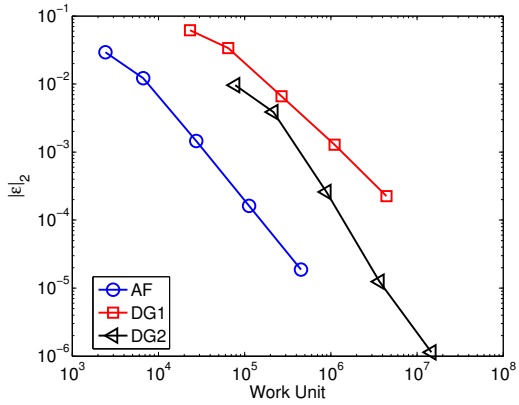
The L_2 norm errors of three schemes are measured against WU in Fig. 7.14. The L_2 norm errors of AF and DG2 exhibited similar convergence shown in Fig. 7.9. However, using the common WU reveals the marked reduction of WU needed to achieve the same level of error for AF compared to DG2, a third-order scheme. On average, one order of magnitude fewer WU is required to achieve the same level of error for AF in comparison to DG2.



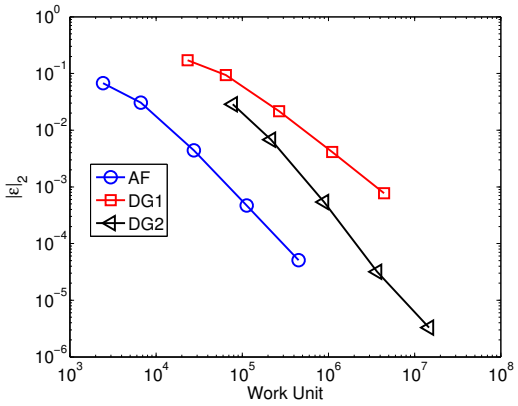
(a) Density error $L_2(\overline{\rho})$ convergence



(b) x-momentum error $L_2(\overline{\rho u})$ convergence



(c) y-momentum error $L_2(\overline{\rho v})$ convergence



(d) Total energy error $L_2(\overline{\rho E})$ convergence

Figure 7.14: $L_2(\overline{\mathbf{u}})$ norm of conservative variable errors as a function of WU on structured grid at $t = 20$. Courant numbers are $\nu = 0.7$, $\nu = 0.3$, and $\nu = 0.2$ for AF, DG1, and DG2.

Tables 7.14, 7.15, and 7.16 summarize WU and errors for the moving vortex problem for AF, DG1 and DG2.

Table 7.14: $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables for AF on structured grid at $t = 20$ and $\nu = 0.7$

DOF	WU	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order	$L_2(\overline{\rho E})$	Order
1.21e+03	2.42e+03	1.78e-02		3.34e-02		6.72e-02	
3.34e+03	6.68e+03	8.25e-03	1.52	1.32e-02	1.84	3.06e-02	1.55
1.37e+04	2.74e+04	1.13e-03	2.82	2.00e-03	2.66	4.40e-03	2.74
5.65e+04	1.13e+05	1.16e-04	3.22	2.19e-04	3.12	4.71e-04	3.16
2.26e+05	4.52e+05	1.25e-05	3.21	2.43e-05	3.18	5.09e-05	3.21

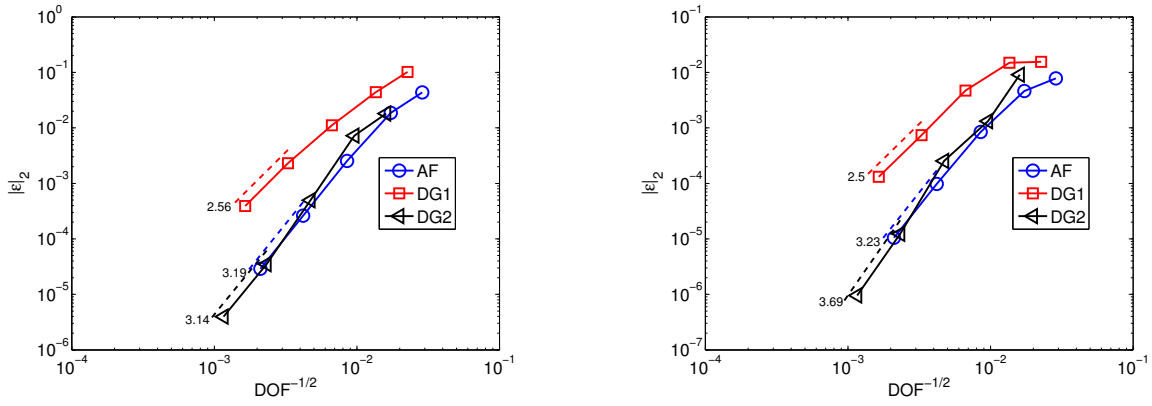
Table 7.15: $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables for DG1 on structured grid at $t = 20$ and $\nu = 0.3$

DOF	WU	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order	$L_2(\overline{\rho E})$	Order
1.94e+03	2.33e+04	4.00e-02		8.56e-02		1.71e-01	
5.40e+03	6.48e+04	2.18e-02	1.18	3.70e-02	1.64	9.35e-02	1.18
2.23e+04	2.68e+05	5.28e-03	2.00	9.21e-03	1.96	2.16e-02	2.07
9.23e+04	1.11e+06	1.03e-03	2.30	1.86e-03	2.26	4.14e-03	2.33
3.69e+05	4.43e+06	1.94e-04	2.41	3.21e-04	2.53	7.71e-04	2.42

Table 7.16: $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables for DG2 on structured grid at $t = 20$ and $\nu = 0.2$

DOF	WU	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order	$L_2(\overline{\rho E})$	Order
3.89e+03	7.78e+04	6.70e-03		1.28e-02		2.87e-02	
1.08e+04	2.16e+05	2.02e-03	2.35	4.61e-03	2.00	6.79e-03	2.82
4.47e+04	8.93e+05	1.19e-04	3.98	2.99e-04	3.85	5.41e-04	3.56
1.85e+05	3.69e+06	6.73e-06	4.05	1.49e-05	4.23	3.18e-05	4.00
7.38e+05	1.48e+07	6.37e-07	3.40	1.34e-06	3.47	3.29e-06	3.27

Additional velocity magnitude and entropy error convergence results for a vortex are shown in Fig. 7.15. The entropy error of DG2 shows, Fig. 7.15b, a convergence rate of $\mathcal{O}(h^{3.69})$. On the other hand, the entropy error of AF converges at the designed rate of $\mathcal{O}(h^{3.23})$. Entropy, which is a scalar output, may be subject to the superconvergence property of DG. As a result, the particular definition of L_2 norm Eq. (7.111) and the choice of basis affect the convergence rate in DG.

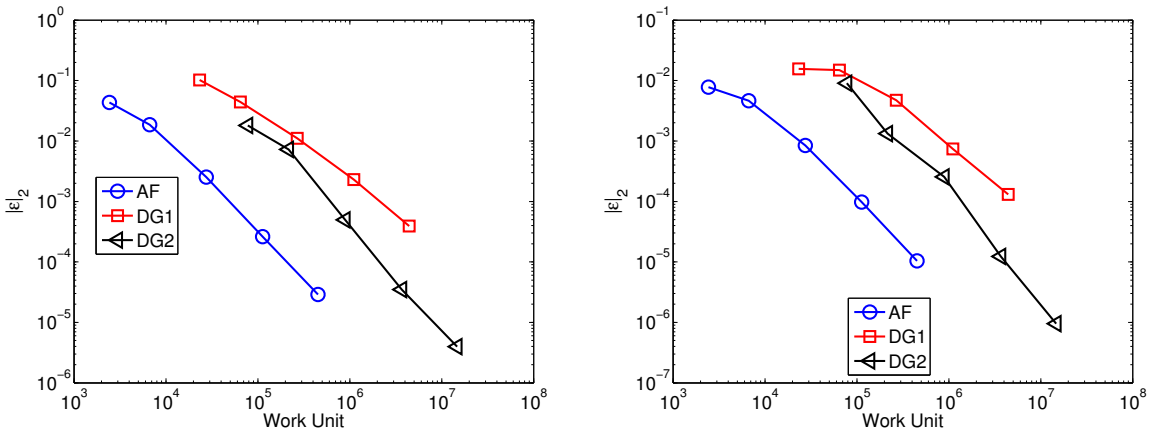


(a) Velocity magnitude error

(b) Entropy error

Figure 7.15: L_2 norm of velocity magnitude and entropy errors on structured grid

The velocity magnitude and entropy error are measured as a function of WU in Fig. 7.16. As shown previously for the conservative variables, the same level of error is achieved with an order of magnitude fewer WU for AF in comparison to DG2. Tables 7.17, 7.18 and 7.19 summarize these error quantities along with WU for AF, DG1 and DG2.



(a) Velocity magnitude error

(b) Entropy error

Figure 7.16: L_2 norm of velocity magnitude and entropy errors as a function of WU on structured grid

Table 7.17: L_2 norm of velocity magnitude and entropy errors for AF on structured grid at $t = 20$ and $\nu = 0.7$

DOF	WU	$L_2(\|\mathbf{v}\|)$	Order	$L_2(s)$	Order
1.212e+03	2.424e+03	4.344e-02		7.755e-03	
3.340e+03	6.680e+03	1.855e-02	1.68	4.631e-03	1.02
1.372e+04	2.745e+04	2.527e-03	2.82	8.417e-04	2.41
5.654e+04	1.131e+05	2.627e-04	3.20	9.800e-05	3.04
2.258e+05	4.517e+05	2.888e-05	3.19	1.044e-05	3.23

Table 7.18: L_2 norm of velocity magnitude and entropy errors for DG1 on structured grid at $t = 20$ and $\nu = 0.3$

DOF	WU	$L_2(\|\mathbf{v}\|)$	Order	$L_2(s)$	Order
1.944e+03	2.333e+04	1.018e-01		1.556e-02	
5.400e+03	6.480e+04	4.404e-02	1.64	1.488e-02	0.09
2.233e+04	2.679e+05	1.107e-02	1.95	4.703e-03	1.62
9.226e+04	1.107e+06	2.299e-03	2.22	7.374e-04	2.61
3.690e+05	4.428e+06	3.909e-04	2.56	1.307e-04	2.50

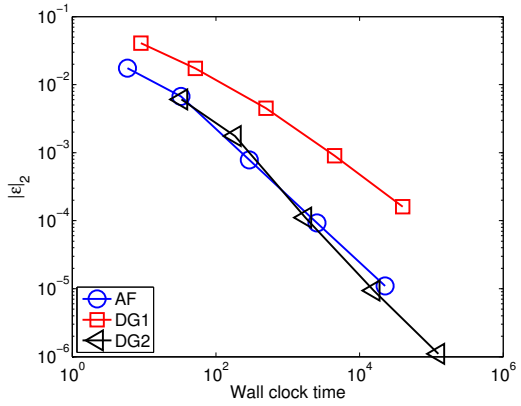
Table 7.19: L_2 norm of velocity magnitude and entropy errors for DG2 on structured grid at $t = 20$ and $\nu = 0.2$

DOF	WU	$L_2(\|\mathbf{v}\|)$	Order	$L_2(s)$	Order
3.888e+03	7.776e+04	1.805e-02		9.058e-03	
1.080e+04	2.160e+05	7.221e-03	1.79	1.318e-03	3.77
4.465e+04	8.930e+05	4.961e-04	3.77	2.541e-04	2.32
1.845e+05	3.690e+06	3.504e-05	3.74	1.232e-05	4.27
7.380e+05	1.476e+07	3.978e-06	3.14	9.525e-07	3.69

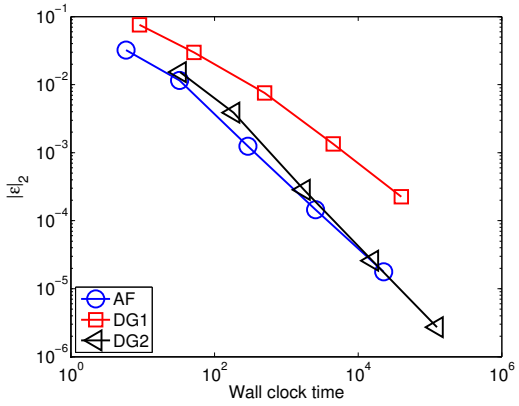
7.8.3 Wall Clock Time Comparison

We have established two different ways to measure and compare error convergence results between AF and DG schemes. The first comparison was deemed inadequate because of the naive comparison using a reference length based only on the mesh size, DOF. The second comparison involved parameters derived by considering theoretical variations of one method in both the spatial and temporal discretization, work unit. Although the work unit was equipped with theoretically sound arguments, it really didn't prove to be a convincing comparison measure. This is mainly due to the lack of a proper computational complexity measure.

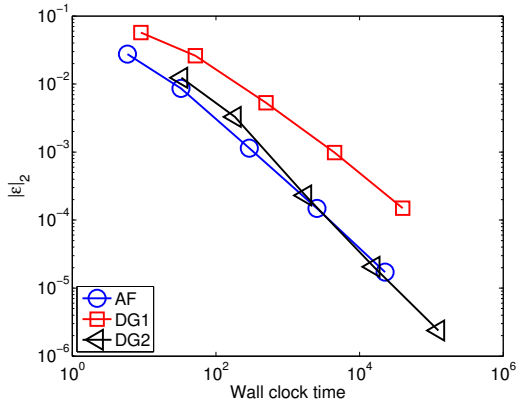
We believe that an objective comparison can be made with wall clock times of simulations. We look at the L_2 norm errors measured as a function of wall clock times in Figs. 7.17 and 7.18 for unstructured grids. The structured grid results are shown in Figs. 7.19 and 7.20. Results suggest that AF and DG2 perform quite closely.



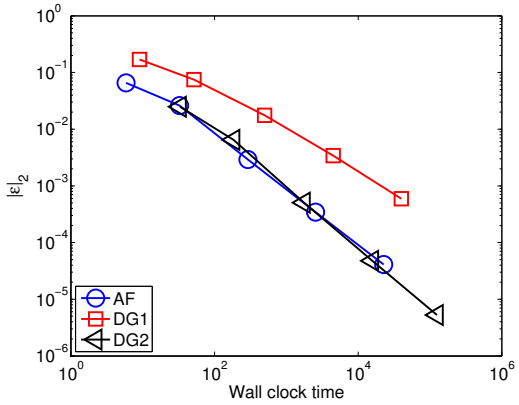
(a) Density error $L_2(\overline{\rho})$ convergence



(b) x-momentum error $L_2(\overline{\rho u})$ convergence

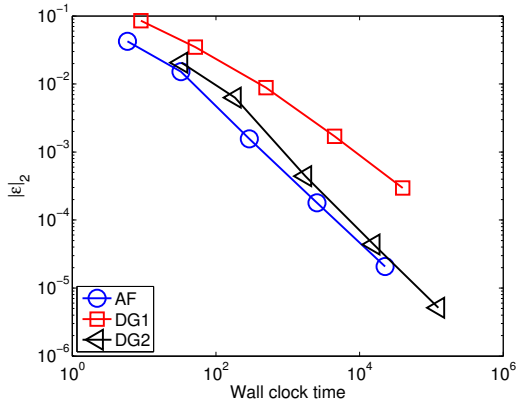


(c) y-momentum error $L_2(\overline{\rho v})$ convergence

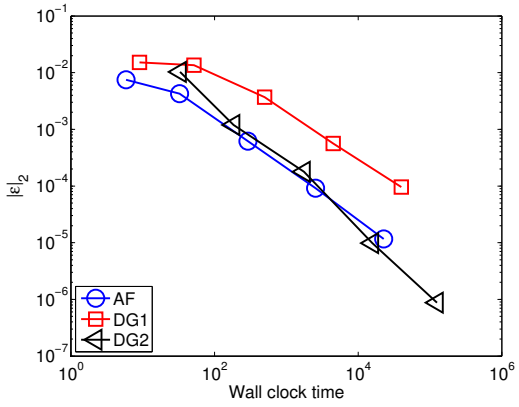


(d) Total energy error $L_2(\overline{\rho E})$ convergence

Figure 7.17: Error convergence comparison as a function of the wall clock time on unstructured grids at $t = 20$. Courant numbers are $\nu = 0.7$, $\nu = 0.3$, and $\nu = 0.2$ for AF, DG1, and DG2.

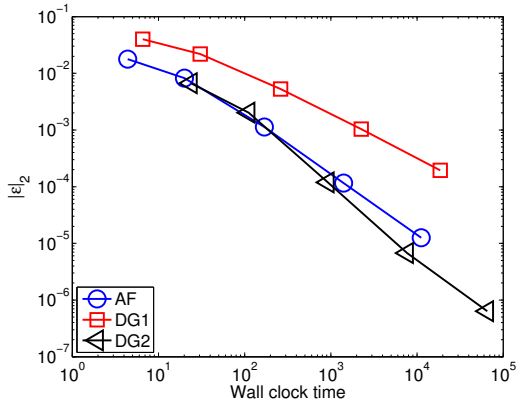


(a) Velocity magnitude error

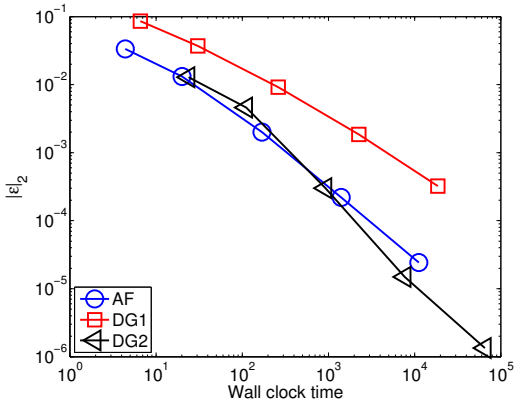


(b) Entropy error

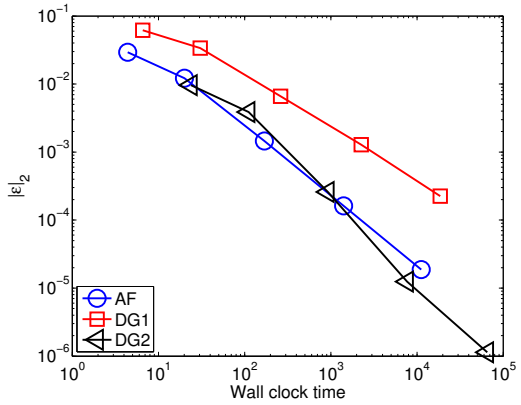
Figure 7.18: L_2 norm of velocity and entropy error as a function of the wall clock time on unstructured grids at $t = 20$. Courant numbers are $\nu = 0.7$, $\nu = 0.3$, and $\nu = 0.2$ for AF, DG1, and DG2.



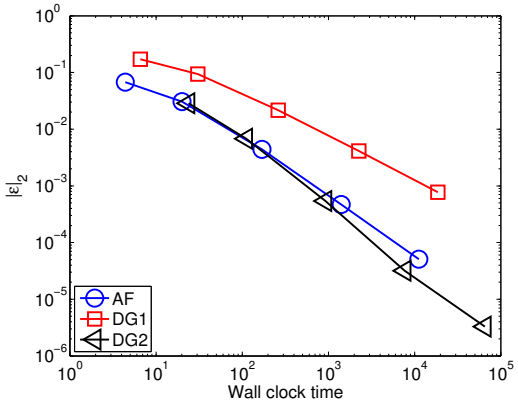
(a) Density error $L_2(\overline{\rho})$ convergence



(b) x-momentum error $L_2(\overline{\rho u})$ convergence

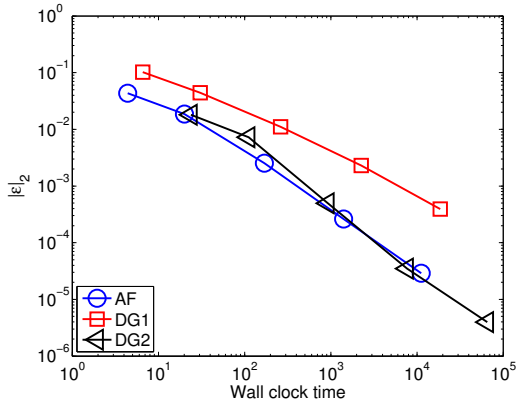


(c) y-momentum error $L_2(\overline{\rho v})$ convergence

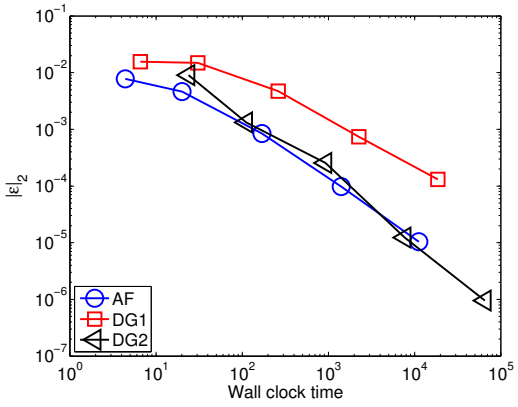


(d) Total energy error $L_2(\overline{\rho E})$ convergence

Figure 7.19: Error convergence comparison as a function of the wall clock time on structured grids at $t = 20$. Courant numbers are $\nu = 0.7$, $\nu = 0.3$, and $\nu = 0.2$ for AF, DG1, and DG2.



(a) Velocity magnitude error



(b) Entropy error

Figure 7.20: L_2 norm of velocity and entropy error as a function of the wall clock time on structured grids at $t = 20$. Courant numbers are $\nu = 0.7$, $\nu = 0.3$, and $\nu = 0.2$ for AF, DG1, and DG2.

Although the wall clock time should provide an objective measure of the computational cost of each numerical method, it is difficult to make the claim in this case. That is because none of the schemes have been thoroughly optimized for the best performance. Therefore, our opinion is that it is difficult to judge the computational efficiency based on these admittedly crude measurements. We, however, believe that further optimizations would benefit AF more than DG2 for the following reasons.

We must discuss the fundamental structures of both schemes to stress why that is. The main difference between AF and DG lies in the manner of design approaches. DG has been largely successful in many applications because it provides a "top-down" mathematical framework. This top-down approach, once established, provides a relatively structured starting point for numerical method designs and implementations. However, this particular feature of DG makes it more difficult to pursue an alternative or an efficient approach if need be.

On the other hand, AF takes a completely opposite design approach where all numerical scheme components are carefully designed leading up to a specific goal. This "bottom-up" design approach of AF can produce intuitive and flexible methods, however, requires more concerted efforts on all aspects of the design. Because there is more than one way to solve the same problem, it can result in many different forms, including an efficient form. We believe this is why AF will come out on top.

CHAPTER 8

Concluding Remarks

8.1 Summary and Conclusion

The AF method, a class of numerical methods for nonlinear hyperbolic conservation laws, was introduced and developed throughout this dissertation. Highlights from chapters are listed as follows.

The fundamental framework for the linear advection scheme was established in Chapter 3 where numerical schemes for one- and two-dimensional problems were presented. Individual numerical components were discussed to describe the process. We presented some mesh alignment related issues in two-dimensional advection problems by analyzing the two-dimensional linear advection scheme. We also provided some recommendations for a fix based on the analysis. In addition, we presented a few comparisons of AF with DG.

We introduced the nonlinear advection scheme in Chapter 4. The nonlinear advection scheme was developed as a simple and general extension of the linear advection scheme. Third-order accuracy was achieved with a simple first-order correction to the characteristic speed in the linear advection scheme. This not only simplified the original AF method (Scheme V) proposed by van Leer for Burgers' equation but also presented us with new perspectives on multidimensional nonlinear advection schemes. Some numerical observations were provided by comparing AF with DG schemes and proved to be competitive.

In Chapter 6, a nonlinear advective system called the pressureless Euler system was introduced for studying the AF method for nonlinear systems. The streamline tracing method, system generalization of the nonlinear characteristic tracing for scalar advection problems, was developed as a result. Since the multidimensional Burgers' equations do not present itself as a conservation system, which AF is designed to solve, an additional mass conservation constraint was introduced to form a conservation system. Studying a conservation system led to the development of a stage called a reconciliation procedure. It was intro-

duced to reconcile two types of updates in AF systems.

We introduced a novel discretization approach for the multidimensional Euler system in Chapter 7. We proposed to decompose the system into two physical processes. We presented individually nonlinear and second-order accurate advective and acoustic numerical methods that were derived from Lax-Wendroff discretization procedures. The advective numerical method developed throughout this dissertation was implemented. A brief summary of the acoustic numerical method was presented. We introduced several nonlinear operator splitting methods that combine these two numerical processes. While we did not successfully execute objective cost comparisons between AF and DG schemes, we presented some useful observations and remarks from a few cases of comparisons, which made us optimistic about the future of AF.

In Chapter 5, AF limiting and design criteria for AF were discussed. Because the numerical approaches in AF are based on distinctions of physical processes, limiting processes are also subject to the same criteria. In particular, the advective process permits a maximum principle where *a priori* bounds are the guiding limiting principle. We introduced a novel way to limit the advective process that obtains and utilizes characteristic information based on the past time history. Limiters based on characteristic information presented promising results. On the other hand, the acoustic process is based on an inherently different physical principle. A principle based on some measure of energy was recommended for acoustic limiting principles. The lack of bounding principles in acoustic problems suggests that an FCT-like limiter that consults from the future time information may be a proper way to limit acoustic problems. Although our understanding on limiting is limited, we believe that these observations and preliminary findings in this thesis can provide new perspectives in the subject of limiting.

In summary, we have discussed the fundamental components of the AF method. CFD has taken a larger role in many aspects of science and engineering fields. As scientific and engineering projects become larger in size and more complex in computations, the importance of good high-order numerical methods became more evident. New high-order methods should not only deliver high-order accuracy but also must present many desirable qualities. Through its unique two stage process, AF achieves a compact, third-order, maximally stable, and conservative multidimensional method that can be adapted to numerous conservation laws. We believe that AF provides a good basis for the next generation of high-order numerical methods.

8.2 Contribution

The major contributions of this dissertation are as follows.

- Extension of van Leer's one-dimensional Scheme V into a general nonlinear framework applicable to nonlinear scalar conservation laws.
- Development of an advection scheme for nonlinear systems of conservation laws.
- Introduction of a reconciliation step in systems of conservation laws.
- Development of a framework for a nonlinear operator splitting method that combines numerical methods for two differently oriented processes.
- Introduction of a framework for the multidimensional limiting.

8.3 Future Work

We have just begun to explore the possibilities of the AF method. There are many improvements to be made. Some future research directions are listed.

- **Robust advection scheme**: A need for a robust advection scheme is evident from the observation that the advection scheme suffers from a mesh alignment issue in multidimensional problems. This mesh alignment issue is related to the characteristic tracing method used for advancing nonconservative solutions. Because the characteristic tracing method inherently lacks numerical diffusion, two solution modes may propagate at different speeds. This becomes more apparent when the advection speed is aligned with a mesh geometry. As a solution, we introduced a couple of numerical remedies for linear problems. These remedies were based on introducing some numerical diffusion by damping bubble functions. However, these methods have not yet been extended to nonlinear problems. Since the method is based on damping all bubble functions in the domain, it would turn out to be a costly procedure. An alternative advection scheme such as the residual distribution method can be suggested. The residual distribution method is not as intuitive as the characteristic tracing method. However, it does not exhibit mesh alignment issues, in fact, it can obtain exact solutions on aligned grids.
- **Genuinely multidimensional limiting**: The importance of a well-behaved solution cannot be stressed enough, especially in multidimensional problems. Our firm belief, which says physics must be apportioned by appropriate processes, provides a

good guiding philosophy for multidimensional limiting. We have introduced *a priori* bound based approach for the advective process limiting. The success of limiting lies on how well one establishes the relation between the current set of information and the characteristic set of information. The characteristic information, derived from the past time information, is based on a physical observation of the domain of dependence of hyperbolic problems. A systematic way to establish the characteristic information needs to be introduced for multidimensional limiting.

The acoustic process calls for an entirely different kind of mechanism for limiting. We believe the limiting criterion based on the quality of the solution, given some measure of determining the solution, via an FCT-type procedure should open up new ways for acoustic process limiting.

- **Nonlinear operator splitting**: A nonlinear system of equations was decomposed into physically distinct components in AF. The nonlinear operator splitting became a necessary and essential aspect of the AF method for combining two nonlinear processes. The nonlinear operator splitting methods presented as a part of this dissertation provide a good starting ground for many future multidimensional nonlinear operator splitting methods. However, one lacking aspect of the current nonlinear operator splitting methods is the production of numerical entropy in cases where it should not be created. This perhaps suggests that there is some deficiency in the current form of the nonlinear operator splitting. Also, some detailed analyses of the operator splitting methods are required in order to understand the best possible method among them. The complete nonlinear operator splitting method option introduced in Section 7.5 must be looked into.

- **Appropriate numerical comparisons**: In this thesis, some results of AF were compared with results obtained from DG schemes. The comparisons are essential parts of validating AF and its performance. Therefore, a meaningful and objective comparison must be provided. An important step toward the objective and meaningful comparison begin with an accurate definition of the computational complexity or cost. The main difference between AF and DG lies in the manner of design approaches. DG and many conventional schemes are categorized in the "top-down" approach where methods are designed around a strong mathematical framework. But AF numerical methods are built from the fundamental components with specific purposes in mind, hence it is categorized as a "bottom-up" approach. We believe that properly addressing differences of the schemes will provide meaningful comparisons.

APPENDIX A

Numerical Verification of Streamtube Area Change Terms in the Two-Dimensional Pressureless Euler System

The numerical method for streamtube area change compensating terms in the density equation was established for the pressureless Euler system in Section 6.1.2. While it is straightforward to understand the one-dimensional implementation, the two-dimensional extension is more complicated, so numerical tests are devised to verify this part of the discretization. This occurs in the nonconservative stage of the update, so we require second-order accuracy. The streamtube area correction term is shown in Eq. (6.29).

$$\det\mathcal{J}|_{\boldsymbol{\xi}_0} = 1 + t\nabla\cdot\mathbf{u}|_{\boldsymbol{\xi}_0} + t^2\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial v}{\partial x}\frac{\partial u}{\partial y}\right)\Bigg|_{\boldsymbol{\xi}_0} \tag{6.29}$$

A.1 Linear Advection

The following initial conditions are considered as a sanity check.

$$\rho_0(x,y) = \exp(-0.25(x^2+y^2)) \tag{A.1}$$

$$u_0(x,y) = -1 \tag{A.2}$$

$$v_0(x,y) = 1 \tag{A.3}$$

Because the velocity field is constant in time, the pressureless Euler equation reduces to linear advection with a constant advection speed $(a,b) = (-1,1)$. The L_2 norm errors in density were obtained at $t = 0.5$ with $\nu = 0.6$ on two types of triangular grids. The solution errors show second-order convergence rates in density for the triangular grid A and unstructured grid type, see Tables A.1 and A.2.

Table A.1: Density error convergence on triangular grid A

DOF	$L_2(\rho)$	Order
3.80e+02	4.07e-03	
1.49e+03	6.88e-04	2.63
5.92e+03	1.54e-04	2.19
2.36e+04	3.65e-05	2.09
9.41e+04	9.23e-06	1.99

Table A.2: Density error convergence on unstructured grid

DOF	$L_2(\rho)$	Order
4.21e+02	3.04e-03	
1.51e+03	8.23e-04	2.06
5.86e+03	2.21e-04	1.95
2.34e+04	5.97e-05	1.90
9.52e+04	1.49e-05	1.98

A.2 Solid Body Rotation

A solid body rotation problem is described by the following initial conditions.

$$\rho_0(x, y) = 1 \tag{A.4}$$

$$u_0(x, y) = -0.5y \tag{A.5}$$

$$v_0(x, y) = 0.5x. \tag{A.6}$$

In this case, a constant density is exposed to the solid body rotation. Density stays uniform in space but the magnitude changes according to the imposed velocity field. The simulation is stopped at $t = 1.0$ with $\nu = 0.6$.

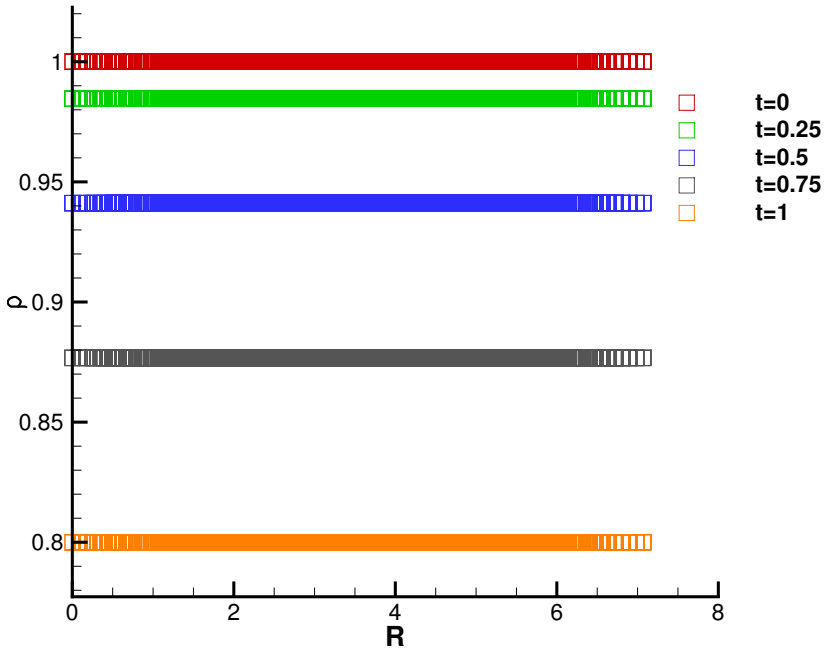


Figure A.1: Density time history on unstructured grid. R is radial direction

As shown in Fig. A.1, density gradually decreases in time according to

$$\rho(t) = \frac{\rho(0)}{1 + t\nabla \cdot \mathbf{u} + t^2 \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial v}{\partial x}\frac{\partial u}{\partial y} \right)} = \frac{1}{1 + 0.25t^2} \tag{A.7}$$

The velocity is linear in space. Therefore, the scheme produces numerically exact results. This test problem is a simple way to check if the numerical scheme behaves as it should. The errors read better than $\mathcal{O}(h^2)$ convergence due to the linear nature of the problem.

A.3 Verifying First-Order Terms in Eq. (6.29)

These initial conditions are considered next.

$$\rho_0(x, y) = \exp(-0.25(x^2 + y^2)) \tag{A.8}$$

$$u_0(x, y) = -\frac{1}{3}\left(\cos\frac{\pi}{5}(x + y) + 2 \right) \tag{A.9}$$

$$v_0(x, y) = \frac{1}{3}\left(\sin\frac{\pi}{5}(x + y) + 2 \right) \tag{A.10}$$

The velocity gradients are derived.

$$\frac{\partial u_0}{\partial x} = \frac{1}{3}\frac{\pi}{5}\sin\frac{\pi}{5}(x+y) \tag{A.11}$$

$$\frac{\partial u_0}{\partial y} = \frac{1}{3}\frac{\pi}{5}\sin\frac{\pi}{5}(x+y) \tag{A.12}$$

$$\frac{\partial v_0}{\partial x} = \frac{1}{3}\frac{\pi}{5}\cos\frac{\pi}{5}(x+y) \tag{A.13}$$

$$\frac{\partial v_0}{\partial y} = \frac{1}{3}\frac{\pi}{5}\cos\frac{\pi}{5}(x+y) \tag{A.14}$$

From these values, we see that the second-order term in Eq. (6.29) cancels out and only the first-order term is in effect. The errors were obtained at $t = 0.5$ with $\nu = 0.8$ on primarily two types of triangular grids. The solution errors show second-order convergence rates in density for the triangular grid A and unstructured grid type, see Tables A.3 and A.4.

Table A.3: Solution error convergence on triangular grid A

DOF	$L_2(\rho)$	Order	$L_2(u)$	Order	$L_2(v)$	Order
3.80e+02	2.89e-03		2.30e-03		2.31e-03	
1.49e+03	4.54e-04	2.73	3.21e-04	2.91	3.21e-04	2.92
5.92e+03	1.02e-04	2.19	4.54e-05	2.86	4.52e-05	2.86
2.36e+04	2.45e-05	2.06	6.16e-06	2.90	6.13e-06	2.90
9.41e+04	6.16e-06	2.00	8.48e-07	2.87	8.47e-07	2.86

Table A.4: Solution error convergence on unstructured grid

DOF	$L_2(\rho)$	Order	$L_2(u)$	Order	$L_2(v)$	Order
4.21e+02	2.77e-03		1.04e-03		1.05e-03	
1.51e+03	6.24e-04	2.35	1.56e-04	2.99	1.69e-04	2.88
5.86e+03	1.76e-04	1.88	2.99e-05	2.46	3.35e-05	2.40
2.34e+04	4.60e-05	1.94	6.55e-06	2.20	6.70e-06	2.33
9.52e+04	1.17e-05	1.95	1.54e-06	2.06	1.53e-06	2.11

A.4 Veryfing Second-Order Terms in Eq. (6.29)

Next, the following initial conditions are considered.

$$\rho_0(x,y) = \exp(-0.25(x^2+y^2)) \tag{A.15}$$

$$u_0(x,y) = \frac{1}{3}\left(\cos\frac{\pi}{5}y + 2\right) \tag{A.16}$$

$$v_0(x,y) = \frac{1}{3}\left(\sin\frac{\pi}{5}x + 2\right) \tag{A.17}$$

Whence

$$\frac{\partial u_0}{\partial x} = 0 \tag{A.18}$$

$$\frac{\partial u_0}{\partial y} = -\frac{1}{3}\frac{\pi}{5}\sin\frac{\pi}{5}y \tag{A.19}$$

$$\frac{\partial v_0}{\partial x} = \frac{1}{3}\frac{\pi}{5}\cos\frac{\pi}{5}x \tag{A.20}$$

$$\frac{\partial v_0}{\partial y} = 0 \tag{A.21}$$

These velocity gradients lead to no first-order term. The errors were obtained at $t = 0.5$ with $\nu = 0.8$ on primarily two types of triangular grids. The solution errors show second-order convergence rates in density for the triangular grid A and unstructured grid type, see Tables A.5 and A.6.

Table A.5: Solution error convergence on triangular grid A

DOF	$L_2(\rho)$	Order	$L_2(u)$	Order	$L_2(v)$	Order
3.80e+02	3.84e-03		5.48e-04		5.61e-04	
1.49e+03	8.02e-04	2.31	1.05e-04	2.44	1.10e-04	2.40
5.92e+03	2.08e-04	1.97	2.59e-05	2.05	2.78e-05	2.01
2.36e+04	5.41e-05	1.95	6.75e-06	1.95	7.22e-06	1.96
9.41e+04	1.39e-05	1.97	1.74e-06	1.96	1.85e-06	1.97

Table A.6: Solution error convergence on unstructured grid

DOF	$L_2(\rho)$	Order	$L_2(u)$	Order	$L_2(v)$	Order
4.21e+02	2.63e-03		4.06e-04		5.02e-04	
1.51e+03	6.19e-04	2.28	1.01e-04	2.20	1.05e-04	2.47
5.86e+03	1.77e-04	1.85	2.70e-05	1.95	3.07e-05	1.83
2.34e+04	4.70e-05	1.92	7.01e-06	1.95	8.23e-06	1.90
9.52e+04	1.18e-05	1.98	1.80e-06	1.95	2.14e-06	1.92

A.5 Complex Velocity

The initial conditions are now more complex.

$$\rho_0(x,y) = \left(1 - \frac{(\gamma-1)}{8\gamma\pi^2}\epsilon^2 \exp\left(1-r^2\right)\right)^{1/(\gamma-1)} \tag{A.22}$$

$$u_0(x,y) = -\frac{\epsilon y}{2\pi} \exp\left(0.5(1-r^2)\right) \tag{A.23}$$

$$v_0(x,y) = \frac{\epsilon x}{2\pi} \exp\left(0.5(1-r^2)\right) \tag{A.24}$$

where $\epsilon = 5$ and $\gamma = 1.4$. And

$$\frac{\partial u_0}{\partial x} = \frac{xy\epsilon}{2\pi} \exp\left(0.5(1-r^2)\right) \tag{A.25}$$

$$\frac{\partial u_0}{\partial y} = \frac{(y^2-1)\epsilon}{2\pi} \exp\left(0.5(1-r^2)\right) \tag{A.26}$$

$$\frac{\partial v_0}{\partial x} = \frac{(1-x^2)\epsilon}{2\pi} \exp\left(0.5(1-r^2)\right) \tag{A.27}$$

$$\frac{\partial v_0}{\partial y} = \frac{-xy\epsilon}{2\pi} \exp\left(0.5(1-r^2)\right) \tag{A.28}$$

Because of the radial symmetry, the first-order term in Eq. (6.29) is eliminated. The errors were obtained at $t = 0.5$ with $\nu = 0.7$ on primarily two types of triangular grids. The solution errors show second-order convergence rates in density for the triangular grid A and unstructured grid type, see Tables A.7 and A.8.

Table A.7: Solution error convergence on triangular grid A

DOF	$L_2(\rho)$	Order	$L_2(u)$	Order	$L_2(v)$	Order
3.80e+02	1.45e-02		8.55e-03		1.20e-02	
1.49e+03	3.10e-03	2.28	1.36e-03	2.72	1.74e-03	2.86
5.92e+03	6.77e-04	2.22	2.53e-04	2.45	2.96e-04	2.59
2.36e+04	1.61e-04	2.09	5.51e-05	2.21	6.39e-05	2.22
9.41e+04	4.10e-05	1.98	1.38e-05	2.01	1.70e-05	1.91

Table A.8: Solution error convergence on unstructured grid

DOF	$L_2(\rho)$	Order	$L_2(u)$	Order	$L_2(v)$	Order
4.21e+02	9.70e-03		9.20e-03		9.42e-03	
1.51e+03	2.36e-03	2.23	1.63e-03	2.73	1.35e-03	3.06
5.86e+03	6.18e-04	1.99	3.70e-04	2.20	3.68e-04	1.93
2.34e+04	1.55e-04	2.00	9.15e-05	2.02	9.28e-05	1.99
9.52e+04	3.94e-05	1.96	2.40e-05	1.91	2.43e-05	1.91

APPENDIX B

Discretization of the Nonconservative Euler System by a Lax-Wendroff Procedure

The nonconservative Euler system is shown.

$$\frac{\partial \rho}{\partial t} + k\mathbf{v} \cdot \nabla \rho + h\rho \nabla \cdot \mathbf{v} \quad = 0 \tag{B.1}$$

$$\frac{\partial \mathbf{v}}{\partial t} + k\mathbf{v} \cdot \nabla \mathbf{v} + h\frac{1}{\rho}\nabla p \quad = 0 \tag{B.2}$$

$$\frac{\partial p}{\partial t} + k\mathbf{v} \cdot \nabla p + h\gamma p \nabla \cdot \mathbf{v} = 0 \tag{B.3}$$

where $\mathbf{v} = (u, v)^T$ is a velocity vector. The coefficients k and h indicate the terms associated with the advective and acoustics processes, respectively. In case $k = 1$ and $h = 0$, the nonlinear advection system is recovered, and vice versa. These coefficients also serve as a mark to denote which terms belong to the advective or acoustic process. We present the discretizations of nonconservative variables.

B.1 Discretization of the Nonconservative Variables by a Lax-Wendroff Procedure

We employ a Lax-Wendroff procedure to discretize the nonconservative Euler equations up to second-order accuracy in time. We begin by evaluating the first- and second-order terms in the entropy equation.

$$s(\mathbf{0}, t) = s_0 + ts_t + \frac{t^2}{2}s_{tt} + \mathcal{O}(t^3) \tag{B.4}$$

$$s_t = -k\left(us_x + vs_y\right) \tag{B.5}$$
$$= -k\left(\mathbf{v}\cdot\nabla s\right) \tag{B.6}$$

$$s_{xt} = -ku(s_{xx}u + s_x u_x + s_{xy}v + s_y v_x) \tag{B.7}$$
$$s_{yt} = -kv(s_{xy}u + s_x u_y + s_{yy}v + s_y v_y) \tag{B.8}$$
$$(\nabla s)_t = -k\left(\nabla(\mathbf{v}\cdot\nabla s)\right) \tag{B.9}$$

$$s_{tt} = -k\left(u_t s_x + v_t s_y + us_{xt} + vs_{yt}\right) \tag{B.10}$$
$$= -k\left(\mathbf{v}_t\cdot\nabla s + \mathbf{v}\cdot(\nabla s)_t\right) \tag{B.11}$$

Entropy transport process is advective. However, the interaction between acoustics and advection processes is implied in velocity field $\mathbf{v}_t$, as shown in the second-order term s_{tt}. Because of the simultaneous nature of advective and acoustic processes, the entropy is affected by both processes. The coupling of the two processes is only evident in the second- and higher-order terms. With this in mind, the rest of nonconservative variables will exhibit some nonlinear coupling of the two processes.

Next, the density terms are evaluated as follows.

$$\rho(\mathbf{0},t) = \rho_0 + t\rho_t + \frac{t^2}{2}\rho_{tt} + \mathcal{O}(t^3) \tag{B.12}$$

$$\rho_t = -\left(k(u\rho_x + v\rho_y) + h\rho(u_x + v_y)\right) \tag{B.13}$$
$$= -\left(k\mathbf{v}\cdot\nabla\rho + h\rho\nabla\cdot\mathbf{v}\right) \tag{B.14}$$

$$\rho_{xt} = -k\left(\rho_{xx}u + \rho_x u_x + \rho_{xy}v + \rho_y v_x\right) - h\left(\rho(u_{xx} + v_{xy}) + \rho_x(u_x + v_y)\right) \tag{B.15}$$
$$\rho_{yt} = -k\left(\rho_{xy}u + \rho_x u_y + \rho_{yy}v + \rho_y v_y\right) - h\left(\rho(u_{xy} + v_{yy}) + \rho_y(u_x + v_y)\right) \tag{B.16}$$
$$(\nabla\rho)_t = -k\left(\nabla(\mathbf{v}\cdot\nabla\rho)\right) - h\left(\rho\nabla\nabla\cdot\mathbf{v} + \nabla\rho(\nabla\cdot\mathbf{v})\right) \tag{B.17}$$

$$\rho_{tt} = -\left(k\left(u_t\rho_x + v_t\rho_y + u\rho_{xt} + v\rho_{yt}\right) + h\left(\rho_t(u_x + v_y) + \rho(u_{xt} + v_{yt})\right)\right) \tag{B.18}$$
$$= -\left(k\left(\mathbf{v}_t\cdot\nabla\rho + \mathbf{v}\cdot(\nabla\rho)_t\right) + h\left(\rho_t\nabla\cdot\mathbf{v} + \rho(\nabla\cdot\mathbf{v})_t\right)\right) \tag{B.19}$$

The velocity terms are shown as follows.

$$\mathbf{v}(\mathbf{0},t) = \mathbf{v}_0 + t\mathbf{v}_t + \frac{t^2}{2}\mathbf{v}_{tt} + \mathcal{O}(t^3) \tag{B.20}$$

$$u_t = -\left(k(uu_x + vu_y) + h\frac{p_x}{\rho}\right) \tag{B.21}$$

$$v_t = -\left(k(uv_x + vv_y) + h\frac{p_y}{\rho}\right) \tag{B.22}$$

$$\mathbf{v}_t = -\left(k\mathbf{v}\cdot\nabla\mathbf{v} + h\frac{\nabla p}{\rho}\right) \tag{B.23}$$

$$u_{xt} = -\left(k(u_xu_x + uu_{xx} + v_xu_y + vu_{xy}) + h\frac{1}{\rho}\left(p_{xx} - \frac{p_x\rho_x}{\rho}\right)\right) \tag{B.24}$$

$$u_{yt} = -\left(k(u_yu_x + uu_{xy} + v_yu_y + vu_{yy}) + h\frac{1}{\rho}\left(p_{xy} - \frac{p_x\rho_y}{\rho}\right)\right) \tag{B.25}$$

$$v_{xt} = -\left(k(u_xv_x + uv_{xx} + v_xv_y + vv_{xy}) + h\frac{1}{\rho}\left(p_{xy} - \frac{p_y\rho_x}{\rho}\right)\right) \tag{B.26}$$

$$v_{yt} = -\left(k(u_yv_x + uv_{xy} + v_yv_y + vv_{yy}) + h\frac{1}{\rho}\left(p_{yy} - \frac{p_y\rho_y}{\rho}\right)\right) \tag{B.27}$$

$$(\nabla\mathbf{v})_t = -k\left(\nabla(\mathbf{v}\cdot\nabla\mathbf{v})\right) - h\frac{1}{\rho}\left(\nabla^2 p - \frac{1}{\rho}\nabla p\otimes\nabla\rho\right) \tag{B.28}$$

$$(\nabla\cdot\mathbf{v})_t = -k\left(\mathbf{I}\cdot\nabla(\mathbf{v}\cdot\nabla\mathbf{v})\right) - h\frac{1}{\rho}\left(\mathbf{I}\cdot\left(\nabla^2 p - \frac{1}{\rho}\nabla p\otimes\nabla\rho\right)\right) \tag{B.29}$$

$$u_{tt} = -\left(k(u_tu_x + v_tu_y + uu_{xt} + vu_{yt}) + h\frac{1}{\rho}\left(p_{xt} - \frac{1}{\rho}p_x\rho_t\right)\right) \tag{B.30}$$

$$v_{tt} = -\left(k(u_tv_x + v_tv_y + uv_{xt} + vv_{yt}) + h\frac{1}{\rho}\left(p_{yt} - \frac{1}{\rho}p_y\rho_t\right)\right) \tag{B.31}$$

$$\mathbf{v}_{tt} = -\left(k\left(\mathbf{v}_t\cdot\nabla\mathbf{v} + \mathbf{v}\cdot(\nabla\mathbf{v})_t\right) + h\frac{1}{\rho}\left((\nabla p)_t - \frac{1}{\rho}\nabla p\rho_t\right)\right) \tag{B.32}$$

Finally, the pressure terms are evaluated as follows.

$$p(\mathbf{0},t) = p_0 + tp_t + \frac{t^2}{2}p_{tt} + \mathcal{O}(t^3) \tag{B.33}$$

$$p_t = -\left(k(up_x + vp_y) + hp\gamma(u_x + v_y)\right) \tag{B.34}$$
$$= -\left(k\mathbf{v}\cdot\nabla p + hp\gamma\nabla\cdot\mathbf{v}\right) \tag{B.35}$$

$$p_{xt} = -k\left(p_{xx}u + p_x u_x + p_{xy}v + p_y v_x\right) - h\gamma\left(p(u_{xx} + v_{xy}) + p_x(u_x + v_y)\right) \tag{B.36}$$
$$p_{yt} = -k\left(p_{xy}u + p_x u_y + p_{yy}v + p_y v_y\right) - h\gamma\left(p(u_{xy} + v_{yy}) + p_y(u_x + v_y)\right) \tag{B.37}$$
$$(\nabla p)_t = -k\left(\nabla(\mathbf{v}\cdot\nabla p)\right) - h\gamma\left(p\nabla\nabla\cdot\mathbf{v} + \nabla p(\nabla\cdot\mathbf{v})\right) \tag{B.38}$$

$$p_{tt} = -\left(k\left(u_t p_x + v_t p_y + up_{xt} + vp_{yt}\right) + h\gamma\left(p_t(u_x + v_y) + p(u_{xt} + v_{yt})\right)\right) \tag{B.39}$$
$$= -\left(k\left(\mathbf{v}_t\cdot\nabla p + \mathbf{v}\cdot(\nabla p)_t\right) + h\gamma\left(p_t\nabla\cdot\mathbf{v} + p(\nabla\cdot\mathbf{v})_t\right)\right) \tag{B.40}$$

In order to satisfy the consistency requirement, we ought to account for all terms found in the Lax-Wendroff expansion. All of the first-order terms belong quite clearly in either the advective or the acoustics processes. However, the second-order terms, denoted by $()_{tt}$, contain not only higher-order terms in respective processes but also nonlinear interaction terms between two processes. Identifying nonlinear interaction terms can be eased by looking for the coefficients kh or hk associated with each term.

APPENDIX C

Derivations of Advective and Acoustic Nonlinear Operators in the Multidimensional Euler System

Derivations of nonlinear advection and acoustics systems are reviewed in detail.

C.1 Nonlinear Advection System

Recall the advection operator.

$$\mathcal{K}\mathbf{q} = k\mathbf{v} \cdot \nabla \mathbf{q} \tag{C.1}$$

where $\mathbf{q} = (\rho, u, v, p)^T$.

Let's review the advective components of the Euler system.

$$\rho_t + k\mathbf{v} \cdot \nabla \rho = \rho_t + \mathcal{K}\rho = 0 \tag{C.2}$$

$$\mathbf{v}_t + k\mathbf{v} \cdot \nabla \mathbf{v} = \mathbf{v}_t + \mathcal{K}\mathbf{v} = 0 \tag{C.3}$$

$$p_t + k\mathbf{v} \cdot \nabla p = p_t + \mathcal{K}p = 0 \tag{C.4}$$

Second-order terms are obtained by taking temporal derivatives of the first-order operators.

The advective second-order terms are listed below.

$$\begin{aligned}
\rho_{tt} &= -(\mathcal{K}\rho)_t = -k(\mathbf{v}_t \cdot \nabla\rho + \mathbf{v} \cdot (\nabla\rho)_t) \\
&= k^2\left((\mathbf{v} \cdot \nabla\mathbf{v}) \cdot \nabla\rho + \mathbf{v} \cdot \nabla(\mathbf{v} \cdot \nabla\rho)\right) \\
&= (\mathcal{K}_t + \mathcal{K}^2)\rho \qquad \text{(C.5)} \\
\mathbf{v}_{tt} &= -(\mathcal{K}\mathbf{v})_t = -k\left(\mathbf{v}_t \cdot \nabla\mathbf{v} + \mathbf{v} \cdot (\nabla\mathbf{v})_t\right) \\
&= k^2\left((\mathbf{v} \cdot \nabla\mathbf{v}) \cdot \nabla\mathbf{v} + \mathbf{v} \cdot \nabla(\mathbf{v} \cdot \nabla\mathbf{v})\right) \\
&= (\mathcal{K}_t + \mathcal{K}^2)\mathbf{v} \qquad \text{(C.6)} \\
p_{tt} &= -(\mathcal{K}p)_t = -k\left(\mathbf{v}_t \cdot \nabla p + \mathbf{v} \cdot (\nabla p)_t\right) \\
&= k^2\left((\mathbf{v} \cdot \nabla\mathbf{v}) \cdot \nabla p + \mathbf{v} \cdot \nabla(\mathbf{v} \cdot \nabla p)\right) \\
&= (\mathcal{K}_t + \mathcal{K}^2)p \qquad \text{(C.7)}
\end{aligned}$$

The second-order terms are comprised of a nonlinear operator $\mathcal{K}_t$ and a linear but quadratic operator $\mathcal{K}^2$.

Finally, the nonlinear advective system can be expresed in terms of the first- and second-order terms.

$$\mathbf{q}(t) = \mathbf{q}_0 - t\mathcal{K}\mathbf{q} + \frac{t^2}{2}\left(\mathcal{K}_t + \mathcal{K}^2\right)\mathbf{q} \qquad \text{(C.8)}$$

C.2 Nonlinear Acoustics System

Let's recall the acoustics components of the Euler system.

$$\mathcal{H}\rho = h\rho\nabla \cdot \mathbf{v}, \quad \mathcal{H}\mathbf{v} = h\frac{1}{\rho}\nabla p, \quad \mathcal{H}p = h\gamma p\nabla \cdot \mathbf{v} \qquad \text{(C.9)}$$

The acoustics components of the Euler system are shown.

$$\rho_t + h\rho\nabla \cdot \mathbf{v} = \rho_t + \mathcal{H}\rho = 0 \qquad \text{(C.10)}$$

$$\mathbf{v}_t + h\frac{1}{\rho}\nabla p = \mathbf{v}_t + \mathcal{H}\mathbf{v} = 0 \qquad \text{(C.11)}$$

$$p_t + h\gamma p\nabla \cdot \mathbf{v} = p_t + \mathcal{H}p = 0 \qquad \text{(C.12)}$$

And second-order terms are similarly evaluated.

$$\begin{aligned}
\rho_{tt} &= -(\mathcal{H}\rho)_t = -h\left(\rho_t \nabla\cdot\mathbf{v} + \rho(\nabla\cdot\mathbf{v})_t\right) \\
&= h^2\left(\nabla^2 p - \frac{1}{\rho}\nabla p\cdot\nabla\rho + \rho(\nabla\cdot\mathbf{v})^2\right) \\
&= (\mathcal{H}_t + \mathcal{H}^2)\rho \qquad \text{(C.13)} \\
\mathbf{v}_{tt} &= -(\mathcal{H}\mathbf{v})_t = -h\left(\left(\frac{1}{\rho}\right)_t \nabla p + \frac{1}{\rho}\nabla p_t\right) \\
&= h^2\left(\frac{\gamma p}{\rho}\nabla\nabla\cdot\mathbf{v} + \frac{(\gamma-1)}{\rho}\nabla p\nabla\cdot\mathbf{v}\right) \\
&= (\mathcal{H}_t + \mathcal{H}^2)\mathbf{v} \qquad \text{(C.14)} \\
p_{tt} &= -(\mathcal{H}p)_t = -h\left(\gamma p_t \nabla\cdot\mathbf{v} + \gamma p\left(\nabla\cdot\mathbf{v}\right)_t\right) \\
&= h^2\left(\frac{\gamma p}{\rho}\nabla^2 p - \frac{\gamma p}{\rho^2}\nabla p\cdot\nabla\rho + \gamma^2 p(\nabla\cdot\mathbf{v})^2\right) \\
&= (\mathcal{H}_t + \mathcal{H}^2)p \qquad \text{(C.15)}
\end{aligned}$$

The second-order terms are comprised of a nonlinear operator $\mathcal{H}_t$ and a linear but quadratic operator $\mathcal{H}^2$.

In summary, the nonlinear acoustics system takes the following form.

$$\mathbf{q}(t) = \mathbf{q}_0 - t\mathcal{H}\mathbf{q} + \frac{t^2}{2}\left(\mathcal{H}_t + \mathcal{H}^2\right)\mathbf{q} \qquad \text{(C.16)}$$

C.3 Nonlinear Operator Splitting Nonlinear Interaction Terms

We first present the exact second-order nonlinear system using operator notations.

$$\begin{aligned}
\mathbf{q}(t) &= \mathbf{q}_0 - t(\mathcal{K}+\mathcal{H})\mathbf{q} + \frac{t^2}{2}\Big((\mathcal{K}+\mathcal{H})\mathbf{q}\Big)_t \qquad \text{(C.17)} \\
&= \mathbf{q}_0 - t(\mathcal{K}+\mathcal{H})\mathbf{q} + \frac{t^2}{2}\Big(\mathcal{K}_t\mathbf{q} + \mathcal{K}\mathbf{q}_t + \mathcal{H}_t\mathbf{q} + \mathcal{H}\mathbf{q}_t\Big) \qquad \text{(C.18)} \\
&= \mathbf{q}_0 - t(\mathcal{K}+\mathcal{H})\mathbf{q} + \frac{t^2}{2}\Big(\underbrace{(\mathcal{K}_t + \mathcal{K}^2)\mathbf{q}}_{\text{nonlinear advection}} + \underbrace{(\mathcal{H}_t + \mathcal{H}^2)\mathbf{q}}_{\text{nonlinear acoustics}} + \mathcal{K}\mathcal{H}\mathbf{q} + \mathcal{H}\mathcal{K}\mathbf{q}\Big) \qquad \text{(C.19)}
\end{aligned}$$

The terms $\mathcal{K}\mathcal{H}\mathbf{q}$ and $\mathcal{H}\mathcal{K}\mathbf{q}$ can be described as the nonlinear interaction terms that arise when two operators act on one another in a particular order. It is worthwhile to take a

moment to list out all nonlinear interaction terms because it tells us the origins of each interaction term.

$$\mathcal{HK}s = 0 \tag{C.20}$$

$$\mathcal{HK}\rho = hk\frac{t^2}{2}\left(\mathbf{v}\cdot\nabla\rho(\nabla\cdot\mathbf{v}) + \rho\nabla\cdot(\mathbf{v}\cdot\nabla\mathbf{v})\right) \tag{C.21}$$

$$\mathcal{HK}u = hk\frac{t^2}{2}\left(-\frac{p_x(u\rho_x + v\rho_y)}{\rho^2} + \frac{up_{xx} + vp_{xy}}{\rho} + \frac{p_x u_x + p_y v_x}{\rho}\right) \tag{C.22}$$

$$\mathcal{HK}u = hk\frac{t^2}{2}\left(-\frac{p_y(u\rho_x + v\rho_y)}{\rho^2} + \frac{up_{xy} + vp_{yy}}{\rho} + \frac{p_x u_y + p_y v_y}{\rho}\right) \tag{C.23}$$

$$\mathcal{HK}p = hk\frac{t^2}{2}\left(\gamma\mathbf{v}\cdot\nabla p(\nabla\cdot\mathbf{v}) + \gamma p\nabla\cdot(\mathbf{v}\cdot\nabla\mathbf{v})\right) \tag{C.24}$$

$$\mathcal{KH}s = kh\frac{t^2}{2}\left(\frac{\nabla p\cdot\nabla s}{\rho}\right) \tag{C.25}$$

$$\mathcal{KH}\rho = kh\frac{t^2}{2}\left(\mathbf{v}\cdot\nabla\rho(\nabla\cdot\mathbf{v}) + \rho\mathbf{v}\cdot(\nabla\nabla\cdot\mathbf{v}) + \frac{\nabla p\cdot\nabla\rho}{\rho}\right) \tag{C.26}$$

$$\mathcal{KH}u = kh\frac{t^2}{2}\left(-\frac{p_x(u\rho_x + v\rho_y)}{\rho^2} + \frac{up_{xx} + vp_{xy}}{\rho} + \frac{p_x u_x + p_y u_y}{\rho}\right) \tag{C.27}$$

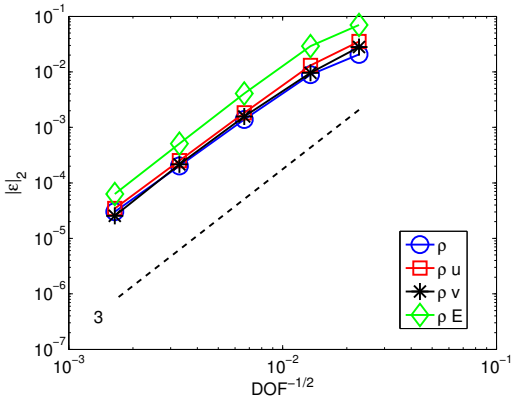
$$\mathcal{KH}u = kh\frac{t^2}{2}\left(-\frac{p_y(u\rho_x + v\rho_y)}{\rho^2} + \frac{up_{xy} + vp_{yy}}{\rho} + \frac{p_x v_x + p_y v_y}{\rho}\right) \tag{C.28}$$

$$\mathcal{KH}p = kh\frac{t^2}{2}\left(\gamma\mathbf{v}\cdot\nabla p(\nabla\cdot\mathbf{v}) + \gamma p\mathbf{v}\cdot(\nabla\nabla\cdot\mathbf{v}) + \frac{\nabla p\cdot\nabla p}{\rho}\right) \tag{C.29}$$

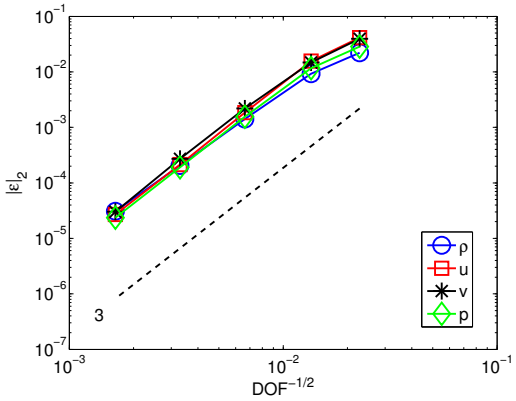
APPENDIX D

Numerical Results of the Two-Dimensional Euler System using Nonsymmetric Operator Splitting B

The isentropic moving vortex problem, which is described in Section 7.7, results using the nonsymmetric operator splitting B method are presented. The L_2 norm errors in conservative variables and solution reconstruction variables are displayed in Figs. D.1 and D.2 for the unstructured and structured grid type. The density error converges at a rate slightly less than $\mathcal{O}(h^3)$ while all other variables converge at the designed accuracy. The L_2 norm errors are summarized in Tables D.1 and D.2 for the unstructured and structured grid type.

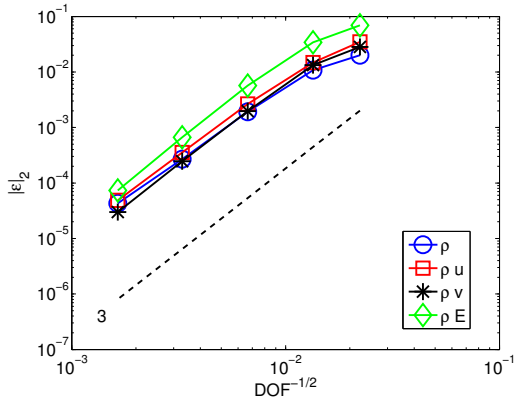


(a) $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables

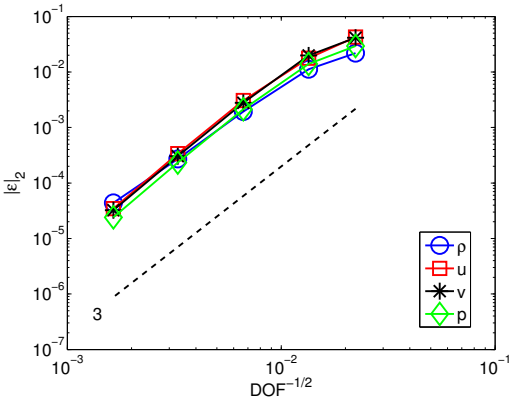


(b) $L_2(\mathbf{q}(\mathbf{x}))$ norm errors of solution reconstructions

Figure D.1: Nonsymmetric OS B error convergence results on unstructured grid at $t = 20$ and $\nu = 0.7$



(a) $L_2(\overline{\mathbf{u}})$ norm errors of conservative variables



(b) $L_2(\mathbf{q}(\mathbf{x}))$ norm errors of solution reconstructions

Figure D.2: Nonsymmetric OS B error convergence results on structured grid at $t = 20$ and $\nu = 0.7$

Table D.1: $L_2(\overline{\mathbf{u}})$ norm of conservative variable error on unstructured grid at $t = 20$ and $\nu = 0.7$

DOF	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order	$L_2(\overline{\rho E})$	Order
1.16e+03	2.07e-02		3.51e-02		6.99e-02	
3.32e+03	8.90e-03	1.60	1.31e-02	1.88	2.92e-02	1.66
1.39e+04	1.39e-03	2.60	1.83e-03	2.75	4.07e-03	2.76
5.62e+04	2.03e-04	2.76	2.51e-04	2.84	5.06e-04	2.98
2.27e+05	3.01e-05	2.73	3.44e-05	2.85	6.29e-05	2.99

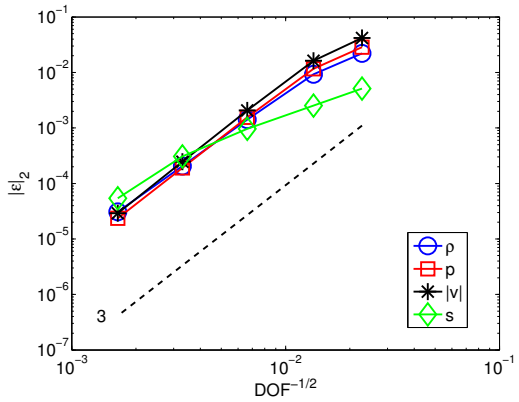
Table D.2: $L_2(\overline{\mathbf{u}})$ norm of conservative variable error on structured grid at $t = 20$ and $\nu = 0.7$

DOF	$L_2(\overline{\rho})$	Order	$L_2(\overline{\rho u})$	Order	$L_2(\overline{\rho E})$	Order
1.21e+03	2.01e-02		3.49e-02		6.93e-02	
3.34e+03	1.08e-02	1.23	1.47e-02	1.71	3.41e-02	1.40
1.37e+04	1.91e-03	2.45	2.65e-03	2.42	5.66e-03	2.54
5.65e+04	2.69e-04	2.77	3.55e-04	2.84	6.63e-04	3.03
2.26e+05	4.29e-05	2.65	4.85e-05	2.87	7.32e-05	3.18

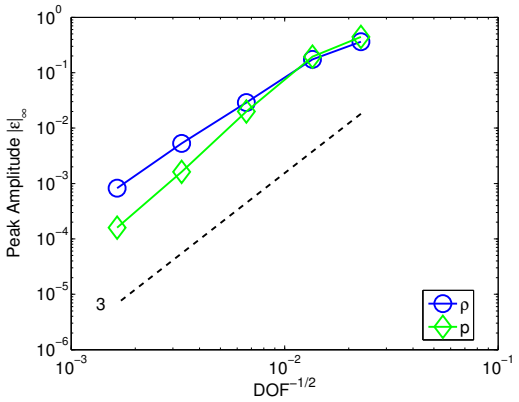
The nonsymmetric operator splitting method B, in fact method A shows the same trend, displays a concerning aspect of the operator splitting approach. The nonlinear interaction terms shown in Eqs. (7.73)-(7.76) in Section 7.4.2 were implemented with only the numerical accuracy in mind. We believe that the improper treatment of the interaction terms has

manifested into low density convergence. It has been known that errors in pressure or divergence give rise to acoustic waves that carry the errors away. However, errors in density or vorticity remain local.

The errors in vortex related quantities are presented in Figs. D.3 and D.4 for the unstructured and structured grid type. Tables D.3 and D.5 show the summary of errors for unstructured and structured grid type. Tables D.4 and D.6 show the summary of peak vortex amplitude errors for unstructured and structured grid types.

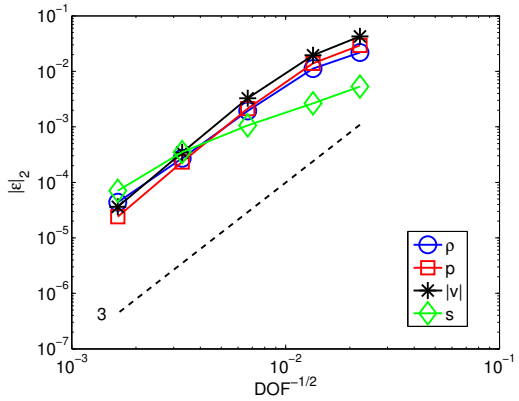


(a) Vortex scalar $L_2(\mathbf{q}(\mathbf{x}))$ norm error. $|\mathbf{v}|$ is a velocity magnitude.

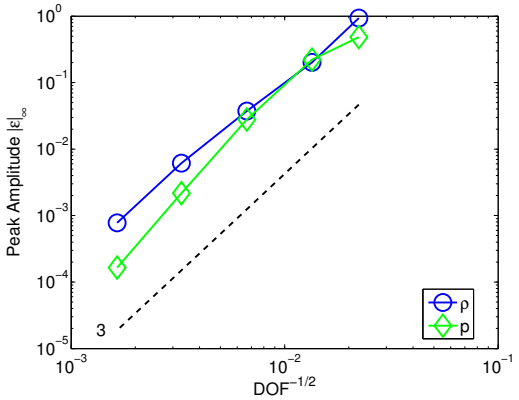


(b) Peak vortex amplitude error in density and pressure

Figure D.3: Vortex scalar quantity error convergence results on unstructured grid at $t = 20$ and $\nu = 0.7$



(a) Vortex scalar $L_2(\mathbf{q}(\mathbf{x}))$ norm error. $|\mathbf{v}|$ is a velocity magnitude.



(b) Peak vortex amplitude error in density and pressure

Figure D.4: Vortex scalar quantity error convergence results on structured grid at $t = 20$ and $\nu = 0.7$

Table D.3: Errors of some vortex scalar quantities on unstructured grid at $t = 20$ and $\nu = 0.7$

DOF	$L_2(\rho)$	Order	$L_2(p)$	Order	$L_2(\|\mathbf{v}\|)$	Order	$L_2(s)$	Order
1.16e+03	2.20e-02		2.85e-02		4.20e-02		5.13e-03	
3.32e+03	9.26e-03	1.64	1.16e-02	1.71	1.62e-02	1.81	2.53e-03	1.34
1.39e+04	1.41e-03	2.63	1.55e-03	2.82	2.08e-03	2.87	9.63e-04	1.35
5.62e+04	2.05e-04	2.76	1.90e-04	3.00	2.46e-04	3.05	3.06e-04	1.64
2.27e+05	3.07e-05	2.72	2.36e-05	2.99	2.91e-05	3.06	5.39e-05	2.49

Table D.4: Peak vortex amplitude errors on unstructured grid at $t = 20$ and $\nu = 0.7$

DOF	$x_{p_{\text{min}}}$	$y_{p_{\text{min}}}$	$L_\infty(\rho)$	Order	$L_\infty(p)$	Order
1.16e+03	-1.34e-01	-5.16e-02	3.63e-01		4.44e-01	
3.32e+03	4.09e-02	-6.22e-02	1.74e-01	1.40	1.94e-01	1.57
1.39e+04	3.21e-03	4.73e-05	2.88e-02	2.52	1.98e-02	3.19
5.62e+04	3.96e-03	2.14e-03	5.30e-03	2.42	1.61e-03	3.59
2.27e+05	-5.55e-04	-9.49e-04	8.21e-04	2.67	1.60e-04	3.32

Table D.5: Errors of some vortex scalar quantities on structured grid at $t = 20$ and $\nu = 0.7$

DOF	$L_2(\rho)$	Order	$L_2(p)$	Order	$L_2(\|\mathbf{v}\|)$	Order	$L_2(s)$	Order
1.21e+03	2.18e-02		2.94e-02		4.27e-02		5.30e-03	
3.34e+03	1.12e-02	1.31	1.40e-02	1.47	1.95e-02	1.55	2.65e-03	1.37
1.37e+04	1.94e-03	2.48	2.13e-03	2.66	3.26e-03	2.53	1.06e-03	1.29
5.65e+04	2.72e-04	2.78	2.32e-04	3.13	3.46e-04	3.17	3.50e-04	1.57
2.26e+05	4.39e-05	2.63	2.40e-05	3.28	3.60e-05	3.27	7.08e-05	2.31

Table D.6: Peak vortex amplitude errors on structured grid at $t = 20$ and $\nu = 0.7$

DOF	$x_{p_{\text{min}}}$	$y_{p_{\text{min}}}$	$L_\infty(\rho)$	Order	$L_\infty(p)$	Order
1.21e+03	-2.64e-01	-1.52e+00	9.32e-01		4.82e-01	
3.34e+03	3.14e-03	1.22e-01	2.01e-01	3.02	2.23e-01	1.53
1.37e+04	-1.71e-04	-7.00e-03	3.74e-02	2.38	2.79e-02	2.94
5.65e+04	3.30e-03	3.30e-03	6.14e-03	2.55	2.17e-03	3.61
2.26e+05	1.00e-03	-1.05e-04	7.75e-04	2.99	1.65e-04	3.72

D.1 Discussion – Numerical Entropy Generation

In the absence of entropy generating mechanisms such as shocks and boundaries in the Euler equations, the entropy level must be maintained. In the isentropic vortex test problem, the initial entropy level is not perturbed by any sources and must be maintained constant in time.

Reducing the level of numerically generated entropy has been of interest for many due to its impact on the drag prediction in sensitive aerodynamic simulations [8]. The residual distribution (RD) schemes have shown good numerical entropy property compared to other numerical methods based on one-dimensional Riemann solvers. More notably, the decomposition of the Euler equations into hyperbolic and elliptic operators [64] have produced a noticeably better-behaved numerical entropy property although it was only available for two-dimensional steady flows.

The AF method, which prides itself on the notion of splitting via physical disturbances, should in principle behave much better in terms of the numerical entropy preservation. In fact, the individual advective and acoustic nonlinear schemes have shown to be entropy preserving. It is more difficult to confirm in the acoustic nonlinear scheme because there is no standalone conservation system for acoustics, which in a strict sense is not the AF scheme, but the nonconservative acoustic system has been confirmed as entropy preserving. The nonlinear advection scheme on the pressureless Euler system does not generate entropy because there is no mechanism for it. However, as shown in Fig. D.3, nonsymmetric operator splitting methods produce additional numerical entropy, especially compared to the results of the staggered operator splitting method see Section 7.7. At the present time, we believe the reason for this is due entirely to the improper discretization and evaluation of the nonlinear interaction terms.

APPENDIX E

Alternative Pressure Equation in Euler System

We try to find an alternative pressure equation formulation in the Euler system. The idea is that the pressure term can be divided into an isentropic and a non-isentropic term. The isentropic term would be directly related to the density equation and do not need to be explicitly evaluated, which would reduce the computational cost. And the non-isentropic pressure term would be obtained by considering entropy generating mechanisms only.

E.1 Alternative Pressure Form in Euler System

We recall Eq. (7.12), and the relation between pressure and density is defined from it.

$$p = Ke^{s/c_v}\rho^\gamma = C\rho^\gamma \tag{E.1}$$

where γ is the ratio of specific heats. The entropy s is constant along a particle path in a smooth flow and the constant C is dependent on the initial condition of entropy. Let us try to define pressure without an explicit assumption of constant entropy.

$$dp = dp|_{s\neq\text{const}} + dp|_{s=\text{const}} \tag{E.2}$$

$$= \frac{K}{c_v}e^{s/c_v}\rho^\gamma ds + Ke^{s/c_v}d(\rho^\gamma) \tag{E.3}$$

$$= \frac{p(s)}{c_v}ds + dp|_{s=\text{const}} \tag{E.4}$$

Notice that this expression reduces to isentropic pressure expression when the entropy is constant along a particle path.

Formerly, we have assumed that each primitive variable can be expanded and advanced in time according to the Taylor approximation.

$$\mathbf{q}(t) = \mathbf{q}_0 + t\mathbf{q}_t + \frac{t^2}{2}\mathbf{q}_{tt} \tag{E.5}$$

We first determine the pressure under the constant entropy assumption, $p|_{s=\text{const}}$.

$$\begin{aligned} p|_{s=\text{const}}(t) &= p_0 + tp_t + \frac{t^2}{2}p_{tt} \\ &= f_0(\rho_0) + tf_1(\rho_t) + \frac{t^2}{2}f_2(\rho_t, \rho_{tt}) \end{aligned} \tag{E.6}$$

where f_0, f_1 and f_2 are functions of density expansion terms. It is clear that functions for the zeroth- and first-order terms are easily defined.

$$f_0(\rho_0) = C\rho_0^\gamma \tag{E.7}$$

$$f_1(\rho_t) = \frac{\gamma p}{\rho}\rho_t = a^2\rho_t \tag{E.8}$$

where $a^2 = \frac{\partial p}{\partial \rho}\big|_{s=\text{const}} = \frac{\gamma p}{\rho}$. We assume that f_2 is derived from f_1.

$$\begin{aligned} f_2(\rho_t, \rho_{tt}) &= (f_1(\rho_t))_t \\ &= (a^2)_t\rho_t + a^2\rho_{tt} \end{aligned} \tag{E.9}$$

The derivative of speed of sound squared is defined as follows.

$$\begin{aligned} (a^2)_t &= \left(\frac{\gamma p}{\rho}\right)_t \\ &= \frac{\gamma p_t}{\rho} - \frac{\gamma p}{\rho^2}\rho_t = a^2\left(\frac{\gamma - 1}{\rho}\right)\rho_t \end{aligned} \tag{E.10}$$

Furthermore, we assume that $f_2 = (f_{2,\text{Adv}} + f_{2,\text{Aco}})$ has two components in advection and acoustics operations. We expand and simplify the advection operator component.

$$\begin{aligned} f_{2,\text{Adv}}(\rho_t, \rho_{tt}) &= (a^2)_t\rho_t + a^2\rho_{tt} \\ &= a^2\left(\frac{\gamma - 1}{\rho}\right)\rho_t^2 + a^2\rho_{tt} \\ &= a^2\left(\frac{\gamma - 1}{\rho}\right)(-\mathbf{u}\cdot\nabla\rho)^2 + a^2\left((\mathbf{u}\cdot\nabla\mathbf{u})\cdot\nabla\rho + \mathbf{u}\cdot\nabla(\mathbf{u}\cdot\nabla\rho)\right) \\ &= a^2\left[u^2\left\{\left(\frac{\gamma - 1}{\rho}\rho_x^2 + \rho_{xx}\right)\right\} + 2uv\left\{\left(\frac{\gamma - 1}{\rho}\rho_x\rho_y + \rho_{xy}\right)\right\} + \right. \\ &\qquad \left. v^2\left\{\left(\frac{\gamma - 1}{\rho}\rho_y^2 + \rho_{yy}\right)\right\} + 2(uu_x + vu_y)\rho_x + 2(uv_x + vv_y)\rho_y\right] \\ &\equiv p_{tt,\text{Adv}} \end{aligned} \tag{E.11}$$

Terms in $\{\}$ represent second-derivatives of the pressure. Similarly, the acoustics operator component is expanded and simplified.

$$\begin{aligned} f_{2,\text{Aco}}(\rho_t, \rho_{tt}) &= (a^2)_t \rho_t + a^2 \rho_{tt} \\ &= a^2 \left(\frac{\gamma - 1}{\rho} \right) \rho_t^2 + a^2 \rho_{tt} \\ &= a^2 \left(\frac{\gamma - 1}{\rho} \right) (-\rho \nabla \cdot \mathbf{u})^2 + a^2 \left(\nabla^2 p - \frac{1}{\rho} \nabla p \cdot \nabla \rho + \rho (\nabla \cdot \mathbf{u})^2 \right) \\ &= a^2 \left[\nabla^2 p - \frac{1}{\rho} \nabla p + \gamma \rho (\nabla \cdot \mathbf{u})^2 \right] \\ &\equiv p_{tt,\text{Aco}} \end{aligned} \tag{E.12}$$

Therefore, f_2 is indeed a function of ρ_t, ρ_{tt} and a^2 for both advection and acoustics operators.

Lastly, the entropy dependent pressure term, $p|_{s \neq \text{const}}$, is discussed. The differential entropy is

$$ds = c_v \left(\frac{dp}{p} - \gamma \frac{d\rho}{\rho} \right). \tag{E.13}$$

This can be thought of as the change or production of entropy due to independent changes in pressure and density.

$$d^2 s = c_v \left(\frac{d^2 p}{p} - \gamma \frac{d^2 \rho}{\rho} - \left(\frac{dp^2}{p^2} - \gamma \frac{d\rho^2}{\rho^2} \right) \right) \tag{E.14}$$

We derive the change in entropy in time. Because density and pressure can independently evolve according to the partial differential equations, any mismatch or disturbance would

be manifested into a generation of the numerical entropy.

$$\frac{ds}{c_v}(t) = ts_t + \frac{t^2}{2}s_{tt} \tag{E.15}$$

$$= t\left(\frac{p_t}{p} - \gamma\frac{\rho_t}{\rho}\right) + \frac{t^2}{2}\left(\frac{p_{tt}}{p} - \gamma\frac{\rho_{tt}}{\rho} - \left(\frac{p_t^2}{p^2} - \gamma\frac{\rho_t^2}{\rho^2}\right)\right) \tag{E.16}$$

$$\begin{aligned} &= tk\left(-\frac{\mathbf{u}\cdot\nabla p}{p} + \gamma\frac{\mathbf{u}\cdot\nabla\rho}{\rho}\right) \\ &\quad + \frac{t^2}{2}\left(hk\left[\frac{\nabla p\cdot\nabla p}{p\rho} - \gamma\frac{\nabla p\cdot\nabla\rho}{\rho^2}\right] + k^2\left[-\frac{(\mathbf{u}\cdot\nabla p)^2}{p^2} + \gamma\frac{(\mathbf{u}\cdot\nabla\rho)^2}{\rho^2}\right] + \right. \\ &\quad + k^2\left[\frac{\mathbf{u}\cdot\nabla(\mathbf{u}\cdot\nabla p)}{p} + \frac{(\mathbf{u}\cdot\nabla\mathbf{u})\cdot\nabla p}{p}\right. \\ &\qquad\qquad \left.\left.-\gamma\left(\frac{\mathbf{u}\cdot\nabla(\mathbf{u}\cdot\nabla\rho)}{\rho} + \frac{(\mathbf{u}\cdot\nabla\mathbf{u})\cdot\nabla\rho}{\rho}\right)\right]\right) \end{aligned} \tag{E.17}$$

We notice that the first and second-order terms would cancel each other out under the isentropic flow condition. Then pressure for nonconstant entropy is defined

$$p|_{s\neq\text{const}}(t) = p_0 + tp_t + \frac{t^2}{2}p_{tt} \tag{E.18}$$

$$= p_0 + \frac{p(s)}{c_v}ds \tag{E.19}$$

$$= p_0 + tp(s)\left(\frac{p_t}{p} - \gamma\frac{\rho_t}{\rho}\right) + \frac{t^2}{2}p(s)\left(\frac{p_{tt}}{p} - \gamma\frac{\rho_{tt}}{\rho} - \left(\frac{p_t^2}{p^2} - \gamma\frac{\rho_t^2}{\rho^2}\right)\right) \tag{E.20}$$

It is evident that if the isentropic condition is assumed, it would result in no change in pressure $p|_{s\neq\text{const}} = p_0$.

$$\begin{aligned} dp &= \frac{t^2}{2}kh\left(2\gamma\mathbf{u}\cdot\nabla p(\nabla\cdot\mathbf{u}) + 2\gamma p(\mathbf{u}\cdot(\nabla\nabla\cdot\mathbf{u})) + \frac{\nabla p\cdot\nabla p}{\rho} + \right. \\ &\qquad\qquad \left.\gamma p(u_x^2 + 2u_y v_x + v_y^2)\right) \\ &= a^2\delta\rho \end{aligned} \tag{E.21}$$

Based on this analysis, it should be possible to construct a version of the AF method having this property. If presented with initial data that is in some discrete sense isentropic, the scheme should return a solution that is isentropic in the same sense.

BIBLIOGRAPHY

[1] M. B. Abbott. *An Introduction to the Method of Characteristics*. New York: American Elsevier, 1966.

[2] Rémi Abgrall. Residual distribution schemes: current status and future trends. *Computers & Fluids*, 35(7):641–669, 2006.

[3] Rémi Abgrall. A review of residual distribution schemes for hyperbolic and parabolic problems: the July 2010 state of the art. *Communications in Computational Physics*, 11(04):1043–1080, 2012.

[4] Rémi Abgrall and Mohamed Mezine. Construction of second order accurate monotone and stable residual distribution schemes for unsteady flow problems. *Journal of Computational Physics*, 188(1):16–55, 2003.

[5] Ryosuke Akoh, Satoshi Ii, and Feng Xiao. A CIP/multi-moment finite volume method for shallow water equations with source terms. *International journal for numerical methods in fluids*, 56(12):2245–2270, 2008.

[6] Ryosuke Akoh, Satoshi Ii, and Feng Xiao. A multi-moment finite volume formulation for shallow water equations on unstructured mesh. *Journal of Computational Physics*, 229(12):4567–4590, 2010.

[7] Steven R. Allmaras, John E. Bussoletti, Craig L. Hilmes, Forrester T. Johnson, Robin G. Melvin, Edward N. Tinoco, Venkat Venkatakrishnan, Laurence B. Wigton, and David P. Young. Algorithm issues and challenges associated with the development of robust CFD codes. In *Variational Analysis and Aerospace Engineering*, pages 1–19. Springer, 2009.

[8] Jérôme Bastin and Gilbert Rogé. A multidimensional fluctuation splitting scheme for the three dimensional euler equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 33(6):1241–1259, 1999.

[9] Jay P. Boris and David L. Book. Flux-corrected transport. I. SHASTA, a fluid transport algorithm that works. *Journal of computational physics*, 11(1):38–69, 1973.

[10] Daniel Bouche, G. Bonnaud, and D. Ramos. Comparison of numerical schemes for solving the advection equation. *Applied mathematics letters*, 16(2):147–154, 2003.

[11] F. Bouchut and G. Bonnaud. Numerical simulation of relativistic plasmas in hydrodynamic regime. *ZEITSCHRIFT FUR ANGEWANDTE MATHEMATIK UND MECHANIK*, 76:287–290, 1996.

[12] P. N. Childs and K. W. Morton. Characteristic Galerkin methods for scalar conservation laws in one dimension. *SIAM Journal on Numerical Analysis*, 27(3):553–594, 1990.

[13] Bernardo Cockburn. Discontinuous Galerkin methods for convection-dominated problems. In *High-order methods for computational physics*, pages 69–224. Springer, 1999.

[14] Bernardo Cockburn and Chi-Wang Shu. Runge–Kutta discontinuous Galerkin methods for convection-dominated problems. *Journal of scientific computing*, 16(3):173–261, 2001.

[15] Phillip Colella. Multidimensional upwind methods for hyperbolic conservation laws. *Journal of Computational Physics*, 87(1):171–200, 1990.

[16] Phillip Colella and Paul R. Woodward. The piecewise parabolic method (PPM) for gas-dynamical simulations. *Journal of computational physics*, 54(1):174–201, 1984.

[17] Richard Courant and David Hilbert. Methods of mathematical physics, vol. II. *Interscience, New York*, 1962.

[18] Richard Courant, Eugene Isaacson, and Mina Rees. On the solution of nonlinear hyperbolic differential equations by finite differences. *Communications on Pure and Applied Mathematics*, 5(3):243–255, 1952.

[19] H. Deconink, R. Struijs, G. Bourgois, and P. L. Roe. Compact advection schemes on unstructured grids. *VKI Lecture Series*, 2004, 1993.

[20] Michael Dumbser, Olindo Zanotti, Raphaël Loubère, and Steven Diot. A posteriori subcell limiting of the discontinuous Galerkin finite element method for hyperbolic conservation laws. *Journal of Computational Physics*, 278:47–75, 2014.

[21] Timothy A. Eymann. *Active Flux Schemes*. PhD thesis, The University of Michigan, 2013.

[22] Timothy A. Eymann and Philip L. Roe. Multidimensional Active Flux schemes. In *21st AIAA computational fluid dynamics conference*, page 2940, 2013.

[23] Doreen Fan and Philip L. Roe. Investigations of a new scheme for wave propagation. In *22nd AIAA Computational Fluid Dynamics Conference*, page 2449, 2015.

[24] Michael Fey. Multidimensional upwinding. Part II. Decomposition of the Euler equations into advection equations. *Journal of Computational Physics*, 143(1):181–199, 1998.

[25] Kurt O. Friedrichs. Symmetric hyperbolic linear differential equations. *Communications on pure and applied Mathematics*, 7(2):345–392, 1954.

[26] Sergei Konstantinovich Godunov. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik*, 89(3):271–306, 1959.

[27] Vasilii Mikhailovich Goloviznin and Aleksander Andreevich Samarskii. Finite difference approximation of convective transport equation with space splitting time derivative. *Matematicheskoe Modelirovanie*, 10(1):86–100, 1998.

[28] Jacques Hadamard. *Lectures on Cauchy's problem in linear partial differential equations*. Courier Corporation, 2014.

[29] Ami Harten, Bjorn Engquist, Stanley Osher, and Sukumar R. Chakravarthy. Uniformly high order accurate essentially non-oscillatory schemes, III. *Journal of computational physics*, 71(2):231–303, 1987.

[30] Charles Hirsch. *Numerical Computation of Internal and External Flows. Volumne 1: Fundamentals of Numerical Discretization*. John Wiley & Sons, 1990.

[31] Changqing Hu and Chi-Wang Shu. Weighted essentially non-oscillatory schemes on triangular meshes. *Journal of Computational Physics*, 150(1):97–127, 1999.

[32] M. E. Hubbard and A. L. Laird. Achieving high-order fluctuation splitting schemes by extending the stencil. *Computers & fluids*, 34(4):443–459, 2005.

[33] M. E. Hubbard and P. L. Roe. Compact high-resolution algorithms for time-dependent advection on unstructured grids. *International journal for numerical methods in fluids*, 33(5):711–736, 2000.

[34] H. T. Huynh. A piecewise-parabolic dual-mesh method for the Euler equations. In *12th Computational Fluid Dynamics Conference*, page 1739, 1995.

[35] H. T. Huynh. A flux reconstruction approach to high-order schemes including discontinuous Galerkin methods. In *18th AIAA Computational Fluid Dynamics Conference*, page 4079, 2007.

[36] H. T. Huynh. High-order methods including discontinuous Galerkin by reconstructions on triangular meshes. *AIAA Paper*, 44:2011, 2011.

[37] Satoshi Ii and Feng Xiao. CIP/multi-moment finite volume method for euler equations: a semi-Lagrangian characteristic formulation. *Journal of Computational Physics*, 222(2):849–871, 2007.

[38] Satoshi Ii and Feng Xiao. High order multi-moment constrained finite volume method. Part I: Basic formulation. *Journal of Computational Physics*, 228(10):3669–3707, 2009.

[39] Nikolaj Nikolaevic Janenko. *The method of fractional steps*. Springer, 1971.

[40] Guang-Shan Jiang and Chi-Wang Shu. Efficient implementation of weighted ENO schemes. *Journal of computational physics*, 126(1):202–228, 1996.

[41] S. A. Karabasov and V. M. Goloviznin. Compact accurately boundary-adjusting high-resolution technique for fluid dynamics. *Journal of Computational Physics*, 228(19):7426–7451, 2009.

[42] David A. Kopriva and John H. Kolias. A conservative staggered-grid Chebyshev multidomain method for compressible flows. *Journal of computational physics*, 125(1):244–261, 1996.

[43] Lilia Krivodonova. Limiters for high-order discontinuous Galerkin methods. *Journal of Computational Physics*, 226(1):879–896, 2007.

[44] Lilia Krivodonova, J. Xin, J. F. Remacle, Nicolas Chevaugeon, and Joseph E. Flaherty. Shock detection and limiting with discontinuous Galerkin methods for hyperbolic conservation laws. *Applied Numerical Mathematics*, 48(3-4):323–338, 2004.

[45] Nipun Kwatra, Jonathan Su, Jón T Grétarsson, and Ronald Fedkiw. A method for avoiding the acoustic time step restriction in compressible flow. *Journal of Computational Physics*, 228(11):4146–4161, 2009.

[46] Peter Lax and Burton Wendroff. Systems of conservation laws. *Communications on Pure and Applied mathematics*, 13(2):217–237, 1960.

[47] Randall J. LeVeque. *Numerical methods for conservation laws*. Springer Science & Business Media, 1992.

[48] Randall J. LeVeque. High-resolution conservative algorithms for advection in incompressible flow. *SIAM Journal on Numerical Analysis*, 33(2):627–665, 1996.

[49] Yingjie Liu, Chi-Wang Shu, Eitan Tadmor, and Mengping Zhang. Central discontinuous Galerkin methods on overlapping cells with a nonoscillatory hierarchical reconstruction. *SIAM Journal on Numerical Analysis*, 45(6):2442–2467, 2007.

[50] Maria Lukácová-Medvid'ová, J Saibertová, and Gerald Warnecke. Finite volume evolution Galerkin methods for nonlinear hyperbolic systems. *Journal of Computational Physics*, 183(2):533–562, 2002.

[51] Maria Lukacova-Medvid'ova and Gerald Warnecke. Lax-Wendroff type second order evolution Galerkin methods for multidimensional hyperbolic systems. *EAST WEST JOURNAL OF NUMERICAL MATHEMATICS*, 8(2):127–152, 2000.

[52] Tyler B. Lung. *Toward a Simple, Accurate Lagrangian Hydrocode*. PhD thesis, The University of Michigan, 2015.

[53] Tyler B. Lung and Philip L. Roe. Toward a reduction of mesh imprinting. *International Journal for Numerical Methods in Fluids*, 76(7):450–470, 2014.

[54] Hong Luo, Joseph D. Baum, and Rainald Löhner. A hermite WENO-based limiter for discontinuous Galerkin method on unstructured grids. *Journal of Computational Physics*, 225(1):686–713, 2007.

[55] J. N. Lyness and D. Jespersen. Moderate degree symmetric quadrature rules for the triangle. *IMA Journal of Applied Mathematics*, 15(1):19–32, 1975.

[56] R. W. MacCormack. The effect of viscosity in hypervelocity impact cratering. *AIAA Paper No. 69-354*, 1969.

[57] K. W. Morton. Shock capturing, fitting and recovery. In *Eighth International Conference on Numerical Methods in Fluid Dynamics*, pages 77–93. Springer, 1982.

[58] K. W. Morton and Philip L. Roe. Vorticity-preserving Lax–Wendroff-type schemes for the system wave equation. *SIAM Journal on Scientific Computing*, 23(1):170–192, 2001.

[59] K. W. Morton, E. Süli, and Peixiong Lin. Characteristic Galerkin methods for hyperbolic problems. In *Nonlinear Hyperbolic Problems: Theoretical, Applied, and Computational Aspects*, pages 430–439. Springer, 1993.

[60] N. C. Nguyen, J. Peraire, and B. Cockburn. Hybridizable discontinuous Galerkin methods. In *Spectral and High Order Methods for Partial Differential Equations*, pages 63–84. Springer, 2011.

[61] Jaime Peraire, Cuong Nguyen, and Bernardo Cockburn. An embedded discontinuous Galerkin method for the compressible Euler and Navier-Stokes equations. In *20th AIAA Computational Fluid Dynamics Conference*, page 3228, 2011.

[62] Per-Olof Persson and Jaime Peraire. Sub-cell shock capturing for discontinuous Galerkin methods. In *44th AIAA Aerospace Sciences Meeting and Exhibit*, page 112, 2006.

[63] Jianxian Qiu and Chi-Wang Shu. Runge–Kutta discontinuous Galerkin method using WENO limiters. *SIAM Journal on Scientific Computing*, 26(3):907–929, 2005.

[64] Mani Rad. *A Residual Ristribution Approach to the Euler Equations that Preserves Potential Flow*. PhD thesis, The University of Michigan, 2001.

[65] W. H. Reed and T. R. Hill. Triangular mesh methods for the neutron transport equation. *Los Alamos Report LA-UR-73-479*, 1973.

[66] Robert D. Richtmyer and K. W. Morton. *Difference methods for initial value problems*. Interscience, 1967.

[67] A. V. Rodionov. Correlation between the discontinuous Galerkin method and MUSCL-type schemes. *Mathematical Models and Computer Simulations*, 8(3):285–300, 2016.

[68] Philip L Roe. Approximate riemann solvers, parameter vectors, and difference schemes. *Journal of computational physics*, 43(2):357–372, 1981.

[69] Philip L. Roe. Fluctuations and signals-a framework for numerical evolution problems. *Numerical methods for fluid dynamics*, 11, 1982.

[70] Philip L. Roe. Sonic flux formulae. *SIAM journal on scientific and statistical computing*, 13(2):611–630, 1992.

[71] Philip L. Roe. Discontinuous solutions to hyperbolic systems under operator splitting. In *Upwind and High-Resolution Schemes*, pages 470–490. Springer, 1997.

[72] Philip L. Roe. Linear bicharacteristic schemes without dissipation. *SIAM Journal on Scientific Computing*, 19(5):1405–1427, 1998.

[73] Philip L. Roe. Did numerical methods for hyperbolic problems take a wrong turning? *In print*, 2017.

[74] Philip L. Roe. Is discontinuous reconstruction reconstruction really a good idea? *In print*, 2017.

[75] Philip L. Roe, Tyler Lung, and Jungyeoul Maeng. New approaches to limiting. In *22nd AIAA Computational Fluid Dynamics Conference*, page 2913, 2015.

[76] Yu G. Rykov. Propagation of shock wave type singularities in equations of two-dimensional zero-pressure gas dynamics. *Mathematical Notes*, 66(5):628–635, 1999.

[77] Chi-Wang Shu. TVB uniformly high-order schemes for conservation laws. *Mathematics of Computation*, 49(179):105–121, 1987.

[78] Joseph L. Steger and R. F. Warming. Flux vector splitting of the inviscid gasdynamic equations with application to finite-difference methods. *Journal of computational physics*, 40(2):263–293, 1981.

[79] Gilbert Strang. On the construction and comparison of difference schemes. *SIAM Journal on Numerical Analysis*, 5(3):506–517, 1968.

[80] Yuzhi Sun, Zhi Jian Wang, and Yen Liu. Spectral (finite) volume method for conservation laws on unstructured grids: VI. Extension to viscous flow. *Journal of Computational Physics*, 215(1):41–58, 2006.

[81] Peter K. Sweby. High resolution schemes using flux limiters for hyperbolic conservation laws. *SIAM journal on numerical analysis*, 21(5):995–1011, 1984.

[82] H. Takewaki, A. Nishiguchi, and T. Yabe. Cubic interpolated pseudo-particle method (CIP) for solving hyperbolic-type equations. *Journal of computational physics*, 61(2):261–268, 1985.

[83] Bram van Leer. Towards the ultimate conservative difference scheme. II. Monotonicity and conservation combined in a second-order scheme. *Journal of computational physics*, 14(4):361–370, 1974.

[84] Bram van Leer. Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection. *Journal of computational physics*, 23(3):276–299, 1977.

[85] Bram van Leer. Towards the ultimate conservative difference scheme. V. A second-order sequel to godunov's method. *Journal of computational Physics*, 32(1):101–136, 1979.

[86] Bram van Leer. Flux-vector splitting for the Euler equation. In *Upwind and High-Resolution Schemes*, pages 80–89. Springer, 1997.

[87] Z. J. Wang and Yen Liu. The spectral difference method for 2d Euler equations on unstructured grids. In *17th AIAA Computational Fluid Dynamics Conference*, page 5112, 2005.

[88] Zhi Jian Wang. Spectral (finite) volume method for conservation laws on unstructured grids: I. Basic formulation. *Journal of Computational Physics*, 178(1):210–251, 2002.

[89] Zhi Jian Wang and Yen Liu. Spectral (finite) volume method for conservation laws on unstructured grids: II. Extension to two-dimensional scalar equation. *Journal of Computational Physics*, 179(2):665–697, 2002.

[90] Zhijian J. Wang, Krzysztof Fidkowski, Rémi Abgrall, Francesco Bassi, Doru Caraeni, Andrew Cary, Herman Deconinck, Ralf Hartmann, Koen Hillewaert, Hung T. Huynh, et al. High-order CFD methods: current status and perspective. *International Journal for Numerical Methods in Fluids*, 72(8):811–845, 2013.

[91] R. F. Warming, Paul Kutler, and Harvard Lomax. Second-and third-order noncentered difference schemes for nonlinear hyperbolic equations. *AIAA J*, 11(2):189–196, 1973.

[92] Feng Xiao, Ryosuke Akoh, and Satoshi Ii. Unified formulation for compressible and incompressible flows by using multi-integrated moments II: Multi-dimensional version for compressible and incompressible flows. *Journal of Computational Physics*, 213(1):31–56, 2006.

[93] Zhiliang Xu, Yingjie Liu, and Chi-Wang Shu. Hierarchical reconstruction for discontinuous Galerkin methods on unstructured grids with a WENO-type linear reconstruction and partial neighboring cells. *Journal of Computational Physics*, 228(6):2194–2212, 2009.

[94] Takashi Yabe, R. Tanaka, T. Nakamura, and Feng Xiao. An exactly conservative semi-Lagrangian scheme (CIP–CSL) in one dimension. *Monthly Weather Review*, 129(2):332–344, 2001.

[95] Yang Yang, Dongming Wei, and Chi-Wang Shu. Discontinuous Galerkin method for Krauses consensus models and pressureless Euler equations. *Journal of Computational Physics*, 252:109–127, 2013.

[96] Helen C. Yee, Neil D. Sandham, and M. Jahed Djomehri. Low-dissipative high-order shock-capturing methods using characteristic-based filters. *Journal of computational physics*, 150(1):199–238, 1999.

[97] Manwai Yuen. Some exact blowup solutions to the pressureless Euler equations in RN. *Communications in Nonlinear Science and Numerical Simulation*, 16(8):2993–2998, 2011.

[98] Steven T. Zalesak. Fully multidimensional flux-corrected transport algorithms for fluids. *Journal of computational physics*, 31(3):335–362, 1979.

[99] Ya B Zel'Dovich. Gravitational instability: An approximate theory for large density perturbations. *Astronomy and astrophysics*, 5:84–89, 1970.

[100] G. C. Zha and E. Bilgen. Numerical solutions of Euler equations by using a new flux vector splitting scheme. *International Journal for Numerical Methods in Fluids*, 17(2):115–144, 1993.

[101] Mengping Zhang and Chi-Wang Shu. An analysis of and a comparison between the discontinuous Galerkin and the spectral finite volume methods. *Computers & fluids*, 34(4):581–592, 2005.

[102] Xiangxiong Zhang and Chi-Wang Shu. On maximum-principle-satisfying high order schemes for scalar conservation laws. *Journal of Computational Physics*, 229(9):3091–3120, 2010.

[103] Xiangxiong Zhang and Chi-Wang Shu. Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 467, pages 2752–2776. The Royal Society, 2011.