ANALYSIS AND DESIGN OF OPTIMAL DISCONTINUOUS FINITE ELEMENT SCHEMES

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF AERONAUTICS AND ASTRONAUTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Kartikey Asthana
June 2016

© 2016 by Kartikey Asthana. All Rights Reserved.

Re-distributed by Stanford University under license with the author.

This work is licensed under a Creative Commons Attribution-Noncommercial 3.0 United States License.
http://creativecommons.org/licenses/by-nc/3.0/us/

This dissertation is online at: http://purl.stanford.edu/bq419xn0538

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Antony Jameson, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Sanjiva Lele

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Peter Pinsky

Approved for the Stanford University Committee on Graduate Studies.

Patricia J. Gumport, Vice Provost for Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.

Abstract

Discontinuous finite-element schemes have the potential to provide highly accurate solutions to differential equations arising in a wide variety of applications ranging across mechanics, electro-magnetics, molecular chemistry, quantitative finance etc. Additionally, their design is intrinsically parallel, facilitating efficient acceleration through Graphical Processing Units. However, the complexity in formulation makes it very challenging to rigorously quantify the accuracy and stability offered by these schemes even for simple linear problems in one-dimension. In this dissertation, a rigorous framework is proposed that utilizes eigensolution analysis to develop consistency and convergence theory in the context of linear, hyperbolic conservation laws. It is proven that the general Flux Reconstruction (FR) formulation is consistent and converges at a time-dependent rate starting from a short-time value associated with polynomial interpolation and eventually asymptoting to a long-time value associated with dispersion and dissipation. This has led to the identification of a new class of schemes, designated Super-convergent FR (SFR), that exhibits an enhanced rate of convergence. Similarly, analysis of error for under-resolved simulations has led to the identification of a new set of high-order schemes, designated Optimal FR (OFR), that minimizes wave propagation errors for the range of resolvable wavenumbers. Finally, a computationally efficient, spectral filtering operation is proposed to stabilize solutions to non-linear, hyperbolic conservation laws. This has enabled accurate simulations of flows involving shock discontinuities even for extremely high polynomial orders. The filtered solution provides reduced total variation, reduced maximum overshoot/undershoot, and allows sub-element shocks to be localized in the interior of an element.

Acknowledgements

The only appropriate beginning to this dissertation is through the expression of my gratitude to Prof. Antony Jameson. Without his supervision and encouragement, this thesis would not have been possible. In fact, many of the ideas pursued herein owe themselves to the emphasis that Prof. Jameson places on independent thought. Even beyond academics, the experiences of his life, ranging from sports and sports cars to economics and airplane design, have been truly inspiring. It has been my pleasure and privilege to be his 50th doctoral student.

I am very grateful to my thesis readers, Prof. Sanjiva Lele and Prof. Peter Pinsky, for their time and patience. The scope of this work has benefited immensely from Prof. Lele's observations and continual guidance. The content, itself, has its foundations in the courses by Prof. Pinsky who provided me with the enriching opportunity to learn the art of teaching as a course assistant. Indeed, their approval of this thesis is a great honor for me.

I would like to sincerely thank Prof. Gerritsen and Prof. Papanicolaou for serving on my oral defense committee. There is no doubt that my appreciation of Applied Mathematics, in all its beauty and breadth, stems from the classrooms of such excellent teachers. It is fair to say that CME302 and MATH238 have been two of my favorite classes at Stanford.

My stay at the university was supported by the Stanford Graduate Fellowship. In particular, I would like to thank the family of Thomas V Jones for their generosity. The importance of this fellowship program for international students cannot be emphasized enough.

My time at the Aerospace Computing Laboratory was made very pleasant by the company of jovial colleagues. I begin by thanking Jerry Watkins for co-authoring not only research papers but also rock songs, and for teaching me how to drive. Thanks to Josh Romero for explaining traditions in the American culture; to Abhishek Sheshadri for his

perspective on pure mathematics and the meaning of life; to Manuel for his positive outlook and practical advice; to Andy for introducing me to the Gewürztraminer and to Jacob, David, Jonathan, Zach and Freddy for begin awesome. Special thanks to Prof. David Williams, Jonathan Bull, Prof. Philipp Birken and Prof. Peter Vincent for their detailed comments and valuable feedback to my research.

Beyond the lab, much of my graduate life has been spent in the splendid company of Abhimanyu Banerjee, Akash Sarma, Akshay Subramaniam and Raunak Borker. I am thankful to them for all the food, drinks, movies, video games and late-night walks that we spent contemplating on matters of zero consequence.

Coming from a culturally entrenched Indian society, I recognize that it is still not common for one to pursue a doctoral degree in science from across the seas. I am eternally grateful to my parents, Kamal and Ranjana, for providing me with their unconditional love and support to break out from the norm. Their faith in my abilities is very empowering. I am blessed to have my elder brother, Siddharth, who has used the wisdom of his hard-earned experiences to serve as the pathfinder for my life. Finally, I am thankful to my girlfriend, Apoorva, for always lending a patient ear to my worries and my troubles in the several years that have gone by.

This acknowledgement cannot be complete without a very special mention of my undergraduate adviser, Prof. Tapan Sengupta who introduced me to computational fluid dynamics. Indeed, my fascination with von Neumann analysis is deeply rooted in my training at the High Performance Computing Laboratory in IIT Kanpur. Prof. Sengupta's lessons on integrity, diligence and persistence continue to motivate me through life. This thesis is dedicated to him.

Preface

The latter half of the twentieth century has witnessed the advent of computational methods for simulating fluid flow. Fueled by the rapid development of computer hardware, finite-volume approximations of conservation laws have relieved the applied physicist from the need for closed-form solutions. Modern researchers now find themselves in the easily accessible field of computational fluid dynamics (CFD), facilitated by the availability of myriad open-source as well as commercial software packages. Owing to classical results on gas dynamics for first-order [53] and second-order [136, 75] schemes, the finite-volume approach is now a mature technology. For the aerospace industry, modern computer codes provide a significantly cheaper, reproducible and reasonably robust alternative to physical experiments conducted in the wind-tunnel. Engineers of the last millennium have employed CFD to aid in the design [74] of most commercial airliners in operation today.

However, the impact of finite-volume schemes has been largely restricted to steady flows such as that during an aircraft in cruise. Other important flight configurations such as pitching, plunging, gusts, buffet and flutter have been tackled using conventional empirical techniques, flight experiments and proprietary wisdom. The presence of excessive numerical dissipation and dispersion in such low-order schemes has also impeded fundamental research in flow physics. For instance, in computational aero-acoustics [129], pressure disturbances originating at solid surfaces need to be propagated across long distances at the correct wave and group velocities and without spurious dissipation. Similarly, in direct numerical solution (DNS) of turbulent flows, numerical dissipation can have an unacceptable adulteration of the energy spectrum [97]. A similar problem manifests in the case of vortex-dominated flows such as that over helicopter blades or flapping wings.

Consequently, the last two decades in computational mechanics have been dedicated to

the development of high-order methods. The principal argument in their support stems from the observation that increasing the order of convergence offers an exponential reduction in error. This is in contrast to the polynomial reduction offered by an increase in resolution of the computational domain. Even in regard to design, high-order methods have the potential to provide the same level of accuracy with considerably reduced computational effort [62] as compared to traditional second-order methods.

Several approaches have been adopted to achieve high-order accuracy. The conventional practice has relied upon compact finite difference schemes [58, 88] that are particularly well suited for Cartesian meshes. These schemes have been responsible for remarkable success in fundamental explorations on rectilinear geometries involving DNS of compressible shear flows [87], hypersonic boundary layers [159], receptivity studies regarding transition [115] etc. Although extensions to curvilinear and deforming meshes have been proposed [145], the applicability of such schemes to industrial problems remains limited due to the difficulty involved in generating structured meshes for complex geometries. Consequently, the development of discontinuous finite-element methods has been a focal point for recent efforts. These methods, which can operate on unstructured meshes, combine the high-order accuracy of continuous finite-element schemes [68] with the localized stencils of finite-volume schemes. Specifically, the numerical solution in each element can be arbitrarily refined on account of its representation in an appropriate, often hierarchal, basis. In contrast to continuous finite-element schemes, the solution is allowed to be discontinuous at the element interfaces which yields an element-local mass matrix that can be easily inverted. Finally, and perhaps most importantly, these methods are amenable to massively parallel algorithms that can be distributed to the level of individual elements, paving the way for accelerators such as multi-core processors and Graphical Processing Units [29, 92, 156].

The most popular example within this class of methods is the Discontinuous Galerkin (DG) [39, 89, 106] formulation. This computationally tractable, high-order method has been used to tackle conservation laws arising in a very wide variety of applications ranging across fluid mechanics [12, 39], solid mechanics [108, 158], electromagnetics [63], semiconductors [33], polymer processing [14], stochastic control [34] etc. Similar to continuous finite-element schemes, the DG formulation approximates the weak form of the

conservation law. However, conservation of the flux is enforced at the element interfaces not by continuity of state variables but by the introduction of Riemann solutions as in finite-volume schemes. The DG formulation comes in two main flavors. The traditional one incorporates full-order integrations of fluxes within each element, often leading to significant computational costs, especially for nonlinear problems. The more recent flavor comprises the collocation-based nodal DG method which projects the flux function onto a set of nodes in the interior of the element. This helps to reduce computational cost while providing a very intuitive, nodal representation of the solution [62].

Another popular instance of discontinuous finite element methods is the Spectral Difference (SD) formulation, proposed initially as a staggered-grid Chebyshev multidomain method [81], and generalized later to simplexes [91]. Similar to DG, the numerical solution and flux in SD are both approximated locally through Lagrangian basis functions, but on different nodes warranting the name 'staggered-grid'. Another importance difference is that the SD formulation directly approximates the strong form of the conservation law leading to a simpler prescription and computational efficiency [90].

In 2007, Huynh [69] proposed a Flux Reconstruction (FR) approach to first-order hyperbolic systems that provides a generalized differential framework for discontinuous finite-element schemes on tensor-product elements. This formulation is unifying in that it recovers several existing high-order methods including the collocation-based nodal DG method and a version of the SD formulation. It also provides a broad scope for the derivation of new linearly stable schemes such as the g_2 scheme by Huynh himself, the Energy Stable FR (ESFR) family by Vincent et al. [142], and the CMFR family by Lopez et al. [93] which offers the promise of extending FR to elastodynamics.

In 2009, Huynh [71] extended the FR formulation to second-order hyperbolic systems on tensor-product elements, enabling the computation of problems with diffusion. Soon after, the ESFR family was extended to advection-diffusion problems [28] as well. The correction procedure was further generalized to triangular [70, 32, 152] and tetrahedral elements [154, 155], and to general non-linear, second-order conservation laws such as the Euler [31] and Navier-Stokes equations [151]. The surge in popularity of FR also motivated other frameworks such as Correction Procedure via Reconstruction (CPR) [49] and Lifting Collocation Penalty (LCP) [147].

Despite the aforementioned advances in FR, which have been concerned primarily with implementation, progress along the lines of fundamental analysis has been rather limited. For instance, there are no published results guaranteeing that the FR formulation is indeed consistent [86]! Along the same line, convergence of the numerical solution, or even stability as per Lax's criterion, is not assured. While these might get dismissed by a casual practitioner as merely theoretical details, there are some serious voids from a practical point of view as well. There are no analytical estimates, published as yet, for the rate of convergence (order of accuracy) of FR for either time-dependent or steady-state problems. Researchers in the past [30, 150] have relied on canonical numerical experiments as a proxy for these rates. Most often, in these experiments, the numerical solution is marched for a fixed amount of time using the DG scheme with fully-upwinded fluxes. Consequently, the corresponding observations fail to capture the long-time super-convergence exhibited by FR schemes. Similarly, the effect of interface flux, location of solution points and the choice of correction function on the accuracy of the scheme is not well understood.

Another major challenge faced by the FR formulation is in regard to stability for nonlinear fluxes. The nodal nature of the formulation naturally introduces aliasing errors [76] when a general flux is projected onto the polynomial space of the numerical solution. These errors can lead to instabilities especially if the weak form of the conservation law admits discontinuous solutions such as in transonic or supersonic flow. Jameson et al. have suggested the use of Gauss-Legendre solution points, or more generally quadrature points [150], to minimize the magnitude of aliasing errors. However, this is often insufficient by itself when the polynomial order of the scheme is high. Consequently, some of the most important problems in aerospace design are still out of reach of high-order discontinuous finite-element schemes.

In this dissertation, we undertake these three major challenges in a rigorous study supported by adequate numerical evidence. Correspondingly, the discussion is distributed in three parts as follows.

We begin with an eigensolution analysis for linear time-dependent problems to prove that the FR formulation is indeed consistent. Moreover, any FR scheme which is stable in the von-Neumann sense is provably convergent. Interestingly, the rate of convergence for such problems is a function of time, starting from a short-time rate associated with

polynomial interpolation and eventually asymptoting to a long-time rate associated with the spatial derivative operator. Both these rates can be inferred from a simple eigenvalue computation for any given FR scheme.

The eigensolution analysis also allows us to examine the effect of polynomial order, correction function, interface flux and solution points on the dispersion and dissipation properties of the formulation. Using this framework, a new set of linearly stable high-order FR schemes is proposed that minimizes wave propagation errors for the range of resolvable wavenumbers. The resulting optimal schemes, designated Optimal ESFR (OESFR) and Optimal FR (OFR), provide a higher resolving efficiency in comparison to the Discontinuous Galerkin (DG) scheme as well as standard high-order compact finite difference schemes.

The second part of the dissertation derives analytical estimates for the rate of convergence of FR for steady-state problems. This has the lead to the derivation of a special class of schemes, designated Super-convergent FR (SFR), which exhibits an enhanced rate of convergence. The DG scheme is recovered as one member of this family while the SD scheme is not. We also show that the rate of convergence for steady-state problems is identical to the short-time rate of convergence for time-dependent problems.

In the third and final part of the dissertation, we prove that stability of the FR solution can be ensured for non-linear fluxes through the addition of adequate artificial dissipation. Unfortunately, a direct application of dissipation would severely limit the rate of convergence of the solution. Towards this end, we pose viscosity as a spectral filtering operation implemented in the physical space via a strictly local convolution integral. Coupled with a discontinuity sensor, this approach provides a computationally efficient method that captures shock discontinuities while preserving accuracy in smooth regions of the solution, even for very, very high polynomial orders such as $P = 119$. The filtered solution provides reduced total variation, reduced maximum overshoot/undershoot, and allows sub-element shocks to be localized in the interior of an element.

Note that the results derived herein for the general FR formulation are indeed directly applicable to DG. This is valuable from the point of view of verification since the DG scheme has been thoroughly investigated in the last two decades. We also note the scope of this work has been restricted to tensor-product elements for analytical simplicity. However,

this is hardly a limitation since any triangular element can either be completely replaced by quadrilaterals [68] or derived from degenerated quadrilaterals with collapsed edges [110]. All computations recorded in this work have been performed using computer codes developed by members of the Aerospace Computing Laboratory.

Contents

Abstract iv

Acknowledgements v

Preface vii

I Consistency and convergence theory for linear, time-dependent problems 1

1 Introduction 2

2 Eigensolution analysis for 1-D advection 7

2.1 FR formulation for 1-D advection 7

2.2 Eigensolution of the semi-discrete conservation law 11

2.2.1 Stability, dispersion and dissipation 13

2.2.2 Modal weights and characteristic time scales of dissipation 16

2.2.3 Evolution of pointwise error 18

2.3 Effect of polynomial order, correction function and solution points 21

2.3.1 Effect of polynomial order 21

2.3.2 Effect of correction function 23

2.3.3 Effect of solution points 27

3 Consistency of the semi-discrete formulation 29

3.1 Eigenvalue decomposition of the polynomial differentiation matrix, D . . . 30

3.2 Asymptotic eigenvalue decomposition of the differentiation matrix, Q . . . 31

4 Convergence of the semi-discrete formulation 37

4.1 Proof of convergence . . . 37
4.2 Rate of convergence of the FR formulation . . . 39
4.2.1 Rate of convergence of pointwise error . . . 40
4.2.2 Rate of convergence of 'relative' error for upwinded schemes . . . 45
4.3 Numerical experiments in 2-D . . . 46
4.3.1 Linear advection-diffusion in 2-D . . . 48
4.3.2 Isentropic Euler vortex in 2D . . . 51

5 Schemes with minimal dispersion and dissipation 57

5.1 Constrained minimization problem in spectral space . . . 58
5.2 Optimal Energy Stable FR schemes (OESFR) . . . 60
5.3 Optimal FR schemes (OFR) . . . 63
5.4 Time integration and CFL restriction . . . 68
5.4.1 Fully-discrete dispersion relation for a multi-stage time scheme . . . 68
5.4.2 CFL restriction . . . 70
5.5 Numerical tests . . . 70
5.5.1 Case 1: Advection of a sharp Gaussian on a uniform grid . . . 72
5.5.2 Case 2: Advection of a high wavenumber packet on a uniform grid . 75
5.5.3 Case 3: Advection of a sharp Gaussian on a non-uniform grid . . . 77
5.5.4 Taylor-Green vortex . . . 77

II Convergence theory for linear, steady-state problems 84

6 Introduction 85

7 Error in solution to steady, advection-diffusion 89

7.1 Problem specification . . . 89
7.2 Error in the steady-state numerical solution . . . 91
7.3 Measures of error . . . 93

7.4 Derivative operators in Flux Reconstruction 94

8 Rate of convergence in the $\mathscr{L}^\delta(\Omega^\delta)$ semi-norm 98

8.1 Error in the first derivative operator 99
8.2 Error in the second derivative operator 109
8.3 Rate of convergence in the $\mathscr{L}^\delta(\Omega^\delta)$ semi-norm 113
8.4 Super-convergent FR (SFR) schemes 114

9 Rate of convergence in the $\ell^2(\Omega^\delta)$ norm 119

9.1 Eigensolution convergence estimate for the first derivative operator 120
9.2 Eigensolution convergence estimate for the second derivative operator . . . 122
9.3 Eigensolution convergence estimate for the steady-state solution 126

10 Numerical experiments 130

10.1 Rate of convergence for derivative operators 131
10.2 Rate of convergence for steady-state solution 134
10.3 Applicability to general multi-dimensional, non-linear, vector conservation laws, and non tensor-product elements 137
10.3.1 Tensor product elements in multiple dimensions 137
10.3.2 Triangles, tetrahedra and prisms 139
10.3.3 Non-linear vector conservation laws in multiple dimensions 140

III Stability theory and spectral filtering for non-linear, time-dependent problems 141

11 Introduction 142

12 Energy bounds and stabilization 147

12.1 Energy bounds for the numerical solution 148
12.2 Stabilization via artificial dissipation 151

13 Design of convolution filters 156

13.1 Artificial viscosity as a Fourier-filtering operation 157

13.2 Filtering through integral operators in physical space 159
13.3 Local filter formulation in 1-D 162
13.4 Local filter formulation in 2-D for tensor product elements 165
13.5 Discontinuity detection via the concentration method 167

14 Numerical experiments **171**
14.1 Linear advection 172
14.1.1 Rate of convergence for filtering without sensing 176
14.2 Inviscid Burgers equation 178
14.3 Sod's shock tube 181
14.4 Shu-Osher shock turbulence interaction 183
14.5 Rigid body rotation 186
14.6 Symmetric Riemann problem 188
14.7 Double Mach reflection 190
14.8 Transonic airfoil 192

A Impact of boundary conditions **198**

B Group velocity and mesh reflection **207**

Bibliography **215**

List of Tables

1.1 Rate of convergence, at the set of solution points, for the DG scheme of polynomial order P for linear, time-dependent problems [9]. 5

2.1 Normalized eigenvalues, half lives, and energy distribution for the DG scheme of order $P = 2$ for $\frac{kh}{P+1} = \frac{\pi}{6}$, unit grid-spacing $h = 1$ and unit wavespeed $c = 1$. 18

3.1 Full eigenvalues for the DG scheme of order $P = 2$ for $\frac{kh}{P+1} = 0$. 33

4.1 Rates of convergence for the physical and the spurious modes for the DG scheme of order $P = 2$ on Gauss Legendre solution points. The rates are calculated using $\frac{kh}{(P+1)\pi} = 0.01$ and $\frac{kh}{(P+1)\pi} = 0.02$. 42

4.2 Rates of convergence of the eigensystem components for DG using Gauss-Legendre solution points obtained from direct eigenvalue computation. . . . 45

5.1 Objective function (relative to DG) and internal zeros for the OESFR and OFR schemes. 61

5.2 Resolving efficiency e_1 for $\varepsilon = 0.01, 0.001$ for DG, OESFR, OFR and the C_+ scheme. 67

5.3 Limiting CFL values for the DG, OESFR, OFR and the C_+ scheme for RK44 and RK45. 71

5.4 Limiting CFL values for DG, OESFR, OFR and the C_+ scheme for RK33 and SSPRK(4,3). The C_+ scheme has only been derived only for RK3, RK44 and RK45. 71

5.5 Details of numerical tests conducted in Chapter 5. 72

8.1 $P = 4$: Parameters for the super-convergent schemes marked in Fig. 8.1b. . 115

9.1 Rates of convergence of the eigensystem components for the fully upwinded, $\alpha = 0$, DG scheme using Gauss-Legendre solution points - obtained via direct eigenvalue computation. 124
9.2 Rates of convergence of the eigensystem components for the centered, $\alpha = \frac{1}{2}$, DG scheme using Gauss-Legendre solution points - obtained via direct eigenvalue computation. 124

10.1 Rate of convergence for the first derivative operator, $\sigma^{(1)}$, in $\ell^2(\Omega^\delta)$ norm using DG. 132
10.2 Rate of convergence for the second derivative operator, $\sigma^{(2)}$, in $\ell^2(\Omega^\delta)$ norm, for isotropic second order operators using DG. 132
10.3 Rate of convergence for the second derivative operator, $\sigma^{(2)}$, in $\ell^2(\Omega^\delta)$ norm, for an anisotropic second order operator using DG. 132
10.4 $\ell^2(\Omega^\delta)$ norm of error in the first derivative operator for $P = 2$ using DG and another SFR scheme. 133
10.5 $\ell^2(\Omega^\delta)$ norm of error in the second derivative operator for $P = 2$ using DG and another SFR scheme. 133
10.6 Rate of convergence of error in the $\mathscr{L}^\delta(\Omega^\delta)$ semi-norm for the steady, forced, advection-diffusion problem using DG. 135
10.7 Rate of convergence of error in the $\ell^2(\Omega^\delta)$ norm for the steady, forced, advection-diffusion problem using DG. 135
10.8 Rate of convergence for the first and second derivative in the $\ell^2(\Omega^\delta)$ norm for the steady, forced, advection-diffusion problem using DG. 135
10.9 $\mathscr{L}^\delta(\Omega^\delta)$ norm of error in the steady-state numerical solution for $P = 2$ using DG and another SFR scheme. 136
10.10 $\ell^2(\Omega^\delta)$ norm of error in the steady-state numerical solution for $P = 2$ using DG and another SFR scheme. 136
10.11 Rate of convergence of error in the $\ell^2(\Omega^\delta)$ norm for the 2-D steady, forced, advection-diffusion problem using DG. 139

13.1 Maximum value of the enhanced kernel in Eqn. (13.43) over the solution points in $[-1,1]$. The functions are normalized with respect to the extrema over the point values. 170

14.1 Error metrics for the numerical solution of the linear advection problem . . 174
14.2 L_2 norm of error for the filtered solution at the crossover points indicated in Fig. 14.3 for the grid with $N = 16$ elements. 177
14.3 Error metrics for the numerical solution of Sod's shock tube problem. Here $M = u/\sqrt{\gamma p/\rho}$ 183
14.4 Error metrics for the numerical solution of the rigid body rotation problem . 188

List of Figures

2.1 Normalized eigenvalues: Real and imaginary parts of λ_p for each of the numerical modes $p = 1,2,3$ for the fully upwinded, $\alpha = 0$, DG scheme of order $P = 2$. 14

2.2 Normalized eigenvalues: Real and imaginary parts of λ_p for each of the numerical modes $p = 1,2,3$ for the centered, $\alpha = 0.5$, DG scheme of order $P = 2$. 15

2.3 Energy distribution: $|\beta|^2$ for the 3 eigenmodes of the DG scheme of order $P = 2$ using Gauss-Legendre solution points. 17

2.4 Evolution of the norm of pointwise error, Eqn. (2.29), against number of time periods, $T = 2\pi/(kc)$, for the DG scheme of order $P = 2$ using Gauss-Legendre solution points, $\frac{kh}{P+1} = \frac{\pi}{6}$, unit grid spacing $h = 1$ and unit wavespeed $c = 1$. 20

2.5 Effective wavenumber: k_p^{eff} for each of the numerical modes $p = 1,2,3$ for the DG scheme of order $P = 2$. 22

2.6 Effect of polynomial order on the dispersive and dissipative properties of the physical mode for the DG scheme of order $P = 2$. 24

2.7 Effect of correction function on the dispersive and dissipative properties of the physical mode for the DG scheme of order $P = 2$. 25

2.8 Effect of polynomial order on the fraction of energy retained in the physical mode for $P = 2$ using Gauss-Legendre solution points. 26

2.9 Effect of correction function on the fraction of energy retained in the physical mode for $P = 2$ using Gauss-Legendre solution points. 26

2.10 Effect of solution points on relative energies: Distribution of energy among numerical modes for DG scheme via FR for $P = 2$. 28

4.1 Rate of convergence for the fully upwinded DG scheme of order $P = 2$ on Gauss Legendre solution points with unit wavespeed $c = 1$. The reference time for relative error is taken to be $t_0 = T$. 43
4.2 Rate of convergence for the centered DG scheme of order $P = 2$ on Gauss Legendre solution points with unit wavespeed $c = 1$. 44
4.3 2-D advection-diffusion: Rate of convergence for the DG scheme of order $P = 2$ on Gauss Legendre solution points. 49
4.4 2-D advection-diffusion: Rate of convergence for the DG scheme of order $P = 3$ on Gauss Legendre solution points. 50
4.5 Isentropic vortex: Evolution of error for the DG scheme of order $P = 2$ with upwinded flux, $\alpha = 0$, on Gauss Legendre solution points. 54
4.6 Isentropic vortex: Evolution of error for the DG scheme of order $P = 3$ with upwinded flux, $\alpha = 0$, on Gauss Legendre solution points. 55
4.7 Isentropic vortex: Rate of convergence for the DG scheme with upwinded flux, $\alpha = 0$, on Gauss Legendre solution points. 56

5.1 OESFR: Variation of the objective function with the ESFR parameter C denoting the presence of a minimum close to the DG value of $C = 0$. 62
5.2 Optimal schemes: Dispersion and dissipation properties of the physical mode for DG, OESFR and OFR schemes for $P = 4$. 64
5.3 Optimal schemes: Dispersion and dissipation properties of the physical mode for DG, OESFR and OFR schemes for $P = 5$. 65
5.4 Left boundary correction functions for DG, OESFR and OFR schemes for (a) $P = 4$ and (b) $P = 5$. 66
5.5 Case 1: Snapshots of the numerical solution for the DG, OESFR, OFR and C_+ schemes at intervals of a quarter-period. 73
5.6 Case 1: Evolution of L^2 norm of the error for the DG, OESFR, OFR and C_+ schemes. 74

5.7 Case 1 with refined grids: Snapshots of the numerical solution for the DG, OESFR schemes using 61 elements, OFR scheme using 45 elements and C_+ scheme using 76 elements at intervals of a quarter-period. 76
5.8 Case 2: Snapshots of the numerical solution for the DG, OESFR, OFR and C_+ schemes at intervals of a quarter-period. 78
5.9 Case 2: Evolution of L^2 norm of the error for the DG, OESFR, OFR and C_+ schemes. 79
5.10 Case 3: (a) Variation of element size h_j across the perturbed, non-uniform grid, (b) Evolution of L^2 norm of the error for the DG, OESFR, OFR and C_+ schemes. 80
5.11 Case 3: Snapshots of the numerical solution for the DG, OESFR, OFR and C_+ schemes at intervals of a quarter-period. 81

8.1 Region of stability in the parameter space, Eqn. (8.43), for the SFR family in the case of pure-advection. 116
8.2 Left-boundary correction function for the superconvergent family of schemes defined in Section 8.4. 117

13.1 Local representation of the piecewise solution obtained by padding interface-information from neighboring elements. The dashed curves / translucent surfaces represent the original piecewise solution; the solid lines / solid surfaces denote the padded solution. 163

14.1 Numerical solution for linear advection at $t = 8.0$ for $P = 8, N = 100$: (a) Direct solution via FR without filtering; (b) enhanced concentration sum, Eqn. (13.43), applied to the direct solution; (c) & (d) Solutions obtained with sensing and filtering operations at each time step. 174
14.2 Grid convergence: Box-filtered solution for linear advection at $t = 8.0$ for $P = 8$ and $N = 6, 12, 25, 50$. 175

14.3 Rate of convergence for linear advection with a sinusoidal initial condition, $P = 1,2,3,4$, $N = 8,16$ at $t = 1$: The dotted lines represent the rates for the unfiltered solution. The solid and dashed curves represent the rates as a function of the filter width parameter when the solution is filtered in each element at every time step. 178
14.4 (a) Numerical solution with filtering, (b) corresponding point-wise error for Burgers equation at $t = 1.0$ for $P = 8$ and $N = 40$. (c),(d) similar results for $P = 7,9$ 179
14.5 (a) Numerical solution with filtering and (b) point-wise error for Burgers equation at $t = 1.0$ for $P = 119$ and $N = 3$. 180
14.6 Numerical solution with filtering for Sod's shock tube problem at $t = 0.4$ for $P = 8$ and $N = 56$ 182
14.7 Numerical solution with filtering for Shu-Osher shock turbulence interaction problem at $t = 2.0$ for $P = 8$ and $N = 56,112,224,448$. The reference solution is obtained via direct solution for $P = 0$, $N = 9600$. 184
14.8 Numerical solution for rigid body rotation at $t = 2\pi$ for $P = 8, N = 56 \times 56$: (a) Direct solution via FR without filtering; (b) enhanced concentration sum, Eqn. (13.43), applied to the direct solution along lines; (c) & (d) Solutions obtained with sensing and filtering operations at each time step. . 187
14.9 Numerical solution with isotropic box filtering for radially symmetric Riemann problem at $t = 0.1$ for $P = 8$ and $N = 56 \times 56$ 189
14.10 Numerical solution with isotropic box filtering for double Mach reflection problem at $t = 0.1$ for $P = 8$ and $N = 56 \times 224$. 191
14.11 Numerical solution with box filtering for Mach 0.8 flow past a NACA0012 airfoil at an AoA of 1.25 degrees. 193
14.12 Distribution of the coefficient of pressure on the surface of the airfoil. The plotted symbols correspond to flux points. 195
14.13 Convergence history of the residual and integrated coefficients. 196
14.14 Comparison of integrated coefficients with the converged solutions of Vassberg and Jameson [139]. The dotted lines represent 'continuum' values and the error bar corresponds to the current FR calculation. 197

A.1 Real part of the numerical and exact solutions in the first element for the fully upwinded DG scheme of order $P = 3$ on Gauss-Legendre points. . . . 200
A.2 $|\beta|^2$ for the 4 eigenmodes of the DG scheme of order $P = 3$ using Gauss-Legendre solution points. 202
A.3 Evolution of the norm of error in the first element for the fully upwinded DG scheme of order $P = 3$ on Gauss-Legendre points. Here, $T = \frac{2\pi}{\omega}$ is the time period of oscillations. 202
A.4 Quasi-steady amplitude of error $|A_p|$ for the 4 numerical modes for the fully upwinded DG scheme of order $P = 3$. 204
A.5 Rate of convergence against the ESFR parameter C for $P = 1, \ldots, 5$. 206

B.1 Energy weighted group velocity, $\hat{v}_g$, for the DG scheme of orders $P = 0, \ldots, 3$ using Gauss-Legendre solution points. 210
B.2 Snapshots of the numerical solution for the centered DG scheme of order $P = 2$. The dotted line represents the exact solution. 212
B.3 Snapshots of the numerical solution for the fully-upwinded, $\alpha = 0$, DG scheme of order $P = 2$. The dotted line represents the exact solution. 213
B.4 Snapshots of the numerical solution for the centered DG scheme of order $P = 3$. The dotted line represents the exact solution. 214

Part I

Consistency and convergence theory for linear, time-dependent problems

Chapter 1

Introduction

The tools for analysis of discontinuous finite element schemes borrow primarily from techniques in functional analysis. In case of the Discontinuous Galerkin (DG) method, results from approximation theory have been employed to derive *a-priori* [45, 78, 89] as well as *a-posteriori* [122, 124] error estimates for linear as well as non-linear time-dependent problems. For the Flux Reconstruction (FR) formulation, the introduction of special broken Sobolev-type norms has enabled proofs of linear stability [142], investigations regarding non-linear stability [76], and their extensions to multiple dimensions [32, 153, 155]. A salient feature of these methods is that they exhibit super-convergence [3, 40] for hyperbolic problems. For instance, Cheng and Shu [36], who studied the grid-convergence of error for DG using upwinded fluxes on arbitrary meshes, have shown that evolution of the L^2 norm of error is bounded by a linear function of time; the constant term converges at the order $P+1$, and the linear term converges at the order $P+3/2$ or higher. Hence, interestingly, the rate of convergence of error at short-time, $t = 0^+$ is different from that at long-time, $t \to \infty$. In this first part of the dissertation, we demonstrate that Fourier analysis can be used to rigorously investigate the super-convergence property for any scheme within the FR formulation. Along the way, we develop a complete theory for consistency and convergence, and explore the effect of interface flux, correction functions and the choice of solution points on the accuracy of the scheme.

Traditionally, Fourier analysis has been utilized to study the dispersive and dissipative properties of numerical schemes. Instances of such studies range across finite difference schemes [58, 129, 88], finite element schemes [130, 5], the DG method [65, 66, 67, 117, 98] etc. Such investigations are motivated by a non-asymptotic measure of accuracy which is based on the fraction of resolvable wavenumbers for which constituent waves can be numerically propagated with negligible dispersion and dissipation. This is of particular relevance for problems involving wave-propagation which appear naturally in electro-magnetics, aero-acoustics and studies of transition to turbulence. Specifically, the non-conservative form of a hyperbolic system provides a functional dependence of the wavespeeds on wavenumber which is denoted as the analytical dispersion relation [149]. Similarly, the discrete numerical system corresponding to a given scheme provides a numerical dispersion relation. With the exception of spectral methods [56], most numerical methods preserve the dispersion relation only for low wavenumbers, and introduce significant dispersion and/or dissipation beyond a critical wavenumber denoted as the resolving efficiency. However, spectral methods are not easily applicable to complex flow geometries. Spectral analysis of the numerical dispersion relation was utilized by Tam et al. [129] to obtain dispersion relation preserving finite difference schemes for aero-acoustics. This facilitated the development of several implicit compact difference schemes [1, 58, 64, 88, 113] with near-spectral resolution.

Spectral analysis of the DG method for wave propagation problems was carried out by Hussaini et al. [67]. It was shown that the P^{th} order DG method, applied to linear advection, results in one physical mode and P spurious or parasitic modes which get damped out quickly on account of relatively high dissipation. Rigorous studies regarding the behavior of dispersion and dissipation errors in the asymptotic limit of vanishing wavenumber have also been conducted. In [65], it was shown that the dispersion error for the physical mode is of order $2P+3$ and the dissipation error is of order $2P+2$. This was further rigorously established by Ainsworth [6, 4] who proved that for a fixed mesh of spacing h, these errors decay at an exponential rate when $2P+1 \approx \kappa kh$, where k is the wavenumber and $\kappa > 1$ is some constant.

Focusing on FR schemes, the first insight into spectral properties was provided by

Huynh himself [69], wherein Fourier analysis applied to 1-D advection, with fully upwinded Riemann flux at the interface, was utilized to recognize the presence of principal and spurious eigenvalues. Taylor expansion of the principal eigenvalue was used to define an order of accuracy in the spectral space and it was proven that this order is independent of the choice of solution points. Additionally, the variation of the principal eigenvalue over the range of resolvable wavenumbers was used to ascertain numerical stability limits for several FR schemes. Another study was conducted by Vincent et. al. [141] who investigated the wave properties of Energy Stable Flux Reconstrcution (ESFR) [142] schemes. In that investigation, the authors employed an automated procedure based on modal analysis to combine all the numerical modes into a single functional relationship between the wavespeed and the wavenumber. It was suggested that this functional relationship is representative of the dispersive and dissipative properties of the scheme and that multiple modes should not be identified distinctly. A highlight of the study was the identification of the 'C_+' scheme that offers the highest CFL limit obtained when the ESFR parameter C is chosen to be equal to C_+ for a given polynomial order.

In this first part of the dissertation, we show that an analytical form of the numerical solution, for the scalar linear advection problem, can be prescribed by posing the numerical conservation law as a linear dynamical system. The limit of asymptotic grid resolution can then be inferred from the limiting state of the eigensystem. We show that an *a-priori* bound on the rate of convergence can be obtained from the convergence rate of the eigenvalues and modal weights of the eigenvectors. The convergence rate is a function of time, and both the short-time, $t = 0^+$, and long-time asymptotic, $t \to \infty$, values are readily determined. The short-time rate of convergence depends on the convergence-rates of the modal weights for the 'spurious' modes, and the long-time rate depends on the convergence-rate of the 'physical' eigenvalue. Table 1.1 records these rates for the DG scheme, recovered via FR, for upwinded and centered fluxes, and odd and even P. In an analogous manner, rates can be determined for any stable scheme within the FR family. We also show that the rate of convergence obtained when wavenumber k is refined for a fixed grid spacing h, is one order higher than that when h is refined for a fixed k.

This part is formatted as follows. In Chapter 2, we describe the model problem under consideration and the eigensystem associated with the numerical solution. We explore the

DG scheme	Short-time $t = 0^+$	Long-time $t \to \infty$
one-sided fluxes	$P+1$	$2P+1$
centered fluxes and even P	$P+2$	$2P+2$
centered fluxes and odd P	P	$2P$

Table 1.1: Rate of convergence, at the set of solution points, for the DG scheme of polynomial order P for linear, time-dependent problems [9].

effect of interface flux, correction function and location of solution points on the dispersion, dissipation properties and the distribution of energy among the numerical eigenmodes. We see that in addition to polynomial order and choice of correction function, the choice of solution points is also important even for linear fluxes. In Chapter 3, we prove that the semi-discrete conservation law in FR is consistent for all admissible choices of interface flux, correction function and solution points. Chapter 4 proves that the rate of convergence for any FR scheme is a function of time and can be obtained directly from the associated eigenvalues and modal weights. Finally, in Chapter 5 we define a minimization problem to obtain FR schemes with minimal wave propagation errors. The semi-discrete as well as fully discrete dispersion relations are derived and the limiting CFL values for these optimal FR (OFR) schemes are compared with those for existing FR schemes. The resolving efficiency of these optimal schemes offer considerable improvement over the DG scheme and compare well with conventional compact finite difference schemes. All theoretical results are accompanied by numerical experiments that validate our findings.

The content of this part if largely drawn from two papers:

- K Asthana, A Jameson: High-order flux reconstruction schemes with minimal dispersion and dissipation. Journal of Scientific Computing, 2015.
- K Asthana, J Watkins, A Jameson: On consistency, convergence and rate of convergence of Flux Reconstruction for time-dependent problems, under review at SIAM Journal of Numerical Analysis.

A preliminary but rigorous discussion regarding the impact of boundary conditions on the rate of convergence is included in Appendix A. Analysis indicates that convergence in

elements that are aligned with active boundaries is limited to the short-time rate associated with interpolation. In other words, the long-time super-convergence exhibited for initial value problems is not exhibited for boundary-driven solutions.

Chapter 2

Eigensolution analysis for 1-D advection

This chapter begins with a brief description of the FR formulation to arrive at the update equation for 1-D scalar advection. We demonstrate that the numerical solution obtained using an FR scheme of order P leads to a linear dynamical system with $P+1$ eigenmodes. The eigenvalues of this system prescribe the numerical dispersion relation and, for upwinded schemes, are associated with characteristic time scales of dissipation. The eigenvectors prescribe the distribution of modal weights or energy among the numerical eigenmodes. These weights are obtained through the expansion of the discrete initial condition in the basis of eigenvectors. Such an eigensystem completely specifies the evolution of error in time. The chapter ends with a qualitative study of the effect of polynomial order, correction function and solution points on the dispersion, dissipation and modal distribution properties of FR schemes as a function of wavenumber.

2.1 FR formulation for 1-D advection

Consider the linear advection of an analytic scalar field, $u : \mathbb{R} \mapsto \mathbb{C}$, in 1-D,

$$\begin{aligned} \frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} &= 0, \quad x \in \mathbb{R},\, t > 0, \\ u(x,0) &= \exp(ikx), \end{aligned} \tag{2.1}$$

where $c \in \mathbb{R}$ is the constant wavespeed and the initial condition represents an isolated Fourier component. Such a choice for $u(x,0)$ is motivated by the linearity of the problem. The solution for a general initial condition, $U_0(x)$, amenable to the Fourier Transform, $U_0(x) = \int_{-\infty}^{\infty} \hat{U}(k)\exp(ikx)\mathrm{d}k$ can be obtained by linear superposition of solutions of Eqn. (2.1) for different k. Moreover, support on the entire real line, $x \in \mathbb{R}$, allows us to investigate the limit of asymptotic spectral resolution, $k \to 0$, in the subsequent chapters.

For a general FR scheme of order P, the domain $\Omega = \mathbb{R}$ is partitioned into non-overlapping elements, $\Omega = \bigcup_n \Omega_n$ where $\Omega_n = [x_n, x_{n+1})$ and the sequence of nodes, $\{x_n\}$, is increasing on the real line. Each element is further discretized with $P+1$ variably distributed, distinct solution points $\boldsymbol{x_n} = [x_{n,1}, x_{n,2}, \ldots, x_{n,P+1}]^T$. We adopt the notation of indicating vectors using boldface symbols as well as sub-script indexing so that $\boldsymbol{x_n} = (x_{n,p})_{p=1,2,\ldots,P+1}$. The numerical solution, $u^\delta(x,t)$, is defined as a piecewise interpolating polynomial such that in the n^{th} element,

$$u^\delta(x,t) = \sum_{p=1}^{P+1} u^\delta_{n,p}(t)\ell_{n,p}(x), \quad x \in \Omega_n, \tag{2.2}$$

where $\ell_{n,p}$ is the p^{th} Lagrange polynomial in the n^{th} element. The solution is allowed to be discontinuous at the element interfaces. Assuming exact integration in time, Eqn. (2.1) leads to the semi-discrete numerical conservation law,

$$\begin{aligned} \frac{d\boldsymbol{u^\delta_n}}{dt} + c\left.\frac{\partial^\delta \boldsymbol{u^\delta}}{\partial x}\right|_n = 0, \quad \forall n,\, t > 0, \\ \boldsymbol{u^\delta_n}(0) = \exp(ik\boldsymbol{x_n}), \end{aligned} \tag{2.3}$$

where $\boldsymbol{u^\delta_n}(t) = (u^\delta_{n,p}(t))_{p=1,2,\ldots,P+1}$ denotes the vector of solution values and $\left.\frac{\partial^\delta \boldsymbol{u^\delta}}{\partial x}\right|_n$ is the vector of numerically evaluated derivative values in the n^{th} element, which constitutes the defining feature of the scheme. In case of FR, definition of the numerical derivative follows from the requirement of conservation of flux. This is explicitly incorporated into the formulation in two steps. First, common interface values are defined between elements. These are computed from the polynomial functions on either side of the interface,

$$u^\delta_I(x_n) = (1-\alpha)u^\delta(x_n^-) + \alpha u^\delta(x_n^+), \tag{2.4}$$

where $\alpha \in \left[0, \frac{1}{2}\right]$ denotes the upwinding coefficient and $(.)^-$, $(.)^+$ denote the left and right hand limits respectively. The discontinuous flux, cu^δ, is then corrected to ensure continuity across elements through the use of special 'correction' functions of degree $P+1$. Towards this end, consider introducing a linear isoparametric map from the physical domain $x \in \Omega_n$ to the parent domain $\xi \in [-1,1)$ such that

$$\xi|_{\Omega_n}(x) = 2\frac{x - x_n}{x_{n+1} - x_n} - 1. \tag{2.5}$$

The left-boundary correction function $g_L(\xi)$ and right-boundary correction function $g_R(\xi)$ both belong to $\mathbb{P}_{P+1}$ and satisfy the following constraints in the parent space,

$$\begin{aligned} g_L(-1) &= g_R(+1) = 1, \\ g_L(+1) &= g_R(-1) = 0. \end{aligned} \tag{2.6}$$

The corrected flux, denoted by $cu^{\delta,c}$,

$$\begin{aligned} cu^{\delta,c}(x,t) = cu^\delta(x,t) &+ (cu_I^\delta(x_n,t) - cu^\delta(x_n^+,t))g_L(\xi(x)) \\ &+ (cu_I^\delta(x_{n+1},t) - cu^\delta(x_{n+1}^-,t))g_R(\xi(x)), \quad x \in \Omega_n, \end{aligned} \tag{2.7}$$

is a piecewise polynomial of degree $P+1$ which is globally continuous on account of corrections emanating from the element interfaces. It acquires the common interface values at the edges but is not necessarily equal to the discontinuous flux values internally.
The numerical derivative $\frac{\partial^\delta u^\delta}{\partial x}$ is then defined as

$$\begin{aligned} \frac{\partial^\delta u^\delta}{\partial x}(x,t) &= \frac{\partial u^{\delta,c}}{\partial x}(x,t) \\ &= \frac{\partial u^\delta}{\partial x}(x,t) + \frac{2}{h}(u_I^\delta(x_n,t) - u^\delta(x_n^+,t))\frac{dg_L}{d\xi}(\xi(x)) \\ &\quad + \frac{2}{h}(u_I^\delta(x_{n+1},t) - u^\delta(x_{n+1}^-,t))\frac{dg_R}{d\xi}(\xi(x)), \quad x \in \Omega_n. \end{aligned} \tag{2.8}$$

A detailed description with schematics and motivation is provided in [69]. For ease of analysis, we will restrict our attention to the case of uniform element size $h = x_{n+1} -$

$x_n, \forall n$ and identical distribution of solution points across elements in the parent space $\boldsymbol{\xi} = (\xi_p)_{p=1,2,\ldots,P+1}$. Plugging Eqn. (2.4) in Eqn. (2.8) and utilizing the exponential displacement relations,

$$u^{\delta}(x_n^-,t) = \exp(-ikh)u^{\delta}(x_{n+1}^-,t), \tag{2.9}$$

$$u^{\delta}(x_{n+1}^+,t) = \exp(ikh)u^{\delta}(x_n^+,t), \tag{2.10}$$

for $x \in \Omega_n$, we get that

$$\begin{aligned} c\frac{\partial^{\delta} u^{\delta}}{\partial x}(x,t) = c\frac{\partial u^{\delta}}{\partial x}(x,t) + c\frac{2}{h}(1-\alpha)(\exp(-ikh)u^{\delta}(x_{n+1}^-,t) - u^{\delta}(x_n^+,t))\frac{dg_L}{d\xi}(\xi(x)) \\ + c\frac{2}{h}\alpha(\exp(ikh)u^{\delta}(x_n^+,t) - u^{\delta}(x_{n+1}^-,t))\frac{dg_R}{d\xi}(\xi(x)), \quad x \in \Omega_n, \end{aligned} \tag{2.11}$$

which can be expressed in the vector form as

$$c\frac{\partial^{\delta} \boldsymbol{u}^{\delta}}{\partial x}\bigg|_n = c\frac{2}{h}\boldsymbol{Q}(kh)\boldsymbol{u}_{\boldsymbol{n}}^{\boldsymbol{\delta}}, \quad x \in \Omega_n, \tag{2.12}$$

where $\boldsymbol{Q}(kh) \in \mathbb{C}^{(P+1)\times(P+1)}$ is the numerical differentiation operator in the parent space, specific to the scheme,

$$\boldsymbol{Q}(kh) = \boldsymbol{D} + (1-\alpha)\boldsymbol{g}_{L,\xi}(\exp(-ikh)\boldsymbol{\ell}_+^T - \boldsymbol{\ell}_-^T) + \alpha\boldsymbol{g}_{R,\xi}(\exp(ikh)\boldsymbol{\ell}_-^T - \boldsymbol{\ell}_+^T). \tag{2.13}$$

Here $\boldsymbol{D} \in \mathbb{R}^{(P+1)\times(P+1)}$ is the polynomial differentiation operator such that

$$D_{p,m} = \frac{\mathrm{d}\ell_m}{\mathrm{d}\xi}(\xi_p), \quad p,m = 1,2,\ldots,P+1, \tag{2.14}$$

where ℓ_m is the m^{th} Lagrange polynomial in the parent domain. $\boldsymbol{g}_{L,\xi}, \boldsymbol{g}_{R,\xi} \in \mathbb{R}^{(P+1)\times 1}$ are the derivatives of the left-boundary and right-boundary correction functions at the solution points, and $\boldsymbol{\ell}_-, \boldsymbol{\ell}_+ \in \mathbb{R}^{(P+1)\times 1}$ are the extrapolated values of the $P+1$ Lagrange basis polynomials,

$$\ell_{+_p} = \ell_p(+1),\ \ell_{-_p} = \ell_p(-1), \quad p = 1,2,\ldots,P+1. \tag{2.15}$$

Eqns. (2.3, 2.12) lead to a purely element-local description of the semi-discrete numerical conservation law,

$$\frac{d\boldsymbol{u}_n^{\boldsymbol{\delta}}}{dt} + c\frac{2}{h}\boldsymbol{Q}(kh)\boldsymbol{u}_n^{\boldsymbol{\delta}} = 0, \quad \forall n,\, t > 0, \tag{2.16}$$
$$\boldsymbol{u}_n^{\boldsymbol{\delta}}(0) = \exp(ikx_n)\boldsymbol{w}_0,$$

where $\boldsymbol{w}_0$ is the collocation projection of the initial condition onto the solution points,

$$\boldsymbol{w}_0 = \left(\exp\left(ik\frac{h}{2}(1+\xi_p)\right)\right)_{p=1,2,\ldots,P+1}. \tag{2.17}$$

2.2 Eigensolution of the semi-discrete conservation law

The element-local semi-discrete numerical conservation law, Eqn. (2.16), represents a simple linear dynamical system that can be analytically integrated in time using matrix factorization. Assuming that the differentiation matrix is diagonalizable,

$$\begin{aligned}\boldsymbol{Q}(kh) &= \boldsymbol{W}(kh)\boldsymbol{\Gamma}(kh)\boldsymbol{W}^{-1}(kh) \\ &= \boldsymbol{W}(kh)ik\frac{h}{2}\boldsymbol{\Lambda}(kh)\boldsymbol{W}^{-1}(kh),\end{aligned} \tag{2.18}$$

where $\boldsymbol{\Gamma}$ is the diagonal matrix of eigenvalues $\gamma_p(kh) \in \mathbb{C}$ for $p = 1,2,\ldots,P+1$, $c\boldsymbol{\Lambda}$ is the diagonal matrix of numerical wavespeeds, $c\lambda_p(kh) \in \mathbb{C}$, and $\boldsymbol{W} \in \mathbb{C}^{P+1\times P+1}$ is the dense matrix containing eigenvectors of the differentiation operator. Hereafter, we shall refer to λ as the normalized eigenvalue and $\gamma = ik\frac{h}{2}\lambda$ as the full eigenvalue. Note that for correction functions with non-trivial derivatives, i.e. $\boldsymbol{g}_{L,\xi} \neq \boldsymbol{0}$, the matrix $\boldsymbol{Q}$ is complex and, empirically, we have not encountered a case where the assumption of diagonalizability has failed to hold.

The initial condition can also be expanded in the basis of the eigenvectors,

$$\boldsymbol{w}_0 = \boldsymbol{W}\boldsymbol{\beta}, \tag{2.19}$$

where $\beta_p \in \mathbb{C}$ is the expansion coefficient along the p^{th} column of $\boldsymbol{W}$, $\boldsymbol{w}_p(kh)$, for $p =$

$1,2,\ldots,P+1$. The solution to Eqn. (2.16) can now be written as

$$\begin{aligned}
\boldsymbol{u_n^\delta}(t) &= \exp\left(-ct\left(\frac{2}{h}\boldsymbol{Q}(kh)\right)\right)\boldsymbol{u_n^\delta}(0) \\
&= \boldsymbol{W}\exp(-ikct\boldsymbol{\Lambda})\boldsymbol{W}^{-1}\exp(ikx_n)\boldsymbol{W}\boldsymbol{\beta} \\
&= \exp(ikx_n)\boldsymbol{W}\exp(-ikct\boldsymbol{\Lambda})\boldsymbol{\beta} \\
&= \exp(ikx_n)\sum_{p=1}^{P+1}\exp(-ikc\lambda_p t)\beta_p\boldsymbol{w_p},
\end{aligned} \tag{2.20}$$

which shows that the numerical solution is a superposition of $P+1$ eigenmodes along the eigenvectors of $\boldsymbol{Q}$ weighted by the expansion coefficients, β_p, of the initial condition. The time evolution of these modes is dictated by the normalized eigenvalues λ_p. We can also write the analytical form of the exact solution on the vector of solution points, $\boldsymbol{x_n} = x_n + \frac{h}{2}(1+\boldsymbol{\xi})$, by sampling the functional form $\exp(ik(x-ct))$,

$$\begin{aligned}
\boldsymbol{u_n}(t) &= \exp(ikx_n)\exp(-ikct)\boldsymbol{W}\boldsymbol{\beta} \\
&= \exp(ikx_n)\exp(-ikct)\sum_{p=1}^{P+1}\beta_p\boldsymbol{w_p}.
\end{aligned} \tag{2.21}$$

We are now ready to evaluate the vector of pointwise error at the set of solution points in the n^{th} element,

$$\begin{aligned}
\boldsymbol{e}_n(t) &= \boldsymbol{u}_n^\delta(t) - \boldsymbol{u}_n(t) \\
&= \left(u_{n,p}^\delta(t) - u_{n,p}(t)\right)_{p=1,2,\ldots,P+1}.
\end{aligned} \tag{2.22}$$

Substituting Eqns. (2.20) and (2.21) in Eqn. (2.22),

$$\begin{aligned}
\boldsymbol{e}_n(t) &= \exp(ikx_n)\exp(-ikct)\sum_{p=1}^{P+1}\left(\exp(-ikc(\lambda_p-1)t)-1\right)\beta_p\boldsymbol{w_p} \\
&= \exp(ikx_n)\exp(-ikct)\sum_{p=1}^{P+1}\left[\exp(kc\lambda_p^{\mathrm{Im}}t)\exp(-ikc(\lambda_p^{\mathrm{Re}}-1)t)-1\right]\beta_p\boldsymbol{w_p}.
\end{aligned} \tag{2.23}$$

2.2.1 Stability, dispersion and dissipation

The eigenvalues associated with the differentiation operator provide a direct measure of the dispersive, dissipative and stability properties of the numerical scheme [8, 67, 141]. Eqn. (2.23) shows that the condition for pointwise stability [69] of the numerical solution can be prescribed as

$$\lambda_p^{\mathrm{Im}}(kh) \leq 0 \text{ for } p = 1, 2, \ldots, P+1, \tag{2.24}$$

which must hold for $k \in [0, \frac{(P+1)\pi}{h}]$ where the upper bound corresponds to the Nyquist limit for the grid.

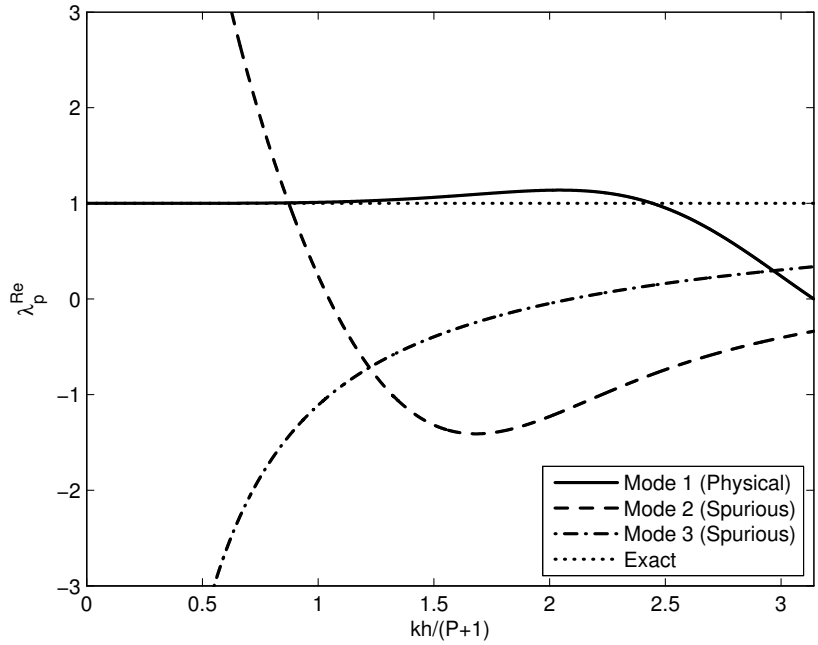
For a given mode indexed by p, the deviation from unity of the real part of the normalized eigenvalue, $\lambda_p^{\mathrm{Re}}(kh) - 1$, leads to numerical dispersion, i.e. the numerical solution propagates at an incorrect speed. The imaginary part of the normalized eigenvalue, $\lambda_p^{\mathrm{Im}}(kh)$, accounts for numerical dissipation, i.e. the numerical solution decays in amplitude with time for a stable, upwinded scheme, ($\alpha < \frac{1}{2}$).

Figures 2.1 and 2.2 plot the real and imaginary parts of λ for the three numerical modes in the case of the fully upwinded and centered DG schemes of order $P = 2$. The plots suggest the definition of a 'physical' eigenmode, indexed at $p = 1$, based on the error in the normalized eigenvalue in the asymptotic limit,

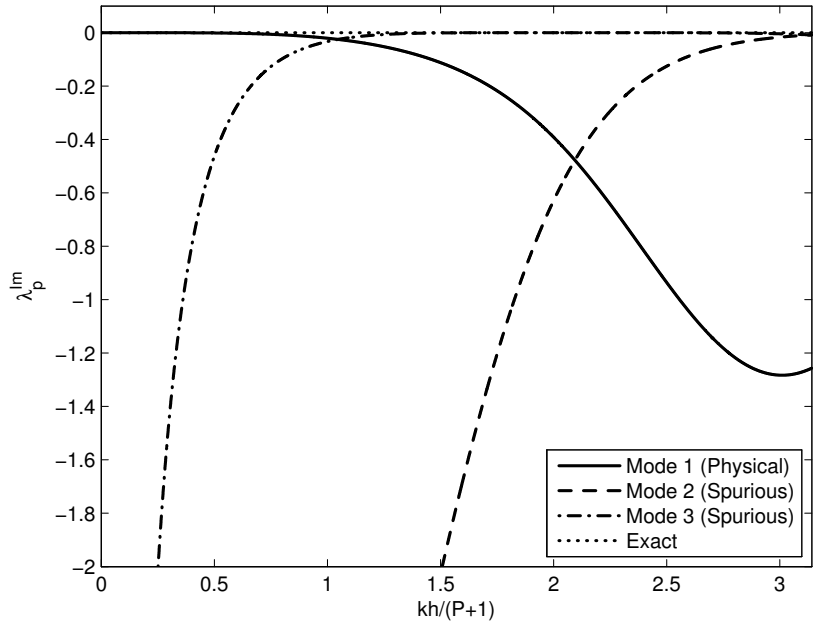
$$\lim_{kh \to 0} \lambda_1(kh) = 1. \tag{2.25}$$

All remaining modes, $p \neq 1$, exhibit excessive dispersion and dissipation errors as $kh \to 0$ and are therefore referred to as 'spurious' modes. The existence of a physical mode and the nature of the spurious modes is rigorously established in Chapter 3.

These results are in agreement with the results of Hussaini et al. [67] wherein modal analysis was performed starting directly from the DG framework. A more recent and detailed exposition for the DG scheme can be found in [98]. Note that outside the asymptotic limit, the spurious modes may exhibit lower dispersion and/or dissipation errors than the physical mode, as can be clearly seen in Figure 2.2a. This motivates the investigation of how the initial condition is distributed among the $P+1$ eigenmodes as a function of kh, and is discussed in the next subsection.

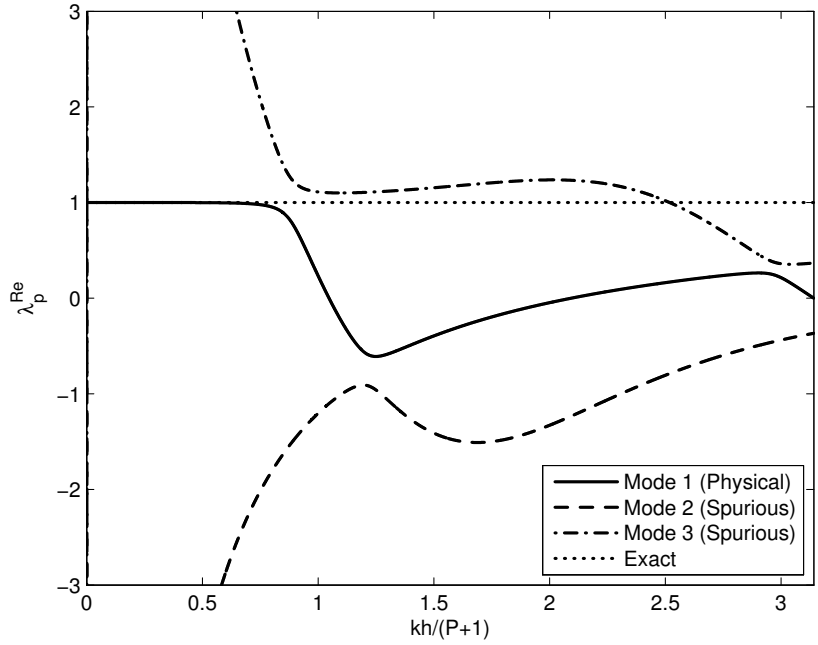


(a) λ^{Re}: measure of dispersion

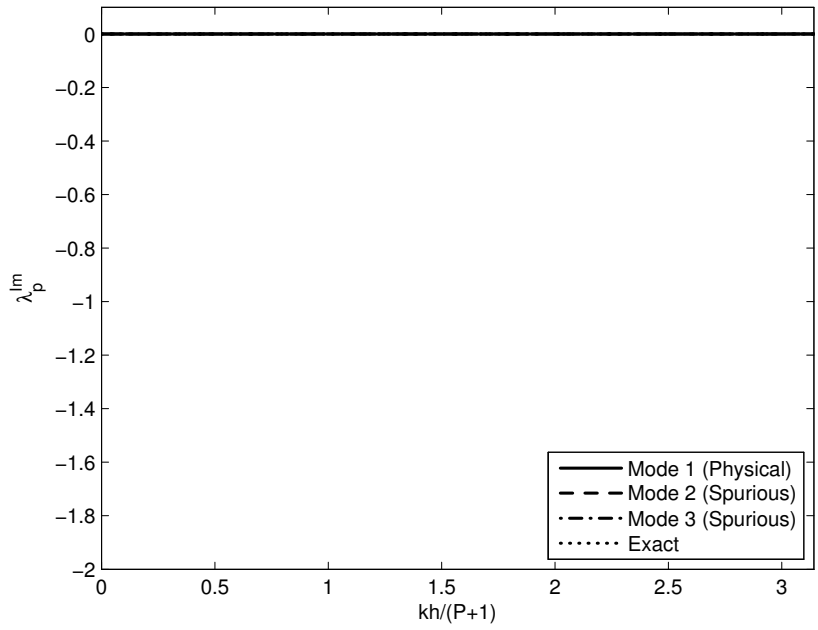


(b) λ^{Im}: measure of dissipation

Figure 2.1: Normalized eigenvalues: Real and imaginary parts of λ_p for each of the numerical modes $p = 1,2,3$ for the fully upwinded, $\alpha = 0$, DG scheme of order $P = 2$.



(a) λ^{Re}: measure of dispersion



(b) λ^{Im}: measure of dissipation

Figure 2.2: Normalized eigenvalues: Real and imaginary parts of λ_p for each of the numerical modes $p = 1,2,3$ for the centered, $\alpha = 0.5$, DG scheme of order $P = 2$.

Note that in the case of a centered flux, $\alpha = \frac{1}{2}$, it can be shown that the numerical solution propagates without any dissipation, $\lambda_p^{\text{Im}} = 0, \forall\, p$, provided that the correction functions obey certain orthogonality conditions. In particular, Vincent et. al. [142] have proven neutral stability in the weak sense that a certain broken Sobolev norm remains invariant with time when $\alpha = \frac{1}{2}$ and the correction functions belong to the ESFR family of functions. However, a complete lack of numerical dissipation is not necessarily beneficial. For instance, spurious modes remain completely undamped for centered fluxes. Unless absorbed by the boundary conditions, these modes would persist perpetually, contributing dispersion errors to the fraction of energy received by them from the initial condition. Similarly, if the mesh is coarse, centered fluxes can result in the generation of spurious numerical waves that propagate upstream (c.f. Appendix B.)

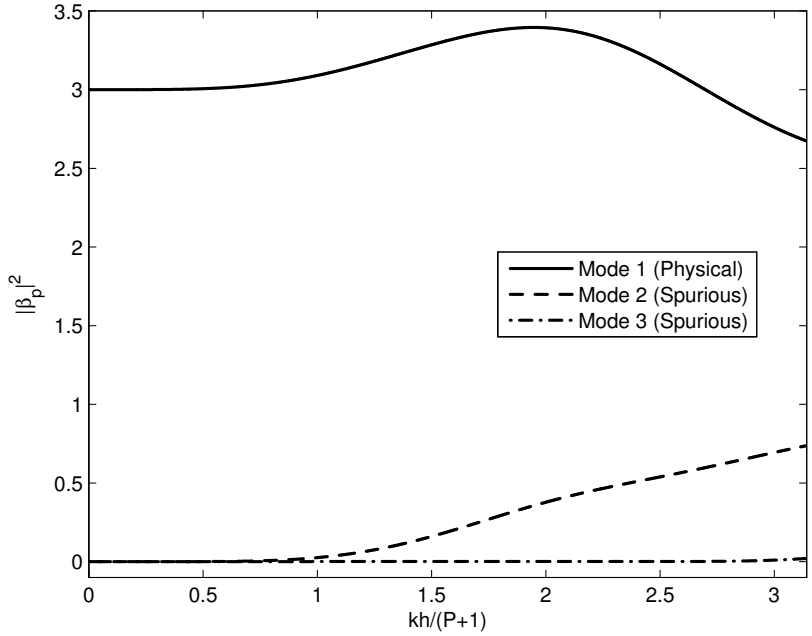
2.2.2 Modal weights and characteristic time scales of dissipation

The eigenvectors of $\boldsymbol{Q}$ dictate the distribution of energy among the physical and the spurious modes. For a given wavenumber k, the initial condition $\boldsymbol{w}_0$ is distributed to the $P+1$ numerical modes with weights given by $\boldsymbol{\beta}$. Figures 2.3a and 2.3b plot $|\beta|^2$ for the 3 eigenmodes of the fully-upwinded and centered DG schemes of order $P = 2$ using Gauss-Legendre solution points. We see that for low and moderate wavenumbers, the physical mode contains a dominant fraction of the energy. Moreover, in the limit of $kh \to 0$, the fraction of energy in the spurious modes vanishes while $|\beta_1|^2$ converges to $P+1$. This aspect will also be rigorously proven in Chapter 3.

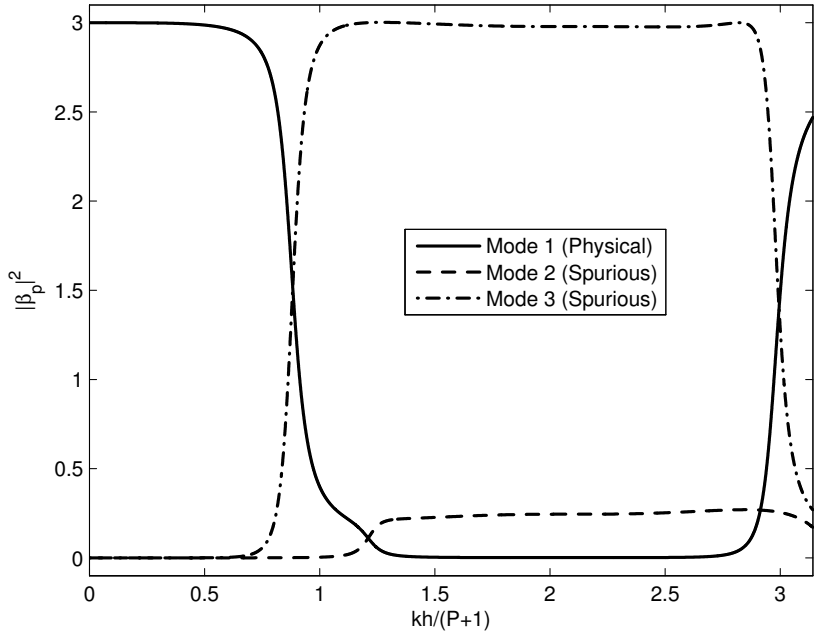
For a stable and upwinded scheme, $\lambda_p^{\text{Im}} < 0$, we can define the half-life for each eigenmode,

$$\tau_p = -\frac{1}{kc\lambda_p^{\text{Im}}} \text{ for } p = 1, 2, \ldots, P+1, \tag{2.26}$$

which is the time after which the amplitude of the mode has decayed by a factor of $\exp(-1)$. Table 2.1 records the half-lives for the 3 modes in the case of the fully upwinded DG scheme of order $P = 2$ for $\frac{kh}{P+1} = \frac{\pi}{6}$, unit grid spacing $h = 1$ and unit wavespeed $c = 1$. These parameter values correspond to a resolution of 12 solution points per wavelength of the sinusoidal initial condition. The tabulated values suggest that the numerical differentiation



(a) Fully upwind interface flux, $\alpha = 0$



(b) Centered interface flux, $\alpha = \frac{1}{2}$

Figure 2.3: Energy distribution: $|\beta|^2$ for the 3 eigenmodes of the DG scheme of order $P = 2$ using Gauss-Legendre solution points.

operator introduces two widely separated time scales,

$$\tau_{\text{physical}} = \tau_1 \gg \tau_{\text{spurious}} = \max_{p=2}^{P+1} \tau_p. \tag{2.27}$$

The energy $|\beta|^2$ associated with the spurious modes $p = 2, \ldots, P+1$ is rapidly dissipated by an upwinded scheme within $\sim \tau_{\text{spurious}}$ time. The remaining energy is dissipated after a much longer time, τ_{physical}. In the case of a centered flux, the energy associated with the spurious modes persists perpetually and such characteristic time scales cannot be defined.

	$\alpha = 0$				$\alpha = \frac{1}{2}$	
	λ_p^{Re}	λ_p^{Im}	τ_p	$\lvert\beta_p\rvert^2$	λ_p	$\lvert\beta_p\rvert^2$
Mode 1 (Physical)	1.00032	-0.00114	556.34096	3.00744	0.99874	2.98377
Mode 2 (Spurious)	4.13266	-5.31987	0.11967	0.00065	4.41886	0.00058
Mode 3 (Spurious)	-3.22312	-0.40856	1.55820	0.00010	-3.50774	0.00003
Exact	1	0	-	-	1	-

Table 2.1: Normalized eigenvalues, half lives, and energy distribution for the DG scheme of order $P = 2$ for $\frac{kh}{P+1} = \frac{\pi}{6}$, unit grid-spacing $h = 1$ and unit wavespeed $c = 1$.

2.2.3 Evolution of pointwise error

Numerical eigenmodes completely specify the evolution of pointwise error in time. To see this, consider the vector ℓ^2 norm of pointwise error at the set of solution points in the n^{th} element,

$$\|\boldsymbol{e}_n(t)\|_2 = \|\boldsymbol{u}_n^{\boldsymbol{\delta}}(t) - \boldsymbol{u}_n(t)\|_2 = \left[\sum_{p=1}^{P+1} \left| u_{n,p}^{\delta}(t) - u_{n,p}(t) \right|^2 \right]^{1/2}. \tag{2.28}$$

This definition is motivated by the nodal nature of FR and is different from the usual functional norm that measures integrated error over a fixed domain. In particular, this simple Euclidean measure is local to the n^{th} element and does not encounter any technical difficulties in the limit of $h \to 0$ which we will explore in the next chapter. Substituting Eqn.

(2.23) in Eqn. (2.28),

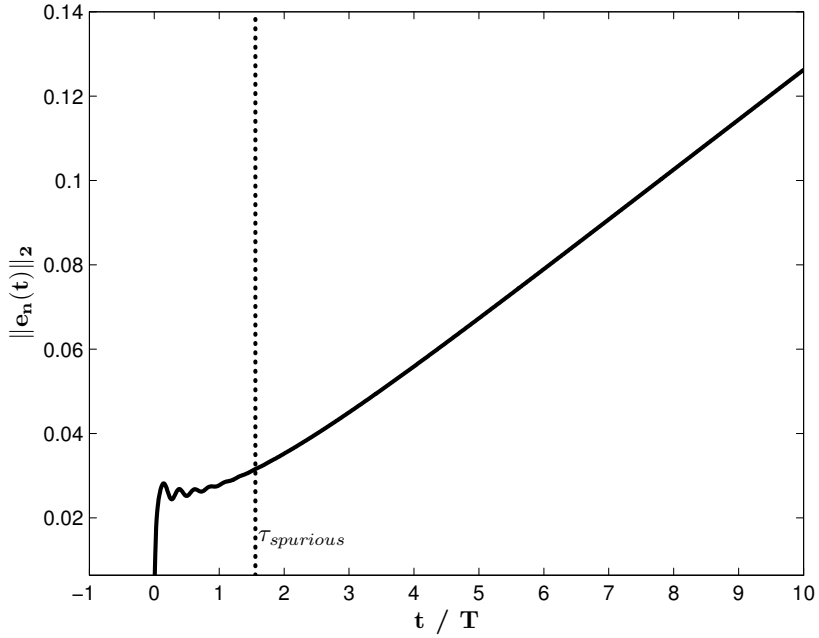
$$\begin{aligned}\|\boldsymbol{e}_n(t)\|_2 &= |\exp(ikx_n)|\,|\exp(-ikct)|\,\|\sum_{p=1}^{P+1}(\exp(-ikc(\lambda_p-1)t)-1)\,\beta_p\boldsymbol{w_p}\|_2 \\ &= \|\sum_{p=1}^{P+1}(\exp(-ikc(\lambda_p-1)t)-1)\,\beta_p\boldsymbol{w_p}\|_2, \end{aligned} \tag{2.29}$$

where we have used that $|\exp(\zeta)| = 1$ if ζ is purely imaginary. Figure 2.4 shows the evolution of this norm against number of time periods, $T = 2\pi/(kc)$, for the fully-upwinded DG scheme of order $P = 2$ using Gauss-Legendre solution points, $\frac{kh}{P+1} = \frac{\pi}{6}$, unit grid spacing $h = 1$ and unit wavespeed $c = 1$ as in Table 2.1. For the upwinded scheme, the error grows rapidly within τ_{spurious} time owing to the loss of energy associated with the decay of spurious modes. This region is marked by oscillations associated with large phase errors. Beyond that time, the growth of error is much slower corresponding to the physical mode and is given by,

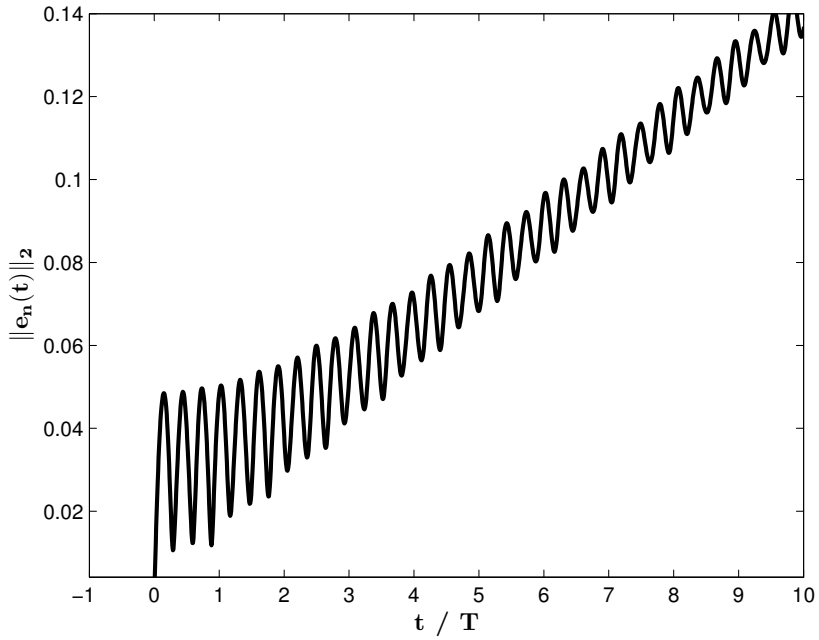
$$\|\boldsymbol{e}_n(t \gg \tau_{\text{spurious}})\|_2 \simeq \|\,(\exp(-ikc(\lambda_1-1)t)-1)\,\beta_1\boldsymbol{w_1} + (w_0 - \beta_1 w_1)\|_2 \tag{2.30}$$

where $w_0 - \beta_1 w_1$ is the constant error introduced due to loss of energy in the spurious modes. On the other hand, in the case of a centered flux, the spurious modes do not decay at all and continue to contribute phase errors to the portion of the energy received by them from the initial condition. This again highlights why there is a prevalence of upwinded fluxes in applications of FR schemes.

Note that typical fluid simulations can be significantly less resolved than an effective wavenumber of $\frac{kh}{P+1} = \frac{\pi}{6}$. Such lack of resolution can result in certain important and interesting numerical phenomena especially for simulations of instabilities such as transition to turbulence. A discussion of such non-asymptotic behavior is included in Appendix B.



(a) Fully upwind interface flux, $\alpha = 0$



(b) Centered interface flux, $\alpha = \frac{1}{2}$

Figure 2.4: Evolution of the norm of pointwise error, Eqn. (2.29), against number of time periods, $T = 2\pi/(kc)$, for the DG scheme of order $P = 2$ using Gauss-Legendre solution points, $\frac{kh}{P+1} = \frac{\pi}{6}$, unit grid spacing $h = 1$ and unit wavespeed $c = 1$.

2.3 Effect of polynomial order, correction function and solution points

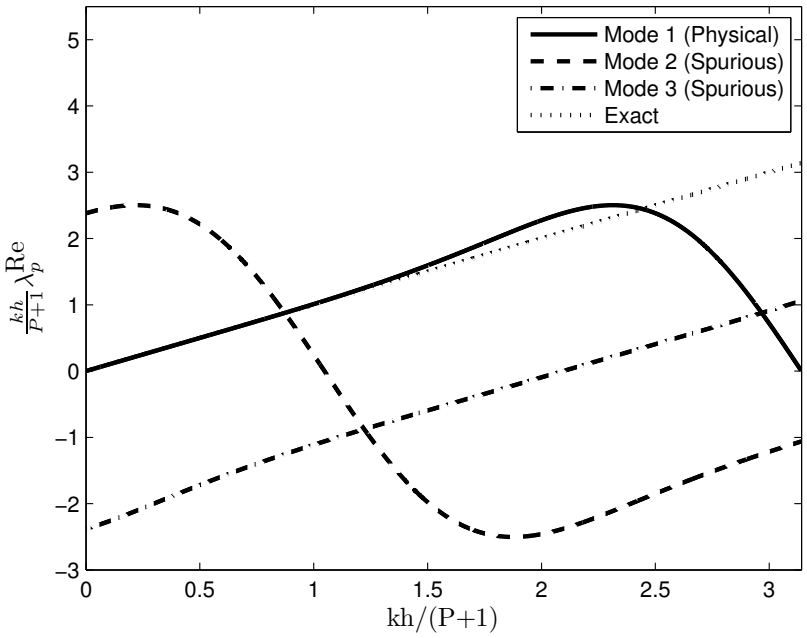
This section provides a qualitative discussion of the effect of polynomial order, correction function and the location of solution points on the dispersive, dissipative and modal distribution properties discussed earlier in this chapter. In order to minimize the total number of illustrations, the interface flux is fully upwinded, $\alpha = 0$. Once again, all quantities of interest are plotted as functions of wavenumber, k, normalized by the number of degrees of freedom per unit-length, $P+1/h$, to have the normalized Nyquist limit as π. Moreover, for comparison with classical studies [88, 67], the real part of the normalized eigenvalue, λ^{Re}, is plotted in the form of an effective wavenumber,

$$k_p^{\mathrm{eff}} = \frac{kh}{P+1}\lambda_p{}^{\mathrm{Re}} = \frac{2}{P+1}\gamma_p{}^{\mathrm{Im}}. \tag{2.31}$$

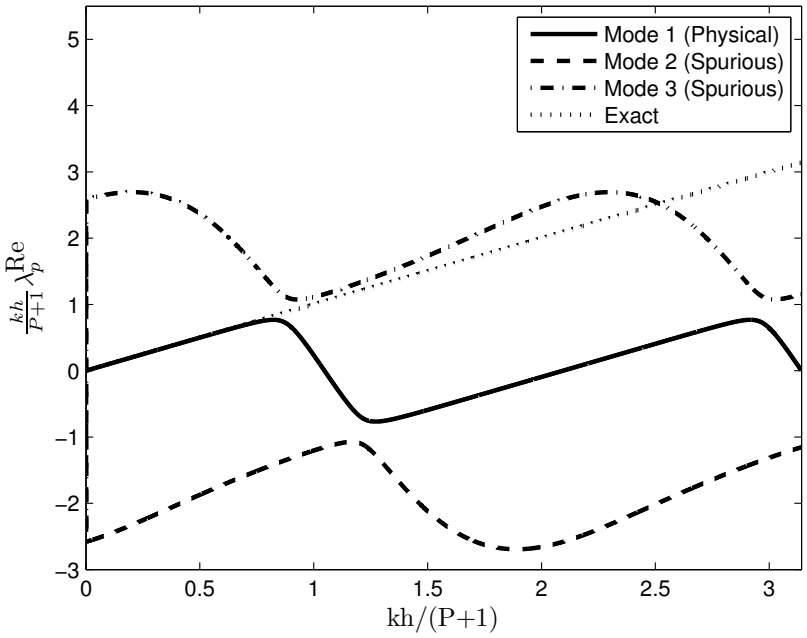
For instance, Figs. 2.5a and b plot the effective wavenumber for the three numerical eigenmodes for the DG scheme of order $P = 2$ for the upwinded and centered fluxes respectively. These are the traditional representations of Figs. 2.1a and 2.2a. Finally, for a meaningful comparison, the modal weights are plotted in the form of relative energy, i. e. $\frac{|\beta_p|^2}{\sum_{p=1}^{P+1}|\beta_p|^2}$, which measures the fraction of energy received by the pth mode from the sinusoidal initial condition $\exp(ikx)$.

2.3.1 Effect of polynomial order

Figures 2.6a and b plot the effective wavenumber and the imaginary part of the normalized eigenvalue corresponding to the physical mode for the DG scheme for $P = 0$ to 5 against the normalized wavenumber. The exact dispersion relation is also indicated for reference. The range of resolvable wavenumbers, defined by the Nyquist limit, increases in proportion to the number of degrees of freedom within each element. Interestingly, even the range of normalized wavenumbers $kh/(P+1)$ over which are the dispersion relation is approximately preserved, increases with polynomial order. While higher polynomial order results in greater overshoot (Fig. 2.6a) suggesting a lower CFL limit, the dominant effect is that



(a) Fully upwind interface flux, $\alpha = 0$



(b) Centered interface flux, $\alpha = \frac{1}{2}$

Figure 2.5: Effective wavenumber: k_p^{eff} for each of the numerical modes $p = 1,2,3$ for the DG scheme of order $P = 2$.

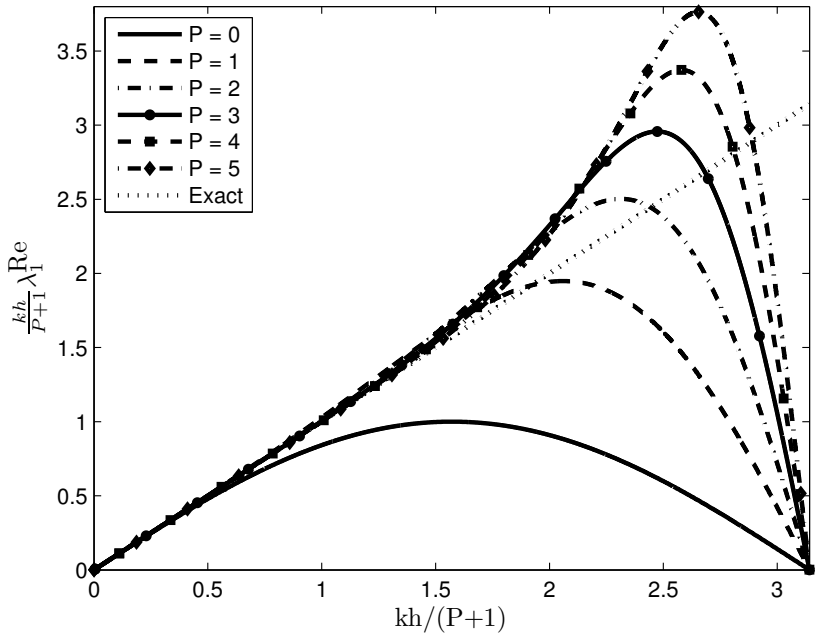
of reduced dissipation (2.6b), leading to an increase in resolving efficiency. These results are again in agreement with those of Hussaini et al. [67] for the DG scheme, and with Ainsworth [6] who shows that the leading coefficient of dispersion and dissipation errors decreases rapidly with increasing polynomial order.

Figure 2.8 plots the fraction of energy received by the physical mode for the DG scheme for $P = 0$ to 5 using Gauss-Legendre solution points. Here, again, a competing effect can be observed with increase in polynomial order. While the point of departure from unity is pushed towards high normalized wavenumbers, the energy content for the highest wavenumbers decreases with increase in P.

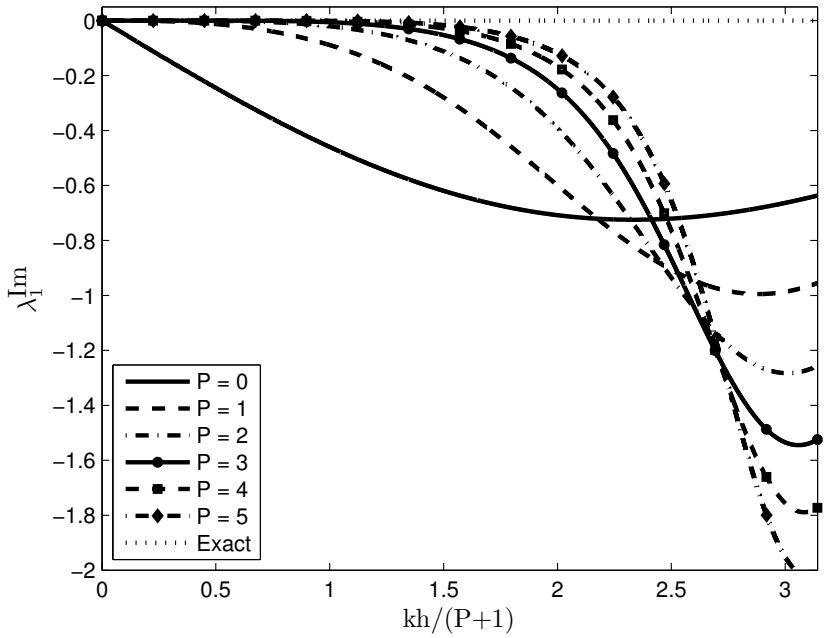
2.3.2 Effect of correction function

Figures 2.7a and b plot the effective wavenumber and the imaginary part of the normalized eigenvalue corresponding to the physical mode for DG, SD and Huynh's g_2 [69] scheme for $P = 2$ against the normalized wavenumber. The choice of correction function has a critical impact on both the real and imaginary parts of the numerical wavespeed for the physical mode. For these three schemes, a competing effect can be observed between the dispersion and dissipation errors. In comparison to DG, for the SD scheme, while the real part of the wavespeed stays nearly accurate for a larger range of wavenumbers, the imaginary part starts deviating earlier, suggesting the possibility of an optimal choice somewhere in between. These results are in agreement with the observations of Vincent et al. [141] wherein the authors used an automated modal energy based sampling procedure, following Van den Abeele [132], to recover a single Fourier mode for any polynomial order. This suggests that the modal reconstruction procedure employed therein, in fact, recovered only the physical mode.

The dependence of energy content on the choice of correction function is relatively weaker as can be seen in Fig. 2.9 which plots $\frac{|\beta_1|^2}{\sum_{p=1}^{P+1} |\beta_p|^2}$ for the DG, SD and Huynh's g_2 correction functions.

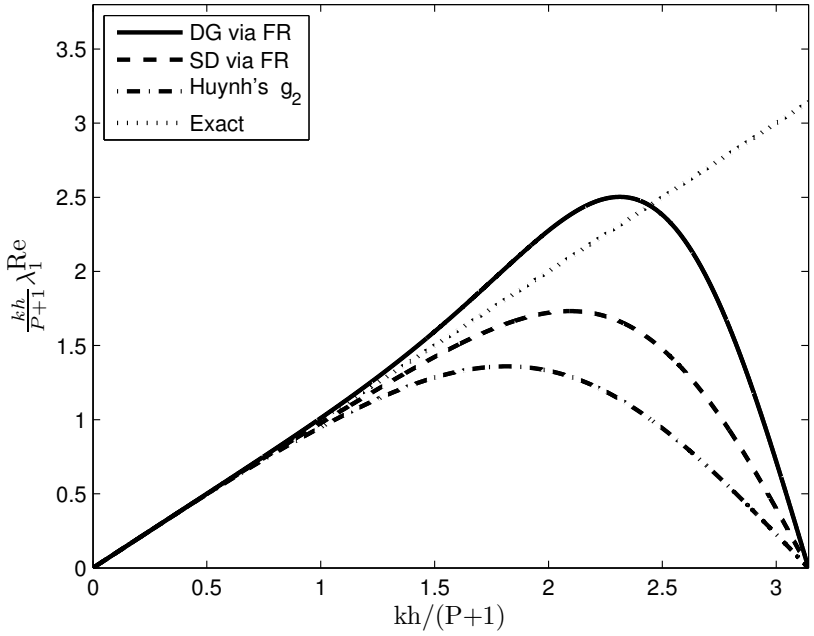


(a) Effective wavenumber, k_p^{eff}

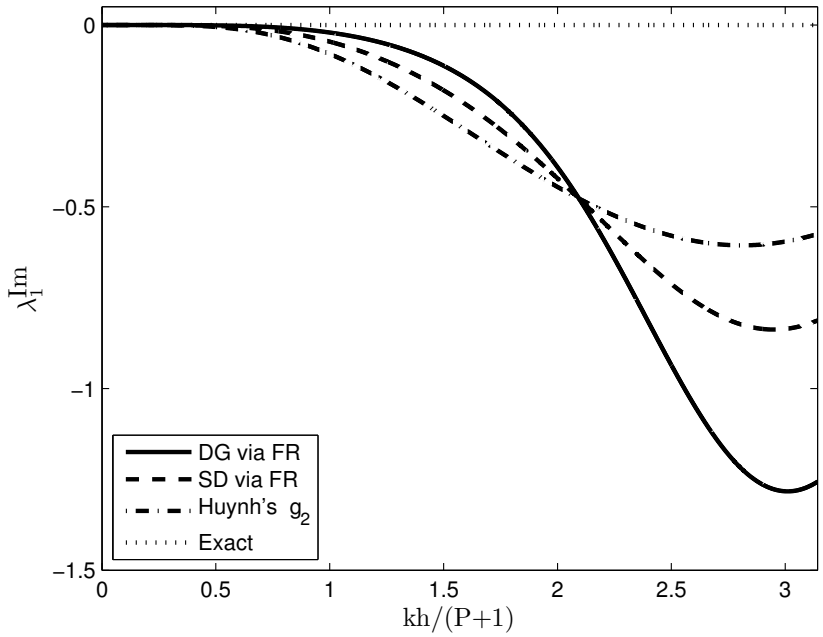


(b) Imaginary part of normalized eigenvalue

Figure 2.6: Effect of polynomial order on the dispersive and dissipative properties of the physical mode for the DG scheme of order $P = 2$.



(a) Effective wavenumber, k_p^{eff}



(b) Imaginary part of normalized eigenvalue

Figure 2.7: Effect of correction function on the dispersive and dissipative properties of the physical mode for the DG scheme of order $P = 2$.

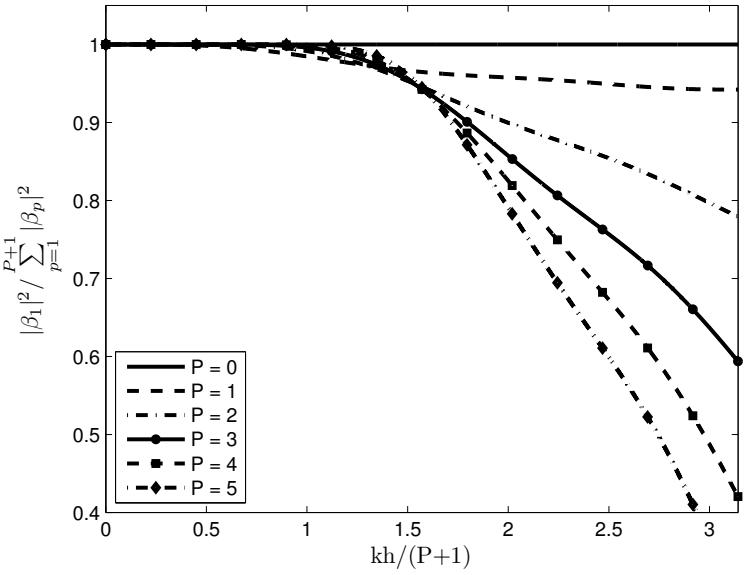


Figure 2.8: Effect of polynomial order on the fraction of energy retained in the physical mode for $P = 2$ using Gauss-Legendre solution points.

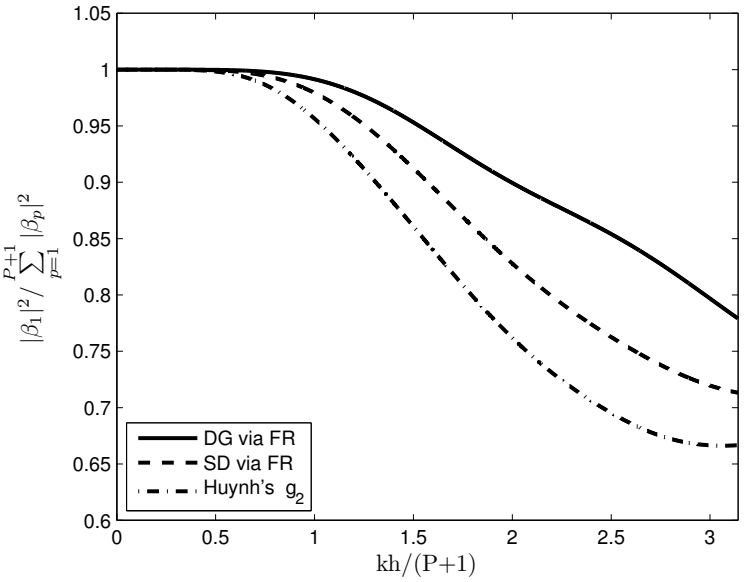
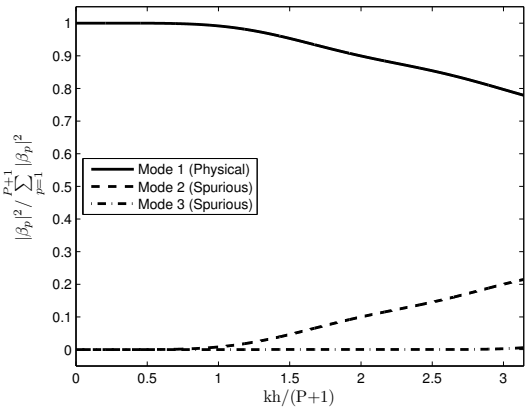


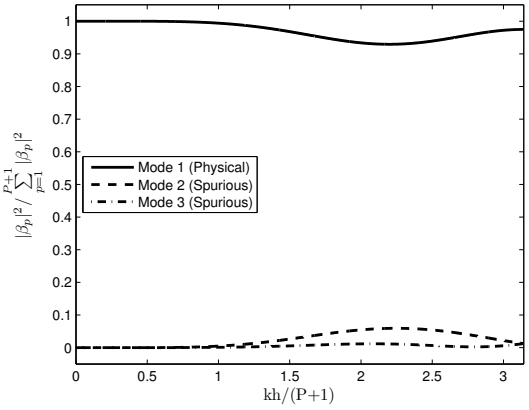
Figure 2.9: Effect of correction function on the fraction of energy retained in the physical mode for $P = 2$ using Gauss-Legendre solution points.

2.3.3 Effect of solution points

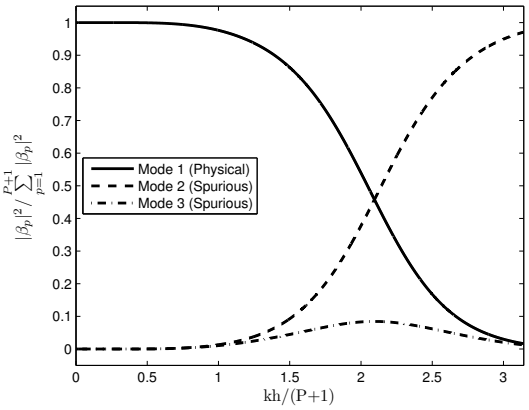
In [69], Huynh showed that the eigenvalues of the FR eigensystem are independent of the choice of solution points by proving similarity of the corresponding differentiation matrices. However, from Figs. 2.10b and c we see that the energy distribution changes significantly when the solution points are chosen to be equidistant or Gauss-Lobatto points respectively. This is expected as the modal weights are linearly dependent on $\boldsymbol{w}_0$ which in turn depends directly on the location of the solution points. While equidistant points do slightly better at preserving maximum energy in the physical mode, their use is still limited due to aliasing errors [76]. In case of Gauss-Lobatto points, one notes the presence of a crossover point whereafter one of the spurious modes captures a higher fraction of energy than the physical mode.



(a) Gauss-Legendre points



(b) Equidistant points



(c) Gauss-Lobatto points

Figure 2.10: Effect of solution points on relative energies: Distribution of energy among numerical modes for DG scheme via FR for $P = 2$.

Chapter 3

Consistency of the semi-discrete formulation

A proof of convergence for stable Flux Reconstruction schemes can be constructed by showing that the numerical conservation law, Eqn. (2.3), is consistent, and that the point-wise error, Eqn. (2.23), vanishes in the limit of asymptotic resolution. For a given FR scheme, i.e. given the correction function, upwinding coefficient, and location of solution points, the only remaining parameter is the non-dimensional wavenumber kh which depends on the initial condition and the grid spacing. Consequently, asymptotic resolution, $kh \to 0$, can be attained either through vanishing grid spacing, $h \to 0$, or through vanishing wavenumber, $k \to 0$. The former mechanism corresponds to classical grid-convergence of numerical methods, while the latter corresponds to wavenumber-convergence of the kind explored by Ainsworth [6, 4]. The two mechanisms are intuitively similar; however, the resulting rates of convergence differ as proven in Chapter 4.

In the results that follow, we prove that for any FR scheme, a unique physical mode exists such that $\lambda_1 \to 1$ and $\beta_1^2 \to P+1$. The modal weights for all other (spurious) modes vanish in the limit. Finally, in Theorem 3.1, we prove that the FR scheme is consistent with respect to the linear advection problem.

3.1 Eigenvalue decomposition of the polynomial differentiation matrix, D

We begin by assembling results on the eigen-decomposition of the polynomial differentiation matrix, $\boldsymbol{D}$, which will be needed in the next section.

Lemma 3.1. *Let $\boldsymbol{v} \in \mathbb{C}^{(P+1)\times 1}$ be the vector obtained from sampling a polynomial v of degree $\leq P$ at $P+1$ distinct points, $\boldsymbol{\xi} = (\xi_p)_{p=1,2,\ldots,P+1}$, in $[-1,1]$. Then, the polynomial differentiation matrix $\boldsymbol{D}$ in Eqn. (2.14) exactly differentiates $\boldsymbol{v}$ to yield the vector obtained from sampling $\frac{dv}{d\xi}$ at the $P+1$ points.*

Proof. Consider expanding $\boldsymbol{v}$ in the discrete monomial basis $\{\mathbf{1}, \boldsymbol{\xi}, \boldsymbol{\xi^2}, \ldots, \boldsymbol{\xi}^P\}$, where $\boldsymbol{\xi^i} = (\xi_p)^i_{p=1,2,\ldots,P+1}$, as $\boldsymbol{v} = \sum_{i=0}^{P} a_i \boldsymbol{\xi^i}$. Then,

$$
\begin{aligned}
\boldsymbol{D}\boldsymbol{v} &= \sum_{i=0}^{P} a_i \boldsymbol{D}\boldsymbol{\xi^i} \\
&= \sum_{i=0}^{P} a_i \left(\sum_{m=0}^{P} \frac{\mathrm{d}\ell_m}{\mathrm{d}\xi}(\xi_p)\xi_m^i \right)_{p=1,2,\ldots,P+1} \\
&= \sum_{i=0}^{P} a_i \left(\frac{\mathrm{d}}{\mathrm{d}\xi} \left(\sum_{m=0}^{P} \ell_m(\xi)\xi_m^i \right)(\xi_p) \right)_{p=1,2,\ldots,P+1} \\
&= \sum_{i=0}^{P} a_i \left(\frac{\mathrm{d}}{\mathrm{d}\xi} \left(\xi^i \right)(\xi_p) \right)_{p=1,2,\ldots,P+1} \\
&= \sum_{i=1}^{P} a_i i \boldsymbol{\xi}^{i-1}, \qquad (3.1)
\end{aligned}
$$

which is the vector obtained from sampling $\frac{dv}{d\xi}$ at the $P+1$ points. □

Lemma 3.2. *For an FR scheme of order P with distinct solution points, the polynomial differentiation matrix $\boldsymbol{D}$ in Eqn. (2.14) has $P+1$ repeated eigenvalues at 0 corresponding to the uni-dimensional eigenspace spanned by $\mathbf{1} = [1, 1, \ldots, 1]^T$.*

Proof. Let $\boldsymbol{v} \in \mathbb{C}^{(P+1)\times 1}$ be an eigenvector of $\boldsymbol{D}$, so that $\boldsymbol{D}\boldsymbol{v} = \gamma\boldsymbol{v}$ and $\|\boldsymbol{v}\|_2 = 1$. Consider expanding $\boldsymbol{v}$ in the discrete monomial basis as $\boldsymbol{v} = \sum_{i=0}^{P} a_i \boldsymbol{\xi}^{\boldsymbol{i}}$. Then, using Lemma 3.1, $\sum_{i=1}^{P} a_i i \boldsymbol{\xi}^{\boldsymbol{i-1}} = \gamma \sum_{i=0}^{P} a_i \boldsymbol{\xi}^{\boldsymbol{i}}$, which can be rewritten as

$$\sum_{i=0}^{P} (a_{i+1}(i+1)(1-\delta_{i,P}) - \gamma a_i)\boldsymbol{\xi}^{\boldsymbol{i}} = \boldsymbol{0}, \tag{3.2}$$

where δ denotes the Kronecker delta function. Let $\boldsymbol{b}$ be such that $b_i = a_{i+1}(i+1)(1-\delta_{i,P}) - \gamma a_i$ for $i = 0, 1, \ldots, P$ and introduce the Vandermonde matrix $\boldsymbol{V} = [\boldsymbol{1}, \boldsymbol{\xi}, \ldots, \boldsymbol{\xi}^{\boldsymbol{P}}]$ so that Eqn. (3.2) can be recognized as $\boldsymbol{V}\boldsymbol{b} = \boldsymbol{0}$, which yields the unique trivial solution $\boldsymbol{b} = \boldsymbol{0}$ since the $P+1$ solution points are assumed to be distinct. This results in the backward recursive relation,

$$\begin{cases} a_i = 0 \quad \forall \quad i & \gamma \neq 0 \\ a_0 = \frac{1}{\sqrt{P+1}}, a_i = 0 \text{ for } i = 1, \ldots, P & \gamma = 0 \end{cases} \tag{3.3}$$

where the factor of $\frac{1}{\sqrt{P+1}}$ arises from the fact that the eigenvectors are unit vectors. Hence, there is a single non-zero eigenvector $\frac{1}{\sqrt{P+1}}\boldsymbol{1}$ and the corresponding eigenvalue is given by $\gamma = 0$.

□

3.2 Asymptotic eigenvalue decomposition of the differentiation matrix, Q

The exact conservation law, Eqn. (2.1), utilizes the exact differentiation operator $\frac{\partial}{\partial \xi} = \frac{h}{2}\frac{\partial}{\partial x}$ which is associated with the eigenvalue $ik\frac{h}{2} \to 0$ as $kh \to 0$. The proof of convergence for the FR formulation is complicated by the observation that the numerical eigenvalues for the spurious modes, $(\gamma_p = ik\frac{h}{2}\lambda_p)_{p\neq 1}$, do not converge to zero, i.e. they are inconsistent. This can be verified directly from Fig. 2.5 or indirectly by noting the normalized eigenvalues in Figs. 2.1, 2.2 diverge as $kh \to 0$. However, it can be shown that when the distribution

of weights among the modes, $\boldsymbol{\beta}$, is taken into account, the inconsistent spurious modes do not receive any energy in the asymptotic limit. Moreover, under suitable conditions, the physical eigenvalue γ_1 converges to $ik\frac{h}{2}$ as well. Consequently, the pointwise error does indeed vanish as required.

The results that follow are directed at an FR scheme of finite order P having $P+1$ distinct solution points. We begin by showing that, although the normalized eigenvalues λ_p may become unbounded, the full eigenvalues, γ_p for $p = 1, 2, \ldots, P+1$, always remain bounded.

Lemma 3.3. *All eigenvalues of the FR differentiation operator $\boldsymbol{Q}(kh)$ remain bounded.*

Proof. This can be shown by noting that the spectral radius of any matrix is bounded by its norm,

$$\begin{aligned}
\max_{p=1}^{P+1} |\gamma_p| &= \max_{p=1}^{P+1} \|\gamma_p \boldsymbol{w}_p\|_2 \\
&= \max_{p=1}^{P+1} \|\boldsymbol{Q}(kh)\boldsymbol{w}_p\|_2 \\
&\leq \max_{\boldsymbol{y}, \|\boldsymbol{y}\|_2=1} \|\boldsymbol{Q}(kh)\boldsymbol{y}\|_2 \\
&= \|\boldsymbol{Q}(kh)\|_2 \\
&= \left\|(1-\alpha)\boldsymbol{g}_{L,\xi}(\exp(-ikh)\boldsymbol{\ell}_+^T - \boldsymbol{\ell}_-^T) + \boldsymbol{D} + \alpha\boldsymbol{g}_{R,\xi}(\exp(ikh)\boldsymbol{\ell}_-^T - \boldsymbol{\ell}_+^T)\right\|_2 \\
&\leq \left((1-\alpha)\|\boldsymbol{g}_{L,\xi}\|_2 + \alpha\|\boldsymbol{g}_{R,\xi}\|_2\right)\left(\|\boldsymbol{\ell}_+\|_2 + \|\boldsymbol{\ell}_-\|_2\right) + \|\boldsymbol{D}\|_F \\
&\leq 2\left(\sqrt{\sum_{p=1}^{P+1} g_{L,\xi_p}^2} + \sqrt{\sum_{p=1}^{P+1} g_{R,\xi_p}^2}\right)\sqrt{\sum_{p=1}^{P+1} \ell_p^2(+1)} + \sqrt{\sum_{p=1}^{P+1}\sum_{q=1}^{P+1} |\frac{\mathrm{d}\ell_q}{\mathrm{d}\xi}(\xi_p)|^2}, \quad (3.4)
\end{aligned}$$

where the subscript F denotes the Frobenius norm. The boundedness of Lagrange polynomials and their derivatives follows from the assumptions that P is finite, the solutions points are distinct and the parent domain is bounded.

□

Table 3.1 shows that the limits of the eigenvalues $\gamma_p(kh \to 0)$ are indeed bounded, for all the $P+1$ modes. The next three results describe the asymptotic behavior of the physical

eigenvalue, γ_1 and also λ_1, the physical eigenvector $\boldsymbol{w}_1$, and the modal weight vector $\boldsymbol{\beta}$.

	$\alpha = 0$		$\alpha = \frac{1}{2}$
	γ_p^{Re}	γ_p^{Im}	γ_p
Mode 1 (Physical)	0.00000	0.00000	0.00000
Mode 2 (Spurious)	1.50000	3.57071	$3.87298\,i$
Mode 3 (Spurious)	1.50000	-3.57071	$-3.87298\,i$
Exact	0	0	0

Table 3.1: Full eigenvalues for the DG scheme of order $P = 2$ for $\frac{kh}{P+1} = 0$.

Lemma 3.4. *In the limit $kh \to 0$, the FR differentiation operator $\boldsymbol{Q}$ has at least one eigenpair which corresponds to the physical mode, $\gamma_1 \to 0$ and $\boldsymbol{w}_1 \to \frac{1}{\sqrt{P+1}}\mathbf{1}$.*

Proof. The proof follows from Eqn. (2.13) by observing that $\mathbf{1}$ belongs to the null space of $\boldsymbol{Q}(0)$,

$$\boldsymbol{Q}(0)\mathbf{1} = \left[\left((1-\alpha)\boldsymbol{g}_{L,\xi} - \alpha\boldsymbol{g}_{R,\xi}\right)(\boldsymbol{\ell}_+^T - \boldsymbol{\ell}_-^T) + \boldsymbol{D}\right]\mathbf{1} = \mathbf{0}, \tag{3.5}$$

where we have used Lemma 3.2, $\boldsymbol{D}\mathbf{1} = \mathbf{0}$, and extrapolation property of the Lagrange operators, $\boldsymbol{\ell}_-^T\mathbf{1} = \boldsymbol{\ell}_+^T\mathbf{1} = 1$.

□

Lemma 3.5. *In the limit $kh \to 0$, if the differentiation operator $\boldsymbol{Q}$ is diagonalizable, then the modal weights β_p, in Eqn. (2.19), converge to $\sqrt{P+1}(\delta_{p1})$ for $p = 1, 2, \ldots, P+1$.*

Proof. Taking limits in Eqn. (2.19) and using Eqn. (2.17),

$$\boldsymbol{W}(0)\boldsymbol{\beta}(0) = \boldsymbol{w}_0(0) = \mathbf{1}. \tag{3.6}$$

Since $\boldsymbol{W}(0)$ is invertible by the assumption of diagonalizability, therefore $\boldsymbol{\beta}(0)$ is unique. Also, from Lemma 3.4, we know that

$$\boldsymbol{W}(0)(\sqrt{P+1}\boldsymbol{e}_1) = \sqrt{P+1}\boldsymbol{w}_1(0) = \mathbf{1}, \tag{3.7}$$

which shows that the one and only solution is given by $\beta_p(0) = \sqrt{P+1}\delta_{p1}$ for $p = 1, 2, \ldots, P+1$. This also shows that there exists exactly one physical mode.

□

The convergence of $\boldsymbol{\beta}$ can also be easily verified from Fig. 2.3 which shows that $\beta_1 \to \sqrt{P+1}$ and $\beta_{p\neq 1} \to 0$.

Lemma 3.6. *In the limit $kh \to 0$, if the differentiation operator $\boldsymbol{Q}$ is diagonalizable, the normalized eigenvalue corresponding to the physical mode converges to unity, $\lambda_1 \to 1$.*

Proof. Consider the numerical derivative vector obtained when the numerical differentiation operator acts on the initial condition, $\boldsymbol{Q}(kh)\exp(ikx_n)\boldsymbol{w}_0$. From Eqns. (2.18), (2.19),

$$\begin{aligned}
\boldsymbol{Q}(kh)\boldsymbol{w}_0 &= \boldsymbol{Q}(kh)\boldsymbol{W}(kh)\boldsymbol{\beta}(kh) \\
&= \boldsymbol{W}(kh)\boldsymbol{\Gamma}(kh)\boldsymbol{\beta}(kh) \\
&= \sum_{p=1}^{P+1} \beta_p(kh)\gamma_p(kh)\boldsymbol{w}_p(kh) \\
&= \beta_1(kh)\gamma_1(kh)\boldsymbol{w}_1(kh) + \sum_{p=2}^{P+1} \beta_p(kh)\gamma_p(kh)\boldsymbol{w}_p(kh), \qquad (3.8)
\end{aligned}$$

so that by Lemmas 3.3, 3.4 and 3.5,

$$\boldsymbol{Q}(kh)\boldsymbol{w}_0 \to \gamma_1(kh)\mathbf{1} = ik\frac{h}{2}\lambda_1(0)\mathbf{1} \text{ as } kh \to 0. \qquad (3.9)$$

Alternatively, consider expanding $\boldsymbol{Q}(kh)\boldsymbol{w}_0$ via the Taylor series, using Eqns. (2.17) and (2.13), and retaining only the first order terms,

$$\begin{aligned}
\boldsymbol{Q}(kh)\boldsymbol{w}_0 &= \left(\boldsymbol{Q}(0) - ikh\left((1-\alpha)\boldsymbol{g}_{L,\xi}\boldsymbol{\ell}_+^T - \alpha\boldsymbol{g}_{R,\xi}\boldsymbol{\ell}_-^T\right)\right)\left(\mathbf{1} + ik\frac{h}{2}(\mathbf{1}+\boldsymbol{\xi})\right) + \mathscr{O}(kh)^2 \\
&= \left(1 + ik\frac{h}{2}\right)\boldsymbol{Q}(0)\mathbf{1} + ik\frac{h}{2}\left[\boldsymbol{Q}(0)\boldsymbol{\xi} - 2\left((1-\alpha)\boldsymbol{g}_{L,\xi}\boldsymbol{\ell}_+^T - \alpha\boldsymbol{g}_{R,\xi}\boldsymbol{\ell}_-^T\right)\mathbf{1}\right] + \mathscr{O}(kh)^2 \\
&= ik\frac{h}{2}\left[\boldsymbol{Q}(0)\boldsymbol{\xi} - 2((1-\alpha)\boldsymbol{g}_{L,\xi} - \alpha\boldsymbol{g}_{R,\xi})\right] + \mathscr{O}(kh)^2, \qquad (3.10)
\end{aligned}$$

where we have used Lemma 3.4. Also, from the properties of $\boldsymbol{\ell}_-, \boldsymbol{\ell}_+$ and $\boldsymbol{D}$, we have, $\boldsymbol{Q}(0)\boldsymbol{\xi} = \left(2((1-\alpha)\boldsymbol{g}_{L,\xi} - \alpha\boldsymbol{g}_{R,\xi}) + \mathbf{1}\right)\mathbb{I}_{P>0}$, so that,

$$\boldsymbol{Q}(kh)\boldsymbol{w}_0 = \begin{cases} ik\frac{h}{2}\mathbf{1} + \mathscr{O}(kh)^2 & P > 0 \\ -ikh((1-\alpha)g_{L,\xi} - \alpha g_{R,\xi}) + \mathscr{O}(kh)^2 & P = 0 \end{cases} \tag{3.11}$$

Now note that the correction functions for $P = 0$ are completely specified by the boundary constraints, Eqn. (2.6), so that $g_L(\xi) = \frac{1}{2}(1-\xi),\ g_{R,\xi} = \frac{1}{2}(1+\xi)$. Hence,

$$\boldsymbol{Q}(kh)\boldsymbol{w}_0 \to ik\frac{h}{2}\mathbf{1} \text{ as } kh \to 0, \forall P. \tag{3.12}$$

Comparing Eqn. (3.12) with Eqn. (3.9), we get $\lambda_1 \to 1$ as $kh \to 0$.

□

The convergence of the physical eigenvalue can be easily verified from Figs. 2.1 and 2.2 which show that $\lambda_1 \to 1$ as $kh \to 0$. The consistency of the Flux Reconstruction formulation can now be proven as follows.

Theorem 3.1. *The FR numerical conservation law, Eqn. (2.3), is consistent if the differentiation operator $\boldsymbol{Q}$ is diagonalizable.*

Proof. For a consistent FR scheme, the numerical conservation law, Eqn. (2.3), must converge to the vectorized form of the exact conservation law, Eqn. (2.1), in the asymptotic limit. Since integration in time is assumed to be perfect, we require that the numerical derivative vector in the parent space, $\boldsymbol{Q}(kh)\boldsymbol{u}_{\boldsymbol{n}}^{\boldsymbol{\delta}}(t)$, converges to the exact derivative values at the solution points,

$$\frac{\partial u_n}{\partial \xi}(\boldsymbol{x}_{\boldsymbol{n}}, t) = \exp(ikx_n)\exp(-ikct)ik\frac{h}{2}\boldsymbol{w}_0 \to \exp(ikx_n)\exp(-ikct)ik\frac{h}{2}\mathbf{1} \text{ as } kh \to 0. \tag{3.13}$$

In order to prove this, we can use Eqns. (2.18) and (2.20), so that

$$\begin{aligned}\boldsymbol{Q}(kh)\boldsymbol{u}_{\boldsymbol{n}}^{\boldsymbol{\delta}}(t) &= \boldsymbol{Q}(kh)\exp(ikx_n)\sum_{p=1}^{P+1}\exp(-ikc\lambda_p t)\beta_p \boldsymbol{w_p} \\ &= \exp(ikx_n)\sum_{p=1}^{P+1}\exp(-ikc\lambda_p t)\beta_p\gamma_p \boldsymbol{w_p}. \end{aligned} \tag{3.14}$$

Next, we pass to the asymptotic limit $kh \to 0$, and utilize the results from Lemmas 3.3, 3.4 and 3.5, so that

$$\boldsymbol{Q}(kh)\boldsymbol{u}_{\boldsymbol{n}}^{\boldsymbol{\delta}}(t) \to \exp(ikx_n)\exp(-ikc\lambda_1 t)\gamma_1 \mathbf{1} \text{ as } kh \to 0. \tag{3.15}$$

Finally, Lemma 3.6 provides us with the limit on λ_1, so that

$$\boldsymbol{Q}(kh)\boldsymbol{u}_{\boldsymbol{n}}^{\boldsymbol{\delta}}(t) \to \exp(ikx_n)\exp(-ikct)ik\frac{h}{2}\mathbf{1} \text{ as } kh \to 0. \tag{3.16}$$

□

Chapter 4

Convergence of the semi-discrete formulation

We are finally prepared to derive results on the convergence of the FR formulation. Note that Theorem 3.1 suggests that any FR scheme that satisfies Lax's stability criterion must be convergent by Lax's equivalence theorem. Nevertheless, we begin this chapter with an explicit proof of convergence which would prepare us for subsequent results on the rate of convergence of FR.

4.1 Proof of convergence

Theorem 4.1. *For any FR scheme, which is stable in the von Neumann sense, Eqn. (2.24), if the differentiation operator $\boldsymbol{Q}$ is diagonalizable, then as $kh \to 0$, the pointwise error vanishes, $\boldsymbol{e}_n \to \boldsymbol{0}$, guaranteeing convergence at the solution points.*

Proof. Using Eqn. (2.23),

$$\begin{aligned}\|\boldsymbol{e}_n(t)\|_2 &= \| \sum_{p=1}^{P+1} \left[\exp(kc\lambda_p^{\mathrm{Im}}t)\exp(-ikc(\lambda_p^{\mathrm{Re}}-1)t)-1\right]\beta_p \boldsymbol{w_p}\|_2 \\ &\leq \sum_{p=1}^{P+1} \left|\exp(kc\lambda_p^{\mathrm{Im}}t)\exp(-ikc(\lambda_p^{\mathrm{Re}}-1)t)-1\right| \, |\beta_p|\end{aligned}$$

$$\leq \left|\exp(kc\lambda_1^{\mathrm{Im}}t)\exp(-ikc(\lambda_1^{\mathrm{Re}}-1)t)-1\right||\beta_1| + \sum_{p=2}^{P+1}\left(|\exp(kc\lambda_p^{\mathrm{Im}}t)\exp(-ikc(\lambda_p^{\mathrm{Re}}-1)t)|+1\right)|\beta_p|. \tag{4.1}$$

Now observe that $|\exp(-ikc(\lambda_p^{\mathrm{Re}}-1)t)|=1$ and $|\exp(kc\lambda_p^{\mathrm{Im}}t)|\leq 1$ by the stability criterion on eigenvalues, Eqn. (2.24), so that by the triangle inequality,

$$\|\boldsymbol{e}_n(t)\|_2 \leq \left|\exp(kc\lambda_1^{\mathrm{Im}}t)\exp(-ikc(\lambda_1^{\mathrm{Re}}-1)t)-1\right||\beta_1| + 2\sum_{p=2}^{P+1}|\beta_p|, \tag{4.2}$$

which can be simplified via the convergence of $\boldsymbol{\beta}$, Lemma 3.5, as

$$\lim_{kh\to 0}\|\boldsymbol{e}_n(t)\|_2 \leq \lim_{kh\to 0}\left|\exp(-ikc(\lambda_1^{\mathrm{Re}}(kh)-1)t)-1\right|\sqrt{P+1}. \tag{4.3}$$

Now, under the conditions of Lemma 3.6, $\lambda_1\to 1$, yielding $\lim_{kh\to 0}\|\boldsymbol{e}_n(t)\|_2 = 0$. □

Theorem 4.1 shows that the numerical solution obtained via FR converges to the exact solution, pointwise, at the set of solution points. Clearly, the use of a spatially discrete numerical conservation law, Eqn. (2.16), has restricted us to pointwise convergence. However, such results have been routinely used for finite-difference schemes, and it is easy to show that pointwise convegence extends to all points in the element owing to the smoothness of the exact and numerical solutions.

Lemma 4.1. *Under the conditions of Theorem 4.1, an FR scheme which converges pointwise at the set of solution points, converges at all points in the element* Ω_n.

Proof. Using the analytical form of the exact solution,

$$\begin{aligned} u(x,t) &= \mathscr{P}(u(x,t)) + (1-\mathscr{P})(\exp(ik(x-ct))), \quad x\in\Omega_n, \\ &= \mathscr{P}(u(x,t)) + \exp(ik(x_n-ct))(1-\mathscr{P})(\exp(ik\frac{h}{2}(1+\xi))), \quad x\in\Omega_n, \end{aligned} \tag{4.4}$$

where $\mathscr{P}$ denotes the collocation projection operator that interpolates its argument at the

set of solution points in the n^{th} element, ξ_p for $p = 1, 2, \ldots, P+1$,

$$\mathscr{P}(u(x,t)) = \sum_{p=1}^{P+1} u_{n,p}(t)\ell_{n,p}(x), \quad x \in \Omega_n. \tag{4.5}$$

From standard error estimates in Lagrange interpolation [123],

$$\begin{aligned}\left|(1-\mathscr{P})(\exp\left(ik\frac{h}{2}(1+\xi)\right))\right| &\leq \sup_{\phi\in[-1,1]}\left|\left(ik\frac{h}{2}\right)^{P+1}\exp\left(ik\frac{h}{2}(1+\phi)\right)\frac{\prod_{q=1}^{P+1}(\xi-\xi_q)}{(P+1)!}\right| \\ &\leq a(P,\boldsymbol{\xi})(kh)^{P+1}. \end{aligned} \tag{4.6}$$

for some constant a that depends on P and $\boldsymbol{\xi}$. Then, from the polynomial form of the numerical solution, Eqn. (2.2), we get

$$|u^{\delta}(x,t) - u(x,t)| = \left|\sum_{p=1}^{P+1}(u^{\delta}_{n,p}(t) - u_{n,p}(t))\ell_{n,p}(x)\right| + \mathscr{O}\left((kh)^{P+1}\right), \quad x \in \Omega_n. \tag{4.7}$$

The proof now follows from Theorem 4.1. □

4.2 Rate of convergence of the FR formulation

Having established convergence, we now turn our attention to the rate of convergence. In this section, we show that, in the limit of asymptotic resolution, the pointwise error is bounded by a linear function of time. The constant term of the function depends on the rates of convergence of the modal weights, β_p, for the spurious modes, and linear term depends on the rate of convergence of the normalized physical eigenvalue, λ_1.

For short times, $t \sim \tau_{\text{spurious}}$, the rate of convergence is largely dictated by that of the spurious modes and is typically $\mathscr{O}((kh)^{P+q})$. Here $q \in \{0,1,2\}$ depends on the type of interface flux and whether P is odd or even. Thereafter, the growth of error is slow corresponding to the physical mode which can be super-convergent. Once the error associated with the physical mode becomes significantly greater than the fraction of energy received

by the spurious modes, the scheme exhibits super-convergence. For instance, in the case of DG, the long-time solution converges at the rate $\mathscr{O}((kh)^{2P+q})$.

Discussion regarding the impact of boundary conditions on convergence is included in Appendix A which rigorously quantifies accuracy in elements close to the boundaries of the domain.

4.2.1 Rate of convergence of pointwise error

Theorem 4.2. *For a convergent FR scheme of order P, the norm of pointwise error, Eqn. (2.29), is bounded above by a term that grows linearly in time,*

$$\|e_n(t)\|_2 \leq 2\,\mathscr{O}\left((kh)^{\min_{p=2}^{P+1} q_p^\beta}\right) + \sqrt{P+1}ck\,\mathscr{O}\left((kh)^{q_1^\lambda}\right)t, \tag{4.8}$$

where $q_1^\lambda > 0$ is the asymptotic rate of convergence of the normalized physical eigenvalue, and $q_p^\beta > 0$ for $p = 1,2,\ldots,P+1$ is that for the modal weights,

$$\lambda_1(kh) = 1 + \mathscr{O}\left((kh)^{q_1^\lambda}\right), \quad kh \to 0, \tag{4.9}$$

$$\beta_p(kh) = \sqrt{P+1}\delta_{1p} + \mathscr{O}\left((kh)^{q_p^\beta}\right), \quad kh \to 0, \tag{4.10}$$

Proof. The definitions in Eqns. (4.9) and (4.10) follow directly from Lemmas 3.6 and 3.5 respectively. Substituting in Eqn. (4.2), we get

$$\begin{aligned}
\|e_n(t)\|_2 &\leq |\exp(-ikc(\lambda_1(kh)-1)t)-1|\,|\beta_1(kh)| + 2\sum_{p=2}^{P+1}|\beta_p(kh)| \\
&= \left|\exp(-ikc\mathscr{O}((kh)^{q_1^\lambda})t)-1\right|\left|\sqrt{P+1}+\mathscr{O}((kh)^{q_1^\beta})\right| + 2\sum_{p=2}^{P+1}|\mathscr{O}((kh)^{q_p^\beta})| \\
&= \sqrt{P+1}\left|kc\mathscr{O}((kh)^{q_1^\lambda})t\right| + 2\sum_{p=2}^{P+1}|\mathscr{O}((kh)^{q_p^\beta})|
\end{aligned}$$

$$= \sqrt{P+1}ck\,\mathscr{O}\left((kh)^{q_1^\lambda}\right)t + 2\,\mathscr{O}\left((kh)^{\min\limits_{p=2}^{P+1} q_p^\beta}\right)\sum_{p=2}^{P+1}|\mathscr{O}((kh)^{q_p^\beta - \min\limits_{p=2}^{P+1} q_p^\beta})|$$

$$= \sqrt{P+1}ck\,\mathscr{O}\left((kh)^{q_1^\lambda}\right)t + 2\,\mathscr{O}\left((kh)^{\min\limits_{p=2}^{P+1} q_p^\beta}\right), \tag{4.11}$$

which yields the desired result. □

Eqn. (4.8) shows that the rate of convergence for FR is bounded by the sum of two terms. The first term depends on the spurious modes and is constant with respect to time. The second term depends on the physical mode and grows linearly in time. For short times, $t \sim \tau_{\text{spurious}}$, the rate of convergence is bounded by $\mathscr{O}\left((kh)^{\min\limits_{p=2}^{P+1} q_p^\beta}\right)$ corresponding to the rate at which the modal weights for the spurious modes vanish. After a long time, $t \gg \tau_{\text{spurious}}$, the rate of convergence is bounded by $\mathscr{O}((kh)^{q_1^\lambda})$ corresponding to the rate at which the normalized physical eigenvalue approaches the exact value. The linearity of the long-time error can be clearly observed in Fig. 2.4.

It is important to note that the long-time convergence rate in Eqn. (4.8) is sensitive to the process through which the limit of $kh \to 0$ is realized. If $k \to 0$ for constant h, the order is one unit higher than if $h \to 0$ for constant k:

$$\sqrt{P+1}ck\,\mathscr{O}\left((kh)^{q_1^\lambda}\right)t = \sqrt{P+1}c\,\mathscr{O}\left(k^{q_1^\lambda+1}\right)\mathscr{O}\left(h^{q_1^\lambda}\right)t. \tag{4.12}$$

This difference of one order can be observed in the definitions of the 'order of accuracy' proposed, somewhat heuristically, in [69, 141]. It is clear now that those definitions correspond to the long-time rate obtained when $h \to 0$ for constant k. Such long-time superconvergence is realized for problems exhibiting an eventual quasi-steady solution which does not depend on the initial condition [10]. However, for most other practical simulations, the realized convergence rate would typically lie somewhere in between the short-time and long-time rates depending on the characteristic time-scale of the problem.

Table 4.1 records the relevant rates of convergence, q_p^λ, q_p^β, for the physical and the spurious modes for the fully upwinded and centered DG schemes of order $P = 2$ on Gauss

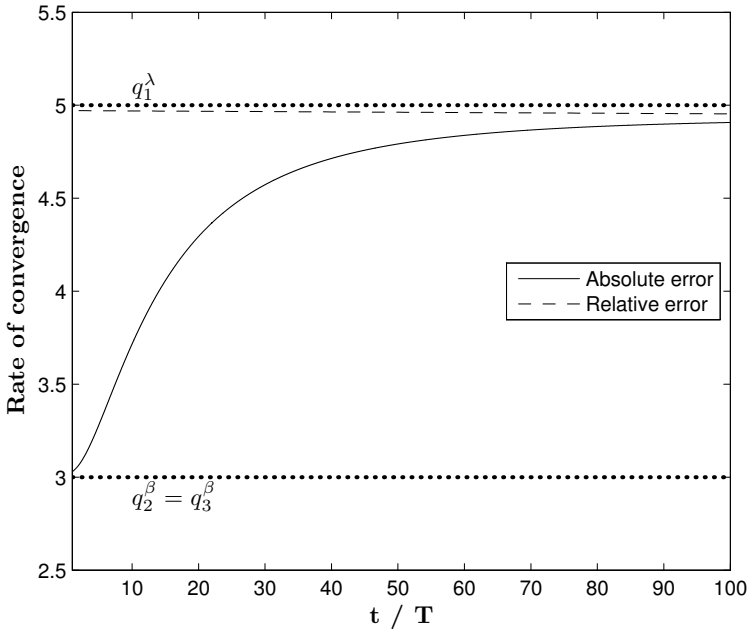
Legendre solution points. In the case of $\alpha = 0$, we expect the rate of convergence to start from $\min_{p=2}^{3} q_p^\beta \simeq 3$ and asymptotically approach $q_1^\lambda \simeq 5$, when h is refined for fixed k. This behavior can be verified from Fig. 4.1a which plots the evolution of the rate of convergence for the fully upwinded DG scheme of order $P = 2$ on Gauss Legendre solution points, $k = \frac{\pi}{2}$, unit wavespeed $c = 1$, using a pair of grid spacings $h = \{0.5, 0.25\}$. Similarly, when refining k for fixed h, on account of Eqn. (4.12), we expect the convergence rate to start from ~ 3 and asymptote to ~ 6. This behavior can be verified from Fig. 4.1b which plots the convergence rate for grid spacing $h = 0.25$ and two wavenumbers $k = \{\pi, \frac{\pi}{2}\}$.

	$\alpha = 0$		$\alpha = \frac{1}{2}$	
	q_p^λ	q_p^β	q_p^λ	q_p^β
Mode 1 (Physical)	4.998	4.002	6.004	4.004
Mode 2 (Spurious)	-	3.037	-	4.057
Mode 3 (Spurious)	-	2.956	-	3.944

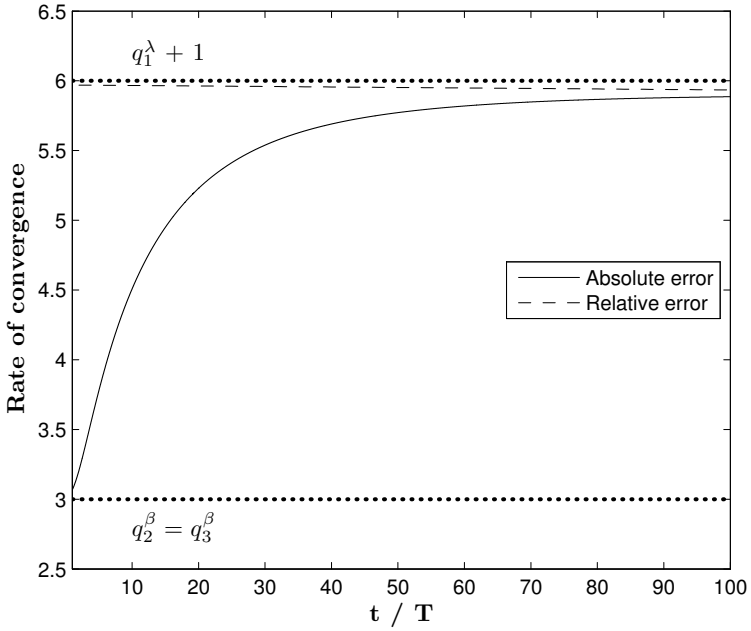
Table 4.1: Rates of convergence for the physical and the spurious modes for the DG scheme of order $P = 2$ on Gauss Legendre solution points. The rates are calculated using $\frac{kh}{(P+1)\pi} = 0.01$ and $\frac{kh}{(P+1)\pi} = 0.02$.

In the case of a centered flux, the absence of dissipation implies that the spurious modes persist forever. The phase error introduced by these modes manifests in the oscillatory rate of convergence that can be seen in Figs. 4.2a and 4.2b. The super-convergent rate is observed only after the error in the physical mode grows to dominate over that of the spurious modes.

Theorem 4.2 provides a semi-analytical procedure for evaluating both the short-time and the long-time rate of convergence for any given scheme within the FR formulation. This procedure involves a direct eigenvalue computation which requires almost no computational effort (instantaneous on modern machines for all practical values of P). This is a significant improvement over grid-convergence studies of the form used until very recently [141] which can potentially require multiple days of computing time for determining the long-time rate of convergence. Table 4.2 records the short-time and long-time rates of convergence for $P = 0$ to 5 for the DG scheme using Gauss-Legendre solution points for both upwinded and centered fluxes.

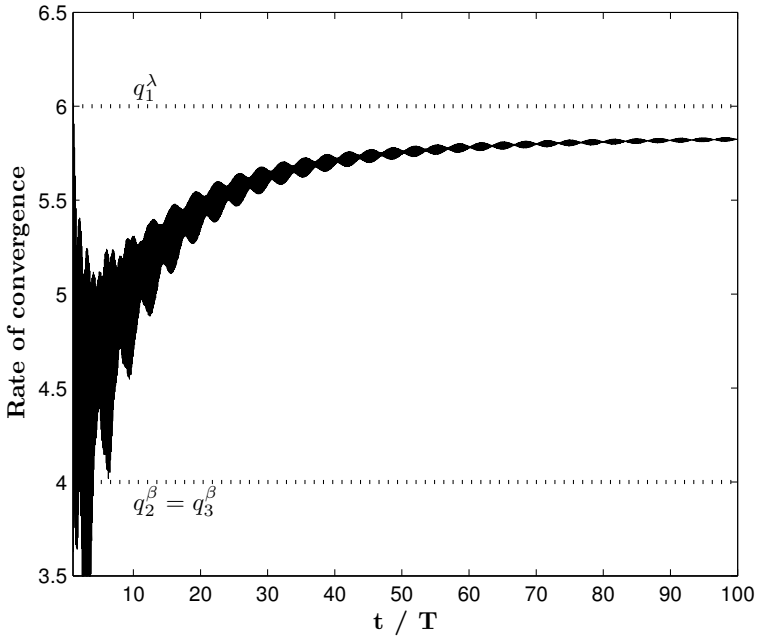


(a) Refinement in grid spacing: $k = \frac{\pi}{2}$, grid spacings $h = \{0.5, 0.25\}$.

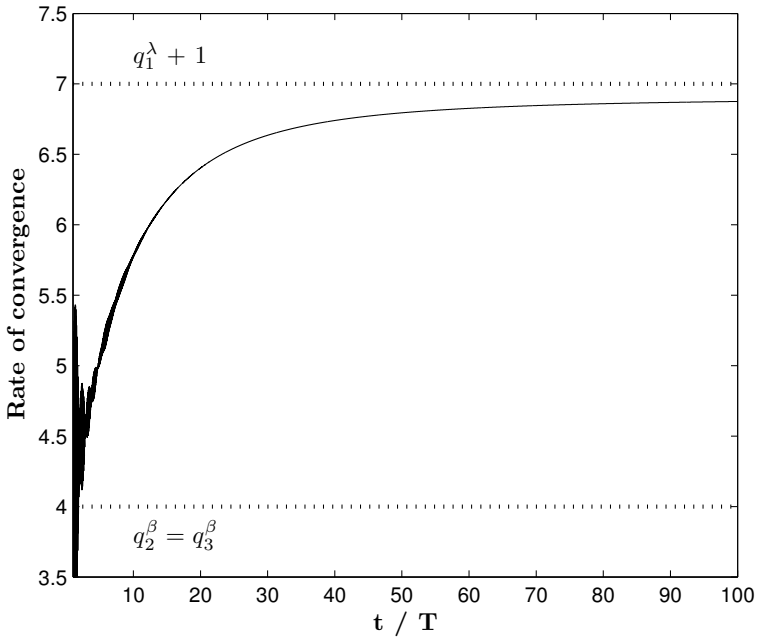


(b) Refinement in wavenumber: $h = 0.25$, wavenumbers $k = \{\pi, \frac{\pi}{2}\}$.

Figure 4.1: Rate of convergence for the fully upwinded DG scheme of order $P = 2$ on Gauss Legendre solution points with unit wavespeed $c = 1$. The reference time for relative error is taken to be $t_0 = T$.



(a) Refinement in grid spacing: $k = \frac{\pi}{2}$, grid spacings $h = \{0.5, 0.25\}$.



(b) Refinement in wavenumber: $h = 0.25$, wavenumbers $k = \{\pi, \frac{\pi}{2}\}$.

Figure 4.2: Rate of convergence for the centered DG scheme of order $P = 2$ on Gauss Legendre solution points with unit wavespeed $c = 1$.

	$\alpha = 0$		$\alpha = \frac{1}{2}$	
P	q_1^λ	$\min_{p=2}^{P+1} q_p^\beta$	q_1^λ	$\min_{p=2}^{P+1} q_p^\beta$
0	1	-	2	-
1	3	2	2	1
2	5	3	6	4
3	7	4	6	3
4	9	5	10	6
5	11	6	10	5
P	$2P+1$	$P+1$	$\begin{cases} 2P+2, & \text{even } P \\ 2P, & \text{odd } P \end{cases}$	$\begin{cases} P+2, & \text{even } P \\ P, & \text{odd } P \end{cases}$

Table 4.2: Rates of convergence of the eigensystem components for DG using Gauss-Legendre solution points obtained from direct eigenvalue computation.

4.2.2 Rate of convergence of 'relative' error for upwinded schemes

From a computational perspective, using a conventional grid convergence study to determine the rate of convergence of the physical mode requires long-time integration while ensuring that the time step is appropriately small to minimize time-stepping errors. This results in an intensive exercise even for simple 1-D cases. In order to reduce the computational effort, the authors in [65, 141] utilized a measure that compares error after an integer number of time-periods to the error at time $t = 0$ which is on account of the collocation projection. Towards this end, define a 'relative' error of the form,

$$\boldsymbol{e}_n^R(t_0 + mT; t_0) = \boldsymbol{u}_n^{\boldsymbol{\delta}}(t_0 + mT) - \boldsymbol{u}_n^{\boldsymbol{\delta}}(t_0), \tag{4.13}$$

where $T = 2\pi/(kc)$ is the time period of the wave, t_0 is a reference time and $m \in \mathbb{Z}^+$ is the number of time periods since the reference time. Such a definition suppresses the error

offset introduced due to the spurious modes outlined in Eqn. (2.30),

$$\begin{aligned}
\|\boldsymbol{e}_n^R(t_0 \gg \tau_{\text{spurious}})\|_2 &= \|\boldsymbol{e}_n(t_0+mT) - \boldsymbol{e}_n(t_0)\|_2 \\
&\simeq \| \left[\exp(-ikc(\lambda_1-1)(t_0+mT)) - \exp(-ikc(\lambda_1-1)t_0)\right] \beta_1 \boldsymbol{w_1} \|_2 \\
&= |\exp(-ikc(\lambda_1-1)t_0)\left(\exp(-ikc(\lambda_1-1)mT)-1\right)| \, |\beta_1| \\
&\leq |\exp(-ikc(\lambda_1-1)mT)-1| \, |\sqrt{P+1} + \mathscr{O}\left((kh)^{q_1^\beta}\right)| \\
&= |-ikc(\lambda_1-1)mT + \mathscr{O}\left(k^2(\lambda_1-1)^2\right)| \sqrt{P+1} \\
&= \sqrt{P+1} ck \, \mathscr{O}\left((kh)^{q_1^\lambda}\right) mT, \qquad (4.14)
\end{aligned}$$

which shows that the rate of convergence of the relative error is bounded only by the rate of convergence of the normalized eigenvalue λ_1. As can be seen from Table 2.1, typically $\tau_{\text{spurious}} < T$, so that a choice of $t_0 = T$ is sufficient. This can be verified from Fig. 4.1 which shows that the rate of convergence computed for the relative error yields the value ~ 5, for h refinement, and ~ 6, for k refinement, for all $m \geq 1$.

4.3 Numerical experiments in 2-D

In this final section, we perform additional numerical experiments to investigate if the theoretical rates of convergence derived for 1-D linear advection also hold for more general conservation laws. The first experiment deals with an advection-diffusion equation over a 2-D domain. Details regarding the application of FR schemes to such second-order equations can be found in [71]. The second experiment deals with a system of non-linear equations in 2-D subjected to a special initial condition at $t = 0$ that allows for a smooth analytical solution for all time $t \geq 0$. The results are plotted in the style of Fig. 4.1.

For generality of interpretation, we measure error with respect to both discrete and functional standard norms. The discrete $\ell^2(\Omega^\delta)$ norm is simply the averaged vector norm of pointwise error at the set of solution points,

$$\|e\|_{\ell^2(\Omega^\delta)}(t) := \left[\frac{1}{N(P+1)^{n_{\text{sd}}}} \sum_{n=1}^{N} \sum_{p=1}^{(P+1)^{n_{\text{sd}}}} \left|e(x_{n,p},t)\right|^2\right]^{1/2}, \qquad (4.15)$$

where $x_{n,p}$ is the pth solution point in the nth element, N is the total number of elements, and n_{sd} is the number of space dimensions. Eqn. (4.15) essentially generalizes the definition introduced in Eqn. (2.28). The functional $L^2(\Omega)$ norm is the functional 2-norm of error over the entire domain,

$$\|e\|_{L^2(\Omega)}(t) := \left[\int_{\Omega} |e(x,t)|^2 \, \mathrm{d}\Omega\right]^{1/2}. \tag{4.16}$$

Note that Theorem 4.2 prescribes the rate of convergence only at the set of solution points, i.e. with respect to the $\ell^2(\Omega^\delta)$ norm. However, the rate of convergence at any other point in the element can at the least be easily bounded. Towards this end, use Eqn. (4.7) in Lemma 4.1, so that for $x \in \Omega_n$,

$$\begin{aligned}
|e(x,t)| &= \left|\sum_{p=1}^{P+1} (u^{\delta}_{n,p}(t) - u_{n,p}(t))\ell_{n,p}(x)\right| + \mathscr{O}\left((kh)^{P+1}\right) \\
&\leq \sum_{p=1}^{P+1} \left|u^{\delta}_{n,p}(t) - u_{n,p}(t)\right| \left|\ell_{n,p}(x)\right| + \mathscr{O}\left((kh)^{P+1}\right) \\
&\leq \max_{p=1}^{P+1} \left|\ell_{n,p}(x)\right| \|\boldsymbol{e}_n(t)\|_1 + \mathscr{O}\left((kh)^{P+1}\right) \\
&\leq \mu \max_{p=1}^{P+1} \left|\ell_{n,p}(x)\right| \|\boldsymbol{e}_n(t)\|_2 + \mathscr{O}\left((kh)^{P+1}\right) \\
&\leq \mu \max_{p=1}^{P+1} \left|\ell_{n,p}(x)\right| \left[2\,\mathscr{O}\left((kh)^{\min_{p=2}^{P+1} q_p^{\beta}}\right) + \sqrt{P+1}ck\,\mathscr{O}\left((kh)^{q_1^{\lambda}}\right)t\right] + \mathscr{O}\left((kh)^{P+1}\right),
\end{aligned} \tag{4.17}$$

where μ is the constant associated with the equivalence of finite-dimensional norms. This suggests that the super-convergence whereby even polynomial orders with centered fluxes converge in short-time at the order $P+2$ (c.f. Table 4.2) may not be exhibited when error is measured using a functional norm.

In the subsections that follow, we show that the results of numerical experiments are in excellent agreement with the theoretical predictions. Convergence in the two norms is nearly identical, except in the case of even polynomial orders when using centered fluxes.

4.3.1 Linear advection-diffusion in 2-D

Consider the linear advection-diffusion of a smooth scalar field, $u : [-1,1]^2 \mapsto \mathbb{R}$, in 2-D,

$$\frac{\partial u}{\partial t} + a_x \frac{\partial u}{\partial x} + a_y \frac{\partial u}{\partial y} - b\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) = 0, \quad (x,y) \in (-1,1)^2,\ t > 0, \tag{4.18}$$

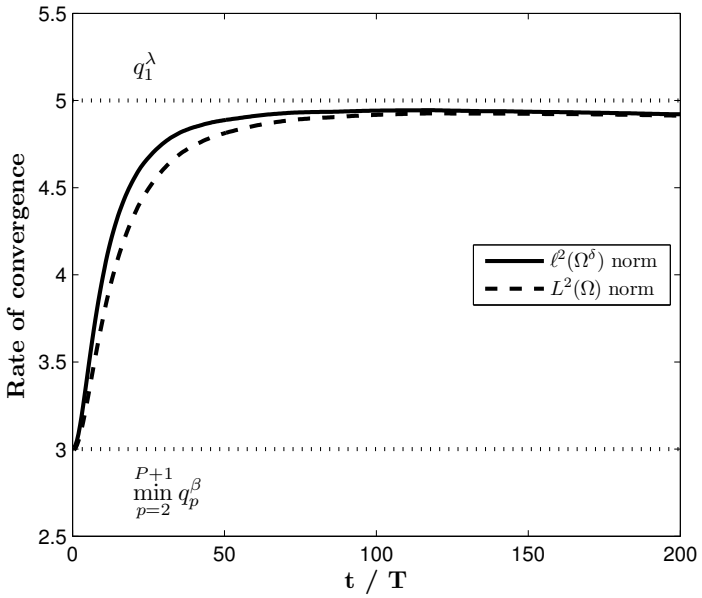
$$u(x,y,0) = \sin(\pi x)\sin(\pi y), \tag{4.19}$$

where $a_x \in \mathbb{R}$, $a_y \in \mathbb{R}$ are the advection speeds, and $b \in \mathbb{R}_+$ is the constant diffusion coefficient. The boundary conditions are assumed to be periodic for simplicity. The analytical solution to Eqns. (4.18 - 4.19) for periodic boundary conditions is given by

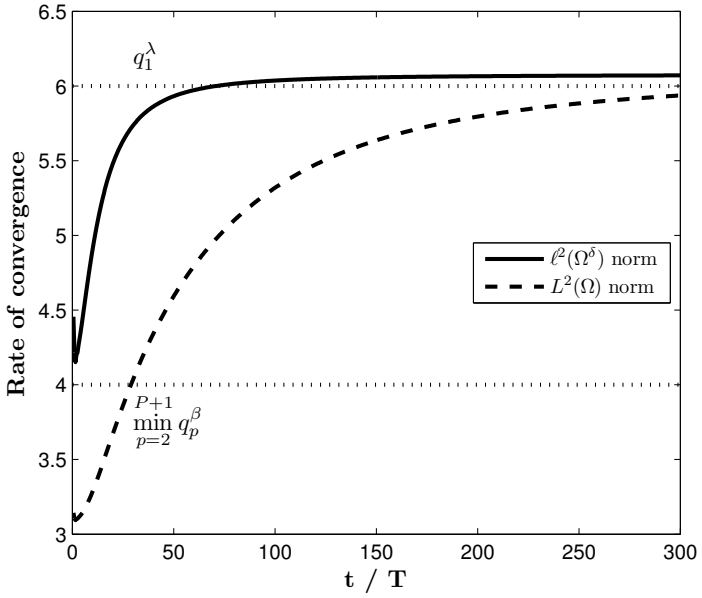
$$u(x,y,t) = \exp\left(-2b\pi^2 t\right)\sin(\pi(x-a_x t))\sin(\pi(y-a_y t)), \quad x \in [-1,1]^2,\ t \geq 0. \tag{4.20}$$

This problem is solved using DG, recovered via FR, for $P = 2,3$ using Gauss-Legendre points as solution points. The numerical solution is marched in time using the standard four-stage fourth order accurate RK4 time stepping scheme for two thousand time-periods in order to capture the long-time rate of convergence. At the same time, the time-step is chosen small enough to ensure that errors due to numerical integration are much smaller than those due to spatial-discretization. Specifically, the time step is fixed at 10^{-3} for $P = 2$ and 10^{-4} for $P = 3$. The longest calculation involves roughly 40 million iterations. The advection speeds are set to unity, $a_x = a_y = 1$. The diffusion coefficients are chosen small enough to ensure that the amplitude of the solution remains much larger than machine precision for the entire calculation. Specifically, $b = 10^{-3}$ for $P = 2$ and 10^{-4} for $P = 3$. The common interface value for the advective flux is calculated using the Lax-Friedrichs or Rusanov formula (parameterized by upwinding parameter α, Eqn. (2.4)), while that for the diffusive flux is calculated using the LDG formula (parameterized by upwinding coefficient β_{LDG}, and jump coefficient τ_{LDG}). Both one-sided, isotropic ($\alpha = 0, \beta_{\text{LDG}} = 1/2, \tau_{\text{LDG}} = 0$) and fully centered ($\alpha = 1/2, \beta_{\text{LDG}} = 0, \tau_{\text{LDG}} = 0$) fluxes are investigated. The rate of convergence is calculated at each time-step using two uniform grids with 8×8 and 16×16 elements respectively.

Figures 4.3 and 4.4 plot the rate of convergence against the number of time-periods

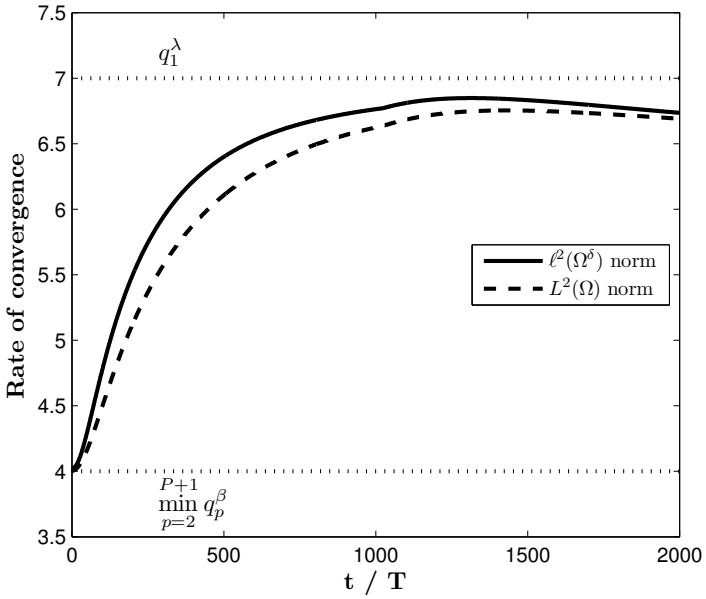


(a) Upwinded interface flux, $\alpha = 0, \beta_{\text{LDG}} = 1/2, \tau_{\text{LDG}} = 0$.

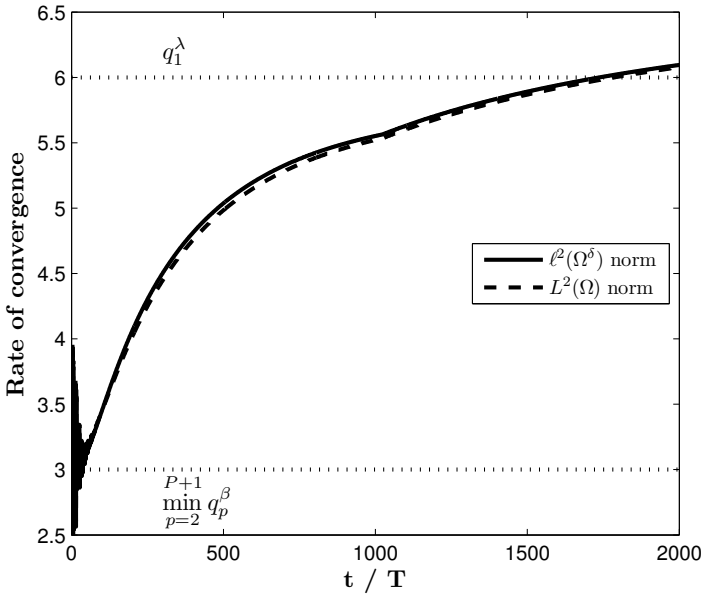


(b) Centered interface flux, $\alpha = 1/2, \beta_{\text{LDG}} = 0, \tau_{\text{LDG}} = 0$.

Figure 4.3: 2-D advection-diffusion: Rate of convergence for the DG scheme of order $P = 2$ on Gauss Legendre solution points.



(a) Upwinded interface flux, $\alpha = 0, \beta_{\text{LDG}} = 1/2, \tau_{\text{LDG}} = 0$.



(b) Centered interface flux, $\alpha = 1/2, \beta_{\text{LDG}} = 0, \tau_{\text{LDG}} = 0$.

Figure 4.4: 2-D advection-diffusion: Rate of convergence for the DG scheme of order $P = 3$ on Gauss Legendre solution points.

for $P = 2$ and $P = 3$ respectively. We see that the results of numerical experiments are in excellent agreement with the theoretical expectations. Convergence in the two norms is nearly identical, except in the case of $P = 2$ when using centered fluxes. In this case, the numerically observed short-time rate of convergence for the functional norm is $P+1$ while that for the discrete norm is $P+2$.

An important empirical observation from this experiment is that the long-time super-convergence rate of $2P+q$ is exhibited by the FR scheme only after multiple thousand time periods depending on the polynomial order. In regard to applications, such a rate is relevant only for problems that either have periodic boundary conditions or exhibit a quasi-steady solution of interest [10]. This has been further validated in Appendix A which shows that convergence in elements close to boundaries of the domain is limited to the short-time rate of $P+q$.

4.3.2 Isentropic Euler vortex in 2D

We now consider the Euler equations for a calorically perfect ideal gas,

$$\frac{\partial}{\partial t}\begin{bmatrix}\rho\\ \rho u\\ \rho v\\ \rho E\end{bmatrix}+\frac{\partial}{\partial x}\begin{bmatrix}\rho u\\ \rho u^2+p\\ \rho uv\\ \rho Hu\end{bmatrix}+\frac{\partial}{\partial y}\begin{bmatrix}\rho v\\ \rho uv\\ \rho v^2+p\\ \rho Hv\end{bmatrix}=\mathbf{0},\quad \text{in } \Omega\subset\mathbb{R}^2,\, t>0, \tag{4.21}$$

$$p=(\gamma-1)\left(\rho E-\frac{1}{2}\rho(u^2+v^2)\right), \tag{4.22}$$

where ρ denotes density, u, v denote the velocity components along x and y respectively, E denotes total internal energy, p denotes pressure, $H = E + p/\rho$ denotes total enthalpy, and γ denotes the ratio of specific heats. For this experiment, we utilize the isentropic vortex test case introduced by Shu [37]. Herein, the initial condition consists of a uniform flow in the y direction superposed with an isentropic vortex centered at the origin. Entropy is constant while the vorticity field decays exponentially in the radial direction. Following the

prescription in [141],

$$\begin{bmatrix} \rho \\ u \\ v \\ p \end{bmatrix}(x,y,0) = \begin{bmatrix} \left[1-\frac{\gamma-1}{2}(\omega MRf)^2\right]^{\frac{1}{\gamma-1}} \\ -\omega fy \\ 1+\omega fx \\ \frac{1}{\gamma M^2}\rho(x,y,0)^{\gamma} \end{bmatrix}, \tag{4.23}$$

$$\omega = \frac{\Gamma}{2\pi R} \tag{4.24}$$

$$f = \exp\left(\frac{1-(x^2+y^2)}{2R^2}\right) \tag{4.25}$$

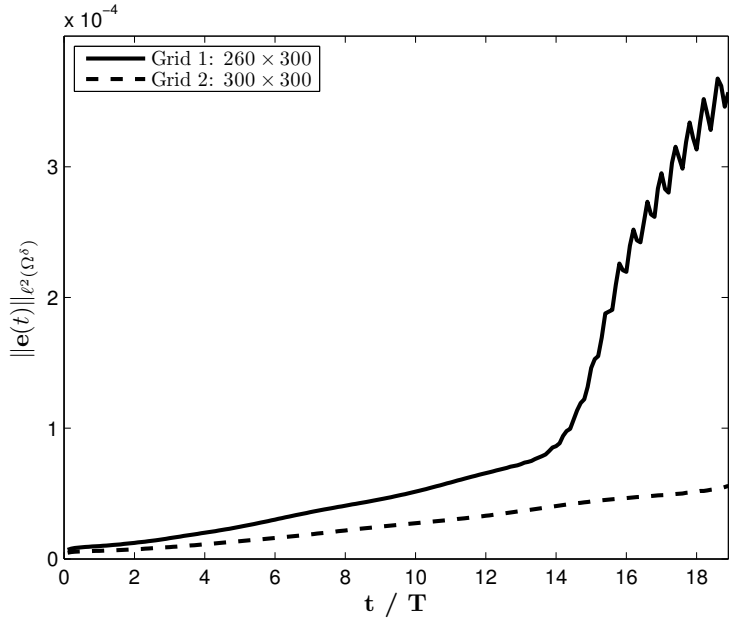
where Γ is the strength of the vortex, M is the free-stream Mach number and R is a measure of the 'radius' of the vortex. For this experiment, we have used the values in [141] so that $\Gamma = 13.5$, $M = 0.4$, $R = 1.5$.

The exact solution for this problem for an infinite domain, $\Omega = \mathbb{R}^2$, is the propagation of the vortex with unit velocity in the y direction. In the numerical simulations, the domain is restricted to a square box, $[-40,40]^2$ with periodic boundary conditions imposed on the boundaries. Note that such periodicity effectively models a lattice of countably infinite vortex systems. While the interaction among vortices is weak, as the numerical solution advances in time, dispersion errors introduce spurious numerical wavefronts that propagate radially outwards, eventually leading to catastrophic non-linear interactions that cause a surge in the rate of error accumulation.

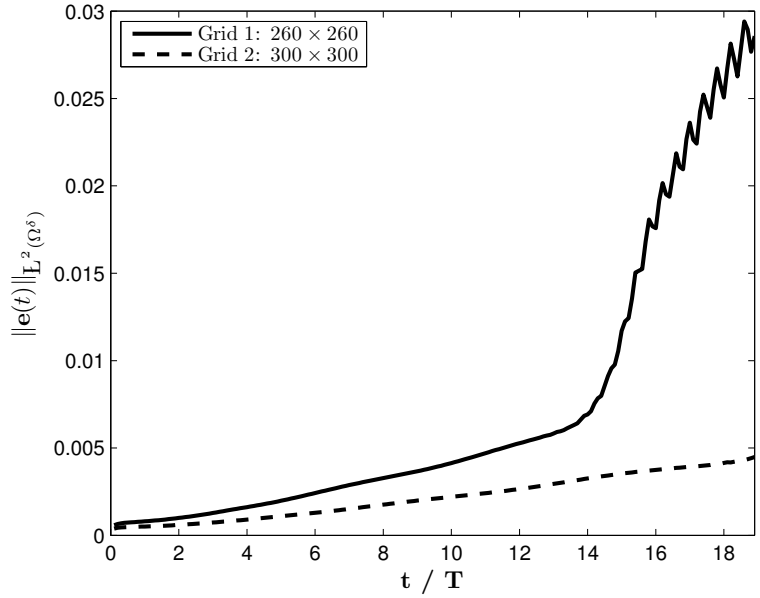
This problem is solved using DG, recovered via FR, for $P = 2,3$ using Gauss-Legendre points as solution points. The numerical solution is marched in time using RK4 and the time-step is chosen small enough to ensure that errors due to numerical integration are much smaller than those due to spatial-discretization. Specifically, the time step is fixed at 2×10^{-3} for $P = 2$ and 10^{-3} for $P = 3$. The common interface value is calculated using the fully-upwinded Lax-Friedrichs or Rusanov formula with $\alpha = 0$. The rate of convergence is calculated at each time-step using two highly refined uniform grids with 260^2 and 300^2 elements for $P = 2$, and 220^2 and 260^2 elements for $P = 3$ respectively.

Figures 4.5 and 4.6 plot the evolution of both the norms of error on both grids for $P = 2$ and $P = 3$ respectively. The plots highlight three phases of error growth. The first phase, visible in Fig. 4.6, is associated with the rapid decay of spurious modes during the initial few iterations leading to a loss of energy and corresponding error accumulation. The second phase, visible in both figures, is associated with the linear growth due to dispersion and dissipation in the physical mode. Both these phases are familiar to us on account of the dynamical system analysis performed in Chapter 2. However, for this non-linear problem, a third phase exists as well. Visible in Fig. 4.5, this phase is associated with the non-linear interaction of spurious wavefronts generated by dispersion errors in each box of the lattice. Such catastrophic growth of error can be delayed either by refining the grid or by increasing the size of the numerical domain.

Evolution of the rate of convergence for the two polynomial orders is plotted in Fig. 4.7. We have confined the time horizon to the first two phases of evolution where the behavior is essentially linear. Once again, the numerical results are in good agreement with the theoretical predictions. In particular, note from Figs. 4.5 and 4.6 that the error for $P = 3$ is several orders of magnitude lower than that for $P = 2$. Consequently, the behavior is essentially linear and the evolution of the rate of convergence is similar to that for linear experiments performed in the previous subsection. This empirical observation further impresses the importance of using high-order methods for modern problems in fluid dynamics.

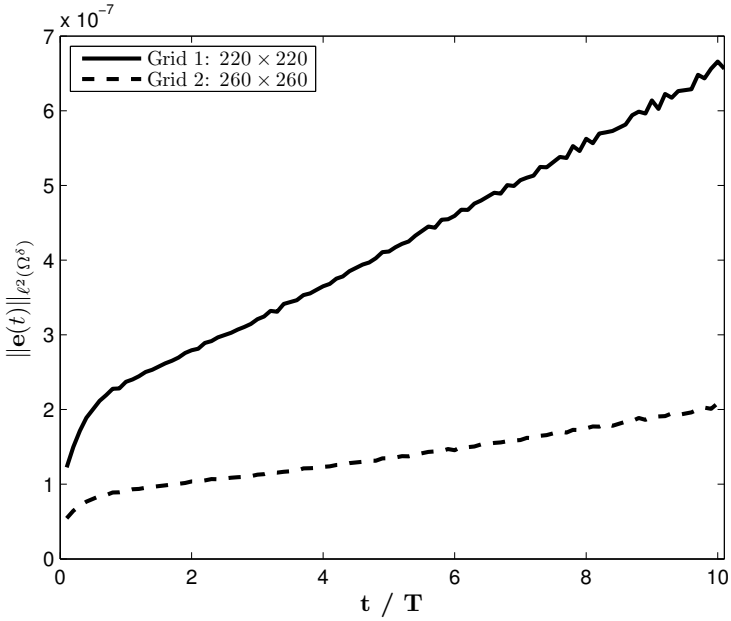


(a) $\ell^2(\Omega^\delta)$ norm

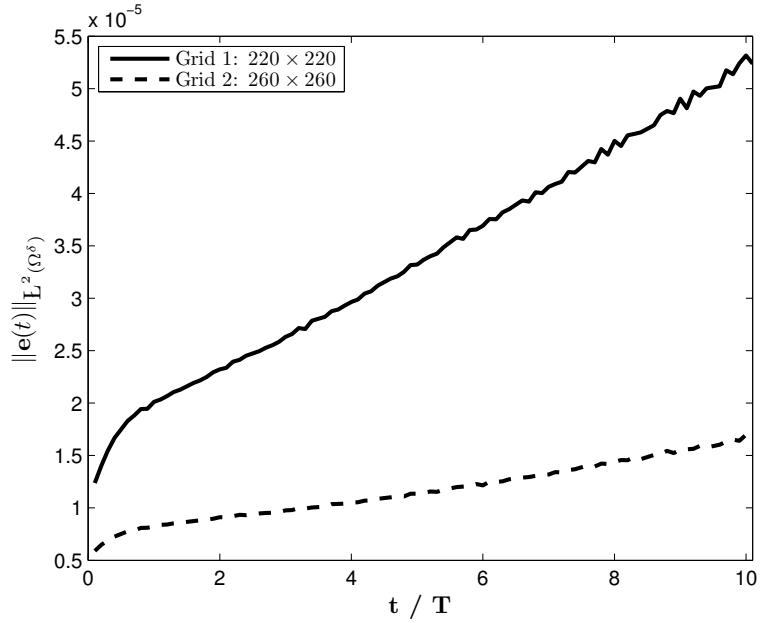


(b) $L^2(\Omega)$ norm

Figure 4.5: Isentropic vortex: Evolution of error for the DG scheme of order $P = 2$ with upwinded flux, $\alpha = 0$, on Gauss Legendre solution points.

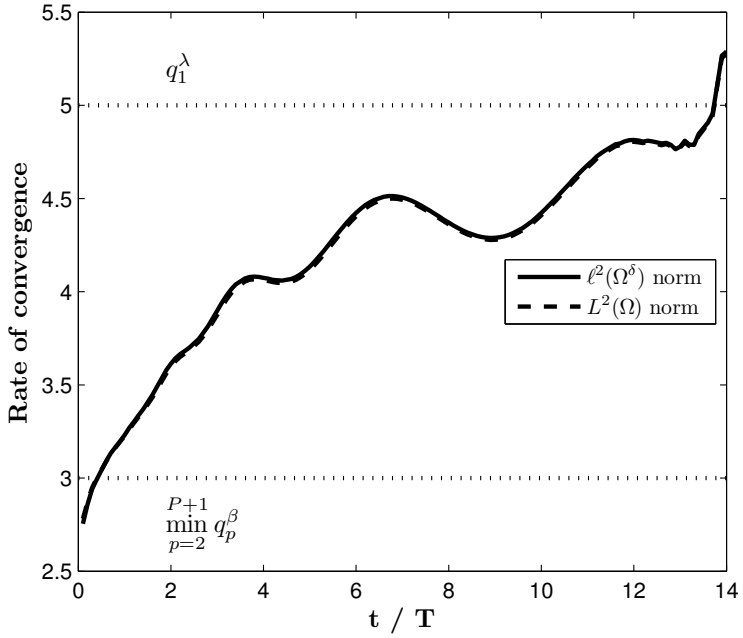


(a) $\ell^2(\Omega^\delta)$ norm

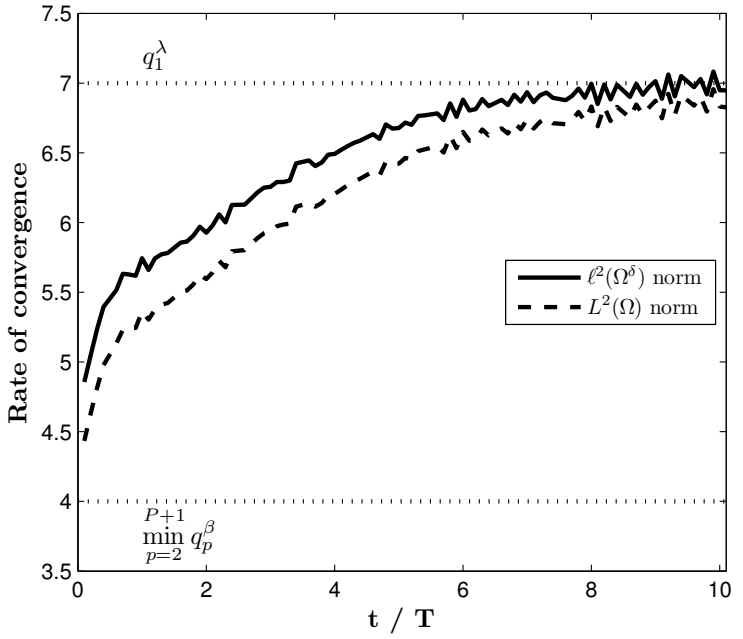


(b) $L^2(\Omega)$ norm

Figure 4.6: Isentropic vortex: Evolution of error for the DG scheme of order $P = 3$ with upwinded flux, $\alpha = 0$, on Gauss Legendre solution points.



(a) $P = 2$



(b) $P = 3$

Figure 4.7: Isentropic vortex: Rate of convergence for the DG scheme with upwinded flux, $\alpha = 0$, on Gauss Legendre solution points.

Chapter 5

Design of FR schemes with minimal dispersion and dissipation

The Flux Reconstruction formulation provides a remarkably broad envelope of discontinuous finite element schemes. The control parameters include (i) the location of solution points ξ, (ii) the upwinding coefficient α or more generally the kind of interface flux, and most importantly (iii) the correction functions $g_L(\xi)$, $g_R(\xi)$ which control the type of scheme recovered. Of these three choices, the first two have little flexibility. For instance, the location of solution points is largely dictated by the need to minimize aliasing errors that arise in the case of non-linear fluxes [76]. Consequently, the solution points are chosen as quadrature points for the unit weighting function, i.e. the Gauss-Legendre points on tensor-product elements. Similarly, the interface flux is usually fully upwinded, $\alpha = 0$, in order to add sufficient artificial dissipation for stabilizing solutions to non-linear conservation laws. The only remaining choice is that of the correction functions. However, here again, the right boundary correction function is implicitly set, $g_R(\xi) = g_L(-\xi)$, in order to preserve symmetry. Moreover, the choice is constrained by the requirement of linear stability of the resulting FR scheme. Nevertheless, several interesting schemes can still be recovered via the formulation [69]. For instance, selecting the left boundary correction as the right Radau polynomial, and vice versa, recovers the DG scheme at least for linear fluxes. Similarly, selecting the correction functions to have zeros at the flux collocation points recovers a version of the Spectral Difference (SD) scheme. An important question

is then to probe if the FR formulation houses other schemes that have superior properties relative to DG, SD etc.

The first investigation of this sort was conducted in [141]. Therein, the authors searched within the class of Energy Stable FR (ESFR) schemes [142] which are linearly stable when the ESFR parameter C is chosen to lie in the half-space $[C_-, \infty)$. For each $P \in \{2,3,4,5\}$, a grid-search was performed over this one-dimensional space to identify the ESFR scheme with the maximum stable time-step for the linear-advection problem. The resulting scheme was denoted as the C_+ scheme.

In this chapter, we are interested in finding FR schemes that provide maximum accuracy for time-dependent problems. Surely, there are several possible objective functions that can be associated with accuracy. The canonical approach of maximizing rate of convergence is ill-suited for a computational setup and will be dealt analytically in Part-II of the dissertation. Instead, our current objective is to minimize wave-propagation error for the entire range of resolvable wavenumbers. Towards this end, we solve an appropriate minimization problem in spectral space. Unlike spectrally optimal finite-difference schemes, the formal order of accuracy or short-time rate of convergence of these schemes is not altered by the optimization procedure. As a simple first case, optimal ESFR schemes are obtained by optimizing over the free parameter C. Thereafter, the general problem of optimizing over all possible choices of correction function is considered. As a result, linearly stable, optimal schemes are recovered that perform favorably compared to existing FR schemes.

5.1 Constrained minimization problem in spectral space

Traditionally, spectrally optimal finite-difference schemes [58, 88, 113] were obtained by relaxing the formal order of accuracy in order to gain free coefficients. These coefficients were then tuned to maximize the region over which the numerical dispersion relation follows the analytical one, within some fixed tolerance. Since most explicit as well as implicit compact schemes have a central (unbiased) stencil, there was no added numerical dissipation and only dispersive errors had to be minimized. Tam [129] suggested that a possible approach is to directly minimize the absolute error in the numerical wavespeed, i.e. given

a polynomial order P,

$$\text{minimize} \quad \int_0^{(P+1)\pi} |\lambda(kh) - 1|^2 \mathrm{d}(kh), \tag{5.1}$$

where λ is the normalized eigenvalue, Eqn. (2.18). However, such an objective function is not suitable for FR, or more generally, discontinuous finite element schemes for two reasons. First, FR schemes exhibit multiple eigenmodes each associated with its own dispersion and dissipation. Clearly, a systematic weighting procedure is needed to accumulate error across the $P+1$ modes. Second, as can be seen from Fig. 2.1 the numerical wavespeed for upwinded FR schemes has an imaginary part. In other words, all practical FR schemes introduce numerical dissipation in addition to dispersion. An objective function of this form would introduce an algebraic weighting of the dispersive and dissipative errors that may lead to a biased optimum. Fortunately, these problems can be easily circumvented by considering the error involved in wave propagation. Towards this end, use Eqn. (2.29) to write

$$\begin{aligned} \|\boldsymbol{e}_n(t)\|_2 &= \| \sum_{p=1}^{P+1} (\exp(-ikc(\lambda_p - 1)t) - 1)\, \beta_p \boldsymbol{w_p} \|_2 \\ &\leq \sum_{p=1}^{P+1} \left| \exp\left(-ikh(\lambda_p - 1)\frac{ct}{h}\right) - 1 \right| |\beta_p| \end{aligned}$$

The non-dimensional factor $\frac{ct}{h}$ is the distance traveled by the wave in time t in terms of the element size h. We now resort to two largely heuristic choices to complete the specification of our optimization problem. First, we are interested in schemes for long-time integration that would provide low errors for propagation across the entire domain. With this intention, we have fixed $t = 100\frac{h}{c}$, where $\frac{h}{c}$ corresponds to the time taken by the wave to travel one element width. Second, instead of working directly with the modal weights, $\boldsymbol{\beta}$, we choose to work with fractional energies,

$$\hat{\beta}_p := \frac{|\beta_p|^2}{\sum_{p=1}^{P+1} |\beta_p|^2}, \tag{5.2}$$

discussed in Section 2.3. The objective function is then given by

$$\eta(g_L;P,\boldsymbol{\xi}) := \frac{1}{(P+1)^2} \int_0^{(P+1)\pi} \sum_{p=1}^{P+1} \left|\exp(-i\,100\,kh\,(\lambda_p(kh)-1))-1\right| \hat{\beta}_p(kh)\mathrm{d}(kh). \quad (5.3)$$

where the factor of $1/(P+1)^2$ is used to normalize the error across polynomial order. The optimization problem with the stability criterion can be stated as follows. Given the polynomial order P, solution points $\boldsymbol{\xi}$,

$$\begin{aligned} &\text{minimize} \quad \eta(g_L;P,\boldsymbol{\xi}) \\ &\text{subject to} \quad \lambda_p^{\mathrm{Im}}(kh) \leq 0 \text{ for } p=1,2,\ldots,P+1,\ kh \in [0,(P+1)\pi], \end{aligned} \quad (5.4)$$

where $g_L(\xi)$ is the left-boundary correction function in the parent domain.

5.2 Optimal Energy Stable FR schemes (OESFR)

We begin by optimizing within the class of ESFR schemes [142] that are linearly stable by construction. For this class, the left boundary correction function is given by

$$g_L(\xi) = \frac{(-1)^P}{2}\left[L_P(\xi) - \left(\frac{\mu_P L_{P-1}(\xi)+L_{P+1}(\xi)}{1+\mu_P}\right)\right],$$

where L_P is the Legendre polynomial of degree P, and

$$\mu_P = C\,\frac{2P+1}{2}\left(\frac{(2P)!}{2^P(P)!}\right)^2, \quad (5.5)$$

where $C \in [C_-,\infty)$ is the free parameter of the scheme. The lower bound, C_-, is the smallest value of C that yields a legitimate energy norm of the form

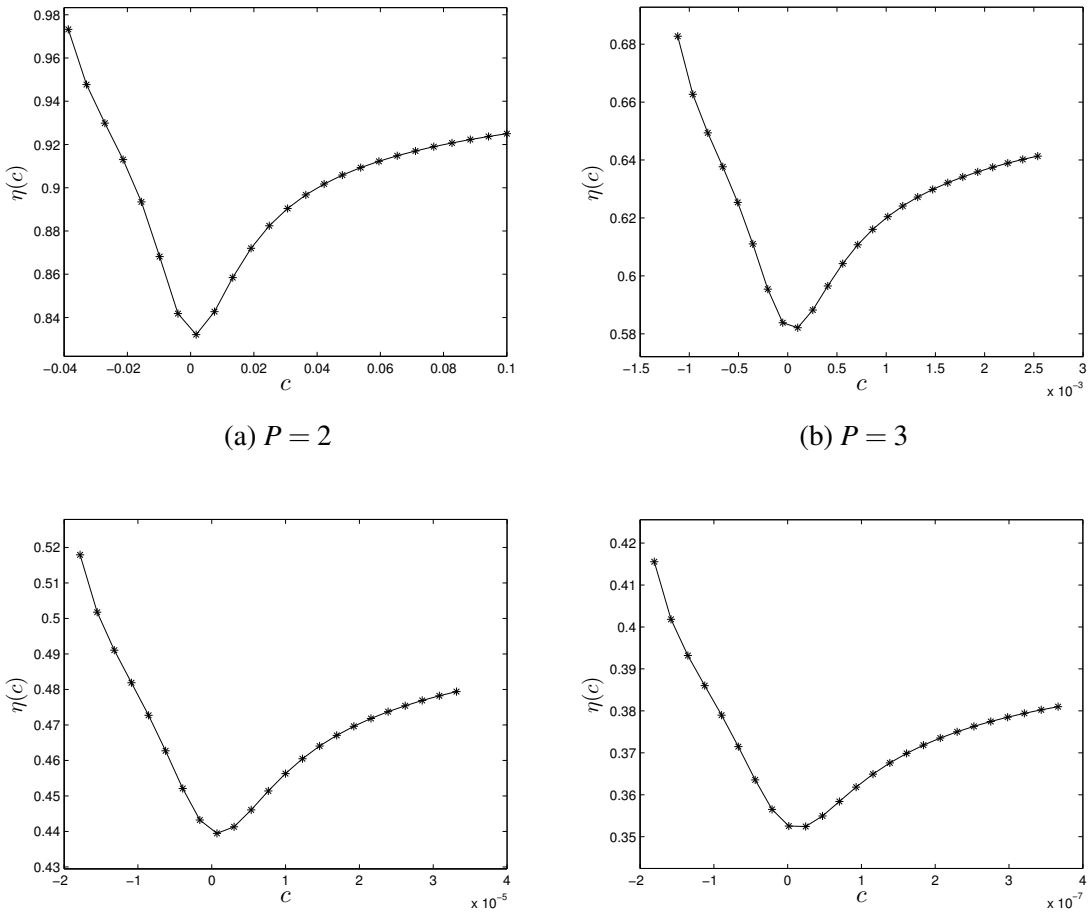
$$\|u^\delta\|_{\bar{\Omega},(0,P)} = \left[\sum_{n=1}^{N}\int_{x_n}^{x_{n+1}}\left((u^\delta(x))^2 + \frac{c}{2}h^{2P}\left(\frac{\partial^P u^\delta}{\partial x^P}(x)\right)^2\right)\mathrm{d}x\right]^{1/2}. \quad (5.6)$$

P	$\frac{\eta_{OESFR}}{\eta_{DG}}$	$\frac{\eta_{OFR}}{\eta_{DG}}$	C_{OESFR}	$\{\zeta_{OESFR}\}$	$\{\zeta_{OFR}\}$
1	0.9997	0.9997	8.40 $\times 10^{-3}$	{-0.324947954}	{-0.324947954}
2	0.9169	0.9168	5.83 $\times 10^{-4}$	{-0.687046768, 0.294921433}	{-0.683006984, 0.302192636}
3	0.9279	0.9183	3.17 $\times 10^{-5}$	{-0.820508633, -0.174051445, 0.580238407}	{-0.839877076, -0.202221672, 0.518569180}
4	0.8806	0.7658	9.68 $\times 10^{-7}$	{-0.883539936, -0.437806887, 0.177747440, 0.725821574}	{-0.856985048, -0.447652425, 0.180019034, 0.638102912}
5	0.8497	0.7216	1.02 $\times 10^{-8}$	{-0.919029563, -0.598291572, -0.115071988, 0.398821159, 0.806299103}	{-0.897887439, -0.577293821, -0.101190260, 0.354120544, 0.760380824 }

Table 5.1: Objective function (relative to DG) and internal zeros for the OESFR and OFR schemes.

The DG scheme is recovered from ESFR by choosing $C = 0$. The choice of $C = C_+$ yields the scheme that offers the highest CFL limit for a given polynomial order.

Vincent et al. [141] have noted that as C is increased beyond 0, a competing effect can be observed regarding the point of departure from the exact dispersion relation. While the departure of the real component of wavespeed gets pushed to higher k, that for the imaginary component gets pulled to lower k. This trade-off suggests the presence of an optimum close to $C = 0$, the DG scheme. Figures 5.1a-d plot the variation of the objective function, Eqn. (5.3), against C for $P = 2, 3, 4$ and 5. We see that the objective is very well behaved and records a discernible minimum close to $C = 0$ as expected. This optimal value of C along with the value of the objective function is tabulated in Table 5.1. The real and imaginary components of the wavespeed for the OESFR scheme are plotted in Figs. 5.2a-b for $P = 4$ and 5.3a-b for $P = 5$ respectively. A detailed discussion of these figures has been deferred to the next section.



(a) $P = 2$ (b) $P = 3$

(c) $P = 4$ (d) $P = 5$

Figure 5.1: OESFR: Variation of the objective function with the ESFR parameter C denoting the presence of a minimum close to the DG value of $C = 0$.

5.3 Optimal FR schemes (OFR)

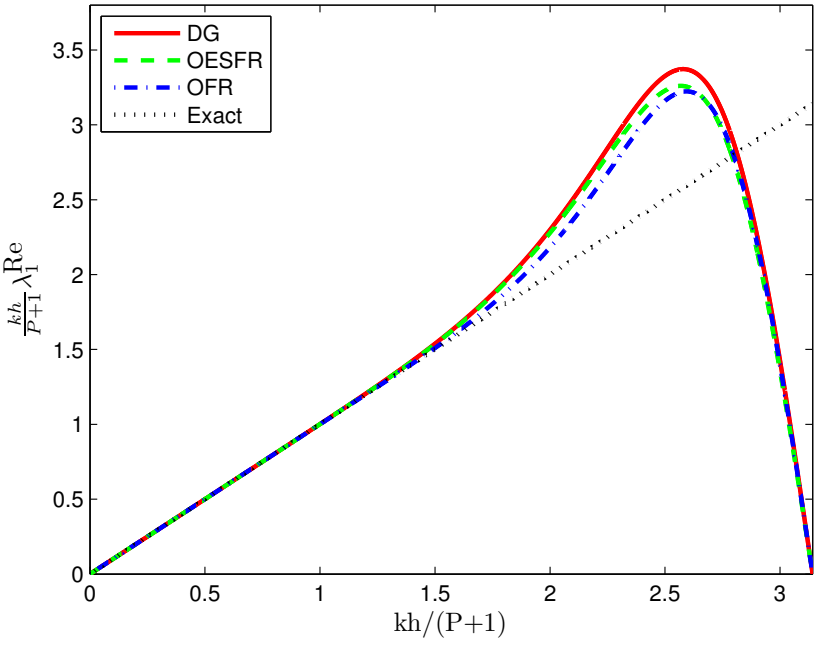
A general form of the left-boundary correction function can be expressed using the internal zeros of the correction polynomial,

$$g_L(\xi) = \prod_{q=1}^{P} \frac{(\xi - \zeta_q)}{1+\zeta_q} \frac{(\xi - 1)}{2}, \tag{5.7}$$

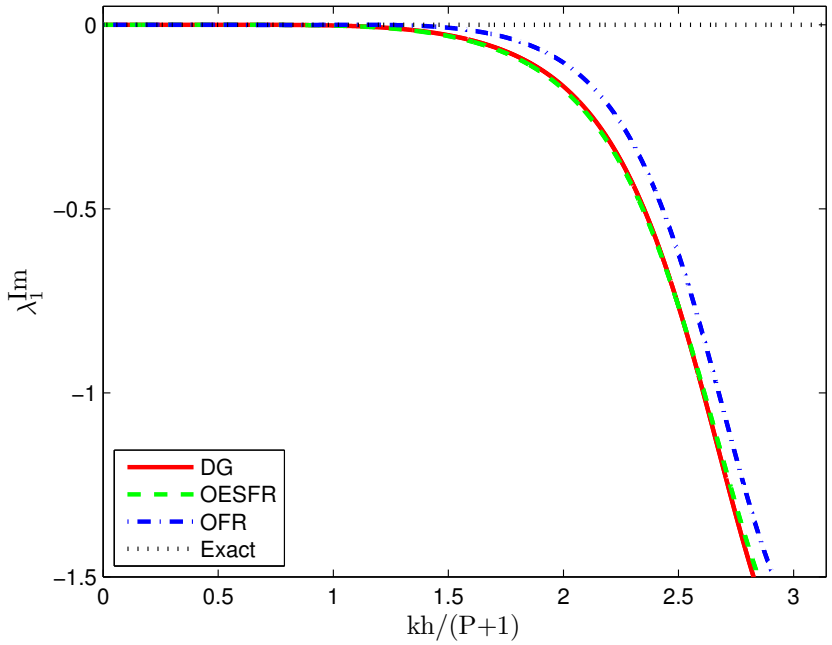
which ensures that the function is unity on the left boundary, vanishes on the right boundary, and also on the set of internal zeros $\boldsymbol{\zeta} = \{\zeta_1, \zeta_2, \ldots, \zeta_P\}$. The solution space is of dimension P and recovers the ESFR family as a curve originating at the set of zeros corresponding to the ESFR scheme with $C = C_-$. The most general optimal FR schemes (OFR) can be obtained by solving Eqn. (5.4) with g_L parameterized by $\boldsymbol{\zeta}$. We employ an interior-point method [20] with the help of MATLAB's optimization toolbox [41] to find local optima. Unfortunately, the constraints are satisfied only within subsets of much lower dimension. This makes a grid search seeding approach [150] infeasible as random initial guesses almost always violate the constraints. As an easy fix, zeros for the initial guess are taken to be the zeros of the corresponding OESFR schemes.

Table 5.1 enlists the internal zeros of the OFR correction function along with the minimum of the objective function. For $P = 1$, the OFR optimization process converges to the zeros of the OESFR correction function, which is expected owing to a single degree of freedom. For $P \geq 2$, the optimization process converges to zeros that are not traced by the ESFR family. Since the constraint is strictly satisfied, we obtain linearly stable FR schemes that have locally minimum wave propagation error. The real and imaginary components of the wavespeed for the OFR scheme are plotted in Figs. 5.2a-b for $P = 4$ and Figs. 5.3a-b for $P = 5$. The contrast between OFR and OESFR or DG grows with polynomial order due to the increase in degrees of freedom.

Figures 5.4a-b plot the left-boundary correction functions for the DG, OESFR and OFR schemes for $P = 4$, 5 respectively. We note that the zeros of the OFR scheme are located in a manner that tends to minimize the deviation of the function from zero. This seems to support Huynh's [69] expectation that a correction function must try to approximate zero

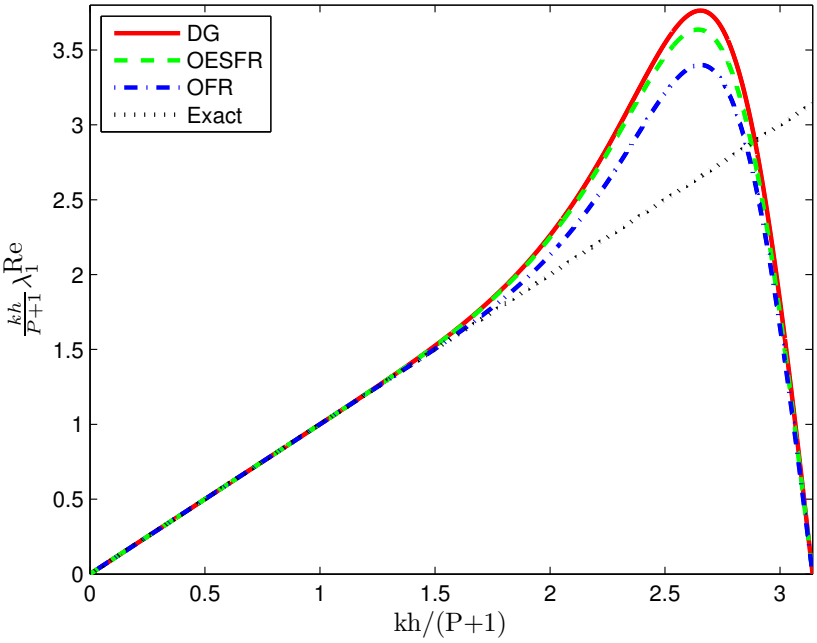


(a) Effective wavenumber, k_p^{eff}

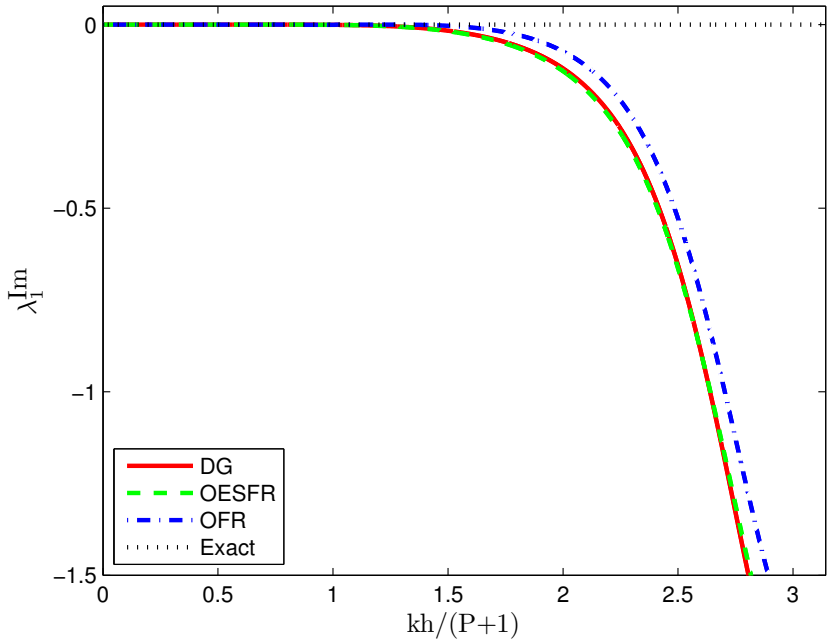


(b) Imaginary part of normalized eigenvalue

Figure 5.2: Optimal schemes: Dispersion and dissipation properties of the physical mode for DG, OESFR and OFR schemes for $P = 4$.



(a) Effective wavenumber, k_p^{eff}



(b) Imaginary part of normalized eigenvalue

Figure 5.3: Optimal schemes: Dispersion and dissipation properties of the physical mode for DG, OESFR and OFR schemes for $P = 5$.

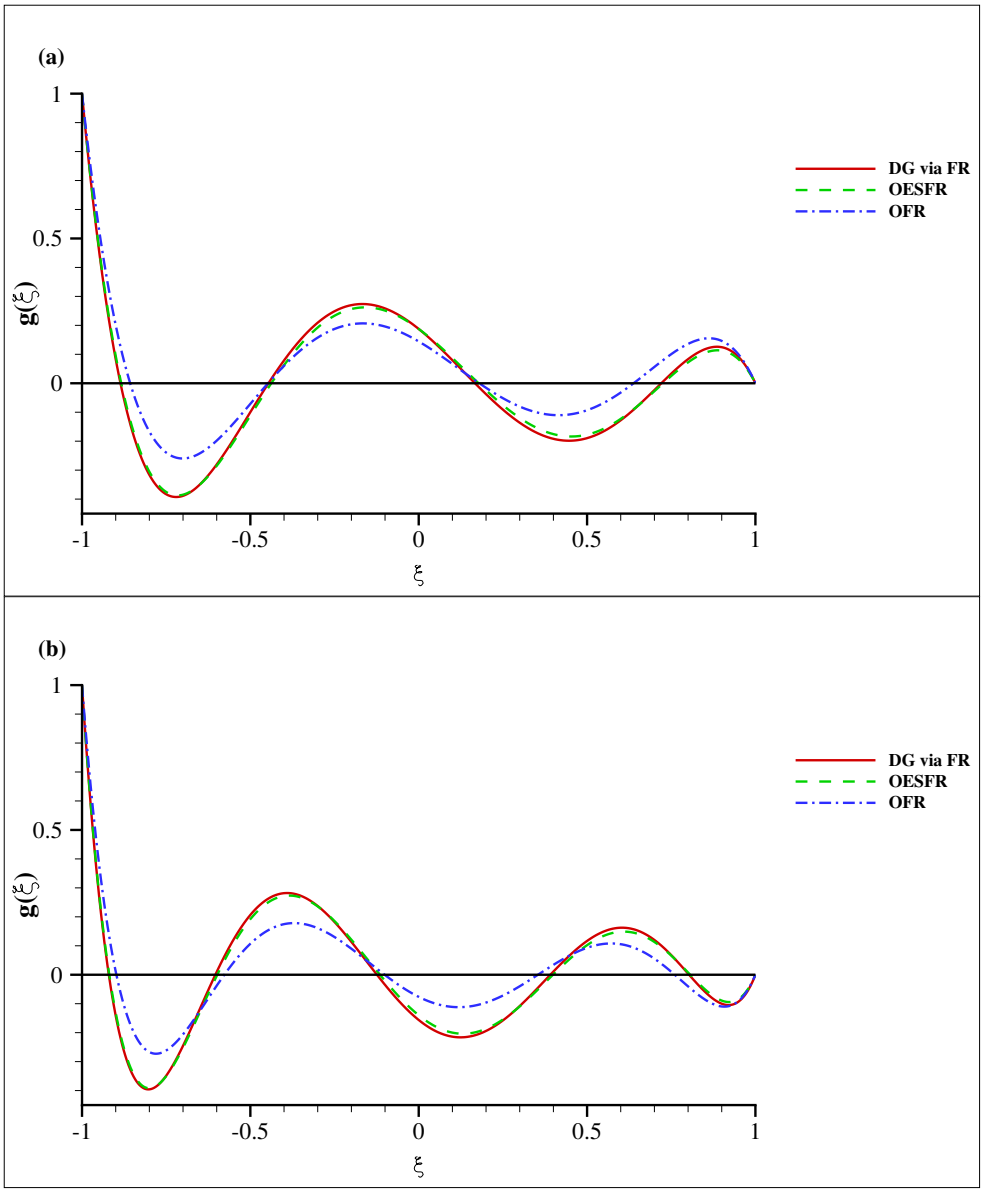


Figure 5.4: Left boundary correction functions for DG, OESFR and OFR schemes for (a) $P = 4$ and (b) $P = 5$.

P	$\varepsilon = 0.01$				$\varepsilon = 0.001$			
	DG	OESFR	OFR	C_+	DG	OESFR	OFR	C_+
1	0.145	0.145	0.145	-	0.066	0.066	0.066	-
2	0.263	0.263	0.263	0.128	0.160	0.160	0.160	0.070
3	0.339	0.339	0.352	0.228	0.233	0.233	0.249	0.149
4	0.391	0.391	0.477	0.300	0.287	0.287	0.409	0.214
5	0.428	0.428	0.511	0.351	0.328	0.328	0.444	0.264

Table 5.2: Resolving efficiency e_1 for $\varepsilon = 0.01, 0.001$ for DG, OESFR, OFR and the C_+ scheme.

in the interior.

The fraction of the spectrum for which the numerical dispersion relation closely approximates the analytical one can be expressed through the 'resolving efficiency' of the physical mode. Following the definition by Lele [88], define

$$e_1 = \frac{(kh)_f}{(P+1)\pi}, \tag{5.8}$$

where $(kh)_f$ is such that

$$|\lambda(kh) - 1| \leq \varepsilon, \text{ for } kh \leq (kh)_f. \tag{5.9}$$

Table 5.2 compares the resolving efficiencies of DG, OESFR, OFR and the C_+ scheme for $P = 1$ to 5. For $P \leq 2$, since the OFR scheme converges to the OESFR scheme, which in turn is only marginally better than DG, the resolving efficiencies for all three schemes are almost identical. However for $P > 2$, the OFR scheme provides higher resolution for both measures of resolving efficiency. An improvement of about 20% is achieved through the use of OFR over DG for $P = 4, 5$. On the other hand, the C_+ scheme suffers from a poor resolving power due to excessive dissipation. In fact, from Table 5.2 we see that the OFR scheme of order 3 provides similar resolution to the C_+ scheme of order 5.

It is interesting to compare the resolving efficiency of OFR schemes with standard Padé

type compact finite difference schemes. Since the time complexity of the 1-D FR formulation is linear in the number of elements, we restrict our attention to compact difference schemes that are at maximum tridiagonal in structure. Table III in [88] lists the resolving efficiencies of 2^{nd} order explicit, 4^{th} order explicit and implicit, 6^{th} order explicit and implicit and 8^{th} order implicit difference schemes, all of which have linear time complexity in evaluating the derivative. On comparing, we see that the OFR scheme provides significantly higher resolution than the explicit difference schemes of the same order. In comparison to the implicit difference schemes, the OFR scheme provides similar resolution for $\varepsilon = 0.01$ but a higher resolution for $\varepsilon = 0.001$. In fact, for $\varepsilon = 0.001$, the OFR scheme of order $P = 5$ provides similar resolution to the 8^{th} order implicit compact difference scheme.

5.4 Time integration and CFL restriction

In this section, we determine CFL restrictions for the proposed schemes and compare them with existing FR schemes. In the process, the numerical time-integration operator $\frac{d^{\delta}}{dt}$ is expanded for the case of a general M-stage explicit Runge-Kutta type scheme which yields the fully-discrete numerical dispersion relation for Flux Reconstruction.

5.4.1 Fully-discrete dispersion relation for a multi-stage time scheme

The semi-discrete numerical conservation law, Eqn. (2.16), assumed exact integration in time. However, during an actual numerical update, the time-derivative operator $\frac{d}{dt}$ is numerically approximated, $\frac{d^{\delta}}{dt}$, which introduces time-stepping errors. The fully-discrete conservation law can be analogously expressed as

$$\begin{aligned} \frac{d^{\delta}}{dt}\boldsymbol{u}_{\boldsymbol{n}}^{\boldsymbol{\delta}} + c\frac{2}{h}\boldsymbol{Q}(kh)\boldsymbol{u}_{\boldsymbol{n}}^{\boldsymbol{\delta}} = 0, \quad \forall n,\, t > 0, \\ \boldsymbol{u}_{\boldsymbol{n}}^{\boldsymbol{\delta}}(0) = \exp(ikx_n)\boldsymbol{w}_0, \end{aligned} \tag{5.10}$$

where the time-integration operator $\frac{d^\delta}{dt}$ can be expanded for a general M-stage explicit Runge-Kutta type scheme so that

$$\boldsymbol{u}_{\boldsymbol{n}}^{\boldsymbol{\delta}}(t+\Delta t) = \boldsymbol{R}(k,h,\Delta t)\boldsymbol{u}_{\boldsymbol{n}}^{\boldsymbol{\delta}}(t), \tag{5.11}$$

$$\boldsymbol{R}(k,h,\Delta t) = \boldsymbol{I} + \sum_{m=1}^{M} \frac{A_m}{m!}\left(-c\frac{2}{h}\mathbf{Q}(kh)\Delta t\right)^m, \tag{5.12}$$

where Δt is the numerical time step, $\boldsymbol{A}$ is the vector of coefficients of the RK scheme, and M is the number of stages. Choosing $A_m = 1$, for $m = 1,2,\ldots,M$, results in the standard M-stage, Mth order accurate RK scheme. This can be easily verified by realizing that $\frac{d^\delta}{dt}$ is a truncated power expansion of the exact operator $\frac{d}{dt}$. Numerical time integration modifies the associated eigensystem to yield the fully-discrete numerical dispersion relation,

$$\begin{aligned}
\boldsymbol{R}(k,h,\Delta t) &= \boldsymbol{I} + \sum_{m=1}^{M} \frac{A_m}{m!}\left(-ikc\boldsymbol{W}(kh)\boldsymbol{\Lambda}(kh)\boldsymbol{W}^{-1}(kh)\Delta t\right)^m \\
&= \boldsymbol{W}(kh)\boldsymbol{W}^{-1}(kh) + \boldsymbol{W}(kh)\left[\sum_{m=1}^{M} \frac{A_m}{m!}\left(-ikc\Delta t\boldsymbol{\Lambda}(kh)\right)^m\right]\boldsymbol{W}^{-1}(kh) \\
&= \boldsymbol{W}(kh)\left[\boldsymbol{I} + \sum_{m=1}^{M} \frac{A_m}{m!}\left(-ikc\Delta t\boldsymbol{\Lambda}(kh)\right)^m\right]\boldsymbol{W}^{-1}(kh).
\end{aligned} \tag{5.13}$$

The eigenvectors of $\boldsymbol{R}$ are identical to those of $\boldsymbol{Q}$, while the eigenvalues depend on the eigenvalues of $\boldsymbol{Q}$, the CFL number $c\Delta t/h$ and the coefficients of the RK scheme, i.e.

$$\boldsymbol{\Gamma}_{\boldsymbol{R}}(kh,\mathrm{CFL}) = \boldsymbol{I} + \sum_{m=1}^{M} \frac{A_m}{m!}\left(-ikh\boldsymbol{\Lambda}(kh)\ \mathrm{CFL}\ \right)^m, \tag{5.14}$$

$$\mathrm{CFL} = \frac{c\Delta t}{h}. \tag{5.15}$$

In the limit of $\Delta t \to 0$, the fully-discrete dispersion relation recovers the semi-discrete relation for any consistent time-stepping scheme.

5.4.2 CFL restriction

Eqn. (5.11) represents a discrete linear dynamical system whose stability criterion can be expressed in terms of the spectral radius ρ,

$$\rho(\mathbf{R}) \leq 1 \quad \forall\, kh \in [0,(P+1)\pi] \tag{5.16}$$

or equivalently,

$$|\text{diag}(\mathbf{\Gamma_R})| \preceq \mathbf{1} \quad \forall\, kh \in [0,(P+1)\pi] \tag{5.17}$$

where $\preceq$ expresses component-wise inequality. Table 5.3 provides limiting CFL values for the DG, OESFR, OFR and C_+ schemes for the standard 4 stage, 4^{th} order RK scheme RK44 and the enhanced stability 5 stage, 4^{th} order scheme RK45 [26]. The corresponding values for 3^{rd} order schemes including the 3 stage standard RK scheme RK33 and the optimal four stage, low storage, strong-stability-preserving scheme SSPRK(4,3) of Kubatko et al. [84] have been provided in Table 5.4. We begin by noting that the CFL limit for the OESFR scheme is higher than DG for all polynomial orders. This is expected as the CFL limit increases with C from $C = 0$ to $C = C_+$ as shown in [141] and that the optimal c value for OESFR was found to be greater than 0. We also note that the CFL limit for the OFR scheme compares well to that for the DG scheme, thereby suggesting that the additional accuracy obtained by tuning the zeros of the correction function does not come at the expense of computational effort! While the C_+ scheme has a much higher CFL limit, it suffers from significantly larger dispersion errors that makes it ill-suited for wave propagation problems. The next section provides numerical examples outlining this aspect.

5.5 Numerical tests

This section presents the results of numerical tests to validate the analytical expectations derived in the previous sections. Since a salient feature of the FR formulation is its applicability to complex geometries with non-uniform grids, it is important to show that the advantage of using the proposed schemes, optimized for a uniform grid, is retained even on arbitrarily stretched non-uniform grids. Table 5.5 lists the various cases that have been

P	RK44				RK45			
	DG	OESFR	OFR	$C+$	DG	OESFR	OFR	$C+$
1	0.464	0.470	0.470	-	0.679	0.686	0.686	-
2	0.235	0.238	0.241	0.688	0.352	0.356	0.361	0.864
3	0.139	0.148	0.126	0.376	0.220	0.224	0.191	0.473
4	0.100	0.103	0.108	0.245	0.152	0.158	0.164	0.311
5	0.068	0.076	0.085	0.174	0.110	0.117	0.128	0.223

Table 5.3: Limiting CFL values for the DG, OESFR, OFR and the C_+ scheme for RK44 and RK45.

P	RK33				optimal SSPRK(4,3)		
	DG	OESFR	OFR	$c+$	DG	OESFR	OFR
1	0.411	0.415	0.415	-	0.594	0.601	0.601
2	0.210	0.212	0.210	0.623	0.308	0.311	0.316
3	0.130	0.133	0.109	0.334	0.191	0.196	0.166
4	0.080	0.091	0.095	0.212	0.133	0.138	0.144
5	0.061	0.068	0.074	0.149	0.098	0.101	0.112

Table 5.4: Limiting CFL values for DG, OESFR, OFR and the C_+ scheme for RK33 and SSPRK(4,3). The C_+ scheme has only been derived only for RK3, RK44 and RK45.

Case	Schemes	P	Soln. pts.	Grid	Initial condition
1	DG, OESFR, OFR, c_+	5	Gauss-Legendre	Uniform	$\mathrm{e}^{-38.6x^2}$
2	DG, OESFR, OFR, c_+	5	Gauss-Legendre	Uniform	$\mathrm{e}^{-0.1x^2}\cos(3\pi x)$
3	DG, OESFR, OFR, c_+	5	Gauss-Legendre	Non-uniform	$\mathrm{e}^{-38.6x^2}$

Table 5.5: Details of numerical tests conducted in Chapter 5.

investigated in this study. The wavespeed is taken to be unity, $c = 1$.

5.5.1 Case 1: Advection of a sharp Gaussian on a uniform grid

The first case deals with linear advection of a Gaussian on a uniform grid. The domain $\Omega = [-10, 10]$ and the number of elements is taken to be 20 which conveniently normalizes the grid spacing to unity. The polynomial order P is 5 and the solution points within each element are chosen to be the Gauss-Legendre points in the parent domain. The initial condition $u_0(x) = \mathrm{e}^{-38.6x^2}$ is a Gaussian with zero mean, unit peak and variance $\sigma^2 = 1/(77.2)$ which is chosen so that $\hat{u_0}(kh = (P+1)\pi)/\hat{u_0}(kh = 0) = 0.1$, where $(\hat{u_0})$ denotes the Fourier transform of $u_0(x)$. This ensures that the initial condition spans the entire range of resolvable wavenumbers with 10% amplitude at the Nyquist limit. This problem is solved using the DG scheme as reference, the proposed OESFR and OFR schemes, and the C_+ scheme. For all schemes, the CFL number is chosen to be half of the maximum allowable CFL number for the DG scheme. The boundary conditions are taken to be periodic and the solution is made to advect for a full period $T = |\Omega|/c = 20$.

Figure 5.5 provides snapshots of the numerical solution for the employed schemes at intervals of a quarter period $T/4$. The exact solution is included for reference. As expected from optimality in spectral space, the OFR scheme provides the least dissipated solution, followed by the OESFR scheme which almost overlaps with the DG solution. This suggests that optimization within the ESFR family provides little advantage over the DG scheme. The C_+ scheme which had the highest stability limit is the most dissipative in this set of schemes.

Figure 5.6 compares the evolution of the L^2 norm of error for the employed schemes.

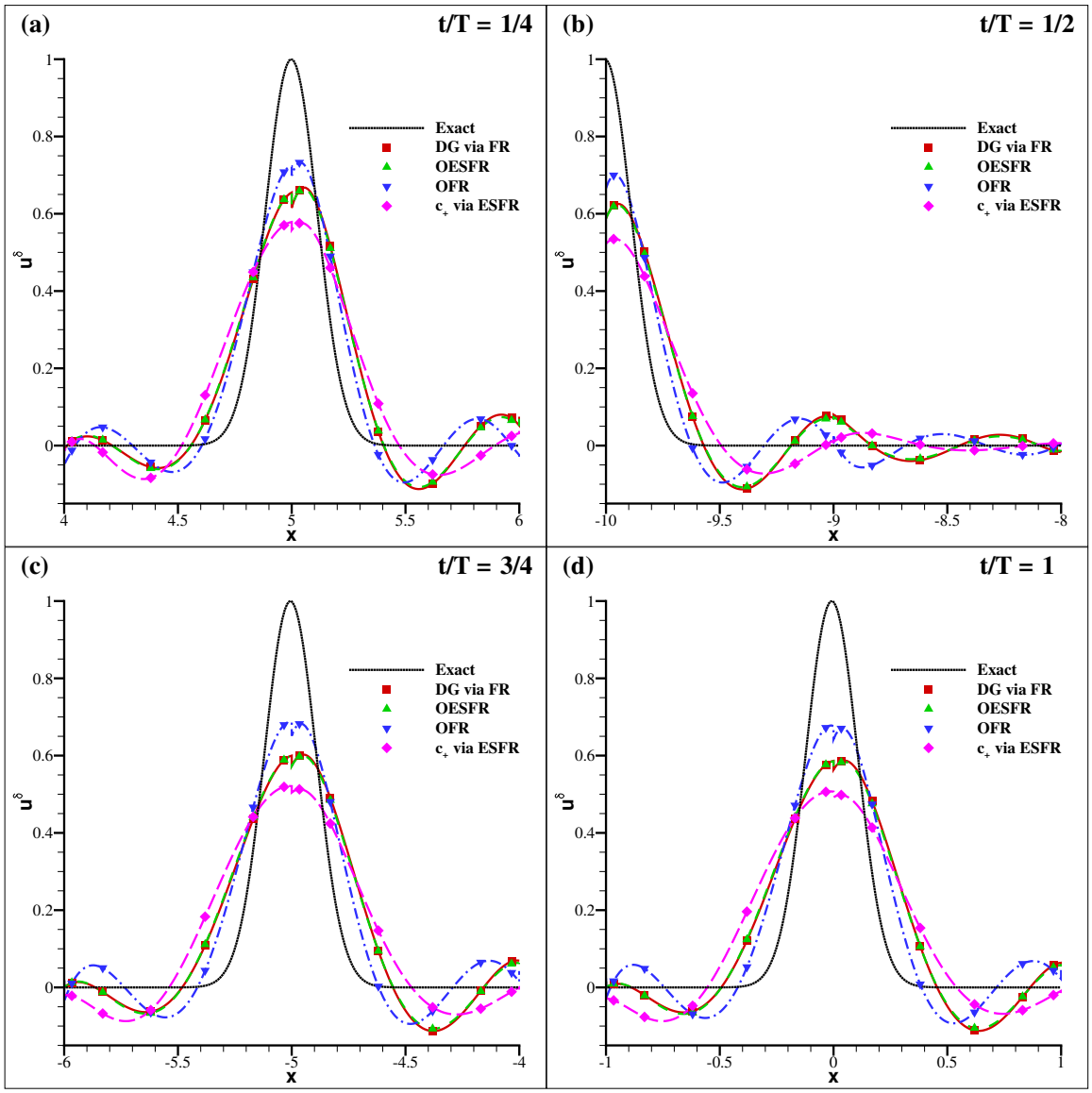


Figure 5.5: Case 1: Snapshots of the numerical solution for the DG, OESFR, OFR and C_+ schemes at intervals of a quarter-period.

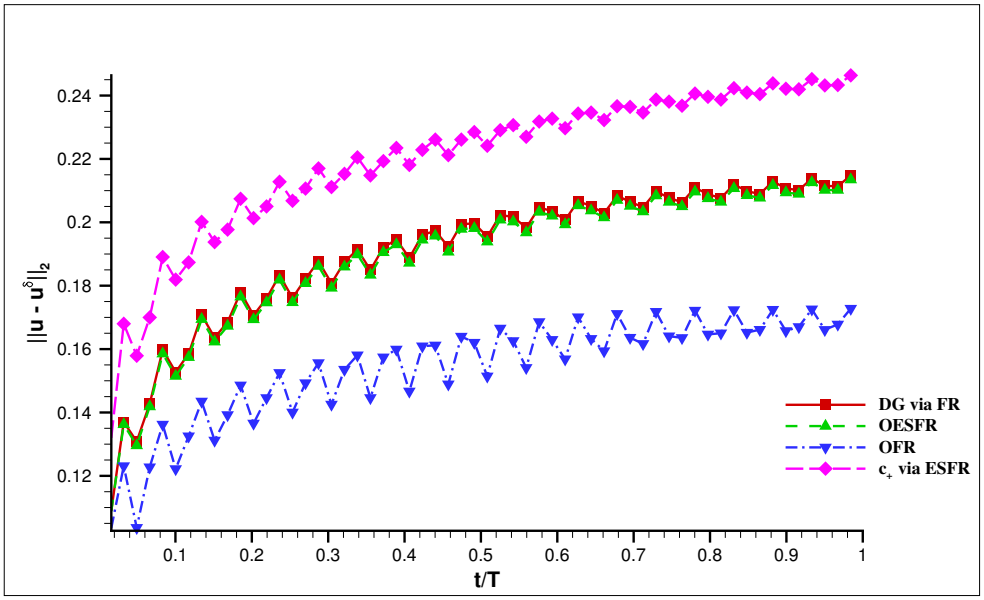


Figure 5.6: Case 1: Evolution of L^2 norm of the error for the DG, OESFR, OFR and C_+ schemes.

As expected, the error for all of the schemes rises with time due to the dispersion and dissipation errors committed at each update. The small scale periodicity in the error is a consequence of the piecewise nature of the FR solution. The errors are lower when the Gaussian peak is located at the interfaces between elements and higher when the peak is centered within any element. The uniformity in the periods is on account of the uniformity in the grid. As expected, we see that the OFR scheme provides a considerably lower error than the DG scheme while the C_+ scheme provides larger ones at all times during the period.

The resolving capabilities of these schemes can alternatively be compared by the grid resolution that each of them requires to accurately advect this steep Gaussian for a full period. Such a measure of resolution is directly related to the resolving efficiency e_1 recorded in Table 5.2. From the table, we see that for the DG scheme, for $P = 5$, to accurately resolve the wavelength at the Nyquist limit for the grid in Case 1, we require $20/e_1|^{DG}_{\varepsilon=0.001,P=5} \simeq 61$ elements. Similarly, for the OFR scheme we require 45 elements, and for the C_+ scheme, we require 76 elements. Note that there would still be some numerical error as the spectrum of the initial condition extends beyond the original Nyquist limit by construction. Figure 5.7 provides snapshots of the numerical solution for the employed schemes at intervals of a quarter period $T/4$ where the number of elements for each scheme is chosen as described above. As can be seen, each of the schemes resolve the Gaussian reasonably well through their respective grids. In this case, the advantage received by using the spectrally optimal OFR scheme is that it provides accurate solutions at coarser grids.

5.5.2 Case 2: Advection of a high wavenumber packet on a uniform grid

The second case is chosen to highlight the advantage of using the proposed scheme in advecting moderately high wavenumber components. For this reason, the initial condition $u_0 = \mathrm{e}^{-0.1x^2}\cos(3\pi x)$ is a packet centered in the spectral space about $kh = (P+1)\pi/2$, half of the Nyquist limit. The remaining parameters of the case are identical to the previous one. Figure 5.8 provides snapshots of the numerical solution for the four schemes. We see that even in one quarter of the period, the excessive dissipation in the C_+ scheme has completely

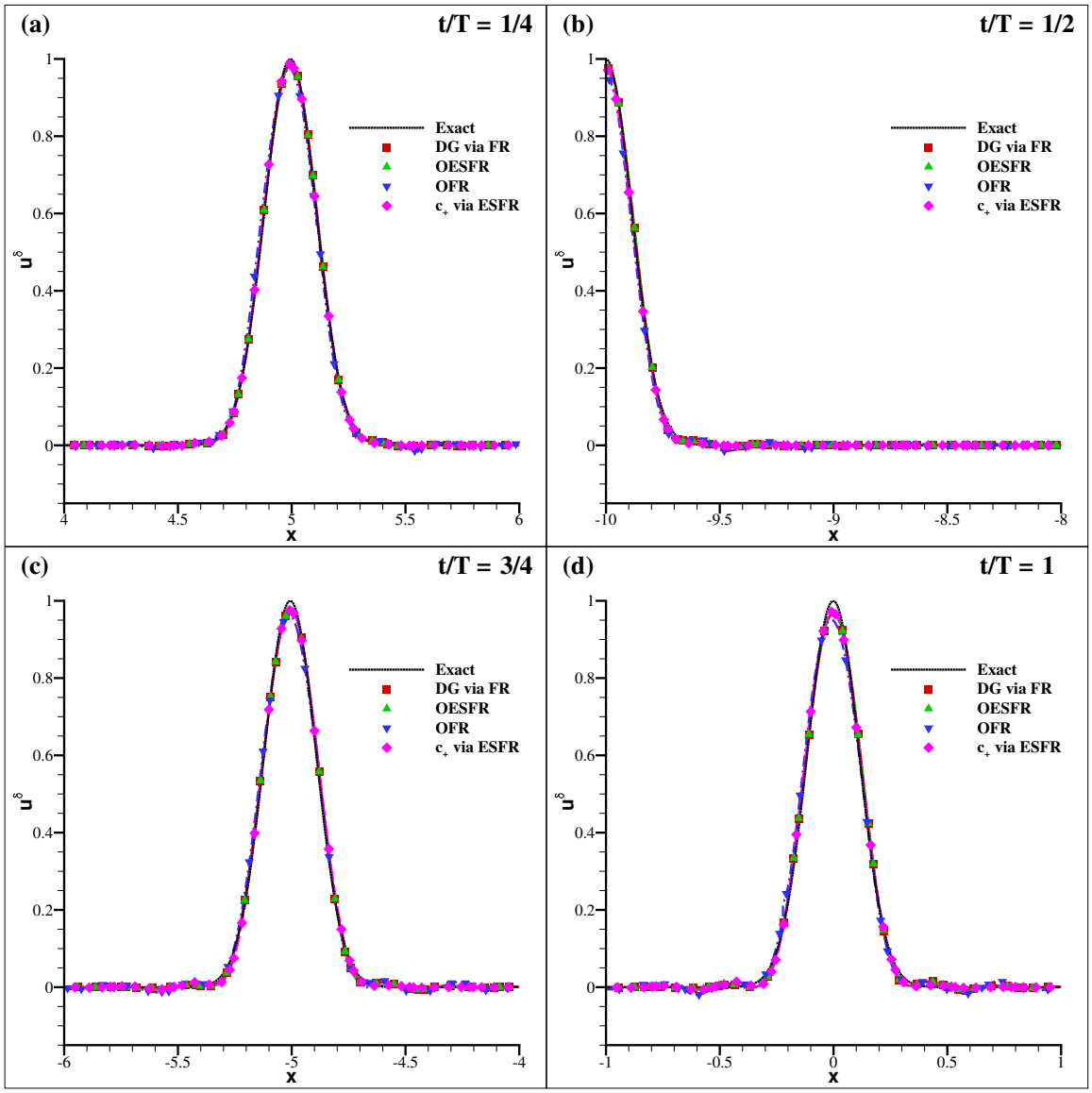


Figure 5.7: Case 1 with refined grids: Snapshots of the numerical solution for the DG, OESFR schemes using 61 elements, OFR scheme using 45 elements and C_+ scheme using 76 elements at intervals of a quarter-period.

dissipated the solution, while the DG and OESFR schemes also cause substantially larger dissipation as well as dispersion errors than the OFR scheme. At the end of one period, the DG and OESFR solutions have decayed completely. Figure 5.9 compares the evolution of error demonstrating the flattening of the norm once the solution has decayed. Note that the apparently small improvement achieved via optimization, in fact, results in a significant reduction in physical error.

5.5.3 Case 3: Advection of a sharp Gaussian on a non-uniform grid

Having confirmed the advantage of optimality on uniform grids, we now deal with the case of a randomly stretched non-uniform grid. The grid for this case is constructed by randomly perturbing the uniform grid in Case 1, $x'_j = x_j + 0.99h\theta$, where x_j is the left interface location of the j^{th} element and θ is uniformly distributed in $[0,1]$. A bar chart of the resulting element widths is given in Fig. 5.10a. All other parameters, including the initial condition are identical to the first case. Figure 5.11 provides similar plots as for the uniform grid case. The results are qualitatively similar but the difference between the schemes is lesser than before. The evolution of the error norm in Fig. 5.10b is directly affected by the non-uniformity in the grid. The solution suffers large errors every time the Gaussian has to pass through an element of relatively large width. While the OFR scheme still provides lower errors than the other FR schemes, the improvement is limited by the presence of highly stretched elements. Snapshots of the numerical solution are provided in Fig. 5.11.

5.5.4 Taylor-Green vortex

The optimal schemes, OESFR and OFR, derived in this chapter have been designed to minimize wave-propagation error for linear advection. This simple case of constant wavespeed allows for a closed-form objective function, Eqn. (5.3), which directly depends on the numerical dispersion relation and the location of solution points. As in the case of optimized finite-difference stencils, the underlying expectation is that FR schemes with superior wave-propagation properties for linear advection would provide reduced error even in the general case of non-linear wavespeeds such as the Navier-Stokes (NS) equations. In

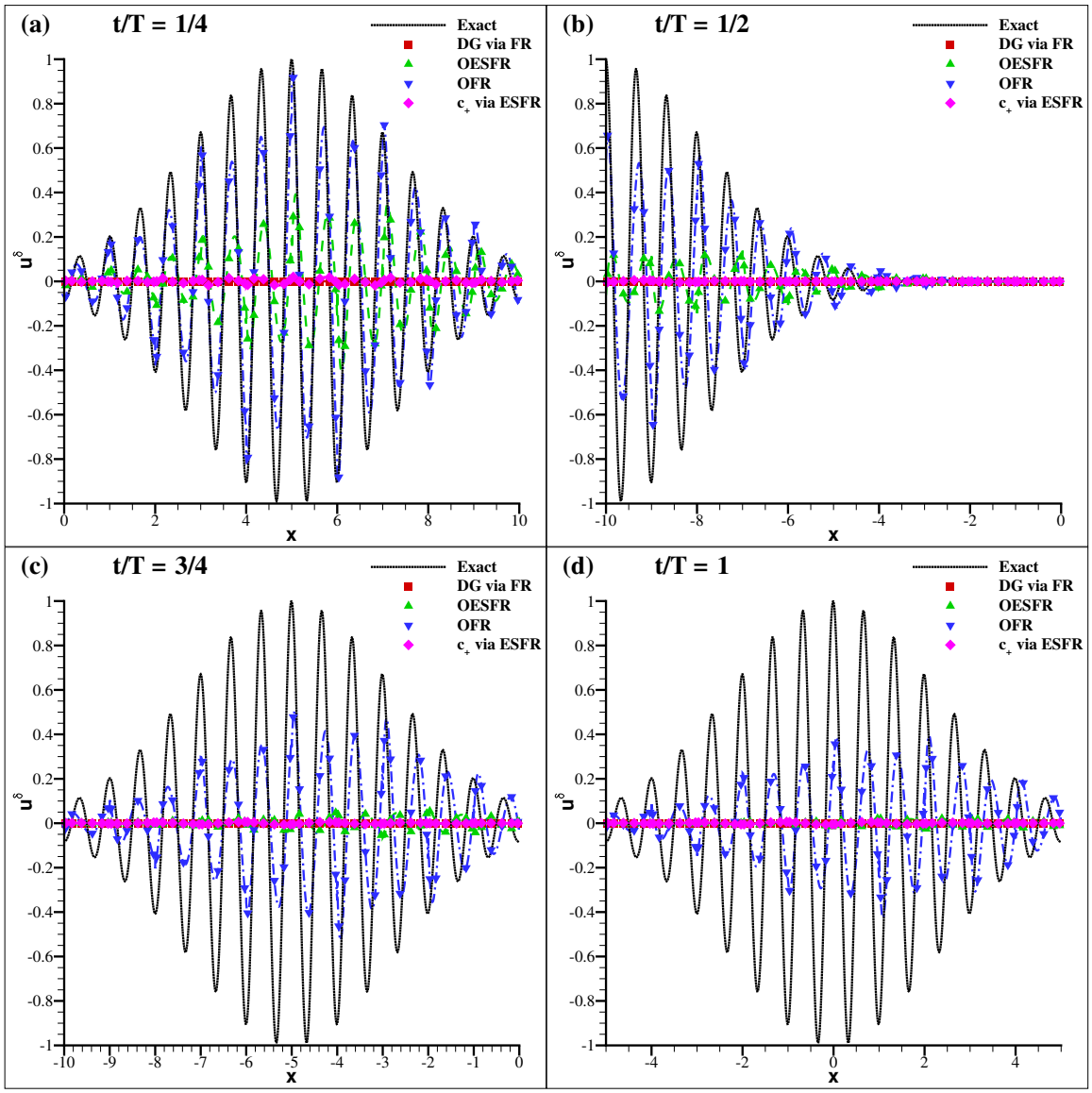


Figure 5.8: Case 2: Snapshots of the numerical solution for the DG, OESFR, OFR and C_+ schemes at intervals of a quarter-period.

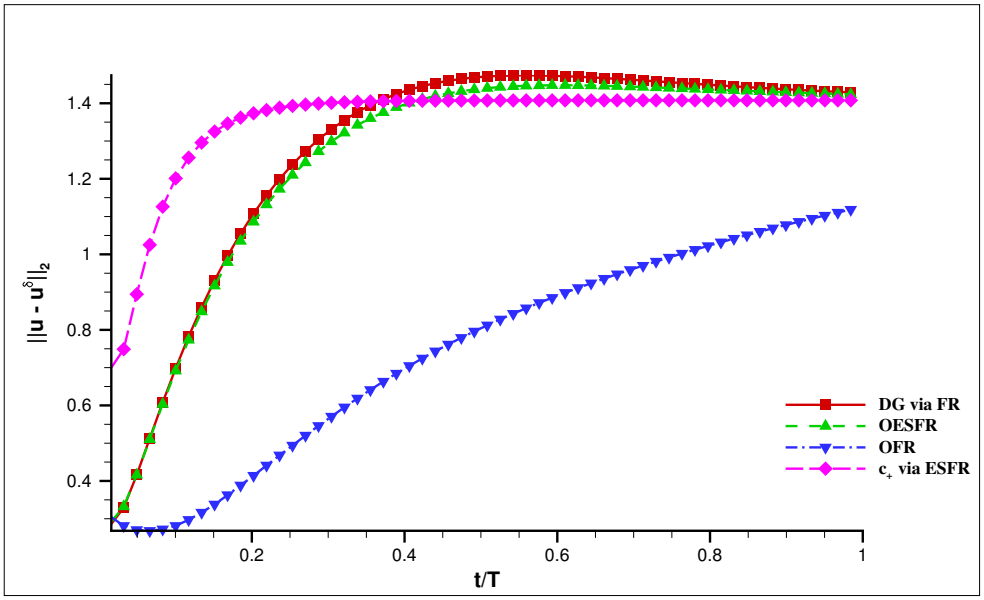


Figure 5.9: Case 2: Evolution of L^2 norm of the error for the DG, OESFR, OFR and C_+ schemes.

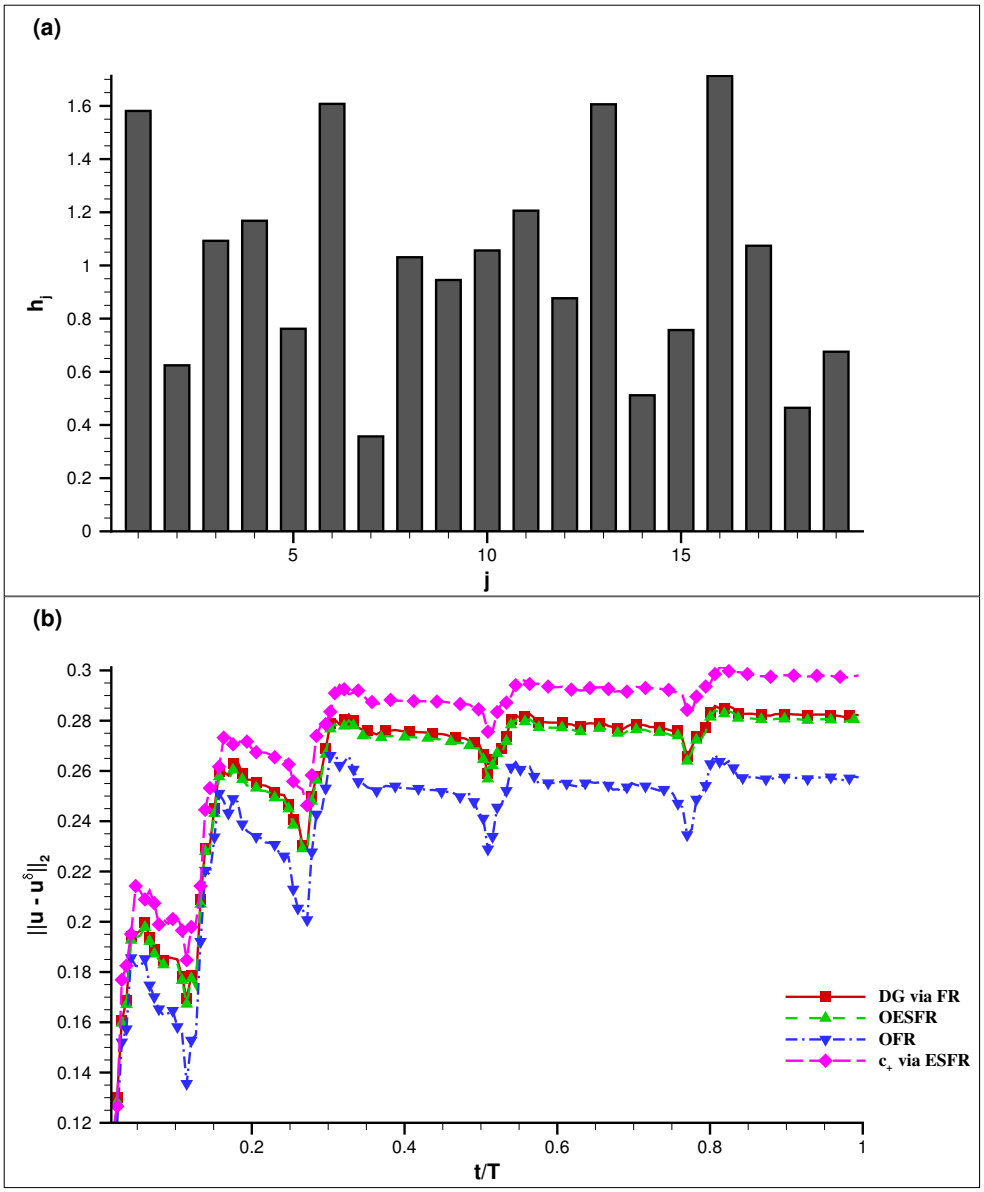


Figure 5.10: Case 3: (a) Variation of element size h_j across the perturbed, non-uniform grid, (b) Evolution of L^2 norm of the error for the DG, OESFR, OFR and C_+ schemes.

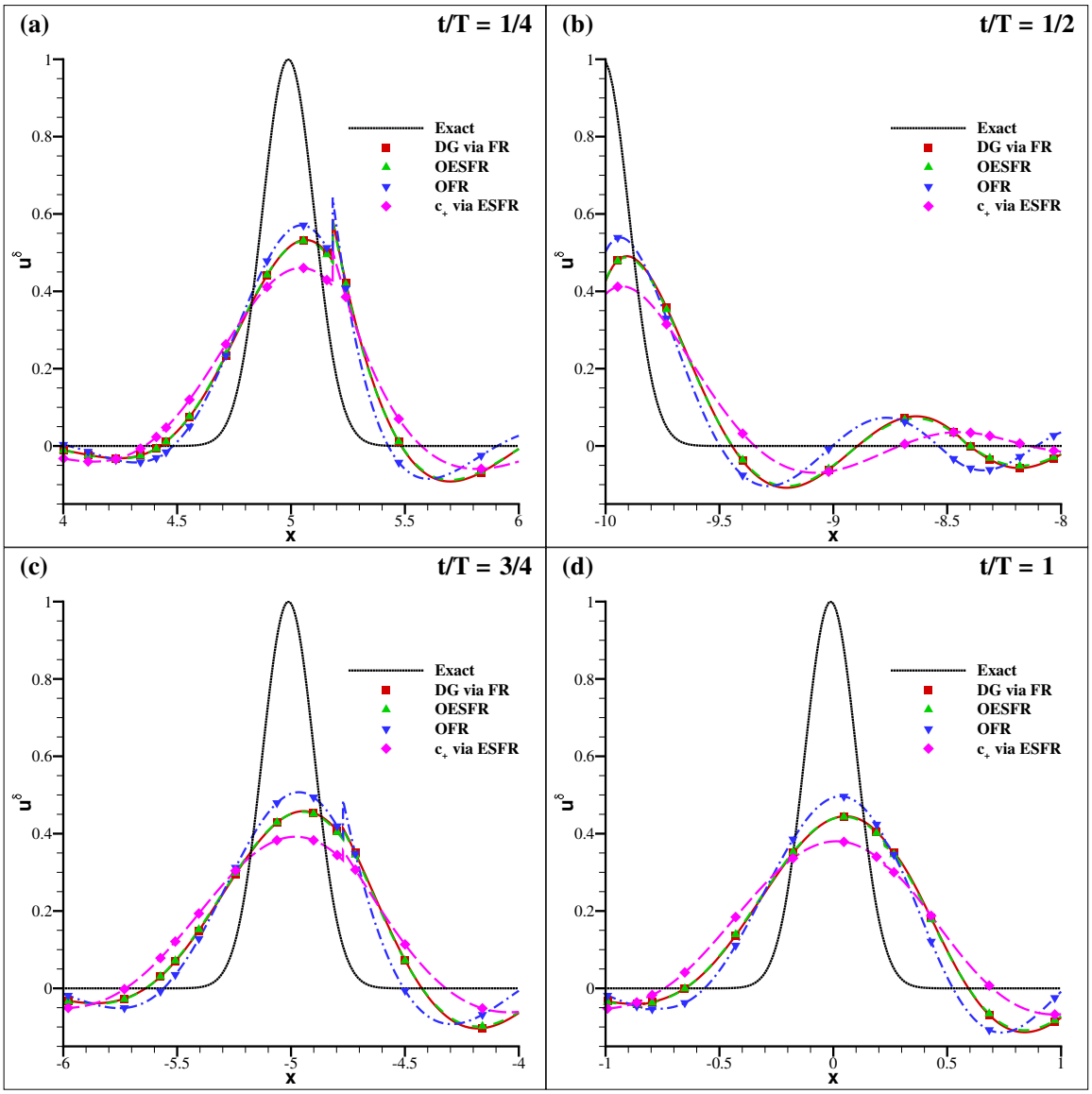


Figure 5.11: Case 3: Snapshots of the numerical solution for the DG, OESFR, OFR and C_+ schemes at intervals of a quarter-period.

their simplest form, NS equations prescribe the conservation of mass, momentum and energy for an isotropic Newtonian fluid under Stokes hypothesis with respect to an inertial frame of reference,

$$\frac{\partial \rho}{\partial t} + \sum_{j=1}^{n_{\text{sd}}} \frac{\partial}{\partial x_j} \left(\rho u_j \right) = 0, \tag{5.18}$$

$$\frac{\partial \rho u_i}{\partial t} + \sum_{j=1}^{n_{\text{sd}}} \frac{\partial}{\partial x_j} \left(\rho u_i u_j + p - \tau_{ij} \right) = \rho G_i, \quad i = 1, 2, \ldots, n_{\text{sd}}, \tag{5.19}$$

$$\frac{\partial \rho E}{\partial t} + \sum_{j=1}^{n_{\text{sd}}} \frac{\partial}{\partial x_j} \left(\rho H u_j - \sum_{i=1}^{n_{\text{sd}}} \tau_{ij} u_i + Q_j \right) = \rho \sum_{j=1}^{n_{\text{sd}}} G_j u_j, \tag{5.20}$$

for $\boldsymbol{x} = [x_1, x_2, \ldots, x_{n_{\text{sd}}}]^T \in \mathbb{R}^{n_{\text{sd}}}$. Here, ρ denotes density, u_i denotes the velocity component along x_i, p denotes pressure, G_i denotes the body force per unit mass along x_i, $E = e + \frac{1}{2} \sum_{j=1}^{n_{\text{sd}}} u_j^2$ denotes total internal energy, and $H = E + p/\rho$ denotes total enthalpy. The constitutive equations for the viscous stress tensor τ_{ij} and heat flux Q_i are given by

$$\tau_{ij} = \mu \left[\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) - \frac{2}{3} \delta_{ij} \sum_{k=1}^{n_{\text{sd}}} \frac{\partial u_k}{\partial x_k} \right], \quad i, j = 1, 2, \ldots, n_{\text{sd}}, \tag{5.21}$$

$$Q_i = -\kappa \frac{\partial u_i}{\partial x_i} = -\mu \frac{C_p}{\text{Pr}} \frac{\partial T_i}{\partial x_i}, \tag{5.22}$$

where μ is the coefficient of dynamic viscosity, C_p denotes the specific heat for constant pressure, T denotes temperature, and Pr denotes the Prandtl number. The various thermodynamic quantities are related by the laws of state for a calorically perfect ideal gas,

$$C_P = \frac{\gamma R}{\gamma - 1}, \tag{5.23}$$

$$p = \rho R T, \tag{5.24}$$

where R is the gas constant and γ denotes the ratio of specific heats.

A canonical solution to Eqns. (5.18 - 5.24) for $n_{\text{sd}} = 3$ is given by the Taylor-Green vortex (TGV) which describes the decay of a vortex in a periodic box on account of vortex stretching. Considered as an important test case for high order methods, TGV has been

featured in the International Workshops on High-Order CFD Methods [146]. As a specific instance, let the domain $\Omega = [0, 2\pi]^3$ and the initial condition be given by

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ p \\ T \end{bmatrix} (x, y, z, 0) = \begin{bmatrix} \sin(x_1)\cos(x_2)\cos(x_3) \\ -\cos(x_1)\sin(x_2)\cos(x_3) \\ 0 \\ 100 + \frac{1}{16}(\cos(2x_1) + \cos(2x_2))(\cos(2x_3) + 2) \\ 300 \end{bmatrix}. \tag{5.25}$$

This problem was simulated using DG, SD, OESFR, as well as OFR schemes by Bull and Jameson [21] for a Reynolds number of 1600. The domain was discretized using uniform hexahedral meshes with 32^3 elements and polynomial orders $P = 3, 4, 5$ and 6. For reference, the computed solutions were compared to those on meshes of 128^3 and 512^3 elements by Debonis [44] who employed Bogey and Bailley's 13-point dispersion-relation-preserving (DRP) scheme [18]. Figure 9 in [21] plots the vorticity-based dissipation rate ε_2 against time for the various schemes and polynomial orders. Bull et al. observed that "All the FR schemes do better than the DRP scheme on the 128^3 mesh despite being computed on a coarser mesh. OFR tends to slightly over-predict ε_2 at $P = 4$ and $P = 5$ while the other schemes, particularly DG, tend to under-predict it. At $P = 6$ OFR lies exactly on the reference DRP solution and DG and SD just below it."

Bull et al. also compare the energy spectra to a pseudo-spectral computation on a 512^3 element mesh computed by Carton de Wiart et al. [27]. Again, in the words of the authors: "It is clear that the OFR scheme achieves much better spectral resolution at higher wavenumbers than either DG or SD. "

The summary of their results can be noted from the following declaration: "The OFR scheme is as stable as, but more accurate than, ESFR recovering the SD and DG schemes. Energy spectra show that the OFR scheme provides superior resolution of the energy in the higher wavenumbers, confirming that the analysis of Asthana and Jameson is applicable to the Navier-Stokes equations in three dimensions."

Part II

Convergence theory for linear, steady-state problems

Chapter 6

Introduction

Among discontinuous finite element schemes, Discontinuous Galerkin (DG) and its variants have enjoyed a continuous rise in popularity and patronage. Particularly in the last two decades, the method has greatly benefited from theoretical studies highlighting super-convergence along nodal [3], functional [35, 36] as well as spectral [6, 4] points of view. In Part I of the dissertation, we proved that such super-convergence extends to the much broader Flux Reconstruction (FR) formulation as well. Specifically, in the case of time-dependent problems, the rate of convergence evolves in time starting from a short-time rate at $t = 0^+$ and ending asymptotically at a long-time rate at $t \to \infty$. These rates, which depend on the correction function, upwinding coefficient and location of solution points, can be readily inferred from the eigenvalues and eigenvectors of the corresponding differentiation operator.

We are now interested in examining the convergence of FR for problems exhibiting steady-state solutions. Such problems appear routinely in aerospace and automobile design applications [74], usually as an internal step which must be solved at each iteration of the design cycle. In the absence of a near-optimal initial guess, the number of function evaluations involved in such optimization problems can be quite large [101]. Consequently, it is vital to have fast steady-state solvers that comply with high-order discontinuous finite element schemes. Instead of tackling the elliptic problem directly, a popular approach is to march the hyperbolic system in pseudo-time starting from uniform or potential flow.

The simulation is deemed converged when the residual of the discretized differential equation falls below a threshold, indicating that a steady-state has been acquired. In certain pathological cases [72], non-linearity of the problem may admit multiple weak-solutions and therefore multiple steady-states, depending on the initial condition. However, for most practical problems, a unique steady-state exists as dictated by boundary data. In regard to fast pseudo-time marching schemes, three schools of thought exist, each with its own merits and limitations. The first school proposes the use of explicit time-stepping schemes coupled with convergence acceleration techniques such as h-multigrid [73, 134], p-multigrid [47], hp-multigrid [99] etc. The second school relies on implicit time stepping schemes [105, 100] coupled with preconditioners [102, 7]. The third school transfers the burden over to distributed, heterogeneous computing on massively parallel machines [156]. Regardless of the technique employed, the converged solution is nearly identical provided that the same convergence criterion is incorporated.

In the chapters that follow, we quantify the accuracy and rate of convergence of the converged, steady-state numerical solution. An important question that naturally arises is whether or not the long-time super-convergence, which exists for time-dependent hyperbolic problems, is relevant for steady-state hyperbolic problems. As one of the primary results, we show that the rate of convergence of steady-state error is identical to the short-time rate of convergence derived for time-dependent problems. This implies that the order $2P+q$ long-time super-convergence of DG has little relevance for aerospace and automobile design applications. Moreover, we prove that a family of infinite schemes exists within the FR formulation that achieves the same short-time rate of convergence as the DG scheme, i. e. order $P+q$. This further bolsters the search for schemes with improved accuracy and/or stability properties relative to DG.

For the DG scheme, several *a-posteriori* error estimates [13, 79] exist for elliptic problems. Arnold et al. [7] provide a comprehensive comparison of several DG type methods, showing that the L^2 norm of the solution converges at the order $P+1$ while the H^1 norm converges at the order P. Similar observations were made by Williams [153] for Couette flow using DG recovered via FR. Similarly, Cockburn et al. [38] have derived *a-priori* error estimates for the Poisson equation on tensor-product elements. The authors show that, with the choice of a special numerical flux, the L^2 norm of the solution converges at the order

$P+1$, while that of the derivative converges at $P+1/2$.

In this second part of the dissertation, we derive estimates for the rate of convergence of FR for a forced, linear, advection-diffusion problem exhibiting a steady-state. Two different measures of error are investigated. The first one accumulates pointwise error over the set of solution points. Denoted by $\ell^2(\Omega^\delta)$, this averaged vector norm is same as the one employed in Chapter 4, Eqn. (4.15). We show that the rate of convergence in this norm is identical to the short-time rate of convergence derived for time-dependent problems recorded in Tables 1.1 and 4.2. The proofs in this regard borrow naturally from the results on the asymptotic eigensystem derived in Chapter 3. We also measure error with respect to the vector semi-norm, $\mathscr{L}^\delta(\Omega^\delta)$, induced by the advection-diffusion operator. This unconventional choice enables us to relate the rate of convergence of the numerical solution with the rate of convergence of numerical derivative operators involved in FR. The value of such a relationship is realized by noting that the FR derivative operators are analytically tractable. Power expansions of these operators can be algebraically manipulated to derive truncation errors of the sort widely used in the analysis of compact finite-difference stencils [88].

We derive estimates for error in the first and second derivative operators as a function of polynomial order, correction functions and upwinding coefficients. For instance, in the case of DG with $P>0$, the first derivative operator converges at the order P unless P is even and fluxes are centered, in which case it converges at order $P+1$. Similarly, the second derivative operator converges at the orders $P-1$ and P respectively. This analysis has led to the prescription of a new class of schemes within the FR family, designated SFR, that are super-convergent for steady-state problems. The DG scheme is recovered as a member within this special class, but the SD scheme is not.

The exposition in this part is written in a self-contained manner to save the reader from the trouble of referring back to Chapter 2 for procedural details of FR. The part is formatted as follows. In Chapter 7, we describe the model problem and introduce the two measures of error, viz. the $\ell^2(\Omega^\delta)$ norm and the $\mathscr{L}^\delta(\Omega^\delta)$ semi-norm. Chapter 8 then derives the rates of convergence of the numerical first and second derivative operators involved in FR. The main results are summarized in Theorems 8.1 and 8.2. These results allow us to obtain the rate of convergence of the numerical solution in the $\mathscr{L}^\delta(\Omega^\delta)$ semi-norm, expressed in

Theorem 8.3. In Chapter 9, we use results from Chapters 3, 4 to express error in terms of the asymptotic eigensolution. This leads to Theorem 9.1 which provides the rate of convergence of the numerical solution in the $\ell^2(\Omega^\delta)$ norm. The part ends with Chapter 10 where we perform simple numerical experiments to verify the derived results.

The content of this part if largely drawn from the paper:

- K Asthana, J Watkins, A Jameson: On the rate of convergence of Flux Reconstruction for steady-state problems, under review at SIAM Journal of Numerical Analysis.

Chapter 7

Error in numerical solution to steady, advection-diffusion

This introductory chapter begins by deriving an equation for error in the FR solution for a forced, scalar, advection-diffusion equation exhibiting a steady state. We define two discrete measures of error that are utilized in the chapters that follow. The first accumulates pointwise error, while the other accumulates pointwise squared derivatives of error over the set of solution points. We show in passing that the rate of convergence of error, with respect to the second measure, i. e. the semi-norm induced by the advection-diffusion operator, is directly dependent on the rates of convergence of numerical spatial derivative operators in FR. We then provide a brief overview of the FR formulation for first and second derivative operators.

7.1 Problem specification

Consider the linear advection-diffusion in $x \in [0,1]$, $t > 0$, for a scalar-valued function, $\hat{u} \in \left(\mathbb{C}^2((0,1)) \times \mathbb{C}^1(\mathbb{R}_{++})\right)$, driven by a steady, sinusoidal forcing term of wavenumber $k = 2\pi m$, $m \in \mathbb{N}$,

$$\frac{\partial \hat{u}}{\partial t}(x,t) + (\mathscr{L}\hat{u})(x,t) = (iak + bk^2)\exp(ikx), \quad x \in (0,1),\ t > 0, \tag{7.1}$$

$$\hat{u}(x,0) = \exp(ik_0 x), \quad x \in [0,1], \tag{7.2}$$

$$\frac{\partial^r \hat{u}}{\partial x^r}(0,t) = \frac{\partial^r \hat{u}}{\partial x^r}(1,t), \quad r = 0,1,\ t \geq 0, \tag{7.3}$$

$$\mathscr{L} := a\frac{\partial}{\partial x} - b\frac{\partial^2}{\partial x^2}, \tag{7.4}$$

where $k_0 = 2\pi m_0,\ m_0 \in \mathbb{N}$, is the wavenumber of the initial condition, and $a \in \mathbb{R}, b \in \mathbb{R}_{++}$ are advection and diffusion coefficients respectively. The boundary conditions are assumed to be periodic for simplicity. The analytical solution to Eqns. (7.1 - 7.4) for $x \in [0,1]$, $t \geq 0$ is given by

$$\hat{u}(x,t) = \exp\left(ik_0(x-at) - bk_0^2 t\right) + \exp(ikx)\left[1 - \exp\left(-iakt - bk^2 t\right)\right], \tag{7.5}$$

where the first term corresponds to the exponential decay of the initial condition and the second one corresponds to the growth of the particular solution driven by the source.

Restriction of the problem to twice continuously differentiable functions allows for numerical procedures that attempt to approximate Eqn. (7.1) directly in strong form. Following the procedure in Chapter 2, the domain $\Omega = [0,1]$ is partitioned into non-overlapping segments $\Omega = \bigcup_{n=1}^{N} \Omega_n$. For a scheme of order P, each segment is further discretized with $P+1$ distinct solution points Ω_n^δ. The numerical solution, $\hat{u}^\delta$, when restricted to Ω_n, resides within a polynomial space $\mathbb{P}_P(\Omega_n)$ consisting of all polynomials of degree P or lower. The numerical conservation law is obtained by satisfying Eqn. (7.1) at the set of solution points, $\Omega^\delta = \bigcup_{n=1}^{N} \Omega_n^\delta$,

$$\frac{\partial^\delta \hat{u}^\delta}{\partial t}(x,t) + \left(\mathscr{L}^\delta \hat{u}^\delta\right)(x,t) = (iak + bk^2)\exp(ikx), \quad x \in \Omega^\delta,\ t > 0, \tag{7.6}$$

$$\hat{u}^\delta(x,0) = (\mathscr{P}\hat{u}_0)(x), \quad x \in [0,1], \tag{7.7}$$

$$\frac{\partial^r \hat{u}^\delta}{\partial x^r}(0,t) = \frac{\partial^r \hat{u}^\delta}{\partial x^r}(1,t), \quad r = 0,1,\ t \geq 0, \tag{7.8}$$

$$\mathscr{L}^{\delta} := a\frac{\partial^{\delta}}{\partial x} - b\frac{\partial^{\delta}}{\partial x}\frac{\partial^{\delta}}{\partial x}, \tag{7.9}$$

where $\mathscr{L}^{\delta}$ is the numerical differentiation operator in space and $\frac{\partial^{\delta}}{\partial t}$ is that in time. $\hat{u}_0(x) = \exp(ik_0x)$ is the initial condition for the exact problem and $\mathscr{P}$ is the collocation-projection operator that maps the initial condition to the space of the numerical solution. The primary quantity of interest is the error in the numerical solution,

$$e_{\hat{u}}(x,t) := \hat{u}^{\delta}(x,t) - \hat{u}(x,t), \quad x \in [0,1],\ t \geq 0. \tag{7.10}$$

A governing equation for the dynamics of error can be obtained by subtracting Eqns. (7.1 - 7.3) from Eqns. (7.6 - 7.8) and adding and subtracting terms, so that after algebraic rearrangement we obtain

$$\frac{\partial^{\delta} e_{\hat{u}}}{\partial t}(x,t) + \left(\mathscr{L}^{\delta} e_{\hat{u}}\right)(x,t) = \left[\left(\frac{\partial}{\partial t} - \frac{\partial^{\delta}}{\partial t}\right) + (\mathscr{L} - \mathscr{L}^{\delta})\right]\hat{u}(x,t), \quad x \in \Omega^{\delta},\ t > 0, \tag{7.11}$$

$$e_{\hat{u}}(x,0) = (\mathscr{P}\hat{u}_0)(x) - \hat{u}_0(x), \quad x \in [0,1], \tag{7.12}$$

$$\frac{\partial^r e_{\hat{u}}}{\partial x^r}(0,t) = \frac{\partial^r e_{\hat{u}}}{\partial x^r}(1,t), \quad r = 0,1,\ t \geq 0, \tag{7.13}$$

which shows that the error itself follows a modified form of the numerical conservation law which is driven by the error in the differentiation operators in space and time. At time $t = 0$, the error is solely on account of the projection of the initial condition on to the space of the numerical solution. Thereafter, it is driven by the inability of the scheme to accurately differentiate the exact solution.

7.2 Error in the steady-state numerical solution

Let u be the unique limit $u(x) := \lim_{t\to\infty} \hat{u}(x,t) = \exp(ikx)$ which satisfies the steady advection-diffusion equation,

$$(\mathscr{L}u)(x) = (iak + bk^2)\exp(ikx), \quad x \in (0,1), \tag{7.14}$$

$$\frac{\partial^r u}{\partial x^r}(0) = \frac{\partial^r u}{\partial x^r}(1), \quad r = 0,1, \tag{7.15}$$

a periodic problem on the unit interval. Similarly, for any stable and consistent numerical scheme, let u^δ be the limit $u^\delta(x) := \lim_{t\to\infty} \hat{u}^\delta(x,t)$ which must satisfy

$$\left(\mathscr{L}^\delta u^\delta\right)(x) = (iak + bk^2)\exp(ikx), \quad x \in \Omega^\delta, \tag{7.16}$$

$$\frac{\partial^r u^\delta}{\partial x^r}(0) = \frac{\partial^r u^\delta}{\partial x^r}(1), \quad r = 0,1. \tag{7.17}$$

Existence and uniqueness of u^δ, in the limit of asymptotic grid resolution, follows from the proof of convergence, Theorem 4.1, for linear, time-dependent problems. Heuristically, this is expected since $\hat{u}^\delta(x,t) \to \hat{u}(x,t)$ for $x \in [0,1]$, at any time $t > 0$. Also, on account of uniqueness, the limit is independent of the time integration scheme implied by the temporal discretization operator $\frac{\partial^\delta}{\partial t}$. In this regard, u^δ can be readily reconciled as a solution to Eqns. (7.16 - 7.17) obtained by marching Eqns. (7.6 - 7.8) in pseudo-time.

With these observations in place, we see from Eqns. (7.11 - 7.17) that the point-wise error also exhibits a unique limit,

$$e_u(x) := \lim_{t\to\infty} e_{\hat{u}}(x,t) = u^\delta(x) - u(x) \tag{7.18}$$

$$= u^\delta(x) - \exp(ikx), \quad x \in [0,1], \tag{7.19}$$

which satisfies

$$\left(\mathscr{L}^\delta e_u\right)(x) = (\mathscr{L} - \mathscr{L}^\delta)\exp(ikx), \quad x \in \Omega^\delta, \tag{7.20}$$

$$\frac{\partial^r e_u}{\partial x^r}(0) = \frac{\partial^r e_u}{\partial x^r}(1), \quad r = 0,1, \tag{7.21}$$

which is a discrete Poisson equation with a forcing term equal to the difference between

the true derivative $\mathscr{L}u$ and the numerical derivative $\mathscr{L}^\delta u$ of the exact solution,

$$(\mathscr{L}-\mathscr{L}^\delta)\exp(ikx) = \left[a\left(ik-\frac{\partial^\delta}{\partial x}\right)+b\left(k^2+\frac{\partial^\delta\partial^\delta}{\partial x\partial x}\right)\right]\exp(ikx),\ \ x\in\Omega^\delta. \tag{7.22}$$

7.3 Measures of error

Traditionally, finite element schemes have been analyzed through the use of functional norms, especially of the Sobolev type which integrate error and its derivatives across the domain. However, owing to the nodal nature of FR, both of our employed measures are discrete in nature, and accumulate over the set of solution points. We define the vector $\ell^2(\Omega^\delta)$ norm of error in a manner analogous to Eqn. (4.15),

$$\|e_u\|_{\ell^2(\Omega^\delta)} := \left[\frac{1}{N(P+1)}\sum_{n=1}^{N}\sum_{p=1}^{P+1}\left|e_u(x_{n,p})\right|^2\right]^{1/2}, \tag{7.23}$$

where $x_{n,p}$ is the pth solution point in the nth element. The rate of convergence of error with respect to this norm would be derived in Chapter 9 using results on the asymptotic eigensolution from Chapter 3.

In order to obtain the second measure, consider the semi-norm which is induced by the advection-diffusion operator,

$$\|e_u\|_{\mathscr{L}^\delta(\Omega^\delta)} := \left[\frac{1}{N(P+1)}\sum_{n=1}^{N}\sum_{p=1}^{P+1}\left|(\mathscr{L}^\delta e_u)(x_{n,p})\right|^2\right]^{1/2} \tag{7.24}$$

where $\mathscr{L}^\delta$ is as defined in Eqn. (7.9). The requirements of non-negativity and triangle inequality follow from the fact that $\|e_u\|_{\mathscr{L}^\delta(\Omega^\delta)} = \|\mathscr{L}^\delta e_u\|_{\ell^2(\Omega^\delta)}$ and that $\mathscr{L}^\delta$ is a linear operator. Let $h := \min_{n=1}^{N}|\Omega_n|$ be the length of the smallest element in the domain. Then, owing to consistency of the FR formulation, we are guaranteed that $\exists\ \sigma^{(1)}>0$, $\sigma^{(2)}>0$

such that

$$\lim_{h\to 0}\left|\left(\frac{\partial^{\delta}}{\partial x}-ik\right)\exp(ikx)\right| \leq c(k,P)\,h^{\sigma^{(1)}}, \tag{7.25}$$

$$\lim_{h\to 0}\left|\left(\frac{\partial^{\delta}\partial^{\delta}}{\partial x\partial x}+k^2\right)\exp(ikx)\right| \leq c(k,P)\,h^{\sigma^{(2)}}, \tag{7.26}$$

for all wavenumbers k, and $x \in \Omega^{\delta}$. Here $c > 0$ is some (k,P)-dependent constant and $\sigma^{(1)}$ and $\sigma^{(2)}$ denote the orders at which the first and second order numerical derivatives converge to the exact derivatives respectively. The rate of convergence of error in $\mathscr{L}^{\delta}$ can now be obtained by substituting Eqns. (7.20), (7.22), (7.25 - 7.26) in (7.24),

$$\begin{aligned}
\|e_u\|_{\mathscr{L}^{\delta}(\Omega^{\delta})} &= \left\|(\mathscr{L}-\mathscr{L}^{\delta})\exp(ikx)\right\|_{\ell^2(\Omega^{\delta})} \\
&= \left\|\left[a\left(ik-\frac{\partial^{\delta}}{\partial x}\right)+b\left(k^2+\frac{\partial^{\delta}\partial^{\delta}}{\partial x\partial x}\right)\right]\exp(ikx)\right\|_{\ell^2(\Omega^{\delta})} \\
&\leq |a|\left\|\left(ik-\frac{\partial^{\delta}}{\partial x}\right)\exp(ikx)\right\|_{\ell^2(\Omega^{\delta})} + b\left\|\left(k^2+\frac{\partial^{\delta}\partial^{\delta}}{\partial x\partial x}\right)\exp(ikx)\right\|_{\ell^2(\Omega^{\delta})} \\
&\leq |a|\max_{x\in\Omega^{\delta}}\left|\left(ik-\frac{\partial^{\delta}}{\partial x}\right)\exp(ikx)\right| + b\max_{x\in\Omega^{\delta}}\left|\left(k^2+\frac{\partial^{\delta}\partial^{\delta}}{\partial x\partial x}\right)\exp(ikx)\right| \\
&\leq c(k,P)\left[|a|h^{\sigma^{(1)}}+bh^{\sigma^{(2)}}\right], \quad \text{as } h\to 0.
\end{aligned} \tag{7.27}$$

This shows that the rate of convergence of error in the $\mathscr{L}^{\delta}$ semi-norm is directly dependent on the rates of convergence of the numerical first and second derivative operators. The utility of this relation lies in the fact that derivative operators in FR are analytically tractable as demonstrated in Chapter 8.

7.4 Derivative operators in Flux Reconstruction

Before proceeding on to error estimates, we provide a self-sustained account of the FR formulation for evaluating the numerical first and second derivatives of a continuous function $f : (0,1) \mapsto \mathbb{C}$. A detailed account of the reconstruction procedure can be found in

[69, 71, 28].

Continuing our discussion from Section 7.1, consider a set of equispaced nodes on the domain such that $x_0 = 0$, $x_n = x_{n-1} + h$ for $n = 1, 2 \ldots, N$, and $x_N = 1$. For this uniform mesh, the elements $\Omega_n = [x_{n-1}, x_n)$ are of equal length h. Introduce a linear map from the nth element $x \in \Omega_n$ to the parent segment $\xi \in [-1, 1)$,

$$\xi|_{\Omega_n}(x) = \frac{2}{h}(x - x_n) - 1, \quad x \in \Omega_n. \tag{7.28}$$

For an FR scheme of order P, let $\boldsymbol{x}_n = (x_{n,p})_{p=1,2,\ldots,P+1}$ denote the set of solution points Ω_n^δ in the nth element, which map to $\boldsymbol{\xi} = (\xi_p)_{p=1,2,\ldots,P+1}$ in the parent space. The numerical approximation, $f^\delta(x)$, when restricted to Ω_n, is defined as the interpolant to the actual function $f(x)$ through $\boldsymbol{x}_n$. For instance, from Eqn. (7.25), we are interested in evaluating the numerical derivative of the Fourier component, $f(x) = \exp(ikx)$. In this case,

$$f^\delta(x) = (\mathscr{P} f)(x) = \sum_{p=1}^{P+1} \exp(ikx_{n,p}) \ell_p(\xi(x)), \quad x \in \Omega_n \tag{7.29}$$

where ℓ_p is the p^{th} Lagrange polynomial in the parent segment. The numerical approximation can equivalently be represented in vector notation as

$$\begin{aligned} \boldsymbol{f}_n^\delta &:= f^\delta(x_{n,p})_{p=1,2,\ldots,P+1} \\ &= \exp(ikx_n) \left(\exp\left(ik\frac{h}{2}(1+\xi_p) \right) \right)_{p=1,2,\ldots,P+1}. \end{aligned} \tag{7.30}$$

The numerical derivative, $\frac{d^\delta f}{dx^\delta}(x \in \Omega_n)$, is obtained by differentiating f^δ after it has been 'corrected' to ensure continuity across the two ends of the element. Towards this end, common interface values are first computed using the polynomial functions on either side of the interface,

$$f_I^\delta(x_n) = (1-\alpha) f^\delta(x_n^-) + \alpha f^\delta(x_n^+), \tag{7.31}$$

where $\alpha \in \left[0, \frac{1}{2}\right]$ denotes the upwinding coefficient and $(\cdot)^-$, $(\cdot)^+$ denote the left and right hand limits. Next, the interface information is conveyed to the interior of the element through the use of correction polynomials, $g_L(\xi)$, $g_R(\xi)$, of degree $P+1$ which satisfy the

constraints

$$g_L(-1) = g_R(+1) = 1, \tag{7.32}$$

$$g_L(+1) = g_R(-1) = 0, \tag{7.33}$$

leading to the construction of the corrected approximation, $f^{\delta,c}$, a piecewise polynomial of degree $P+1$, which is globally continuous,

$$\begin{aligned} f^{\delta,c}(x) = f^{\delta}(x) &+ (f_I^{\delta}(x_n) - f^{\delta}(x_n^+))g_L(\xi(x)) \\ &+ (f_I^{\delta}(x_{n+1}) - f^{\delta}(x_{n+1}^-))g_R(\xi(x)), \quad x \in \Omega_n. \end{aligned} \tag{7.34}$$

Plugging Eqns. (7.29) and (7.31) in Eqn. (7.34) and taking derivatives, we get that the vector of derivative values in the n^{th} element can be expressed compactly [9] as

$$\left(\frac{\partial^{\delta}}{\partial x}\exp(ikx)\right)(\boldsymbol{x}_n) = \left(\frac{\partial f^{\delta,c}}{\partial x}\right)(\boldsymbol{x}_n) = \frac{2}{h}\exp(ikx_n)\boldsymbol{Q}(kh;g,\alpha)\boldsymbol{f}_n^{\delta}, \tag{7.35}$$

where $\boldsymbol{f}_n^{\delta}$ is defined in Eqn. (7.30) and $\boldsymbol{Q} \in \mathbb{C}^{(P+1)\times(P+1)}$ is the numerical differentiation operator in the parent space, Eqn. (2.13),

$$\boldsymbol{Q}(kh;g,\alpha) = \boldsymbol{D} + (1-\alpha)\boldsymbol{g}_{L,\xi}(\exp(-ikh)\boldsymbol{\ell}_+^T - \boldsymbol{\ell}_-^T) + \alpha\boldsymbol{g}_{R,\xi}(\exp(ikh)\boldsymbol{\ell}_-^T - \boldsymbol{\ell}_+^T). \tag{7.36}$$

Here $\boldsymbol{D} \in \mathbb{R}^{(P+1)\times(P+1)}$ is the polynomial differentiation operator, Eqn. (2.14),

$$D_{p,m} = \frac{\mathrm{d}\ell_m}{\mathrm{d}\xi}(\xi_p), \quad p,m = 1,2,\ldots,P+1, \tag{7.37}$$

$\boldsymbol{g}_{L,\xi}, \boldsymbol{g}_{R,\xi} \in \mathbb{R}^{(P+1)\times 1}$ are the derivatives of the left-boundary and right-boundary correction functions at the solution points, and $\boldsymbol{\ell}_-, \boldsymbol{\ell}_+ \in \mathbb{R}^{(P+1)\times 1}$ are the extrapolated values of the $P+1$ Lagrange basis polynomials, Eqn. (2.15),

$$\ell_{+_p} = \ell_p(+1), \;\; \ell_{-_p} = \ell_p(-1), \quad p = 1,2,\ldots,P+1. \tag{7.38}$$

The second-order derivative is recursively defined as the derivative of the first-order derivative after it has been corrected to ensure continuity in derivative across the two ends of the element. Hence,

$$\left(\frac{\partial^{\delta}}{\partial x}\frac{\partial^{\delta}}{\partial x}\exp(ikx)\right)(\boldsymbol{x}_n) = \frac{4}{h^2}\exp(ikx_n)\boldsymbol{Q}(kh;g^{(2)},\alpha^{(2)})\boldsymbol{Q}(kh;g^{(1)},\alpha^{(1)})\boldsymbol{f}_n^{\boldsymbol{\delta}}, \qquad (7.39)$$

where $g^{(1)}, g^{(2)}$ denote the correction functions and $\alpha^{(1)}$, $\alpha^{(2)}$ denote the upwinding coefficients for evaluating the interface fluxes when correcting the polynomial approximation and the numerical first derivative respectively.

Chapter 8

Rate of convergence in the $\mathscr{L}^{\delta}(\Omega^{\delta})$ semi-norm

In Chapter 7 we have shown that the rate of convergence of error in the $\mathscr{L}^{\delta}$ semi-norm is directly dependent on the rates of convergence of the numerical first and second derivative operators. This chapter employs power expansions of the error vector to quantify the accuracy of derivative operators in FR. We show that, for an FR scheme of polynomial order P, the order of convergence of the first derivative operator depends on whether P is odd or even. For odd P, the order is exactly P. For even P, the order can be $P+1$ under special conditions on the correction function and upwinding coefficient. These conditions are satisfied by a family of linearly stable schemes with $P/2$ free parameters. The DG scheme is recovered as a member of this family, while the SD scheme is not. The order of convergence of the second derivative operator is dependent on the orders of convergence of constituent first derivative operators. Coupled with Eqn. (7.27), these results on derivative operators yield Theorem 8.3 which returns the discussion to error in the steady-state numerical solution, Eqn. (7.18), by providing a bound on the rate of convergence in the $\mathscr{L}^{\delta}$ semi-norm.

8.1 Error in the first derivative operator

We begin by deriving the rate at which the numerical first derivative of the Fourier component, $\exp(ikx)$, converges to the exact derivative $ik\ \exp(ikx)$ with refinement in grid spacing h. Using Eqns. (7.30) and (7.35), pointwise error in the derivative at the vector of solution points in the n^{th} element is given by

$$\begin{aligned}\left[\left(\frac{\partial^{\delta}}{\partial x}-ik\right)\exp(ikx)\right](\boldsymbol{x}_n) &= \frac{2}{h}\boldsymbol{Q}(kh;g,\alpha)\exp(ik\boldsymbol{x}_n)-ik\exp(ik\boldsymbol{x}_n)\\ &= \frac{2}{h}\exp(ikx_n)\boldsymbol{E}(kh;g,\alpha),\end{aligned} \tag{8.1}$$

where $\boldsymbol{E}(kh;g,\alpha)$ represents the difference between the numerical and exact derivative values at the set of solution points in the parent space,

$$\boldsymbol{E}(kh;g,\alpha)=\left(\boldsymbol{Q}(kh;g,\alpha)-ik\frac{h}{2}\boldsymbol{I}\right)\exp\left(ik\frac{h}{2}(1+\boldsymbol{\xi})\right). \tag{8.2}$$

The rate of convergence of $\boldsymbol{E}$ with refinement in the non-dimensional wavenumber kh can be obtained from Eqns. (7.25) and (8.1) as

$$\begin{aligned}|\boldsymbol{E}(kh;g,\alpha)| &\preceq \left|\frac{h}{2}\exp(-ikx_n)\left[\left(\frac{\partial^{\delta}}{\partial x}-ik\right)\exp(ikx)\right](\boldsymbol{x}_n)\right|\\ &\preceq c(P)(kh)^{\sigma^{(1)}+1}\,\mathbf{1},\end{aligned} \tag{8.3}$$

where $\preceq$ denotes a component-wise inequality, c is some scheme dependent constant and $\mathbf{1}=[1,1,\ldots,1]^T$. In order to proceed with an asymptotic analysis of this term, consider expanding $\boldsymbol{Q}(kh;g,\alpha)$ in its Taylor series about $kh=0$ using Eqn. (7.36),

$$\boldsymbol{Q}(kh)=\boldsymbol{Q}(0)+\boldsymbol{H}_L\sum_{m=1}^{\infty}\frac{(-ikh)^m}{m!}+\boldsymbol{H}_R\sum_{m=1}^{\infty}\frac{(ikh)^m}{m!}, \tag{8.4}$$

where the explicit parameterization with respect to g,α is suppressed for convenience of notation, and the component matrices are given by

$$\boldsymbol{Q}(0) = \left((1-\alpha)\boldsymbol{g}_{L,\xi} - \alpha\boldsymbol{g}_{R,\xi}\right)(\boldsymbol{\ell}_{+}^{T} - \boldsymbol{\ell}_{-}^{T}) + \boldsymbol{D}, \tag{8.5}$$

$$\boldsymbol{H}_{L} = (1-\alpha)\boldsymbol{g}_{L,\xi}\boldsymbol{\ell}_{+}^{T}, \tag{8.6}$$

$$\boldsymbol{H}_{R} = \alpha\boldsymbol{g}_{R,\xi}\boldsymbol{\ell}_{-}^{T}. \tag{8.7}$$

Similarly, expand the Fourier component in a power series about $kh = 0$,

$$\exp\left(ik\frac{h}{2}(1+\boldsymbol{\xi})\right) = \sum_{r=0}^{\infty}\frac{(ikh)^{r}}{r!}\boldsymbol{s}^{r}, \tag{8.8}$$

where, $\boldsymbol{s}^{r}$ is the sampling of the r^{th} degree polynomial,

$$s^{r}(\xi) = \left(\frac{1+\xi}{2}\right)^{r}, \tag{8.9}$$

on the set of solution points, $\boldsymbol{s}^{r} = (s^{r})(\xi_{p})_{p=1,2,\ldots,P+1}$. Plugging Eqns. (8.5 - 8.8) in Eqn. (8.2), the derivative error in the parent space can be written as

$$\begin{aligned}\boldsymbol{E}(kh) &= \left(\boldsymbol{Q}(0) + \boldsymbol{H}_{L}\sum_{m=1}^{\infty}\frac{(-ikh)^{m}}{m!} + \boldsymbol{H}_{R}\sum_{m=1}^{\infty}\frac{(ikh)^{m}}{m!} - ik\frac{h}{2}\boldsymbol{I}\right)\sum_{r=0}^{\infty}\frac{(ikh)^{r}}{r!}\boldsymbol{s}^{r}, \\ &= \boldsymbol{Q}(0)\boldsymbol{1} + \sum_{j=1}^{\infty}(ikh)^{j}\boldsymbol{v}_{j},\end{aligned} \tag{8.10}$$

where for $j \geq 1$,

$$\boldsymbol{v}_{j} = \boldsymbol{Q}(0)\frac{1}{j!}\boldsymbol{s}^{j} - \frac{1}{2(j-1)!}\boldsymbol{s}^{j-1} + \sum_{m=1}^{j}\frac{1}{m!(j-m)!}\left((-1)^{m}\boldsymbol{H}_{L} + \boldsymbol{H}_{R}\right)\boldsymbol{s}^{j-m}. \tag{8.11}$$

In order to evaluate the component terms, $\boldsymbol{v}_{j}$, we recognize that the matrix operators $\boldsymbol{Q}(0), \boldsymbol{H}_{L}, \boldsymbol{H}_{R}$ in Eqn. (8.5 - 8.7) are dependent on primitive operators, $\boldsymbol{l}_{+}^{T}, \boldsymbol{l}_{-}^{T}, \boldsymbol{D}$, Eqns. (7.37), (7.38), which are intrinsically related to polynomial interpolation. We can now enlist the action of aforementioned primitive operators on the sampled polynomial defined in

Eqn. (8.9),

$$\boldsymbol{l}_{+}^{T}\boldsymbol{s}^{j} = \begin{cases} 1 & j \leq P \\ (\mathscr{P}s^{j})(\xi = 1) & j > P \end{cases} \tag{8.12}$$

$$\boldsymbol{l}_{-}^{T}\boldsymbol{s}^{j} = \begin{cases} \mathrm{I}_{\{j=0\}} & j \leq P \\ (\mathscr{P}s^{j})(\xi = -1) & j > P \end{cases} \tag{8.13}$$

$$\boldsymbol{D}\boldsymbol{s}^{j} = \begin{cases} \frac{j}{2}\boldsymbol{s}^{j-1} & j \leq P \\ \frac{d(\mathscr{P}s^{j})}{d\xi}(\boldsymbol{\xi}) & j > P \end{cases} \tag{8.14}$$

where I denotes the indicator function and $\mathscr{P}$ is the collocation projection operator, Eqn. (7.29). These relations follow from the observation that $s^{j}(\xi)$ is a polynomial of degree j and for $j \leq P$ resides within $\mathbb{P}_{P}$ which is the space of the numerical solution.

In the results that follow, we derive the rate of convergence mentioned in Eqn. (8.3) to determine the dependence of $\sigma^{(1)}$ on the polynomial order P, the upwinding coefficient α and the correction function $g_{L}(\xi)$.

Lemma 8.1. *For any FR scheme of polynomial order P, i.e. for all choices of correction function and upwinding coefficient, $\boldsymbol{E}(kh)$ defined in Eqn. (8.2) vanishes at an order higher than or equal to $P+1$,*

$$\boldsymbol{E}(kh) = \sum_{j=P+1}^{\infty} (ikh)^{j}\boldsymbol{v}_{j}, \tag{8.15}$$

where $\boldsymbol{v}_{j}$ is as defined in Eqn. (8.11).

Proof. We begin by noting that $\boldsymbol{Q}(0)\mathbf{1} = \mathbf{0}$ from Lemma 3.4. Let $1 \leq j \leq P$ in Eqn. (8.11). Then, using the properties in Eqns. (8.12 - 8.14),

$$\begin{aligned} \boldsymbol{v}_{j} &= \frac{1}{j!}\left((1-\alpha)\boldsymbol{g}_{L,\xi} - \alpha\boldsymbol{g}_{R,\xi}\right) + \frac{j}{2j!}\boldsymbol{s}^{j-1} - \frac{1}{2(j-1)!}\boldsymbol{s}^{j-1} \\ &\quad + \sum_{m=1}^{j}\frac{(-1)^{m}}{m!(j-m)!}(1-\alpha)\boldsymbol{g}_{L,\xi} + \frac{1}{j!}\alpha\boldsymbol{g}_{R,\xi} \\ &= (1-\alpha)\boldsymbol{g}_{L,\xi}\left(\frac{1}{j!} + \sum_{m=1}^{j}\frac{(-1)^{m}}{m!(j-m)!}\right) \end{aligned}$$

$$
\begin{aligned}
&= (1-\alpha)\boldsymbol{g}_{L,\xi}\left(\frac{1}{j!}+\sum_{m=0}^{j}\frac{(-1)^m}{m!(j-m)!}-\frac{1}{j!}\right)\\
&= \mathbf{0},
\end{aligned}
$$

where we have used binomial expansion so that $\sum\limits_{m=0}^{j}\frac{(-1)^m}{m!(j-m)!}=\frac{1}{j!}(1+(-1))^j=0$. The proof now follows from Eqn. (8.10). □

Lemma 8.1 and Eqn. (8.3) show that $\sigma^{(1)}\geq P$ for any convergent FR scheme of order P. In order to investigate the possibility of super-convergence, we need to simplify terms in Eqn. (8.11) for $j>P$. As polynomials of degree $j>P$ cannot be interpolated exactly using only $P+1$ points, we require a couple of preliminary results adapted from error estimates in Lagrange interpolation.

Lemma 8.2. *Given a set of $P+1$ distinct points, ξ_p for $p=1,2,\ldots,P+1$, arranged symmetrically in $[-1,1]$ and the collocation projection operator $\mathscr{P}$, Eqn. (7.29), we have the following identities:*

1. $(\mathscr{P}s^{P+1})(\xi=1)-(\mathscr{P}s^{P+1})(\xi=-1)=1+\frac{1}{2^{P+1}}\left((-1)^{P+1}-1\right)\prod\limits_{q=1}^{P+1}(1-\xi_q)$.
2. $\frac{d(\mathscr{P}s^{P+1})}{d\xi}(\xi_p)=\frac{P+1}{2}\left(\frac{1+\xi_p}{2}\right)^P-\frac{1}{2^{P+1}}\prod\limits_{\substack{q=1\\q\neq p}}^{P+1}(\xi_p-\xi_q)$.

Proof. Let $\phi(\xi)=s^{P+1}=\left(\frac{1+\xi}{2}\right)^{P+1}$. Then from the error estimate in Lagrange interpolation, for $\xi\in[-1,1]$,

$$
\begin{aligned}
\phi(\xi)-(\mathscr{P}\phi)(\xi) &= \frac{1}{(P+1)!}\frac{\partial^{P+1}\phi}{\partial\xi^{P+1}}\prod_{q=1}^{P+1}(\xi-\xi_q)\\
&= \frac{1}{2^{P+1}}\prod_{q=1}^{P+1}(\xi-\xi_q). \qquad (8.16)
\end{aligned}
$$

Now it is easy to see that,

$$(\mathscr{P}s^{P+1})(\xi=1)-(\mathscr{P}s^{P+1})(\xi=-1)$$

$$
\begin{aligned}
&= \phi(1) - \frac{1}{2^{P+1}} \prod_{q=1}^{P+1} (1-\xi_q) - \phi(-1) + \frac{1}{2^{P+1}} \prod_{q=1}^{P+1} (-1-\xi_q) \\
&= 1 + \frac{1}{2^{P+1}} \left((-1)^{P+1} \prod_{q=1}^{P+1} (1+\xi_q) - \prod_{q=1}^{P+1} (1-\xi_q) \right) \\
&= 1 + \frac{1}{2^{P+1}} \left((-1)^{P+1} - 1 \right) \prod_{q=1}^{P+1} (1-\xi_q), \qquad (8.17)
\end{aligned}
$$

where we have used the symmetry assumption on the set of points,

$$
\prod_{q=1}^{P+1} (1+\xi_q) = \prod_{q=1}^{P+1} (1-\xi_q). \qquad (8.18)
$$

Similarly, we can obtain an error estimate for the derivative in Lagrange interpolation,

$$
\begin{aligned}
\frac{d\phi}{d\xi}(\xi_p) - \frac{d(\mathscr{P}\phi)}{d\xi}(\xi_p) &= \frac{1}{(P+1)!} \left[\frac{\partial^{P+2}\phi}{\partial \xi^{P+2}}(\xi_p) \prod_{q=1}^{P+1} (\xi_p - \xi_q) + \frac{\partial^{P+1}\phi}{\partial \xi^{P+1}}(\xi_p) \sum_{r=1}^{P+1} \prod_{\substack{q=1 \\ q\neq r}}^{P+1} (\xi_p - \xi_q) \right] \\
&= \frac{1}{(P+1)!} \frac{\partial^{P+1}\phi}{\partial \xi^{P+1}}(\xi_p) \prod_{\substack{q=1 \\ q\neq p}}^{P+1} (\xi_p - \xi_q) \\
&= \frac{1}{2^{P+1}} \prod_{\substack{q=1 \\ q\neq p}}^{P+1} (\xi_p - \xi_q), \qquad (8.19)
\end{aligned}
$$

for $p = 1, 2, \ldots, P+1$. Rearranging terms and substituting $\frac{d\phi}{d\xi}(\xi_p) = \frac{P+1}{2} \left(\frac{1+\xi_p}{2} \right)^P$ into Eqn. (8.19) yields the desired identity. □

Lemma 8.2 explores the effect of extrapolation and differentiation operators on a polynomial that is one degree higher than the dimension of the polynomial basis. This allows us to simplify $\boldsymbol{v}_{P+1}$ in this next result:

Lemma 8.3. *For a given FR scheme of order P, if the solution points $\boldsymbol{\xi}$ are distinct and distributed symmetrically, then, in the limit $kh \to 0$, $\boldsymbol{E}(kh; g, \alpha)$ vanishes at an order*

- $\leq P+1$ *for odd P.*

- $\geq P+2$ *for even* P *if and only if*

$$(1-\alpha)g_{L,\xi_p} - \alpha g_{R,\xi_p} = -\frac{1}{2}\frac{\prod_{\substack{q=1\\q\neq p}}^{P+1}(\xi_p - \xi_q)}{\prod_{q=1}^{P+1}(1-\xi_q)}, \quad p = 1,2,\ldots,P+1. \tag{8.20}$$

where $g_{L,\xi_p} = \frac{dg_L}{d\xi}(\xi_p)$ *and* $g_{R,\xi_p} = \frac{dg_R}{d\xi}(\xi_p)$ *respectively.*

Proof. Proceeding as in the proof of Lemma 8.1, let $j = P+1$ in Eqn. (8.11). Then using the properties in Eqns. (8.12 - 8.14) and Lemma 8.2,

$$\begin{aligned}
\boldsymbol{v}_{P+1} &= \frac{1}{(P+1)!}\left((1-\alpha)\boldsymbol{g}_{L,\xi} - \alpha\boldsymbol{g}_{R,\xi}\right)\left((\mathscr{P}s^{P+1})(\xi=1) - (\mathscr{P}s^{P+1})(\xi=-1)\right) \\
&\quad + \frac{1}{(P+1)!}\frac{d(\mathscr{P}s^{P+1})}{d\xi}(\boldsymbol{\xi}) - \frac{1}{2P!}\boldsymbol{s}^P \\
&\quad + \sum_{m=1}^{P+1}\frac{(-1)^m}{m!(P+1-m)!}(1-\alpha)\boldsymbol{g}_{L,\xi} + \frac{1}{(P+1)!}\alpha\boldsymbol{g}_{R,\xi} \\
&= \frac{1}{(P+1)!}\left((1-\alpha)\boldsymbol{g}_{L,\xi} - \alpha\boldsymbol{g}_{R,\xi}\right)\left(1 + \frac{1}{2^{P+1}}\left((-1)^{P+1}-1\right)\prod_{q=1}^{P+1}(1-\xi_q)\right) \\
&\quad + \frac{P+1}{2(P+1)!}\boldsymbol{s}^P - \frac{1}{2^{P+1}(P+1)!}\boldsymbol{\theta} - \frac{1}{2P!}\boldsymbol{s}^P \\
&\quad + \sum_{m=1}^{P+1}\frac{(-1)^m}{m!(P+1-m)!}(1-\alpha)\boldsymbol{g}_{L,\xi} + \frac{1}{(P+1)!}\alpha\boldsymbol{g}_{R,\xi} \\
&= \frac{1}{(P+1)!}\Bigg[\left((1-\alpha)\boldsymbol{g}_{L,\xi} - \alpha\boldsymbol{g}_{R,\xi}\right)\left(1 + \frac{1}{2^{P+1}}\left((-1)^{P+1}-1\right)\prod_{q=1}^{P+1}(1-\xi_q)\right) \\
&\quad - \frac{1}{2^{P+1}}\boldsymbol{\theta} - \left((1-\alpha)\boldsymbol{g}_{L,\xi} - \alpha\boldsymbol{g}_{R,\xi}\right)\Bigg] \\
&= \frac{1}{2^{P+1}(P+1)!}\Bigg[\left((1-\alpha)\boldsymbol{g}_{L,\xi} - \alpha\boldsymbol{g}_{R,\xi}\right)\left((-1)^{P+1}-1\right)\prod_{q=1}^{P+1}(1-\xi_q) - \boldsymbol{\theta}\Bigg],
\end{aligned}$$

where

$$\boldsymbol{\theta} = \left(\prod_{\substack{q=1 \\ q \neq p}}^{P+1} (\xi_p - \xi_q) \right)_{p=1,2,\ldots,P+1}. \tag{8.21}$$

Simplifying in the case of odd and even P,

$$-2^{P+1}(P+1)!\, \boldsymbol{v}_{P+1} = \begin{cases} \boldsymbol{\theta}, & \text{odd } P \\ 2\left((1-\alpha)\boldsymbol{g}_{L,\xi} - \alpha \boldsymbol{g}_{R,\xi}\right) \prod_{q=1}^{P+1} (1-\xi_q) + \boldsymbol{\theta}, & \text{even } P \end{cases} \tag{8.22}$$

Since $\boldsymbol{\theta} \neq \mathbf{0}$, it is certain that $\boldsymbol{v}_{P+1} \neq \mathbf{0}$ for odd P which proves the first half of the Lemma. However, in the case of even P, if and only if

$$\left((1-\alpha)\boldsymbol{g}_{L,\xi} - \alpha \boldsymbol{g}_{R,\xi}\right) \prod_{q=1}^{P+1} (1-\xi_q) = -\frac{1}{2}\boldsymbol{\theta}, \tag{8.23}$$

then $\boldsymbol{v}_{P+1} = \mathbf{0}$ and the second part of the Lemma is proven as well. □

Lemma 8.3 establishes an upper bound on the rate of convergence of the derivative for odd P. Coupled with the result from Lemma 8.1 it implies that $\sigma^{(1)} = P$ if P is odd. At the same time, it shows that if P is even, then under special circumstances the rate of convergence can be higher, i. e. $\sigma^{(1)} \geq P+1$. An upper bound for even P can be derived by simplifying the term $\boldsymbol{v}_{P+2}$. This is illustrated next for the special case of $P = 0$.

Lemma 8.4. *For the FR scheme of order $P = 0$, in the limit $kh \to 0$, $\boldsymbol{E}(kh;\alpha)$ vanishes at the rate*

- $(kh)^2$ *if* $\alpha \neq \frac{1}{2}$,
- $(kh)^3$ *if* $\alpha = \frac{1}{2}$.

Proof. For the case of $P = 0$, the single solution point is located at the origin, $\boldsymbol{\xi} = [0]$, and the only correction functions that respect the boundary constraints are given by $g_L(\xi) =$

$\frac{1}{2}(1-\xi)$, $g_R(\xi)=\frac{1}{2}(1+\xi)$, so that $\boldsymbol{g}_{L,\xi}=[-\frac{1}{2}]=-\boldsymbol{g}_{R,\xi}$. The monomial $s^j(\xi)$ is approximated by its constant interpolant through $\xi=0$, i.e. $(\mathscr{P}s^j)=\frac{1}{2^j}$. This yields

$$(\mathscr{P}s^j)(\xi=1)=\frac{1}{2^j},\quad (\mathscr{P}s^j)(\xi=-1)=\frac{1}{2^j},\quad \frac{d(\mathscr{P}s^j)}{d\xi}=0. \tag{8.24}$$

Proceeding as in the proof of Lemma 8.3, for any $j\geq 1$,

$$\begin{aligned}
v_j &= -\frac{1}{2(j-1)!}\frac{1}{2^{j-1}}+\sum_{m=1}^{j}\frac{1}{m!(j-m)!}\left((-1)^m(1-\alpha)\frac{-1}{2}+\alpha\frac{1}{2}\right)\frac{1}{2^{j-m}}\\
&= -\frac{1}{2^j(j-1)!}-\frac{1}{2}\sum_{m=1}^{j}\frac{1}{m!(j-m)!}\left((-1)^m(1-\alpha)-\alpha\right)\frac{1}{2^{j-m}}\\
&= -\frac{1}{2^j(j-1)!}-\frac{1}{2}\left[\sum_{m=0}^{j}\frac{1}{m!(j-m)!}\left((-1)^m(1-\alpha)-\alpha\right)\frac{1}{2^{j-m}}-(1-2\alpha)\frac{1}{2^j j!}\right]\\
&= -\frac{1}{2^j(j-1)!}-\frac{1}{2j!}\left[(1-\alpha)\left(-1+\frac{1}{2}\right)^j-\alpha\left(1+\frac{1}{2}\right)^j-(1-2\alpha)\frac{1}{2^j}\right]\\
&= -\frac{1}{2^j(j-1)!}-\frac{1}{2^{j+1}j!}\left[(-1)^j-1+\left((-1)^{j+1}-3^j+2\right)\alpha\right]\\
&= -\frac{1}{2^{j+1}j!}\left[2j+(-1)^j-1+\left((-1)^{j+1}-3^j+2\right)\alpha\right]
\end{aligned}$$

so that

$$-2^{j+1}j!\,v_j=\begin{cases}0, & j=1\\ 4-8\alpha, & j=2\\ 4-24\alpha, & j=3\end{cases} \tag{8.25}$$

which shows that $v_1\equiv 0$, $v_2=0$ if and only if $\alpha=\frac{1}{2}$, and $v_3\neq 0$ if $\alpha=\frac{1}{2}$. □

Lemma 8.4 shows that the simple finite-volume scheme enjoys an enhanced rate of convergence. For $\alpha\neq\frac{1}{2}$, $\sigma^{(1)}=1$ and for $\alpha=\frac{1}{2}$, $\sigma^{(1)}=2$. However, such additional superconvergence is not retained for any other even P. We are now ready to state the final result of this section.

Theorem 8.1. *For an FR scheme of order P, having a symmetric set of distinct solution points* $\boldsymbol{\xi} \in [-1,1]^{P+1}$*, and symmetric correction functions* $g_{R,\xi}(\xi) = -g_{L,\xi}(-\xi)$*, the numerical first derivative converges to the exact first derivative at the order* $\sigma^{(1)}$*,*

$$\lim_{h\to 0}\left|\left(\frac{\partial^{\delta}}{\partial x} - ik\right)\exp(ikx)\right| \leq c_1\, k^{\sigma^{(1)}+1}\, h^{\sigma^{(1)}},\ \ x \in \Omega^{\delta}, \tag{8.26}$$

where h denotes the element width, $c_1(P,g,\alpha)$ *is a scheme dependent constant,* Ω^{δ} *is the set of solution points and,*

1. $\sigma^{(1)} = \begin{cases} 1, & \alpha \neq \frac{1}{2} \\ 2, & \alpha = \frac{1}{2} \end{cases}$ *for* $P = 0$*,*
2. $\sigma^{(1)} = P$ *for odd P,*
3. $\sigma^{(1)} \geq P+1$ *for even P provided, for* $p = 1,2,\ldots,P+1$*,*

$$\begin{cases} g_{L,\xi_p} = g_{L,\xi_{P+2-p}} = -\frac{1}{2}\dfrac{\prod\limits_{\substack{q=1\\ q\neq p}}^{P+1}(\xi_p - \xi_q)}{\prod\limits_{q=1}^{P+1}(1-\xi_q)}, & \alpha \neq \frac{1}{2} \\ g_{L,\xi_p} + g_{L,\xi_{P+2-p}} = -\dfrac{\prod\limits_{\substack{q=1\\ q\neq p}}^{P+1}(\xi_p - \xi_q)}{\prod\limits_{q=1}^{P+1}(1-\xi_q)}, & \alpha = \frac{1}{2} \end{cases} \tag{8.27}$$

where $g_{L,\xi_p} = \frac{dg_L}{d\xi}(\xi_p)$*, or* $\sigma^{(1)} = P$ *otherwise.*

Proof. 1. $P = 0$: The proof follows directly from Lemma 8.4 and Eqn. (8.3).

2. Odd P: The proof follows directly from Lemmas 8.1, 8.3 and Eqn. (8.3).

3. Even P: In the case of even P, begin by noting that the right hand side of Eqn. (8.20)

is symmetric with respect to p, i. e.

$$\prod_{\substack{q=1\\q\neq p}}^{P+1}(\xi_p-\xi_q)=(-1)^P\prod_{\substack{q=1\\q\neq p}}^{P+1}(\xi_p-\xi_q)=\prod_{\substack{q=1\\q\neq P+2-p}}^{P+1}(\xi_{P+2-p}-\xi_q), \tag{8.28}$$

owing to the symmetry of the solution points and that P is even. Hence,

$$(1-\alpha)g_{L,\xi_p}-\alpha g_{R,\xi_p}-(1-\alpha)g_{L,\xi_{P+2-p}}+\alpha g_{R,\xi_{P+2-p}}=0. \tag{8.29}$$

Additionally, from the symmetry of the correction functions, $g_{R,\xi_p}=-g_{L,\xi_{P+2-p}}$ for $p=1,2,\ldots,P+1$, so that

$$\begin{aligned}(1-\alpha)g_{L,\xi_p}-\alpha g_{R,\xi_p}-(1-\alpha)g_{L,\xi_{P+2-p}}+\alpha g_{R,\xi_{P+2-p}}&\\ =(1-\alpha)g_{L,\xi_p}+\alpha g_{L,\xi_{P+2-p}}-(1-\alpha)g_{L,\xi_{P+2-p}}-\alpha g_{L,\xi_p}&\\ =(1-2\alpha)(g_{L,\xi_p}-g_{L,\xi_{P+2-p}}).&\end{aligned} \tag{8.30}$$

From Eqns. (8.29) and (8.30),

$$(1-2\alpha)(g_{L,\xi_p}-g_{L,\xi_{P+2-p}})=0, \tag{8.31}$$

which shows that either $\alpha=\frac{1}{2}$ or $g_{L,\xi_p}=g_{L,\xi_{P+2-p}}=-g_{R,\xi_p}$. The result then follows from substitution into Eqn. (8.20).

□

8.2 Error in the second derivative operator

The second derivative in Flux Reconstruction is defined recursively as the derivative of the corrected first derivative which itself is the derivative of the corrected polynomial approximation. The procedure involves two corrections, one for the polynomial approximation, with correction function $g_L^{(1)}$ and upwinding coefficient $\alpha^{(1)}$, and one for the first derivative, with correction function $g_L^{(2)}$ and upwinding coefficient $\alpha^{(2)}$. Using Eqn. (7.39), the pointwise error in evaluation of the second derivative at the vector of solution points in the nth element is given by

$$\begin{aligned}
&\left[\left(\frac{\partial^{\delta}}{\partial x}\frac{\partial^{\delta}}{\partial x}+k^2\right)\exp(ikx)\right](\boldsymbol{x}_n)\\
&=\frac{4}{h^2}\boldsymbol{Q}(kh;g^{(2)},\alpha^{(2)})\boldsymbol{Q}(kh;g^{(1)},\alpha^{(1)})\exp(ik\boldsymbol{x}_n)+k^2\exp(ik\boldsymbol{x}_n)\\
&=\frac{4}{h^2}\exp(ikx_n)\boldsymbol{F}(kh;g^{(2)},g^{(1)},\alpha^{(2)},\alpha^{(1)}),
\end{aligned}\tag{8.32}$$

where $\boldsymbol{F}$ represents the difference between the numerical and exact second derivative values at the set of solution points in the parent space,

$$\begin{aligned}
&\boldsymbol{F}(kh;g^{(2)},g^{(1)},\alpha^{(2)},\alpha^{(1)})\\
&=\left[\boldsymbol{Q}(kh;g^{(2)},\alpha^{(2)})\boldsymbol{Q}(kh;g^{(1)},\alpha^{(1)})+\frac{h^2k^2}{4}\right]\exp\left(ik\frac{h}{2}(1+\boldsymbol{\xi})\right).
\end{aligned}\tag{8.33}$$

Once again, from Eqns. (7.26) and (8.32), it follows that

$$|\boldsymbol{F}(kh;g^{(2)},g^{(1)},\alpha^{(2)},\alpha^{(1)})|\preceq c(P)(kh)^{\sigma^{(2)}+2}\,\mathbf{1},\tag{8.34}$$

for some scheme dependent constant c. As in the previous subsection, an asymptotic analysis of this term can be performed through a power series expansion in kh about 0. However, the algebraic accounting becomes exceptionally tedious. Nevertheless, a lower bound on the rate of convergence can be readily achieved by substituting Eqn. (8.2) in Eqn. (8.33). Denoting $\boldsymbol{Q}(kh;g^{(m)},\alpha^{(m)})$ as $\boldsymbol{Q}^{(m)}$ for $m=1,2$ and suppressing parameterization where

apparent, we get that

$$\begin{aligned}
\boldsymbol{F} &= \boldsymbol{Q}^{(2)}\left(\boldsymbol{Q}^{(1)}\exp\left(ik\frac{h}{2}(1+\boldsymbol{\xi})\right)\right)+\frac{h^2k^2}{4}\exp\left(ik\frac{h}{2}(1+\boldsymbol{\xi})\right)\\
&= \boldsymbol{Q}^{(2)}\left(\boldsymbol{E}(kh;g^{(1)},\alpha^{(1)})+ik\frac{h}{2}\exp\left(ik\frac{h}{2}(1+\boldsymbol{\xi})\right)\right)+\frac{h^2k^2}{4}\exp\left(ik\frac{h}{2}(1+\boldsymbol{\xi})\right)\\
&= \boldsymbol{Q}^{(2)}\boldsymbol{E}(kh;g^{(1)},\alpha^{(1)})+ik\frac{h}{2}\left(\boldsymbol{E}(kh;g^{(2)},\alpha^{(2)})+ik\frac{h}{2}\exp\left(ik\frac{h}{2}(1+\boldsymbol{\xi})\right)\right)\\
&\qquad\qquad +\frac{h^2k^2}{4}\exp\left(ik\frac{h}{2}(1+\boldsymbol{\xi})\right)\\
&= \boldsymbol{Q}^{(2)}\boldsymbol{E}(kh;g^{(1)},\alpha^{(1)})+ik\frac{h}{2}\boldsymbol{E}(kh;g^{(2)},\alpha^{(2)}),
\end{aligned}\tag{8.35}$$

which shows that error in the second derivative operator can be determined from the error in the two constituent first derivative operators. As a specific example, consider the case of $P=0$ for which the following result can be established.

Lemma 8.5. *For an FR scheme of order $P=0$, in the limit $kh\to 0$, $\boldsymbol{F}(kh)$ vanishes at the rate*

- $(kh)^3$ *if* $\alpha^{(1)}+\alpha^{(2)}\neq 1$,
- $(kh)^4$ *if* $\alpha^{(1)}+\alpha^{(2)}=1$.

Proof. As a continuation of the proof for Lemma 8.4, note that for $P=0$ the differentiation operator is a scalar given by Eqn. (8.4),

$$Q(kh;\alpha)=\frac{1}{2}\sum_{m=1}^{\infty}\frac{(ikh)^m}{m!}\left((-1)^{m+1}(1-\alpha)+\alpha\right).$$

Then, using Eqns. (8.10) and (8.35), the Taylor expansion of $F(kh;\alpha^{(2)},\alpha^{(1)})$ can be written as

$$F=Q^{(2)}\sum_{j=1}^{\infty}(ikh)^j v_j(\alpha^{(1)})+ik\frac{h}{2}\sum_{j=1}^{\infty}(ikh)^j v_j(\alpha^{(2)})$$

$$
\begin{aligned}
&= \sum_{j=1}^{\infty} (ikh)^j \left(Q^{(2)} v_j(\alpha^{(1)}) + ik\frac{h}{2} v_j(\alpha^{(2)}) \right) \\
&= \sum_{j=1}^{\infty} (ikh)^j \left(\frac{1}{2} \sum_{m=1}^{\infty} \frac{(ikh)^m}{m!} \left((-1)^{m+1}(1-\alpha^{(2)}) + \alpha^{(2)} \right) v_j(\alpha^{(1)}) + ik\frac{h}{2} v_j(\alpha^{(2)}) \right) \\
&= \frac{1}{2} \sum_{j=1}^{\infty} (ikh)^{j+1} \left[v_j(\alpha^{(1)}) + v_j(\alpha^{(2)}) + \sum_{m=2}^{\infty} \frac{(ikh)^{m-1}}{m!} \left((-1)^{m+1}(1-\alpha^{(2)}) + \alpha^{(2)} \right) v_j(\alpha^{(1)}) \right] \\
&= \frac{1}{2}(ikh)^2 \left(v_1(\alpha^{(1)}) + v_1(\alpha^{(2)}) \right) + \frac{1}{2}(ikh)^3 \left(v_2(\alpha^{(1)}) + v_2(\alpha^{(2)}) + \left(\alpha^{(2)} - \frac{1}{2} \right) v_1(\alpha^{(1)}) \right) \\
&\quad + \frac{1}{2}(ikh)^4 \left(v_3(\alpha^{(1)}) + v_3(\alpha^{(2)}) + \frac{1}{6} v_1(\alpha^{(1)}) + \left(\alpha^{(2)} - \frac{1}{2} \right) v_2(\alpha^{(1)}) \right) + \mathscr{O}\left((kh)^5 \right)
\end{aligned}
$$

Substituting $v_1 = 0$, $v_2(\alpha) = 4 - 8\alpha$ and $v_3 = 4 - 24\alpha$ from Eqn. (8.25), we get

$$
\begin{aligned}
F = &\frac{1}{2}(ikh)^3 \left(8(1 - \alpha^{(1)} - \alpha^{(2)}) \right) \\
&+ (ikh)^4 \left(3 - 10\alpha^{(1)} - 10\alpha^{(2)} - 4\alpha^{(1)}\alpha^{(2)} \right) + \mathscr{O}\left((kh)^5 \right)
\end{aligned}
$$

which completes the proof. □

Lemma 8.5 highlights the importance of having an isotropic operator. In order to derive an asymptotic bound for $P > 0$, let $\sigma^{(1),1}$ and $\sigma^{(1),2}$ be the rates of convergence associated with the two first derivative operators involved. From Eqns. (8.3) and (8.10), this can be expressed as

$$
\boldsymbol{E}(kh; g^{(m)}, \alpha^{(m)}) = \sum_{j=\sigma^{(1),m}+1}^{\infty} (ikh)^j \boldsymbol{v}_j^{(m)}, \quad m = 1, 2. \tag{8.36}
$$

We can now state the following result regarding a lower bound for the rate of convergence of the second derivative.

Theorem 8.2. *For an FR scheme having a symmetric set of distinct solution points and symmetric correction functions, the numerical second derivative converges to the exact second derivative at the order* $\sigma^{(2)}$*,*

$$\lim_{h\to 0}\left|\left(\frac{\partial^{\delta}}{\partial x}\frac{\partial^{\delta}}{\partial x}+k^2\right)\exp(ikx)\right| \leq c_2\, k^{\sigma^{(2)}+2}\, h^{\sigma^{(2)}},\ \ x\in\Omega^{\delta}, \tag{8.37}$$

where h *denotes the element width,* $c_2(P,g^{(2)},g^{(1)},\alpha^{(2)},\alpha^{(1)})$ *is a scheme dependent constant,* Ω^{δ} *is the set of solution points, and*

1. $\sigma^{(2)}=\begin{cases}1, & \alpha^{(1)}+\alpha^{(2)}\neq 1\\ 2, & \alpha^{(1)}+\alpha^{(2)}=1\end{cases}$ *for* $P=0$*,*
2. $\sigma^{(2)}\geq \min\left(\sigma^{(1),1}-1,\sigma^{(1),2}\right)$ *for* $P>0$*,*

where $\sigma^{(1),1},\sigma^{(1),2}$ *are the rates of convergence of the two first derivative operators involved in differentiating the polynomial approximation and the first derivative respectively.*

Proof.

1. $P=0$: The proof follows directly from Lemma 8.5 and Eqn. (8.34).
2. For $P>0$, define $\sigma=\min(\sigma^{(1),1}-1,\sigma^{(1),2})$. Then, substituting Eqn. (8.36) in Eqn. (8.35) yields

$$\begin{aligned}\boldsymbol{F} &= \boldsymbol{Q}^{(2)}\sum_{j=\sigma^{(1),1}+1}^{\infty}(ikh)^j\boldsymbol{v}_j^{(1)}+ik\frac{h}{2}\sum_{j=\sigma^{(1),2}+1}^{\infty}(ikh)^j\boldsymbol{v}_j^{(2)}\\ &= (ikh)^{\sigma+2}\left[\sum_{j=\sigma^{(1),1}-1-\sigma}^{\infty}(ikh)^j\boldsymbol{Q}^{(2)}\boldsymbol{v}_j^{(1)}+\frac{1}{2}\sum_{j=\sigma^{(1),2}-\sigma}^{\infty}(ikh)^j\boldsymbol{v}_{j-1}^{(2)}\right]\end{aligned} \tag{8.38}$$

Now, by definition, $\sigma \leq \sigma^{(1),1} - 1$ and $\sigma < \sigma^{(1),2}$. So, from the triangle inequality,

$$\lim_{kh \to 0} |\boldsymbol{F}| \leq c(g^{(1)}, g^{(2)}, \alpha^{(1)}, \alpha^{(2)})(kh)^{\sigma+2}. \tag{8.39}$$

The proof then follows from Eqn. (8.34). □

8.3 Rate of convergence in the $\mathscr{L}^{\delta}(\Omega^{\delta})$ semi-norm

Theorems 8.1 and 8.2 combined with Eqn. (7.27) readily yield the following lower bound for the rate of convergence in the semi-norm induced by the advection-diffusion operator.

Theorem 8.3. *For any stable FR scheme, the order of convergence of error in the $\mathscr{L}^{\delta}(\Omega^{\delta})$ semi-norm, Eqn. (7.24), is given by σ,*

$$\begin{aligned} \lim_{h \to 0} \|e_u\|_{\mathscr{L}^{\delta}(\Omega^{\delta})} &\leq C' \left[|a| \, k^{\sigma^{(1)}+1} \, h^{\sigma^{(1)}} + b \, k^{\sigma^{(2)}+2} \, h^{\sigma^{(2)}} \right] \\ &\leq C \, h^{\sigma}, \end{aligned} \tag{8.40}$$

where $C(P, g^{(2)}, g^{(1)}, \alpha^{(2)}, \alpha^{(1)}, a, b, k)$ is a scheme, equation dependent constant and

$$\sigma \geq \min(\sigma^{(1)}, \sigma^{(2)}) \geq P - 1, \tag{8.41}$$

where $\sigma^{(1)}$, $\sigma^{(2)}$ are the rates of convergence of respective derivative operators as described in Theorems 8.1 and 8.2.

Theorem 8.3 provides a closed-form, functional dependence of the rate of convergence on the polynomial order, correction function and upwinding coefficient. We can now identify FR schemes that exhibit nodal super-convergence for steady-state problems, as discussed in the next section.

Note that such tractability has been procured at the cost of an unconventional measure for error. In Chapter 9, we will relax the need for closed-form expressions to obtain the rate of convergence in the $\ell^2(\Omega^{\delta})$ norm using the asymptotic eigensolution.

8.4 Super-convergent FR (SFR) schemes

Theorem 8.1 shows that the rate of convergence of the numerical first derivative operator depends on whether P is odd or even. For odd P, the order is exactly P irrespective of the choice of the correction function or upwinding coefficient. However, for even P, super-convergence can be achieved under special conditions. In particular, for such super-convergence, the correction function is identified uniquely if the interface flux is not centered, $\alpha \neq \frac{1}{2}$, i. e.

$$g_{L,\xi_p} = -\frac{1}{2}\frac{\prod_{\substack{q=1 \\ q\neq p}}^{P+1}(\xi_p - \xi_q)}{\prod_{q=1}^{P+1}(1-\xi_q)}, \ p = 1,2,\ldots,P+1, \tag{8.42}$$

where $g_{L,\xi_p} = \frac{dg_L}{d\xi}(\xi_p)$. However, if the fluxes are centered, $\alpha = \frac{1}{2}$, then a family of correction functions exists. This family is parameterized by $P/2$ free parameters. One possible choice for the parameter set is the correction function derivatives at the solution points in the left half of the parent space, i.e. $p = 1,2,\ldots,\frac{P}{2}$. The remaining derivatives are then prescribed by Eqn. (8.27) as

$$g_{L,\xi_{P+2-p}} = -g_{L,\xi_p} - \frac{\prod_{\substack{q=1 \\ q\neq p}}^{P+1}(\xi_p - \xi_q)}{\prod_{q=1}^{P+1}(1-\xi_q)}, \ p = 1,2,\ldots,\frac{P}{2}, \tag{8.43}$$

$$g_{L,\xi_{\frac{P}{2}+1}} = -\frac{1}{2}\frac{\prod_{\substack{q=1 \\ q\neq p}}^{P+1}(\xi_p - \xi_q)}{\prod_{q=1}^{P+1}(1-\xi_q)}. \tag{8.44}$$

The specification in Eqn. (8.42) or Eqns. (8.43-8.44) ensures that the first derivative operator is super-convergent, i. e. converges at an order higher than P. Moreover, if the numerical second derivative is evaluated using two super-convergent first-derivative operators, then

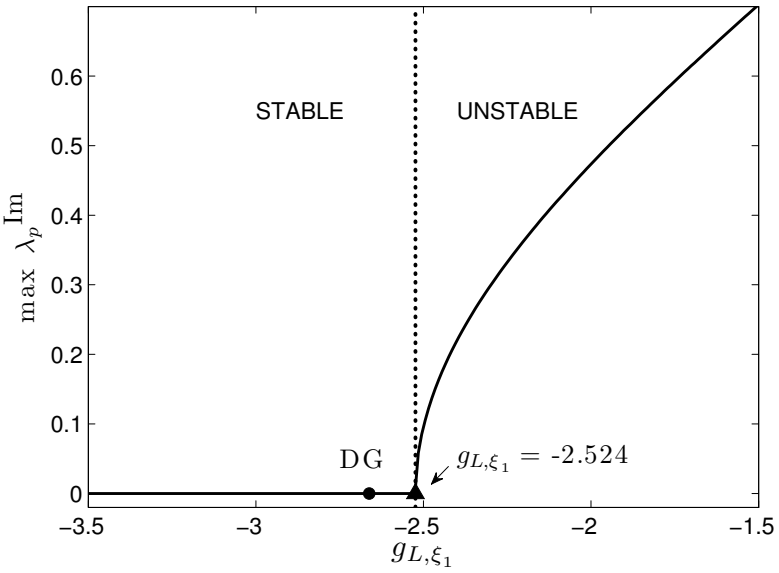
Theorem 8.2 guarantees that the second derivative operator would be super-convergent as well, i. e. converges at an order higher than $P-1$. Finally, Theorem 8.3 then ensures that the resulting FR scheme would be super-convergent in the $\mathscr{L}^{\delta}(\Omega^{\delta})$ semi-norm. We temporarily designate this family as SFR which stands for 'Super-convergent Flux Reconstruction'.

Note that this result on super-convergence is provisional on stability. We can test linear stability of any SFR scheme by computing eigenvalues of the corresponding update matrix. In the case of pure advection, the condition for numerical stability [69] is given by Eqn. (2.24). The presence of a diffusion term would further relax this criterion due to added physical dissipation. Hence, the stability criterion in Eqn. (2.24) also serves as a conservative estimate for the more general advection-diffusion problem. On performing the eigenvalue computation, we find that, in the case of $\alpha \neq \frac{1}{2}$, the unique super-convergent scheme in Eqn. (8.42) is, in fact, unstable at least for $P=2$ and $P=4$. However for $\alpha = \frac{1}{2}$, a family of stable schemes exist. Figure 8.1a plots $\max_{p=1}^{P+1} \lambda_p^{\mathrm{Im}}(kh)$ for $P=2$ against the free parameter g_{L,ξ_1}. We see that all schemes for which $g_{L,\xi_1} < -2.524$ are stable. Figure 8.1b plots the region of stability for this family for $P=4$. We see that the scheme is stable when $\left(g_{L,\xi_1}, g_{L,\xi_2}\right)$ is chosen from within a convex subset of $\mathbb{R}^2$. A vertex of this region is located at $g_{L,\xi_1} = -5.906$, $g_{L,\xi_2} = 1.422$.

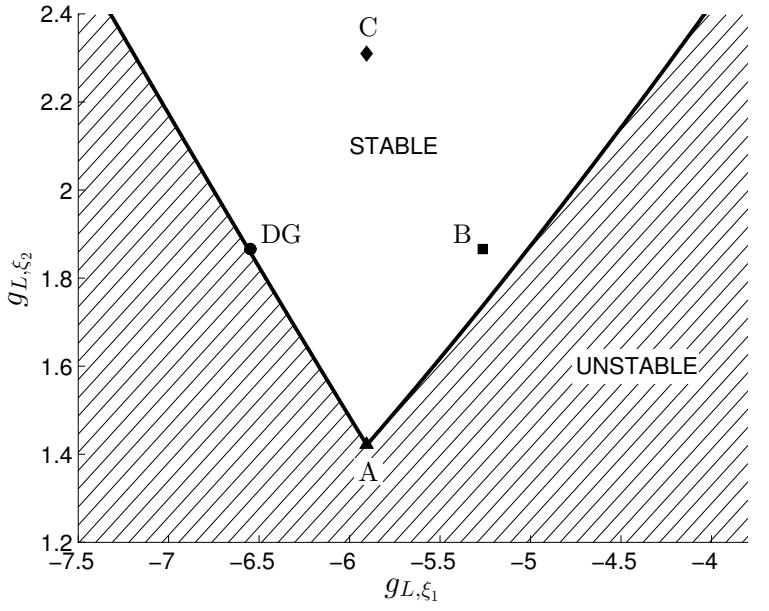
Figure 8.2a plots the correction function $g_L(\xi)$ for $\xi \in [-1,1]$ for four stable members of the SFR family, viz. the boundary scheme with $g_{L,\xi_1} = -2.524$, the DG scheme, and the schemes with g_{L,ξ_1} as $-3.0, -4.0$ respectively. We see that as the parameter g_{L,ξ_1} decreases, the first root of the correction function moves to the left while the second moves to the right. A similar plot for $P=4$ is recorded in Fig. 8.2b for the four schemes marked in Fig. 8.1b and recorded in Table 8.1.

Scheme	g_{L,ξ_1}	g_{L,ξ_2}
A	-5.9060	1.4220
DG	-6.5480	1.8661
B	-5.2640	1.8661
C	-5.9060	2.3102

Table 8.1: $P=4$: Parameters for the super-convergent schemes marked in Fig. 8.1b.

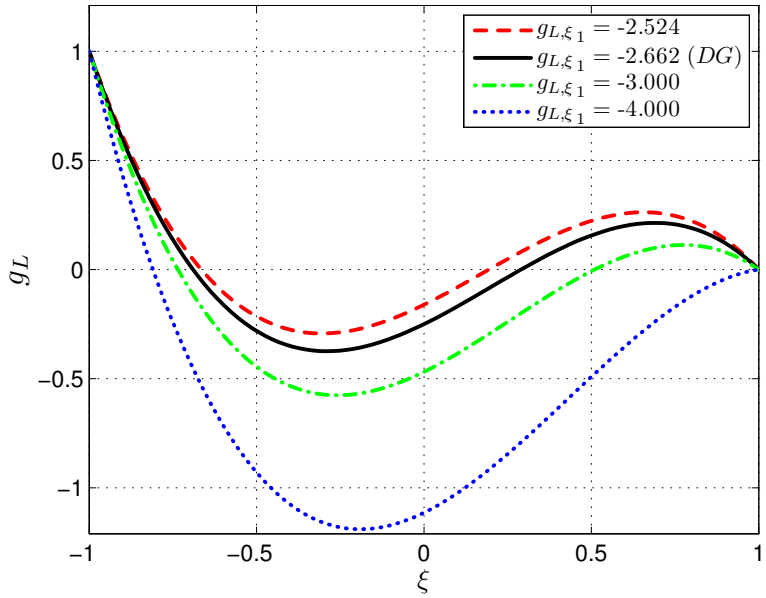


(a) $P = 2$: Maximum imaginary eigenvalue against the free parameter g_{L,ξ_1}.

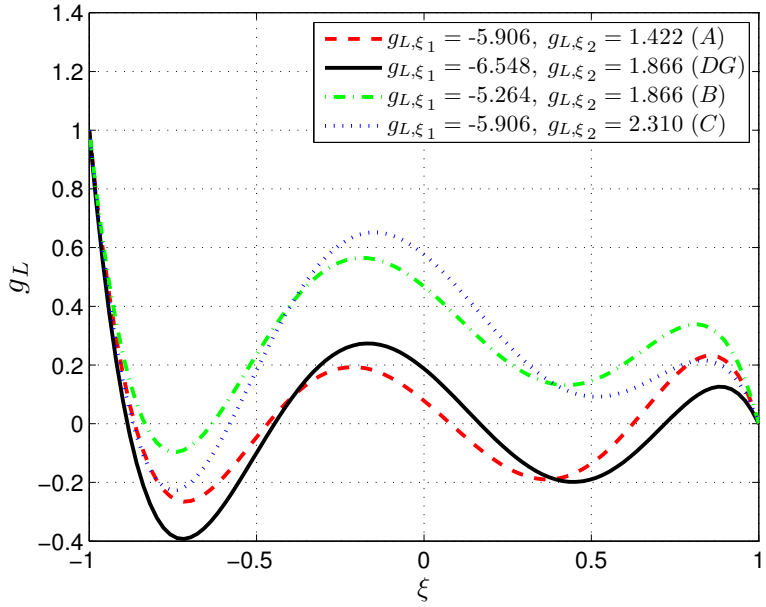


(b) $P = 4$: Region of stability in the space of the two parameters $(g_{L,\xi_1}, g_{L,\xi_2})$.

Figure 8.1: Region of stability in the parameter space, Eqn. (8.43), for the SFR family in the case of pure-advection.



(a) P=2



(b) P=4

Figure 8.2: Left-boundary correction function for the superconvergent family of schemes defined in Section 8.4.

Interestingly, the DG scheme is recovered as one member of this family, while the FR version of the SD scheme [69] is not. Also, for both $P = 2$ and $P = 4$, the DG scheme lies close to the stability boundary but not on it.

This space of linearly stable schemes can potentially be utilized as the feasible set for an optimization problem to maximize accuracy. The corresponding optimal scheme would be stable with respect to linear advection, super-convergent at the set of solution points, and yield the highest accuracy for steady-state problems with respect to the $\mathscr{L}^\delta$ semi-norm. For instance, we can pose the following problem for $P = 2$. Given centered fluxes, $\alpha = 1/2$, $\beta_{\text{LDG}} = 0$, $\tau_{\text{LDG}} = 0$, solution points $\boldsymbol{\xi}$, advection coefficient a, diffusion coefficient b, and element width h,

$$\begin{aligned} &\text{minimize} \quad \left\| e_u(g_{L,\xi_1}) \right\|_{\mathscr{L}^\delta(\Omega^\delta)} \\ &\text{subject to} \quad g_{L,\xi_1} \leq -2.524, \end{aligned} \tag{8.45}$$

where e_u is defined in Eqn. (7.18), g_{L,ξ_1} is the free parameter of the corresponding super-convergent family, and the constraint ensures linear stability with respect to pure advection. As a specific example, let $a = b = 1$, $k = 4\pi$, $k_0 = 2\pi$ in Eqn. (7.6), $h = 1/64$, and fix the solution points as the Gauss-Legendre points. Then, the solution to problem (8.45) lies on the stability boundary, $g^*_{L,\xi_1} = -2.524$. This optimal scheme produces 15% less error compared to DG while providing identical rate of convergence. In fact, if we relax the constraint of stability from pure-advection to the case of this advection-diffusion problem, then the optimal parameter value is attained at $g^{**}_{L,\xi_1} = -1.783$ which provides an 85% reduction in error!

In a similar manner, optimal SFR schemes can be derived for any advection-diffusion data. As the polynomial order increases, the dimension of the parameter space grows, and so should the improvement achieved via optimization as in Chapter 5. We will resume the discussion of SFR schemes in Chapter 10 which quantifies the accuracy of the boundary scheme identified above, relative to the DG scheme, in both the $\mathscr{L}^\delta(\Omega^\delta)$ semi-norm as well as the $\ell^2(\Omega^\delta)$ norm.

Chapter 9

Rate of convergence in the $\ell^2(\Omega^\delta)$ norm

In Chapter 7, we had seen that error in the steady-state numerical solution, Eqn. (7.18), satisfies a discrete Poisson equation with a forcing term. The elliptic nature of this governing equation complicates the derivation of closed-form estimates for the rate of convergence in the $\ell^2(\Omega^\delta)$ norm. Nevertheless, analytical estimates can still be derived using Fourier analysis in a manner similar to the linear dynamical system approach employed in Part I. In this chapter, we introduce the eigensolution to the steady-state error equation, Eqns. (7.20-7.21), under the special condition that the same correction function and upwinding coefficient is employed for all flux corrections. This allows us to express the point-wise error in terms of the eigenvalues and eigenvectors associated with the numerical differentiation operator $\boldsymbol{Q}$. The rate of convergence of steady-state error can then be obtained by utilizing results derived in Chapter 3. Along the way, we derive eigensolution estimates for the rate of convergence of the first and second order derivative operators as well. Combined with Theorems 8.1 and 8.2, this yields further results on the asymptotic behavior of the eigensystem of $\boldsymbol{Q}$.

9.1 Eigensolution convergence estimate for the first derivative operator

We begin by demonstrating that the rate of convergence of the first derivative operator $\boldsymbol{Q}(kh; g, \alpha)$, derived explicitly in Theorem 8.1, can also be expressed in terms of the asymptotic eigensystem associated with $\boldsymbol{Q}$. Towards this end, use the eigenvalue decomposition, Eqn. (2.18), and the expansion of the initial condition in the basis of eigenvectors, Eqn. (2.19), to rewrite the derivative error in the parent space, Eqn. (8.2), as

$$\begin{aligned}\boldsymbol{E}(kh) &= \left(\boldsymbol{W}(kh) ik\frac{h}{2}\boldsymbol{\Lambda}(kh)\boldsymbol{W}^{-1}(kh) - ik\frac{h}{2}\boldsymbol{I}\right)\boldsymbol{W}(kh)\boldsymbol{\beta}(kh)\\ &= ik\frac{h}{2}\boldsymbol{W}(kh)\left(\boldsymbol{\Lambda}(kh) - \boldsymbol{I}\right)\boldsymbol{\beta}(kh)\\ &= ik\frac{h}{2}\sum_{p=1}^{P+1}\left(\lambda_p(kh) - 1\right)\beta_p(kh)\boldsymbol{w}_p(kh),\end{aligned} \tag{9.1}$$

which shows that error in the first derivative operator is a sum of the errors in the eigenvalues, $\lambda_p(kh) - 1$, weighted by the modal weights β_p for the $P+1$ modes. Now, from Lemma 3.6, the asymptotic rate of convergence of the normalized physical eigenvalue, $q_1^\lambda > 0$, can be defined as

$$|\lambda_1(kh) - 1| \leq c_1^\lambda\,(kh)^{q_1^\lambda} \quad \text{as } kh \to 0, \tag{9.2}$$

for some constant c_1^λ. Similarly, from Lemma 3.5, the asymptotic rates of convergence of modal weights, $q_p^\beta > 0$, can be defined as

$$\left|\beta_p(kh) - \sqrt{P+1}\delta_{1p}\right| \leq c_p^\beta\,(kh)^{q_p^\beta}, \quad p = 1,\ldots,P+1, \text{ as } kh \to 0, \tag{9.3}$$

for some constants c_p^β, $p = 1,\ldots,P+1$. These definitions provide the first result of this chapter.

Lemma 9.1. *For an FR scheme of order P, the numerical first derivative converges to the exact first derivative, at the set of solution points, at the order $\sigma^{(1)}$, defined in Eqn. (8.26), where*

$$\sigma^{(1)} \geq \min_{p=2}^{P+1} \left(q_1^\lambda, q_p^\beta, q_p^{\lambda\beta} \right), \tag{9.4}$$

and q_1^λ, q_p^β are as defined in Eqns. (9.2), (9.3) and $q_p^{\lambda\beta} > 0$ for $p = 2, \ldots, P+1$, are the asymptotic rates of convergence of the product of eigenvalues and modal weights,

$$\left| \lambda_p(kh)\beta_p(kh) \right| \leq c_p^{\lambda\beta} (kh)^{q_p^{\lambda\beta}}, \quad p = 2, \ldots, P+1, \ \text{as } kh \to 0. \tag{9.5}$$

Proof. Plugging Eqns. (9.2), (9.3) in Eqn. (9.1) and using the triangle inequality, the Euclidean vector norm of derivative error can be bounded as follows.

$$\begin{aligned}
\|\boldsymbol{E}(kh; g, \alpha)\|_2 &\leq k\frac{h}{2} \sum_{p=1}^{P+1} |\lambda_p(kh) - 1||\beta_p(kh)| \\
&\leq k\frac{h}{2} \left[|\lambda_1(kh) - 1||\beta_1(kh)| + \sum_{p=2}^{P+1} (|\lambda_p(kh)||\beta_p(kh)| + |\beta_p(kh)|) \right] \\
&\leq k\frac{h}{2} \left[c_1^\lambda (kh)^{q_1^\lambda} \left(\sqrt{P+1} + c_1^\beta (kh)^{q_1^\beta} \right) + \sum_{p=2}^{P+1} \left(c_p^{\lambda\beta} (kh)^{q_p^{\lambda\beta}} + c_p^\beta (kh)^{q_p^\beta} \right) \right] \\
&\leq \max_{p=2}^{P+1} \left(\sqrt{P+1} c_1^\lambda, c_p^{(\lambda\beta)}, c_p^\beta \right) (kh)^{1 + \min_{p=2}^{P+1} \left(q_1^\lambda, q_p^{\lambda\beta}, q_p^\beta \right)} \ \text{as } kh \to 0.
\end{aligned} \tag{9.6}$$

Now, for any finite P, the equivalence of vector norms [54] implies that

$$\begin{aligned}
\max_{p=1}^{P+1} |E_p(kh; g, \alpha)| &= \|\boldsymbol{E}(kh; g, \alpha)\|_\infty \\
&\leq c_{\infty,2} \|\boldsymbol{E}(kh; g, \alpha)\|_2
\end{aligned} \tag{9.7}$$

for some constant $c_{\infty,2}$. The proof now follows from Eqn. (8.3). □

9.2 Eigensolution convergence estimate for the second derivative operator

In the case of the second derivative operator, the possibility of different correction functions and/or upwinding coefficients in the two corrections significantly impedes an eigensystem approach. This is because the composite operator, Eqn. (7.39), consists of a product of two matrices and the corresponding eigensystem is not analytically accessible. However, in the special case that both corrections are performed using the same g and α, i. e.

$$\boldsymbol{Q}(kh; g^{(2)}, \alpha^{(2)}) = \boldsymbol{Q}(kh; g^{(1)}, \alpha^{(1)}) = \boldsymbol{Q}(kh; g, \alpha), \tag{9.8}$$

the rate of convergence can be derived in a manner analogous to the previous section. Error in the parent space, Eqn. (8.33), can then be written as

$$\begin{aligned}
\boldsymbol{F}(kh) &= \left[\boldsymbol{W}(kh)\left(ik\frac{h}{2}\right)^2 \boldsymbol{\Lambda}^2(kh)\boldsymbol{W}^{-1}(kh) + \frac{h^2k^2}{4}\boldsymbol{I}\right]\boldsymbol{W}(kh)\boldsymbol{\beta}(kh) \\
&= -\frac{k^2h^2}{4}\boldsymbol{W}(kh)\left[\boldsymbol{\Lambda}^2(kh) - \boldsymbol{I}\right]\boldsymbol{\beta}(kh) \\
&= -\frac{k^2h^2}{4}\sum_{p=1}^{P+1}\left(\lambda_p^2(kh) - 1\right)\beta_p(kh)\boldsymbol{w}_p(kh).
\end{aligned} \tag{9.9}$$

This provides us with the next result of this chapter.

Lemma 9.2. *For an FR scheme of order P that uses the same correction function and upwinding coefficient for the two corrections, the numerical second derivative converges to the exact second derivative, at the set of solution points, at the order $\sigma^{(2)}$, defined in Eqn. (8.37), where*

$$\sigma^{(2)} \geq \min_{p=2}^{P+1}\left(q_1^\lambda, q_p^\beta, 2q_p^{\lambda\beta} - q_p^\beta\right), \tag{9.10}$$

and q_1^λ, q_p^β and $q_p^{\lambda\beta}$ are defined in Eqns. (9.2), (9.3) and (9.5) respectively.

Proof. The rate of convergence of λ_1^2 can be obtained from Eqn. (9.5) by observing that

$$|\lambda_1^2(kh) - 1| = |\lambda_1(kh) - 1||\lambda_1(kh) + 1| \leq 2c_1^\lambda (kh)^{q_1^\lambda}, \text{ as } kh \to 0. \tag{9.11}$$

Similarly, the rate of convergence of $\lambda^2\beta$ can be obtained from Eqns. (9.3) and (9.5),

$$|\lambda_p^2(kh)\beta_p(kh)| = \frac{|\lambda_p(kh)\beta_p(kh)|^2}{|\beta_p(kh)|} \leq \frac{(c_p^{\lambda\beta})^2}{c_p^\beta}(kh)^{2q_p^{\lambda\beta} - q_p^\beta}, \text{ as } kh \to 0. \tag{9.12}$$

Now, plugging Eqns. (9.11), (9.12) and (9.3) in Eqn. (9.9) and using the triangle inequality, the Euclidean norm of error in the second derivative can be bounded as follows.

$$\begin{aligned}
\|\boldsymbol{F}(kh; g, \alpha)\|_2 &\leq \frac{k^2h^2}{4} \sum_{p=1}^{P+1} |\lambda_p^2(kh) - 1||\beta_p(kh)| \\
&\leq \frac{k^2h^2}{4} \left[|\lambda_1^2(kh) - 1||\beta_1(kh)| + \sum_{p=2}^{P+1} (|\lambda_p^2(kh)||\beta_p(kh)| + |\beta_p(kh)|) \right] \\
&\leq \frac{k^2h^2}{4} \left[2c_1^\lambda (kh)^{q_1^\lambda} \left(\sqrt{P+1} + c_1^\beta (kh)^{q_1^\beta} \right) \right. \\
&\qquad \left. + \sum_{p=2}^{P+1} \left(\frac{(c_p^{\lambda\beta})^2}{c_p^\beta} (kh)^{2q_p^{\lambda\beta} - q_p^\beta} + c_p^\beta (kh)^{q_p^\beta} \right) \right] \\
&\leq \max_{p=2}^{P+1} \left(2\sqrt{P+1} c_1^\lambda, \frac{(c_p^{\lambda\beta})^2}{c_p^\beta}, c_p^\beta \right) (kh)^{2 + \min_{p=2}^{P+1} \left(q_1^\lambda, 2q_p^{\lambda\beta} - q_p^\beta, q_p^\beta \right)},
\end{aligned} \tag{9.13}$$

as $kh \to 0$. The proof now follows from Eqns. (9.7) and (8.34). □

The various rates of convergence involved in Lemmas 9.1 and 9.2 for the DG scheme are recorded in Tables 9.1 and 9.2 for polynomial orders $P = 0$ to 5. The tabulated values suggest that the rate of convergence for the product of eigenvalues and modal weights is

P	q_1^λ	$\min_{p=2}^{P+1} q_p^\beta$	$\min_{p=2}^{P+1} q_p^{\lambda\beta}$	$\min_{p=2}^{P+1} 2q_p^{\lambda\beta} - q_p^\beta$
0	1	-	-	-
1	3	2	1	0
2	5	3	2	1
3	7	4	3	2
4	9	5	4	3
5	11	6	5	4
$P>0$	$2P+1$	$P+1$	P	$P-1$

Table 9.1: Rates of convergence of the eigensystem components for the fully upwinded, $\alpha = 0$, DG scheme using Gauss-Legendre solution points - obtained via direct eigenvalue computation.

P	q_1^λ	$\min_{p=2}^{P+1} q_p^\beta$	$\min_{p=2}^{P+1} q_p^{\lambda\beta}$	$\min_{p=2}^{P+1} 2q_p^{\lambda\beta} - q_p^\beta$
0	2	-	-	-
1	2	1	1	1
2	6	4	3	2
3	6	3	3	2
4	10	6	5	4
5	10	5	5	4
$P>0$	$\begin{cases} 2P+2, & \text{even } P \\ 2P, & \text{odd } P \end{cases}$	$\begin{cases} P+2, & \text{even } P \\ P, & \text{odd } P \end{cases}$	$\begin{cases} P+1, & \text{even } P \\ P, & \text{odd } P \end{cases}$	$\begin{cases} P, & \text{even } P \\ P-1, & \text{odd } P \end{cases}$

Table 9.2: Rates of convergence of the eigensystem components for the centered, $\alpha = \frac{1}{2}$, DG scheme using Gauss-Legendre solution points - obtained via direct eigenvalue computation.

the most limiting among the quantities involved in Eqn. (9.4), i. e. for $P > 0$,

$$\begin{aligned}
\min_{p=2}^{P+1}\left(q_1^\lambda, q_p^\beta, q_p^{\lambda\beta}\right) &= \min_{p=2}^{P+1} q_p^{\lambda\beta} \\
&= \begin{cases} P, & \alpha = 0, \\ \begin{cases} P+1, & \text{even } P, \\ P, & \text{odd } P, \end{cases} & \alpha = \frac{1}{2} \end{cases} \\
&= \begin{cases} P+1, & \text{even } P \text{ and } \alpha = \frac{1}{2}, \\ P, & \text{otherwise.} \end{cases}
\end{aligned} \tag{9.14}$$

Comparing with Theorem 8.1, we see that the eigensystem approach provides results identical to the explicit derivation using power expansions, and with significantly reduced algebraic effort. However, Eqn. (9.14) is specific to the DG scheme. Such an eigensystem approach is lacking in that the expressions are not available in closed form, and could not have lead to the discovery of the SFR family.

Similarly, the tables also suggest that for $P > 0$,

$$\begin{aligned}
\min_{p=2}^{P+1}\left(q_1^\lambda, q_p^\beta, 2q_p^{\lambda\beta} - q_p^\beta\right) &= \min_{p=2}^{P+1} 2q_p^{\lambda\beta} - q_p^\beta \\
&= \begin{cases} P-1, & \alpha = 0, \\ \begin{cases} P, & \text{even } P, \\ P-1, & \text{odd } P, \end{cases} & \alpha = \frac{1}{2} \end{cases} \\
&= \begin{cases} P, & \text{even } P \text{ and } \alpha = \frac{1}{2}, \\ P-1, & \text{otherwise.} \end{cases}
\end{aligned} \tag{9.15}$$

Once again, comparing with Theorem 8.2, we see that the eigensystem approach provides results identical to the explicit derivation in Chapter 8. We summarize these two results in the following corollary.

Corollary 9.1. *For the DG scheme with polynomial order $P > 0$,*

$$q_p^{\lambda\beta} = \begin{cases} P+1, & \text{even } P \text{ and } \alpha = \frac{1}{2}, \\ P, & \text{otherwise,} \end{cases} \tag{9.16}$$

where $q_p^{\lambda\beta}$ is as defined in Eqn. (9.5), and

$$2q_p^{\lambda\beta} - q_p^{\beta} = \begin{cases} P, & \text{even } P \text{ and } \alpha = \frac{1}{2}, \\ P-1, & \text{otherwise,} \end{cases} \tag{9.17}$$

where q_p^{β} is as defined in Eqn. (9.3).

9.3 Eigensolution convergence estimate for the steady-state solution

We now derive the rate of convergence of error in the $\ell^2(\Omega^\delta)$ norm using the asymptotic eigensystem associated with the differentiation operator $Q(kh)$. As in Section 9.2, we only consider the special case when all three involved corrections use the same correction function g and upwinding coefficient α. The left hand side of Eqn. (7.20) can then be written in vector notation at the set of solution points in the nth element as

$$\begin{aligned} (\mathscr{L}^\delta e_u)(\boldsymbol{x_n}) &= \left(\frac{2}{h} a\boldsymbol{Q}(kh) - \frac{4}{h^2} b\boldsymbol{Q}^2(kh)\right) e_u(\boldsymbol{x}_n) \\ &= \boldsymbol{W}(kh)(iak\boldsymbol{\Lambda}(kh) + bk^2\boldsymbol{\Lambda}^2(kh))\boldsymbol{W}^{-1}(kh) e_u(\boldsymbol{x}_n) \\ &= \boldsymbol{W}(kh)\boldsymbol{\Pi}(k,h)\boldsymbol{W}^{-1}(kh) e_u(\boldsymbol{x}_n), \end{aligned} \tag{9.18}$$

where $\boldsymbol{\Pi}$ is the diagonal matrix of eigenvalues, $\pi_p(k,h) = iak\lambda(kh) + bk^2\lambda^2(kh)$, $p = 1,\ldots,P+1$, of the numerical advection-diffusion operator. Similarly, the right hand side of Eqn. (7.20) can be written in vector form using Eqns. (7.22), (8.1), (8.32), (9.1) and (9.9)

at the set of solution points in the nth element as

$$\begin{aligned}\left((\mathscr{L}-\mathscr{L}^\delta)\exp(ikx)\right)(\boldsymbol{x_n}) &= -a\frac{2}{h}\exp(ikx_n)\boldsymbol{E}(kh)+b\frac{4}{h^2}\exp(ikx_n)\boldsymbol{F}(kh)\\ &= -\exp(ikx_n)\boldsymbol{W}(kh)\left[\boldsymbol{\Pi}(k,h)-(iak+bk^2)\boldsymbol{I}\right]\boldsymbol{\beta}(kh).\end{aligned} \tag{9.19}$$

Substituting Eqns. (9.18) and (9.19) into (7.20), and noting that $\pi_p(k,h)\neq 0$, we get

$$\begin{aligned}e_u(\boldsymbol{x}_n) &= -\exp(ikx_n)\boldsymbol{W}(kh)\boldsymbol{\Pi}^{-1}(k,h)\left[\boldsymbol{\Pi}(k,h)-(iak+bk^2)\boldsymbol{I}\right]\boldsymbol{\beta}(kh)\\ &= \exp(ikx_n)\sum_{p=1}^{P+1}\left((iak+bk^2)\pi_p^{-1}(k,h)-1\right)\beta_p(kh)\boldsymbol{w}_p(kh),\end{aligned} \tag{9.20}$$

which brings us to the last result of this chapter.

Theorem 9.1. *For an FR scheme of order P that uses the same correction function and upwinding coefficient for all three corrections, the order of convergence of error in the $\ell^2(\Omega^\delta)$ norm, Eqn. (7.23), is given by μ,*

$$\lim_{h\to 0}\|e_u\|_{\ell^2(\Omega^\delta)} \leq C\,h^\mu \tag{9.21}$$

where $C(k,a,b,P)$ is some scheme, equation dependent constant and

$$\mu \geq \min_{p=2}^{P+1}\left(q_1^\lambda, q_p^\beta\right), \tag{9.22}$$

where q_1^λ and q_p^β are as defined in Eqns. (9.2) and (9.3) respectively.

Proof. From Chapter 2, we know that λ_p for $p\neq 1$ diverges in the limit of $kh\to 0$. Then, from $\pi_p(k,h)=iak\lambda(kh)+bk^2\lambda^2(kh)$, for a fixed k,

$$|\pi_p^{-1}(k,h)|\to 0,\quad p=2,\ldots,P+1,\ \text{ as } h\to 0. \tag{9.23}$$

Moreover, using Eqns. (9.2) and (9.11) the convergence rate of π_1 can be obtained as

$$\begin{aligned}|\pi_1(k,h) - (iak - bk^2)| &\leq |ak||\lambda_1(kh) - 1| + |bk^2||\lambda_1^2(kh) - 1| \\ &\leq c(k,a,b)h^{q_1^\lambda}, \text{ as } h \to 0, \end{aligned} \tag{9.24}$$

for some k-dependent constant c. This immediately yields

$$|1 - (iak - bk^2)\pi_1^{-1}(k,h)| \leq \frac{c(k,a,b)}{|(iak - bk^2)|} h^{q_1^\lambda}, \text{ as } h \to 0. \tag{9.25}$$

Then, from the triangle inequality, the Euclidean norm of pointwise error in the nth element can be bounded using Eqn. (9.20) as follows.

$$\begin{aligned}\|e_u(\boldsymbol{x}_n)\|_2 &\leq \sum_{p=1}^{P+1} \left|1 - (iak - bk^2)\pi_p^{-1}(k,h)\right| |\beta_p(kh)| \\ &\leq \left|1 - (iak - bk^2)\pi_1^{-1}(k,h)\right| |\beta_1(kh)| + \sum_{p=2}^{P+1} |\beta_p(kh)| \\ &\leq \frac{c(k,a,b)}{|(iak - bk^2)|} h^{q_1^\lambda} \left(\sqrt{P+1} + c_1^\beta (kh)^{q_1^\beta}\right) + \sum_{p=2}^{P+1} c_p^\beta (kh)^{q_p^\beta} \\ &\leq C'(k,a,b,P)\, h^{\min_{p=2}^{P+1}\left(q_1^\lambda, q_p^\beta\right)} \text{ as } h \to 0. \end{aligned} \tag{9.26}$$

where C' is some scheme and equation dependent constant. The proof is completed by observing that

$$\begin{aligned}\|e_u\|_{\ell^2(\Omega^\delta)} &= \left[\frac{1}{N(P+1)} \sum_{n=1}^{N} \sum_{p=1}^{P+1} \left|e_u(x_{n,p})\right|^2\right]^{1/2} \\ &= \left[\frac{1}{N(P+1)} \sum_{n=1}^{N} \|e_u(\boldsymbol{x}_n)\|_2^2\right]^{1/2} \\ &\leq \frac{C'(k,a,b,P)}{P+1}\, h^{\min_{p=2}^{P+1}\left(q_1^\lambda, q_p^\beta\right)} \text{ as } h \to 0. \end{aligned} \tag{9.27}$$

□

Theorem 9.1 suggests that the rate of convergence of error for steady-state problems, in the standard vector norm, is identical to the short-time rate of convergence for unsteady problems derived in Part I, Theorem 4.2. Remarkably, this rate is largely insensitive to the choice of correction function except for the case of even polynomial order and centered fluxes. This behavior is somewhat expected from Theorem 8.1 which showed that the choice of correction function is important only in special cases. This situation is in direct contrast with that of time-dependent problems where the long-time rate of convergence is highly sensitive to the choice of correction function [141]. Rates for the DG scheme are recorded in Table 1.1.

Consequently, the choice of an FR scheme for steady-state problems should be based on the actual magnitude of error rather than the rate of convergence which is largely invariant. From a practical point of view, the long-time super-convergence is of limited value for most aerospace and automotive design applications.

Chapter 10

Numerical experiments

This chapter provides numerical validation for the main results derived in Chapters 8 and 9. The discussion is divided in two parts. The first part conducts a simple test to verify the analytical estimates for rates of convergence of derivative operators derived in Theorems 8.1, 8.2, and their eigensystem counterparts in Lemmas 9.1, 9.2. Observed numerical values compare well with theoretical expectations. We see that the bound in Theorem 8.1 is tight for all polynomial orders and fluxes, while that in Theorem 8.2 is tight for all cases except one. The bounds obtained via eigensystem analysis are tight as well. In fact, even though Lemma 9.2 has strictly been proven only when the correction functions and upwiding coefficients are identical, it holds true for all cases when $P > 0$.

In the second part, we solve steady, forced, linear, advection-diffusion problems in 1-D and 2-D to verify the estimates for rate of convergence of error in the $\mathscr{L}^\delta(\Omega^\delta)$ semi-norm, Theorem 8.3, as well as the $\ell^2(\Omega^\delta)$ norm, Theorem 9.1. Once again, we find that the numerical values compare well with theoretical expectations. The bounds expressed in Theorem 9.1 are tight for all cases, while those in Theorem 8.3 are tight for all cases except one.

We end this chapter with a brief discussion regarding the applicability of such theoretical results to general multi-dimensional, non-linear problems such as the Navier-Stokes equations, as well as non-tensor-product finite elements such as simplexes.

10.1 Rate of convergence for derivative operators

We begin by employing the first and second order FR operators to differentiate the simple sinusoid, $u(x) = \cos(2\pi x)$, in $[0, 1]$. The left and right correction functions are chosen to be the right and left Radau polynomials, respectively, in order to recover the DG scheme. The solution points in each element are chosen to be the Gauss-Legendre points in the parent space. Rates are obtained for $P = 0$ to 5 for both one-sided and centered fluxes. For the first derivative, the upwind flux is given by $\alpha = 0$ and the centered flux by $\alpha = \frac{1}{2}$. For the second derivative, the one-sided flux is given by $\alpha^{(1)} = 0$, $\alpha^{(2)} = 1$ in the isotropic case and by $\alpha^{(1)} = 0$, $\alpha^{(2)} = 0$ in the anisotropic case; the centered flux is given by $\alpha^{(1)} = \alpha^{(2)} = \frac{1}{2}$. In terms of LDG, the one-sided flux is given by $\beta_{\text{LDG}} = \frac{1}{2}$, $\tau_{\text{LDG}} = 0$ in the isotropic case and by $\beta_{\text{LDG}} = -\frac{1}{2}$, $\tau_{\text{LDG}} = 0$ in the anisotropic case; the centered flux is given by $\beta_{\text{LDG}} = 0$, $\tau_{\text{LDG}} = 0$. Error is measured using the $\ell^2(\Omega^\delta)$ norm for both the first and second order derivatives. The rate of convergence is calculated numerically by using two uniform grids with spacings $h_1 = 1/32$ and $h_2 = 1/64$. The reported values have been rounded to 3 decimal places.

Table 10.1 records the result for the first derivative operator. We see that the numerical convergence rates are in exact agreement with bounds on the theoretical convergence rates stated in Theorem 8.1 and Lemma 9.1. Results for the second derivative are tabulated in Table 10.2 for isotropic operators and Table 10.3 for an anisotropic operator. We see that the eigensystem estimate, Lemma 9.2, is tight for all cases except for $P = 0$ with the one-sided, isotropic flux. This is not surprising since Lemma 9.2 has been proven to hold only when the same upwinding coefficient is used for the two corrections. In fact, the exactness of this estimate for all cases with $P > 0$ is remarkable! Similarly, the bound in Theorem 8.2 is tight for all cases except for $P = 1$ with centered fluxes. A dedicated treatment of this edge case, in a manner similar to Lemma 8.5, would yield an exact match.

Continuing our discussion from Chapter 8, we can also verify convergence for SFR schemes other than DG. For instance, Tables 10.4 and 10.5 record the error in differentiation for $P = 2$ using the SFR scheme with $g_{L,\xi_1} = -2.524$. The corresponding values for DG are also recorded for comparison. We see that this SFR scheme provides not only the same rate of convergence as DG but also higher accuracy for all choices of flux for both

	Theorem 8.1		Lemma 9.1		Grid-convergence study	
P	$\alpha = 0$	$\alpha = \frac{1}{2}$	$\alpha = 0$	$\alpha = \frac{1}{2}$	$\alpha = 0$	$\alpha = \frac{1}{2}$
0	1	2	1	2	0.999	1.998
1	1	1	1	1	0.997	0.994
2	2	3	2	3	2.004	3.004
3	3	3	3	3	2.998	2.995
4	4	5	4	5	4.002	5.000
5	5	5	5	5	4.999	4.995

Table 10.1: Rate of convergence for the first derivative operator, $\sigma^{(1)}$, in $\ell^2(\Omega^\delta)$ norm using DG.

	Theorem 8.2		Lemma 9.2		Grid-convergence study	
	$\alpha^{(1)} = 0$	$\alpha^{(1)} = \frac{1}{2}$	$\alpha^{(1)} = 0$	$\alpha^{(1)} = \frac{1}{2}$	$\alpha^{(1)} = 0$	$\alpha^{(1)} = \frac{1}{2}$
P	$\alpha^{(2)} = 1$	$\alpha^{(2)} = \frac{1}{2}$	$\alpha^{(2)} = 1$	$\alpha^{(2)} = \frac{1}{2}$	$\alpha^{(2)} = 1$	$\alpha^{(2)} = \frac{1}{2}$
0	2	2	1	2	1.999	1.994
1	0	0	0	1	-0.003	0.988
2	1	2	1	2	1.005	2.014
3	2	2	2	2	1.998	2.041
4	3	4	3	4	3.003	4.002
5	4	4	4	4	3.998	4.010

Table 10.2: Rate of convergence for the second derivative operator, $\sigma^{(2)}$, in $\ell^2(\Omega^\delta)$ norm, for isotropic second order operators using DG.

	Theorem 8.2	Lemma 9.2	Grid-convergence study
	$\alpha^{(1)} = 0$	$\alpha^{(1)} = 0$	$\alpha^{(1)} = 0$
P	$\alpha^{(2)} = 0$	$\alpha^{(2)} = 0$	$\alpha^{(2)} = 0$
0	1	1	0.997
1	0	0	-0.008
2	1	1	1.014
3	2	2	1.996
4	3	3	3.005
5	4	4	3.997

Table 10.3: Rate of convergence for the second derivative operator, $\sigma^{(2)}$, in $\ell^2(\Omega^\delta)$ norm, for an anisotropic second order operator using DG.

the first and second derivative. This shows that the right and left Radau polynomials are not the best correction functions for computing derivatives through reconstruction. Simple numerical optimizations can be conducted to identify members of the SFR family that minimize error in differentiation, or equivalently, the $\mathscr{L}^{\delta}(\Omega^{\delta})$ semi-norm of error. However, for most practical purposes, we are interested in the $\ell^2(\Omega^{\delta})$ norm. This brings us to the next section where we see that the advantage of using this SFR scheme over DG is conditional on the type of interface flux.

	DG		SFR: $g_{L,\xi_1} = -2.524$	
h	$\alpha = 0$	$\alpha = \frac{1}{2}$	$\alpha = 0$	$\alpha = \frac{1}{2}$
1/32	5.437×10^{-3}	4.002×10^{-4}	4.803×10^{-3}	3.381×10^{-4}
1/64	1.355×10^{-3}	4.990×10^{-5}	1.195×10^{-3}	4.204×10^{-5}
Order	2.004	3.004	2.007	3.008

Table 10.4: $\ell^2(\Omega^{\delta})$ norm of error in the first derivative operator for $P = 2$ using DG and another SFR scheme.

h	$\alpha^{(1)} = 0$ $\alpha^{(2)} = 1$	$\alpha^{(1)} = \frac{1}{2}$ $\alpha^{(2)} = \frac{1}{2}$	$\alpha^{(1)} = 0$ $\alpha^{(2)} = 0$
	DG		
1/32	1.591	8.908×10^{-2}	1.604
1/64	7.929×10^{-1}	2.206×10^{-2}	7.945×10^{-1}
Order	1.005	2.014	1.014
	SFR: $g_{L,\xi_1} = -2.524$		
1/32	1.337	7.561×10^{-2}	1.352
1/64	6.648×10^{-1}	1.861×10^{-2}	6.666×10^{-1}
Order	1.009	2.023	1.020

Table 10.5: $\ell^2(\Omega^{\delta})$ norm of error in the second derivative operator for $P = 2$ using DG and another SFR scheme.

10.2 Rate of convergence for steady-state solution

Having established the validity of theoretical results for derivative operators, we now return to the steady-state problem in Eqns. (7.1 - 7.3). For the experiments that follow, the advection and diffusion coefficients are set to unity, $a = 1$, $b = 1$, the wavenumber for the source term is $k = 4\pi$ and that for the initial condition is $k_0 = 2\pi$. We are interested in determining the rate of convergence of error in both measures, viz. the $\mathscr{L}^{\delta}(\Omega^{\delta})$ semi-norm and the $\ell^2(\Omega^{\delta})$ norm. Eqns. (7.1 - 7.3) are solved using FR in $[0,1]$ for $P = 0$ to 5 with one-sided, isotropic ($\alpha = 0, \alpha^{(1)} = 0, \alpha^{(2)} = 1$) as well as fully centered ($\alpha = 1/2, \alpha^{(1)} = 1/2, \alpha^{(2)} = 1/2$) fluxes. The corresponding LDG fluxes are as mentioned in the previous section. Once again, Gauss-Legendre points are chosen to be the solution points and the correction functions are chosen to recover the DG scheme. The numerical solution is marched in time using a Backward-Euler time stepping scheme till the equation residual falls below a tolerance of 10^{-10}. The rate of convergence is calculated using two grids with spacings $h_1 = 1/1024$ and $h_2 = 1/2048$ for $P = 0, 1$, $h_1 = 1/64$ and $h_2 = 1/128$ for $P = 2, 3$, and $h_1 = 1/16$ and $h_2 = 1/32$ for $P = 4, 5$ to avoid saturation due to finite precision.

Table 10.6 records the rate of convergence of error in the $\mathscr{L}^{\delta}(\Omega^{\delta})$ semi-norm. Alongside, the analytical expectation from Theorem 8.2 is included for comparison. We see that the estimate matches excellently with the results of the grid convergence study for all polynomial orders and fluxes except for the case of $P = 1$ with centered flux. This is the same case for which the bound in Theorem 8.2 is not tight as discussed in the previous section (c. f. Table 10.2). A dedicated treatment of this edge case, in a manner similar to Lemma 8.5, would yield an exact match.

Table 10.7 records the rate of convergence of the numerical solution in the $\ell^2(\Omega^{\delta})$ norm. Alongside, the analytical expectation from Theorem 9.1 is included for comparison. We see that theoretical expectation matches excellently with the results of the grid convergence study for all polynomial orders. As in the previous section, even though our proof for Theorem 9.1 requires that the same upwinding coefficient be used for all three corrections, the theoretical predictions hold for both choices of interface fluxes.

As an empirical exercise, we also examine the rate of convergence of the first and second derivative of the solution in the $\ell^2(\Omega^{\delta})$ norm. The observed values are recorded in

	Theorem 8.2		Grid convergence study	
P	One-sided	Centered	One-sided	Centered
0	1	2	0.980	2.000
1	0	0	0.000	1.000
2	1	2	1.006	2.014
3	2	2	1.996	2.041
4	3	4	3.055	4.017
5	4	4	3.959	4.092

Table 10.6: Rate of convergence of error in the $\mathscr{L}^{\delta}(\Omega^{\delta})$ semi-norm for the steady, forced, advection-diffusion problem using DG.

	Theorem 9.1		Grid-convergence study	
P	One-sided	Centered	One-sided	Centered
0	1	2	0.981	2.000
1	2	1	2.000	1.000
2	3	4	2.999	4.007
3	4	3	3.999	2.999
4	5	6	4.993	6.085
5	6	5	5.983	4.968

Table 10.7: Rate of convergence of error in the $\ell^2(\Omega^{\delta})$ norm for the steady, forced, advection-diffusion problem using DG.

	Grid-convergence study			
	1st derivative		2nd derivative	
P	One-sided	Centered	One-sided	Centered
0	1.000	2.000	1.000	2.000
1	2.000	1.000	2.000	1.000
2	2.999	4.004	2.999	4.004
3	3.999	2.998	3.998	2.998
4	4.994	6.063	4.994	6.062
5	5.991	4.961	5.892	4.957

Table 10.8: Rate of convergence for the first and second derivative in the $\ell^2(\Omega^{\delta})$ norm for the steady, forced, advection-diffusion problem using DG.

Table 10.8. We see that both the first and the second derivative converge at the same rate as the solution. This is a salient feature of FR lent by the recursive nature of the correction procedure which couples all the elements in the domain. The resulting estimates for the derivatives are therefore global and not element-local.

We finish this section by verifying convergence for SFR schemes other than DG. Following the example from the previous section, Table 10.9 records the $\mathscr{L}^\delta(\Omega^\delta)$ semi-norm of error for $P = 2$ using the SFR scheme with $g_{L,\xi_1} = -2.524$. The corresponding values for DG are recorded for comparison. As expected from Theorem 8.3, this SFR scheme provides the same rate of convergence as DG. Moreover, as discussed in the previous section, the scheme provides almost 20% higher accuracy for both one-sided and centered fluxes.

	DG		SFR: $g_{L,\xi_1} = -2.524$	
h	One-sided	Centered	One-sided	Centered
1/64	9.008	5.039×10^{-1}	7.573	4.277×10^{-1}
1/128	4.485	1.248×10^{-1}	3.761	1.053×10^{-1}
Order	1.006	2.014	1.010	2.023

Table 10.9: $\mathscr{L}^\delta(\Omega^\delta)$ norm of error in the steady-state numerical solution for $P = 2$ using DG and another SFR scheme.

	DG		SFR: $g_{L,\xi_1} = -2.524$	
h	One-sided	Centered	One-sided	Centered
1/64	2.748×10^{-5}	2.038×10^{-6}	2.424×10^{-5}	2.500×10^{-6}
1/128	3.437×10^{-6}	1.267×10^{-7}	3.031×10^{-6}	1.558×10^{-7}
Order	2.999	4.007	2.999	4.004

Table 10.10: $\ell^2(\Omega^\delta)$ norm of error in the steady-state numerical solution for $P = 2$ using DG and another SFR scheme.

We can also compare the $\ell^2(\Omega^\delta)$ norm of error for the two schemes. The error values and rates of convergence are recorded in Table 10.10. Once again, we see that this SFR scheme provides the same rate of convergence as DG. In particular, the super-convergence

for centered fluxes is exhibited as well. However, the error values for the two schemes are not directly comparable. The SFR scheme provides higher accuracy than DG for one-sided fluxes, while the DG scheme is better for centered fluxes. Returning to the discussion at the end of Chapter 8, we see that the DG scheme is not necessarily optimal for steady-state problems, especially since most CFD practitioners employ one-sided interface fluxes. However, optimality of a certain scheme within the SFR family is dependent on the specifics of the conservation law, especially the ratio of advection and diffusion coefficients. Coupled with the observation that the improvement in accuracy is not exceptional, the DG scheme still remains a reliable choice for design problems using discontinuous finite element methods.

10.3 Applicability to general multi-dimensional, non-linear, vector conservation laws, and non tensor-product elements

The results in this second part of the dissertation have been derived under several simplifying assumptions to facilitate analysis. The model problem in Chapter 7 specifies a scalar conservation law with one-dimensional variations, linear fluxes and periodic boundary conditions. This section provides a brief discussion regarding the relaxation of these assumptions to determine the scope of such theoretical expectations.

10.3.1 Tensor product elements in multiple dimensions

The extension of FR to an arbitrary number of spatial dimensions can be readily achieved through the use of tensor-product elements. Owing to the discontinuous nature of the formulation, the representation of geometry is essentially decoupled from that of the numerical solution. Specifically, the type of element and its completeness are chosen based on the impact of the boundary condition on the physics of the problem, and the polynomial order of the solution is chosen based on desired accuracy. This independence allows for

very general unstructured meshes using quadrilaterals in 2-D and hexes in 3-D. Each element in the physical space is mapped to a bi-unit hypercube in the parent space, and the generalized reconstruction procedure involves flux corrections along lines of this hypercube. Consequently, the theoretical results derived in the previous chapters port directly to tensor-product elements.

In order to provide numerical evidence in support of this generalization, consider the two-dimensional analogue of the model problem, Eqns. (7.1 - 7.3),

$$\left(\frac{\partial \hat{u}}{\partial t} + (\mathscr{L}\hat{u})\right)(x,y,t) = \left(i(a_x+a_y)k + 2bk^2\right)\exp\left(ik(x+y)\right), \quad (x,y) \in (0,1)^2,\ t > 0, \tag{10.1}$$

$$\hat{u}(x,y,0) = \exp\left(ik_0(x+y)\right), \quad (x,y) \in [0,1]^2, \tag{10.2}$$

$$\mathscr{L} := a_x\frac{\partial}{\partial x} + a_y\frac{\partial}{\partial y} - b\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right), \tag{10.3}$$

where $k = 4\pi$ is the wavenumber of the forcing term, $k_0 = 2\pi$ is that of the initial condition, and $a_x = a_y = 1, b = 1$ are advection and diffusion coefficients respectively. The boundary conditions are periodic as before. The analytical solution to Eqns. (10.1 - 10.3) for $(x,y) \in [0,1]^2$, $t \to \infty$ is given by

$$\hat{u}(x,y) = \exp\left(ik(x+y)\right). \tag{10.4}$$

Eqns. (10.1 - 10.3) are solved using FR on quadrilaterals for $P = 0$ to 5 with one-sided, isotropic as well as fully centered fluxes. The solution points are chosen as the tensor-product of Gauss-Legendre points, and the 1-D correction functions are chosen to recover the DG scheme. The numerical solution is marched in time using a Backward-Euler time stepping scheme till the equation residual falls below a tolerance of 10^{-10}. The rate of convergence is calculated using uniform grids, $h_x = h_y = h$, with spacings $h_1 = 1/256$ and $h_2 = 1/512$ for $P = 0$, and $h_1 = 1/16$ and $h_2 = 1/32$ for $P = 1, \ldots, 5$.

Table 10.11 records the rate of convergence of the numerical solution in the $\ell^2(\Omega^\delta)$ norm. Alongside, the analytical expectation from Theorem 9.1 is included for comparison. As expected, the results are similar to those in Table 10.7 for the 1-D problem.

P	Theorem 9.1		Grid-convergence study	
	One-sided	Centered	One-sided	Centered
0	1	2	0.918	2.001
1	2	1	1.966	1.009
2	3	4	2.964	4.099
3	4	3	3.981	2.989
4	5	6	4.984	6.082
5	6	5	5.988	4.970

Table 10.11: Rate of convergence of error in the $\ell^2(\Omega^\delta)$ norm for the 2-D steady, forced, advection-diffusion problem using DG.

10.3.2 Triangles, tetrahedra and prisms

The extension of FR to non-tensor-product elements requires the prescription of appropriate correction fields [32, 153]. Experiments on triangles have been reported by Castonguay in [30]. Therein, the author solved the 2-D advection equation on a periodic bi-unit square with unit wavepseeds in both spatial coordinates. The initial condition was sinusoidal and the numerical solution was marched for half a time period using fully upwinded ESFR for spatial derivatives and RK54 for temporal ones. Error with respect to the $\ell^2(\Omega^\delta)$ norm was calculated on a sequence of regular as well as irregular grids. The calculated rates of convergence showed that the scheme with polynomial order P converged at the order $P+1$. A similar numerical investigation was conducted by Williams [150] for the linear advection-diffusion equation with advection wavespeeds $a_x = \cos\pi/6$, $a_y = \sin\pi/6$ and diffusion coefficient $b = 0.1$. Therein, the boundary conditions were periodic, the initial condition was a sinusoidal tensor-product, and the interface fluxes were one-sided. As before, the solution was marched in time for half a time period. The grid convergence study was performed with respect to the functional $L^2(\Omega)$ norm. Once again, calculated rates of convergence showed that the scheme with polynomial order P converged at the order $P+1$. Similar observations have been reported for the case of prisms [30] and tetrahedra [150] as well.

These results are all in line with the theoretical expectations derived in the last several chapters. Since the solution was marched for only one half of a time period, the recorded

rate is indeed the short-time rate of convergence (c. f. Table 1.1 and Fig. 4.3). Moreover, from Theorem 9.1, this short-time rate is identical to the rate of convergence for linear steady-state problems (c. f. Table 9.1.) Hence, the results of this part of the dissertation seem to extend to non-tensor-product elements as well.

10.3.3 Non-linear vector conservation laws in multiple dimensions

The application of FR to conservation laws with non-linear fluxes is marred with issues related to aliasing [76]. The problem is particularly severe in the case of hyperbolic problems that admit distributions as weak solutions. In such cases, errors arising due to aliasing evolve into numerical instabilities through non-linear steepening. A mechanism for controlling such numerical divergence is proposed in the next part of the dissertation.

From the point of view of convergence, Godunov's theorem [53] ensures that, in the presence of discontinuities, finite element schemes can locally be at the most first order accurate. Hence, in the context of the present chapter, we can restrict our attention to non-linear conservation laws with smooth steady-state solutions. A canonical example of this sort is given by Couette flow [148], which, in 3D, deals with fluid flow between two infinite plates translating at unequal fixed velocities. Owing to symmetry in the plane of the plates, the Navier-Stokes equations, Eqns. (5.18 - 5.24), admit solutions with closed-form expressions for variations in the wall-normal direction. A detailed description of the problem including the boundary conditions and analytical solution can be found in [30].

Grid convergence studies on Couette flow have been conducted using structured as well as unstructured grids in 2D using quadrilaterals and triangles [30, 110], and in 3-D using tetrahedra [150]. In all reported experiments, the fluxes were one-sided, and the numerical solution was marched in time until the equation residual converged to machine precision. As in the previous section, the order of convergence in the L^2 norm for a scheme with polynomial order P was found to be $P+1$. Once again, these results are exactly as expected theoretically from Theorem 9.1 and Table 9.1. This provides numerical evidence that the analysis, performed under the simplifying assumptions mentioned before, indeed applies to more general non-linear vector conservation laws in multiple dimensions and non-tensor-product elements.

Part III

Stability theory and spectral filtering for non-linear, time-dependent problems

Chapter 11

Introduction

The stability of discontinuous finite-element schemes, and particularly the Discontinuous Galerkin (DG) method, has been studied extensively over the last two decades. Several notions of stability have been introduced over time, often to facilitate analysis. Consider the following hyperbolic equation that specifies the conservation of a scalar field $u : \mathbb{R} \mapsto \mathbb{R}$ in 1-D,

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \quad x \in \Omega \subset \mathbb{R},\, t > 0, \tag{11.1}$$

$$u(x,0) = u_0(x). \tag{11.2}$$

where $f : \mathbb{R} \mapsto \mathbb{R}$ is the flux and u_0 prescribes the initial data on the domain Ω. We are interested in physically motivated problems such that the existence and uniqueness of the solution to Eqns. (11.1 - 11.2) for $t \geq 0$ can be readily assumed. Moreover, we expect the solution to remain bounded in an appropriate functional or discrete norm, $\|u\|(t) < \infty$, $t \geq 0$. Consequently, any numerical scheme that approximates the solution to Eqns. (11.1 - 11.2) is also required to be stable in the same sense, $\|u^\delta\|(t) < \infty$, $t \geq 0$, where u^δ denotes the numerical solution. Note that this definition of stability is weaker than Lax's stability criterion. Nevertheless, such a guarantee of stability is of immense practical value in that the computation is guarded against numerical divergence.

In the case of linear conservation laws, stability can be easily proven [62] by showing that the L^2 norm of the numerical solution is monotonically non-increasing for consistent

choices of the interface flux. Alternatively, von Neumann analysis of the form conducted in Chapter 2 can be utilized to show L^2 stability on account of Parceval's theorem. However, for a general non-linear conservation law, stability of the numerical solution cannot be guaranteed, especially if the governing equations admit a solution with jump discontinuities. One possible strategy to handle such cases is to use a deformable mesh in conjunction with adaptive finite-element algorithms [59, 107, 133]. However, even though these algorithms appear to be robust for the reported simulations of 2-D and 3-D Euler equations [46], there is no rigorous assurance of stability. The alternative strategy to tackle non-linear fluxes involves the intelligent addition of artificial dissipation.

The simultaneous requirements of stability and accuracy naturally fork two broad ideas. The first one suggests the addition of a variable amount of dissipation in the entire domain. An instance of this approach is the use of limiters [136, 16, 22]. The idea was initially proposed in the context of finite-volume schemes to capture shock discontinuities in supersonic inviscid fluid flow [136, 75]. The action of the limiter is to detect local extrema and limit their growth by comparing the solution and/or derivative in an element with the range of values in the neighboring elements. Hierarchical limiting strategies [82, 85] have been proposed more recently where derivatives are progressively limited, and it has been argued that the limited solution retains the highest possible accuracy. However, such limiting procedures are sensitive to the order and extent of limiting and may require special treatment to handle negative pressure values for the Euler equations. Additionally, repeated action of the limiter can lead to diffusion of sharp features that would not have necessarily resulted in shocks.

Another instance of this approach is the incorporation of an explicit filtering operation applied periodically to dampen high frequency components that would potentially result in numerical instabilities. High-order finite-difference compact schemes have been used in conjunction with anti-aliasing filters [144, 116] to suppress instabilities arising from mesh non-uniformity. The action of such filters is similar to that of low pass filters used in Large Eddy Simulation [138, 17, 111] which serve as a sink of energy for parts of the wavenumber spectrum that cannot be accurately resolved. Filters designed explicitly for shock-capturing have been proposed [2, 19, 135] as well. In the case of finite-element schemes, spectral viscosity can be added in the Fourier space [128] or a convenient Jacobi

space such as those of Legendre polynomials [61, 60]. Such spectral viscosity can be equivalently posed as the spectral expansion of differential operators in the physical space [94]. These approaches benefit from rigorous studies [137, 57, 55] which prove that multiplication with an appropriate filter kernel can accelerate the convergence of Fourier expansions of piecewise discontinuous functions. Note that the use of periodic filtering, where the filter is applied after every few timesteps, differs from the approach where filtering is used once as a post-processing operation [96] to reduce errors and improve convergence.

The other approach towards adding artificial dissipation proceeds by first detecting numerical instabilities through a sensor and then selectively attending to troubled elements [160]. The detection of discontinuities partially relieves the subsequent filtering or limiting action from the constraints of preserving order of accuracy, as smooth regions of the solution remain unchanged. Detectors have been proposed that utilize super-convergence of DG-type methods at the outflow boundaries of elements [83], or the expected rate of decay of modal coefficients in Legendre expansions [103]. The treatment of troubled cells on structured and unstructured grids has been carried out through WENO reconstruction [23, 120] of the solution polynomial using the neighboring polynomials, or through the explicit addition of artificial viscosity [103, 11] in the physical space. The addition of artificial diffusivity for shock capturing has also been proposed for high-order finite-difference schemes [43, 80].

In regard to flow physics, a consistent shock capturing scheme must not only satisfy the weak form of the conservation equation, but also respect entropy behavior dictated by the second law of thermodynamics. Entropy-conservative as well as entropy-stable second-order finite-difference approximations have been derived by Tadmor for both semi-discrete [126] and fully-discrete [127] representations. Such entropy-stability theory can be used to distinguish between shock discontinuities and rarefaction waves within a single computational cell. Non-linear entropy-stability can be ensured for higher order difference schemes [125] through addition of suitable diffusion operators. More recently, entropy-stable ENO schemes [48] as well as arbitrary order spectral collocation schemes [25] have also been obtained that provide entropy-stability of the compressible Navier Stokes equations.

In the case of Flux Reconstruction (FR), investigations regarding stability have been

largely focused on linear fluxes. Huynh [69], in his original paper, utilized von Neumann analysis to demonstrate that the formulation is linearly stable for some very specific choices of correction functions. These functions recovered DG, Spectral Difference (SD), and Huynh's g_2 scheme respectively. A more elaborate study was conducted by Vincent et al. [142] who proved the formulation is linearly stable in a broken Sobolev-type energy norm for a one-parameter family of correction functions designated as Energy Stable FR (ESFR). This result was expanded in [143] to identify a multi-parameter, linearly stable family where the number of parameters depend on the polynomial order. Remarkably, even this expanded family is not exhaustive as it does not recover the OFR schemes derived in Chapter 5.

As in the case of DG, the FR formulation is not guaranteed to be stable for non-linear conservation laws. In [76], Jameson et al. showed that the ESFR family of schemes may not be stable for non-linear fluxes due to aliasing errors. These errors arise since the flux of the numerical solution, $f(u^\delta)$, is not contained within the polynomial basis $\mathbb{P}_P$ in which the approximated flux, f^δ, resides. Consequently, the location of solution points becomes critical when the numerical flux is computed using a collocation projection. In this regard, Jameson et al. recommend the use of Gauss-Legendre points for tensor-product elements. Quadrature points for triangles and tetrahedra [150] have also been proposed. Alternatively, Witherden [156] advocates the use of orthogonal / L^2 projections for determining the numerical flux polynomial.

In this third part of the dissertation, we build upon the result of Jameson et al. [76] to obtain energy bounds for the numerical solution via FR and show that aliasing errors may not vanish even in the asymptotic limit on polynomial order, $P \to \infty$, if the analytical flux function is not sufficiently smooth. We then prove that the addition of adequate artificial dissipation can ensure non-linear stability for any polynomial order. We pose this addition of artificial viscosity as a Fourier filtering operation which is implemented in the physical space via convolution. A notable departure from previous filtering strategies is in the incorporation of edge information from neighboring elements in order to control inter-element discontinuities. Moreover, the filter is selectively applied to troubled cells as indicated by a sensor based on the concentration method [50, 118]. Numerous numerical tests have been performed to validate the theoretical results.

This part is formatted as follows. Chapter 12 is concerned with the theoretical aspects of non-linear stability. We begin by deriving energy bounds on the numerical solution for a general flux, and then derive sufficient conditions for stability through addition of artificial dissipation. Chapter 13 describes the formulation of convolution filters in 1-D and 2-D on tensor-product elements, and provides a brief overview of the concentration sensor used for detecting discontinuities. Finally, in Chapter 14, we demonstrate performance of the sensor-filter combination in several test cases in 1-D and 2-D. The 1-D cases comprise linear advection, Burger's equation, Sod's shock-tube and the Shu-Osher shock-disturbance interaction [119]. In 2-D, we show results for rigid body rotation, the symmetric Riemann problem, double Mach reflection case of Woodward and Colella [157], and transonic flow over a NACA0012 airfoil at an angle of attack.

The content of this part if largely drawn from the paper:

- K Asthana, M R López-Morales, A Jameson: Non-linear stabilization of high-order Flux Reconstruction schemes via Fourier-spectral filtering. Journal of Computational Physics, 2014.

Chapter 12

Energy bounds and stabilization for non-linear conservation laws

This chapter investigates the stability of FR for the 1-D, non-linear, scalar conservation law introduced in Eqns. (11.1 - 11.2). In order to focus on the effect of non-linear fluxes, we restrict our attention to the ESFR [142] family which is provably stable for linear fluxes. For this family, Jameson et al. [76] have shown that rate of growth of a certain energy norm of the numerical solution is dependent on aliasing errors. These errors arise since the flux of the numerical solution, $f(u^\delta)$, is not contained within the polynomial basis $\mathbb{P}_P$ in which the approximated flux, f^δ, resides. We build upon their work to show that stability of the solution depends on the polynomial order P of the scheme and the degree of local continuity of $f(u^\delta)$. For the scheme with $P = 0$, the solution is guaranteed to be stable for all possible flux functions. However, for $P \geq 1$, stability cannot be guaranteed. In such cases, the rate of growth of energy vanishes at the rate h^σ where h is a measure of the grid spacing and σ depends on the degree of local continuity of $f(u^\delta)$. If the flux is smooth, then $\sigma = P + 1$, and the growth of energy also vanishes in the limit of $P \to \infty$.

We then prove that the addition of adequate artificial dissipation can ensure non-linear stability for any polynomial order, in the asymptotic limit, provided that the flux has at least square integrable first derivatives, locally in each element. This artificial dissipation vanishes with grid spacing at the rate h^a where $a < 2$, and with polynomial order at the rate $(P+1)^b$ where $b > -1$. Alternatively, the solution is guaranteed to be stable beyond a

critical grid resolution which depends on the polynomial order, the type of FR scheme, and the flux f.

As expected, the guarantee of stability is acquired at the cost of introducing artificial dissipation which limits the order of accuracy to $a < 2$. Moreover, the introduction of diffusion terms in the governing equation significantly reduces the maximum stable time-step for explicit time-integration schemes. Therefore, in the next chapter, we will relax this guarantee to provide an engineering solution where dissipation is posed in spectral space through a convolution and accuracy is controlled by the use of a discontinuity sensor.

12.1 Energy bounds for the numerical solution

A detailed account of the FR framework for non-linear conservation laws can be found in [69] and [71]. As in the case of linear fluxes, the domain $\Omega \subseteq \mathbb{R}$ is partitioned into elements, $\Omega = \bigcup_{n=1}^{N} \Omega_n = \bigcup_{n=1}^{N} \left(\bar{\Omega}_n + \partial\Omega_n\right)$ where $\bar{\Omega}_n = (x_n, x_{n+1})$ and $\partial\Omega_n = \{x_n, x_{n+1}\}$. For a general FR scheme of order P, each element is further discretized with $P+1$ variably distributed, distinct solution points $\boldsymbol{x_n} = [x_{n,1}, x_{n,2}, \ldots, x_{n,P+1}]^T = (x_{n,p})_{p=1,2,\ldots,P+1}$. The numerical solution, $u^\delta(x,t)$, is defined as a piecewise interpolating polynomial such that in the n^{th} element,

$$u^\delta(x,t) = \sum_{p=1}^{P+1} u^\delta_{n,p}(t)\ell_{n,p}(x), \quad x \in \Omega_n, \tag{12.1}$$

$$f^\delta(x,t) = \sum_{p=1}^{P+1} f(u^\delta_{n,p})\ell_{n,p}(x), \quad x \in \Omega_n, \tag{12.2}$$

where $\ell_{n,p}$ is the p^{th} Lagrange polynomial in the n^{th} element, and the superscript δ denotes a numerically evaluated quantity. The numerical solution and flux are both allowed to be discontinuous at the element interfaces $\partial\Omega = \bigcup_{n=1}^{N} \partial\Omega_n$. Each element in the physical domain $x \in \Omega_n$ is mapped isoparametrically to the parent domain $\xi \in [-1,1]$ as

$$\xi|_{\Omega_n}(x) = 2\frac{x - x_n}{x_{n+1} - x_n} - 1. \tag{12.3}$$

The stability of a given FR scheme of order P can potentially depend on (i) location of solution points $\boldsymbol{\xi}$, (ii) correction function g, and (iii) upwinding coefficient α. In the case of linear fluxes, the choice of solution points has no effect on the stability of the eigensolution. As established in Part I of the dissertation, $\boldsymbol{\xi}$ only affects the interpolation error associated with the collocation projection of the initial condition onto the polynomial space of the solution. The effect of correction function on linear-stability was studied by Vincent et al. [142] who derived a one-parameter family of correction functions for which the broken Sobolev-type norm defined as

$$\|u^{\delta}\|_{\bar{\Omega},(0,P)} = \left[\sum_{n=1}^{N}\int_{x_n}^{x_{n+1}}\left((u^{\delta}(x))^2 + \frac{c}{2}(h_n)^{2P}\left(\frac{\partial^P u^{\delta}}{\partial x^P}(x)\right)^2\right)\mathrm{d}x\right]^{1/2} \tag{12.4}$$

is non-increasing under exact time-integration. Here, $h_n = x_{n+1} - x_n$ is the size of the nth element, and $\bar{\Omega} = \bigcup_{n=1}^{N}\bar{\Omega}_n$ is the union of the element interiors. Their result holds for all admissible upwinding coefficients $\alpha \in [0, \frac{1}{2}]$. The same authors have also demonstrated [141] using an eigenvalue computation that these schemes retain linear stability under numerical time integration when using standard three-stage, third-order (RK33) and four-stage, fourth-order (RK44) Runge-Kutta methods for non-vanishing CFL numbers that depend on P.

However, in the more general case of non-linear fluxes, this ESFR family is not guaranteed to be stable. Jameson et al. [76] have shown that the growth rate of the norm in Eqn. (12.4), for a general flux, is dependent on aliasing errors,

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|u^{\delta}\|_{\bar{\Omega},(0,P)} &= \Theta + \sum_{n=1}^{N}\int_{\bar{\Omega}_n}(f^{\delta} - f(u^{\delta}))\frac{\partial u^{\delta}}{\partial x}\mathrm{d}x \\ &\leq \sum_{n=1}^{N}\int_{\bar{\Omega}_n}(f^{\delta} - f(u^{\delta})\frac{\partial u^{\delta}}{\partial x}\mathrm{d}x,\end{aligned} \tag{12.5}$$

since Θ is provably non-positive. The bounding terms in Eqn. (12.5), referred to as 'aliasing errors', are associated with the collocation projection of the non-linear flux corresponding to the numerical solution, $f(u^{\delta})$, onto the polynomial space of the numerical flux, f_n^{δ}.

Consequently, the choice of solution points is critical for non-linear conservation laws. Jameson et al. suggest the use of Gauss-Legendre quadrature points to reduce aliasing driven instabilities. In this section, we establish an upper bound on the rate of growth of the energy norm. Towards this end, let $\tilde{f} := f(u^\delta)$ for convenience of notation. The broken integral in Eqn. (12.5) can then be rewritten as

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|u^\delta\|_{\bar{\Omega},(0,P)} &\leq \int_{\bar{\Omega}} (f^\delta - \tilde{f})\frac{\partial u^\delta}{\partial x}\mathrm{d}x \\ &= \int_{\bar{\Omega}} (\mathscr{P}(\tilde{f}) - \tilde{f})\frac{\partial u^\delta}{\partial x}\mathrm{d}x, \end{aligned} \tag{12.6}$$

where $\mathscr{P}$ denotes the piecewise collocation projector, such that for $x \in \bar{\Omega}_n$,

$$\begin{aligned}\mathscr{P}(\tilde{f}(x)) &= \sum_{p=1}^{P+1} \tilde{f}(x_{n,p})\ell_{n,p}(x) \\ &= \sum_{p=1}^{P+1} f(u^\delta_{n,p})\ell_{n,p}(x) \\ &= f^\delta(x). \end{aligned} \tag{12.7}$$

Since the sign of the right hand side in Eqn. (12.6) cannot be determined *a-priori*, we are interested in establishing an upper bound for its magnitude. Using the Cauchy-Schwarz inequality, we get

$$\begin{aligned}\int_{\bar{\Omega}} (\mathscr{P}(\tilde{f}) - \tilde{f})\frac{\partial u^\delta}{\partial x}\mathrm{d}x &\leq \left[\int_{\bar{\Omega}} (\mathscr{P}(\tilde{f}) - \tilde{f})^2\mathrm{d}x \int_{\bar{\Omega}} \left(\frac{\partial u^\delta}{\partial x}\right)^2 \mathrm{d}x\right]^{1/2} \\ &= \|\mathscr{P}(\tilde{f}) - \tilde{f}\|_{L^2(\bar{\Omega})} \left\|\frac{\partial u^\delta}{\partial x}\right\|_{L^2(\bar{\Omega})}, \end{aligned} \tag{12.8}$$

where $\|.\|_{L^2(\bar{\Omega})}$ denotes the H^0 Sobolev norm over $\bar{\Omega}$. Let $\tilde{f} \in H^{q_f}(\bar{\Omega}_n)$, $q_f \geq 0$ for $n = 1,2,\ldots,N$, where H denotes the standard Sobolev space. Then, $\tilde{f} \in H^{q_f}(\bar{\Omega})$, and by results from approximation theory [62, 24, 15] on the error bound for collocation projection onto

Gauss-Legendre points, we get

$$\|\mathscr{P}(\tilde{f}) - \tilde{f}\|_{L^2(\bar{\Omega})} \leq C_1 h^{\sigma} \left\| \frac{\partial^{\sigma} \tilde{f}}{\partial x^{\sigma}} \right\|_{L^2(\bar{\Omega})}, \tag{12.9}$$

$$\sigma = \min(P+1, q_f), \tag{12.10}$$

where $h = \max_{n=1}^{N}(h_n)$, and C_1 is some constant. Substituting Eqns. (12.8 - 12.10) in Eqn. (12.6), we obtain an upper bound of the form

$$\frac{1}{2}\frac{d}{dt}\|u^{\delta}\|_{\bar{\Omega},(0,P)} \leq C_1 h^{\sigma} \left\| \frac{\partial^{\sigma} \tilde{f}}{\partial x^{\sigma}} \right\|_{L^2(\bar{\Omega})} \left\| \frac{\partial u^{\delta}}{\partial x} \right\|_{L^2(\bar{\Omega})}, \tag{12.11}$$

which shows that the growth rate of the broken Sobolev-type norm, Eqn. (12.4), can be bounded above by a term that depends on the magnitude of the σ-derivative of $\tilde{f}$ and the first derivative of the numerical solution u^{δ}.

For $P = 0$, the scheme is guaranteed to be non-linearly stable for any flux function since $\frac{\partial u^{\delta}}{\partial x} \equiv 0$ over $\bar{\Omega}$. However, for $P \geq 1$, a non-zero rate of growth may exist for any flux function with a non-trivial Jacobian, i. e. $\frac{\partial f(u^{\delta})}{\partial u^{\delta}} \neq 0$. In such cases, the bound on the right hand side vanishes with grid spacing at the rate h^{σ} where σ depends on the degree of local continuity of $f(u^{\delta})$. If $\tilde{f}$ is smooth, $q_f \to \infty$, then $\sigma = P+1$, and the growth term is bounded by the usual interpolation errors $\sim h^{P+1}\|\tilde{f}^{(P+1)}\|$. Consequently, the bound also vanishes in the limit of $P \to \infty$ for $h < 1$. However, when $\tilde{f}$ has a discontinuous derivative, $q_f < \infty$, then the bound does not vanish in the limit of $P \to \infty$. In such a case, convergence of the bound can only be achieved by refining the grid spacing.

12.2 Stabilization via artificial dissipation

We now describe a mechanism for adding artificial dissipation to ensure that the right hand side of Eqn. (12.11) is non-positive. Heuristically, we must add sufficient dissipation to counteract the growth of energy induced by aliasing errors. Similar to the approach in classical finite volume schemes [75], we begin by introducing a series of artificial diffusion

terms in Eqn. (11.1) such that the numerical conservation law is given by

$$\frac{\partial u^\delta}{\partial t} + \frac{\partial \mathscr{P}(\tilde{f})}{\partial x} = \sum_{s=1}^{\infty} (-1)^{s+1} \varepsilon_s \frac{\partial^{2s} u^\delta}{\partial x^{2s}}, \quad x \in \Omega,\ t > 0, \tag{12.12}$$

where $\varepsilon_s \in \mathbb{R}$ for $s = 1, 2, \ldots, \infty$, and we have used Eqn. (12.7) to replace f^δ with $\mathscr{P}(\tilde{f})$. Non-vanishing terms in the summation are obtained for $s \leq P/2$ for even P, and for $s \leq (P-1)/2$ for odd P. Let the infinite sum in the artificial diffusion term be denoted by $\mathscr{A}_n$ in the n^{th} element,

$$\mathscr{A}_n := \sum_{s=1}^{\infty} (-1)^{s+1} \varepsilon_s \frac{\partial^{2s} u^\delta}{\partial x^{2s}}, \quad x \in \bar{\Omega}_n. \tag{12.13}$$

Then, following the derivation in [76], the corresponding contribution to the bound in Eqn. (12.6) can be obtained as

$$\begin{aligned}
\sum_{n=1}^{N} \int_{x_n}^{x_{n+1}} & \left(u^\delta \mathscr{A}_n + \frac{c}{2} (h_n)^{2P} \frac{\partial^P u^\delta}{\partial x^P} \frac{\partial^P \mathscr{A}_n}{\partial x^P} \right) \mathrm{d}x \\
&= \int_{\bar{\Omega}} u^\delta \mathscr{A}_n \mathrm{d}x + \int_{\bar{\Omega}} \frac{c}{2} (h_n)^{2P} \frac{\partial^P u^\delta}{\partial x^P} \left(\sum_{s=1}^{\infty} (-1)^{s+1} \varepsilon_s \frac{\partial^{P+2s}}{\partial x^{P+2s}} u^\delta \right) \mathrm{d}x \\
&= \int_{\bar{\Omega}} u^\delta \left(\sum_{s=1}^{\infty} (-1)^{s+1} \varepsilon_s \frac{\partial^{2s} u^\delta}{\partial x^{2s}} \right) \mathrm{d}x \\
&= \sum_{s=1}^{\infty} \int_{\bar{\Omega}} (-1)^{s+1} \varepsilon_s u^\delta \frac{\partial^{2s} u^\delta}{\partial x^{2s}} \mathrm{d}x,
\end{aligned} \tag{12.14}$$

where we have used the fact that $\frac{\partial^{P+2s} u^\delta}{\partial x^{P+2s}} = 0$ for $s > 0$ since u^δ is a polynomial of degree P. Hence, the artificial dissipation term results in the new energy bound

$$\frac{1}{2} \frac{d}{dt} \| u^\delta \|_{\bar{\Omega},(0,P)} \leq \int_{\bar{\Omega}} (\mathscr{P}(\tilde{f}) - \tilde{f}) \frac{\partial u^\delta}{\partial x} \mathrm{d}x + \sum_{s=1}^{\infty} \int_{\bar{\Omega}} (-1)^{s+1} \varepsilon_s u^\delta \frac{\partial^{2s} u^\delta}{\partial x^{2s}} \mathrm{d}x. \tag{12.15}$$

For the artificial component, integrating by parts the sth term, s times, and using the periodicity of boundary conditions we get

$$\int_{\bar{\Omega}} (-1)^{s+1} u^{\delta} \frac{\partial^{2s} u^{\delta}}{\partial x^{2s}} \mathrm{d}x = -\left\| \frac{\partial^{s} u^{\delta}}{\partial x^{s}} \right\|^{2}_{L^2(\bar{\Omega})}. \tag{12.16}$$

Summing over the series and substituting into Eqn. (12.15), we finally obtain the modified bound on the rate of growth of energy,

$$\frac{1}{2}\frac{d}{dt}\|u^{\delta}\|_{\bar{\Omega},(0,P)} \leq C_1 h^{\sigma} \left\| \frac{\partial^{\sigma} \tilde{f}}{\partial x^{\sigma}} \right\|_{L^2(\bar{\Omega})} \left\| \frac{\partial u^{\delta}}{\partial x} \right\|_{L^2(\bar{\Omega})} - \sum_{s=1}^{\infty} \varepsilon_s \left\| \frac{\partial^{s} u^{\delta}}{\partial x^{s}} \right\|^{2}_{L^2(\bar{\Omega})}. \tag{12.17}$$

The stabilizing effect of artificial dissipation can now be observed from the following result.

Theorem 12.1. *The numerical solution to Eqn. (12.12) by an ESFR scheme of order P is monotonically decreasing in the norm given by Eqn. (12.4) in the asymptotic limits of $h \to 0$ or $P \to \infty$ provided that the flux function is smooth enough, $f(u^{\delta}) \in H^1(\bar{\Omega})$ and the artificial diffusion coefficients scale as*

$$\varepsilon_1 > 0, \;\; \varepsilon_j \geq 0 \;\; j = 2, 3, \ldots, P, \tag{12.18}$$

$$\varepsilon_1 \propto \frac{1}{\Delta t_{CFL}} \frac{h^{a}}{(P+1)^{b}}, \;\; a < 2, \; b < 1, \tag{12.19}$$

where Δt_{CFL} denotes the maximum time step for linear stability for the chosen time integration scheme.

Proof. In the case of $f(u^{\delta}) \in H^1(\bar{\Omega})$, Eqn. (12.17) yields

$$\frac{1}{2}\frac{d}{dt}\|u^{\delta}\|_{\bar{\Omega},(0,P)} \leq C_1 h \left\| \frac{\partial \tilde{f}}{\partial x} \right\|_{L^2(\bar{\Omega})} \left\| \frac{\partial u^{\delta}}{\partial x} \right\|_{L^2(\bar{\Omega})} - \sum_{s=1}^{\infty} \varepsilon_s \left\| \frac{\partial^{s} u^{\delta}}{\partial x^{s}} \right\|^{2}_{L^2(\bar{\Omega})}.$$

Using $\varepsilon_j \geq 0$ for $j = 2, \ldots, P$,

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\|u^\delta\|_{\bar{\Omega},(0,P)} &\leq C_1 h \left\| \frac{d\tilde{f}}{du^\delta}\frac{\partial u^\delta}{\partial x} \right\|_{L^2(\bar{\Omega})} \left\| \frac{\partial u^\delta}{\partial x} \right\|_{L^2(\bar{\Omega})} - \varepsilon_1 \left\| \frac{\partial u^\delta}{\partial x} \right\|^2_{L^2(\bar{\Omega})} \\
&\leq C_1 h \|w^\delta\|_{L^2(\bar{\Omega})} \left\| \frac{\partial u^\delta}{\partial x} \right\|^2_{L^2(\bar{\Omega})} - \varepsilon_1 \left\| \frac{\partial u^\delta}{\partial x} \right\|^2_{L^2(\bar{\Omega})} \\
&\leq C_1 h |\Omega|^{1/2} w^\delta_{\max} \left\| \frac{\partial u^\delta}{\partial x} \right\|^2_{L^2(\bar{\Omega})} - \varepsilon_1 \left\| \frac{\partial u^\delta}{\partial x} \right\|^2_{L^2(\bar{\Omega})},
\end{aligned} \tag{12.20}$$

where $w^\delta(x)$ denotes the numerical wavespeed, $w^\delta_{\max}$ denotes the maximum numerical wavespeed in the domain and we have used $\|w\|_{L^2(\bar{\Omega})} \leq \sqrt{|\Omega|} w^\delta_{\max}$. The CFL restriction on the time-step requires [8]

$$\Delta t \leq \Delta t_{\text{CFL}} = C_2 \frac{h}{w^\delta_{\max}(P+1)}, \tag{12.21}$$

where $C_2 \sim \mathscr{O}(1)$ is only weakly dependent on P. Substituting Eqn. (12.21) in Eqn. (12.20),

$$\frac{1}{2}\frac{d}{dt}\|u^\delta\|_{\bar{\Omega},(0,P)} \leq h w^\delta_{\max} \left\| \frac{\partial u^\delta}{\partial x} \right\|^2_{L^2(\bar{\Omega})} \left(C_1 |\Omega|^{1/2} - \frac{\varepsilon_1 \Delta t_{\text{CFL}}(P+1)}{C_2 h^2} \right). \tag{12.22}$$

This suggests that the artificial dissipation term dominates the right hand side with respect to convergence in grid spacing $h \to 0$ as well as polynomial order $P \to \infty$, provided that $\varepsilon_1 > 0$ and

$$\varepsilon_1 \propto \frac{1}{\Delta t_{\text{CFL}}} \frac{h^a}{(P+1)^b}, \quad a < 2,\ b < 1, \tag{12.23}$$

which completes the proof. □

Theorem 12.1 deals with stability in the asymptotic limits of grid spacing and polynomial order. Alternatively, a more practical result can be readily derived which shows that, for any given polynomial order, the solution is guaranteed to be stable for all grids finer than a critical resolution.

Theorem 12.2. *Under the conditions of Theorem 12.1, there exists a critical grid resolution,*

$$h_{crit}(P) \propto (P+1)^{\frac{1-b}{2-a}}, \quad a < 2,\ b < 1, \tag{12.24}$$

such that the numerical solution of Eqn. (12.12) by an FR scheme of polynomial order P is monotonically decreasing in the norm given by Eqn. (12.4) for any $h < h_{crit}$.

Proof. Admitting $\varepsilon_1 = C_3 \frac{1}{\Delta t_{\text{CFL}}} \frac{h^a}{(P+1)^b}$ in Eqn. (12.22) yields

$$\frac{1}{2}\frac{d}{dt}\|u^\delta\|_{\bar{\Omega},(0,P)} \leq h w^\delta_{\max} \left\| \frac{\partial u^\delta}{\partial x} \right\|^2_{L^2(\bar{\Omega})} \left(C_1 |\Omega|^{1/2} - C_2^{-1} C_3 \frac{h^{a-2}}{(P+1)^{b-1}} \right).$$

Hence, $\frac{1}{2}\frac{d}{dt}\|u^\delta\|_{\bar{\Omega},(0,P)} \leq 0$ for $h < h_{\text{crit}}$ where

$$\begin{aligned} h_{\text{crit}} &= \left(\frac{C_3 (P+1)^{1-b}}{C_1 C_2 |\Omega|^{1/2}} \right)^{\frac{1}{2-a}} \\ &= C_4 (P+1)^{\frac{1-b}{2-a}}, \end{aligned} \tag{12.25}$$

which completes the proof. □

Theorem 12.2 links the magnitude of the dissipation terms with a critical grid resolution. The constant C_2 can be determined from numerical eigenvalue computation of the CFL number for a given time integration scheme as in Chapter 5. However, the constant C_1 is not directly available. Thus, h_{crit} cannot be determined exactly. However, this result shows that an unstable numerical solution can be stabilized by either increasing the strength of filtering or decreasing the grid spacing. We can also see that the critical grid size, h_{crit}, increases with increasing polynomial order, as expected.

Chapter 13

Design of convolution filters

Theorems 12.1 and 12.2 show that the numerical solution via ESFR to a general non-linear conservation law can be stabilized through the addition of adequate artificial dissipation. However, if added directly to the governing equation, such artificial viscosity would introduce numerical errors of order $a < 2$ with respect to the grid spacing h. This would severely limit the high-order rate of convergence associated with FR schemes when $P \geq 2$. Moreover, the introduction of diffusion terms would also restrict the maximum stable time step for high-resolution meshes. Therefore, in this chapter, we propose a mechanism for adding only a limited amount of artificial dissipation which sacrifices the guarantee of stability to preserve the rate of convergence. This is achieved by relaxing the condition in Eqn. (12.18) to generate power expansions of well-known filter kernels in Fourier space. Consequently, addition of artificial dissipation can be posed as a Fourier-filtering operation, which circumvents the aforementioned time-step restriction as well.

The treatment contained herein borrows from the spectral viscosity method [128, 61] in that artificial dissipation is essentially introduced in Fourier space. However, instead of applying an explicit filter on the Fourier coefficients, the filtering operator is cast as a convolution in physical space similar to deconvolution strategies [111, 2, 137] used in Large Eddy Simulation. This avoids the computational complexity involved in performing Fourier transforms of piecewise polynomial functions.

Note that viscosity can be introduced in the Jacobi-spectral space as well [60, 94, 55].

Such an approach would offer convenience since spectral expansions in orthogonal polynomial bases are computationally efficient and well-conditioned. However, it would amount to a purely local filtering operation where the solution in each element is filtered independent of that in the neighbors. On the other hand, the use of a convolution integral automatically leads to a coupling between neighboring elements. Such coupling has been shown to perform favorably in the case of WENO [160] based limiting strategies for discontinuous Galerkin methods.

The last three sections of this chapter are attributed to heuristics that improve accuracy and computational tractability in adding such artificial dissipation. In particular, we propose an element-local restriction of the convolution integral which avoids the complex procedure of filtering over an unstructured grid. This restriction also provides a direct and efficient coupling of the solutions in neighboring elements. Along the same line, we propose to apply these local-filters only to troubled cells where a discontinuity has been detected by the concentration sensor of Gelb and Tadmor [51].

13.1 Artificial viscosity as a Fourier-filtering operation

We begin by demonstrating that the addition of artificial viscosity to the conservation law is approximately equivalent to a post-filtering operation applied in the Fourier space. This can be achieved by time-splitting Eqn. (12.12) into natural and artificial terms [55, 94],

$$u^{\delta}(x,t+\Delta t) = u^{\delta}(x,t) - \int_{t}^{t+\Delta t} \frac{\partial \mathscr{P}(\tilde{f})}{\partial x}(x,\tau)\mathrm{d}\tau + \int_{t}^{t+\Delta t} \sum_{s=1}^{\infty} (-1)^{s+1} \varepsilon_s \frac{\partial^{2s} u^{\delta}}{\partial x^{2s}}(x,\tau)\mathrm{d}\tau \quad (13.1)$$

$$= u'^{\delta}(x,t+\Delta t) + \int_{t}^{t+\Delta t} \sum_{s=1}^{\infty} (-1)^{s+1} \varepsilon_s \frac{\partial^{2s} u^{\delta}}{\partial x^{2s}}(x,\tau)\mathrm{d}\tau. \quad (13.2)$$

where the contribution from the natural term is denoted as

$$u'^{\delta}(x,t+\Delta t) := u^{\delta}(x,t) - \int_{t}^{t+\Delta t} \frac{\partial \mathscr{P}(\tilde{f})}{\partial x}(x,\tau)\mathrm{d}\tau. \quad (13.3)$$

In this time-split form, the addition of artificial dissipation can be alternatively regarded as a post-processing operation,

$$\Delta u^{\delta}(x,t+\Delta t) := u^{\delta}(x,t+\Delta t) - u'^{\delta}(x,t+\Delta t) \tag{13.4}$$

$$= \int_{t}^{t+\Delta t} \sum_{s=1}^{\infty} (-1)^{s+1} \varepsilon_s \frac{\partial^{2s} u^{\delta}}{\partial x^{2s}}(x,\tau)\mathrm{d}\tau \tag{13.5}$$

$$= \sum_{s=1}^{\infty} (-1)^{s+1} \varepsilon_s \frac{\partial^{2s} u^{\delta}}{\partial x^{2s}}(x,t+\Delta t)\Delta t + \mathscr{O}(\Delta t^2). \tag{13.6}$$

Consider introducing the Fourier transforms of u^{δ} and Δu^{δ},

$$u^{\delta}(x,t) = \int_{\mathbb{R}} U^{\delta}(k,t)\exp(ikx)\mathrm{d}k, \tag{13.7}$$

$$\Delta u^{\delta}(x,t) = \int_{\mathbb{R}} \Delta U^{\delta}(k,t)\exp(ikx)\mathrm{d}k. \tag{13.8}$$

Then, substituting Eqns. (13.7 - 13.8) in (13.6) results in the linearized update given by

$$\begin{aligned}\int_{\mathbb{R}} \Delta U^{\delta}(k,t)\exp(ikx)\mathrm{d}k &= \sum_{s=1}^{\infty} (-1)^{s+1}\varepsilon_s \Delta t \int_{\mathbb{R}} (ik)^{2s} U^{\delta}(k,t)\exp(ikx)\mathrm{d}k \\ &= \sum_{s=1}^{\infty} (-1)^{2s+1}\varepsilon_s \Delta t \int_{\mathbb{R}} k^{2s} U^{\delta}(k,t)\exp(ikx)\mathrm{d}k \\ &= -\sum_{s=1}^{\infty} (\varepsilon_s \Delta t) \int_{\mathbb{R}} k^{2s} U^{\delta}(k,t)\exp(ikx)\mathrm{d}k, \end{aligned} \tag{13.9}$$

where only the first order terms in time have been retained. This, in turn, yields the update equation for a single Fourier coefficient as a function of wavenumber k,

$$\Delta U^{\delta}(k,t) = -\sum_{s=1}^{\infty} (\varepsilon_s \Delta t) k^{2s} U^{\delta}(k,t). \tag{13.10}$$

Eqn. (13.10) confirms that this post-processing update is indeed equivalent to a filtering operation in Fourier space,

$$\bar{U}(k,t) = G(k)\, U^{\delta}(k,t), \tag{13.11}$$

$$G(k) = 1 - \sum_{s=1}^{\infty} (\varepsilon_s \Delta t) k^{2s}, \tag{13.12}$$

where $\bar{U}$ denotes the Fourier transform of the filtered numerical solution $\bar{u}$, and $G : \mathbb{R} \mapsto \mathbb{R}$ is the transfer function of the filter. Hence, the numerical solution to the artificial conservation law, Eqn. (12.12), is approximately equivalent to a filtered solution to the original conservation law, Eqn. (11.1).

13.2 Filtering through integral operators in physical space

The direct application of a Fourier-filter of this form requires a transform of the piecewise polynomial solution at each time-step which would be prohibitively expensive. We can avoid this computational complexity by casting filtering as an integral operation. Towards this end, transform both sides of Eqn. (13.11) and employ the Convolution Theorem so that

$$\begin{aligned} \bar{u}(x,t) &:= \mathscr{F}^{-1}(\bar{U}(k,t)) \\ &= (u^{\delta} * g)(x,t) \\ &= \int_{\mathbb{R}} u^{\delta}(\xi,t) g(x-\xi) \mathrm{d}\xi, \end{aligned} \tag{13.13}$$

where $g(x) = \mathscr{F}^{-1}(G(k))$ is the inverse Fourier transform of the filter kernel. This kernel is parameterized by the artificial dissipation coefficients, $\boldsymbol{\epsilon} = (\varepsilon_s)_{s=1,2,\ldots,\infty}$, as shown in Eqn. (13.12). Before proceeding on to specific instances of the kernel, we outline certain desirable features in the choice of $\boldsymbol{\epsilon}$.

From a computational point of view, observe that if the series in Eqn. (13.12) is truncated at any level S, i. e. $\alpha_s = 0$, $s > S$, then $g(x) = \delta(x) - \sum_{s=1}^{S} \alpha_s (-1)^{s+1} \delta^{(2s)}(x)/\pi^{2s}$

where $\delta(x)$ is the Dirac-delta distribution. The presence of distributions and their derivatives would defeat the purpose of casting artificial dissipation as a convolution. Hence, we are interested in an infinite series ($\forall\ S,\ \exists\ s > S$ such that $\varepsilon_s \neq 0$) which converges to a closed-form sum.

From the point of view of stability, Theorem 12.1 suggests that the choice of ϵ is restricted by the criteria in Eqns. (12.18 - 12.19). However, as mentioned in the beginning of this chapter, such a restriction would severely limit the order of accuracy of the scheme. Moreover, most filter kernels used in mathematical physics tend to have coefficients which alternate in sign, which violates the positivity constraint in Eqn. (12.18).

In line with these two observations, we now explore a set of alternative conditions for suitability of filter kernels. These conditions are less restrictive in that the coefficients of higher order derivatives, $\varepsilon_j,\ j > 1$, can assume negative values as well. Borrowing the results of Sengupta et al. [116], it can be shown that the application of an explicit filter modifies the numerical amplification factor A of the fully discretized differential equation,

$$A_{\text{eff}}(k^*, \zeta) = A(k^*, \zeta)\ G(k^*), \tag{13.14}$$

where $k^* = kh/(P+1)$ is the non-dimensional wavenumber, $\zeta = \Delta t/\Delta t_{\text{CFL}}$ is the non-dimensional timestep, and A_{eff} is the effective amplification factor. The amplitude of A_{eff} is directly affected by the filter, $|A_{\text{eff}}| = |A||G|$. Moreover, if $G(k^*)$ is complex, it can lead to modification of the numerical phase and group velocity as well, which can possibly worsen numerical dispersion. Hence we are interested in purely real, non-negative filter kernels bounded above by 1 for the entire range of resolvable wavenumbers,

$$0 \leq G(k^*) \leq 1,\ \text{ for } k^* \leq \pi. \tag{13.15}$$

Additionally, it is desirable to have minimal modification of low wavenumber components. This is satisfied by Eqn. (13.12) implicitly in the limit of $k^* \to 0$ since

$$\frac{d^s G}{dk^{*s}}(k^* = 0) = 0,\ s = 0, 1. \tag{13.16}$$

Note that it is not necessary for the kernel to vanish at the Nyquist limit.

These conditions are satisfied by well known kernels such as the Gaussian and Box filters. For instance, by choosing

$$\varepsilon_s \Delta t = (-1)^{s+1} \frac{\Delta^{2s}}{2^s s!}, \quad s = 1, 2, \ldots, \infty, \tag{13.17}$$

we get

$$G(k;\Delta) = \exp\left(-\frac{k^2\Delta^2}{2}\right) = \exp\left(-\frac{k^{*2}\Delta^{*2}}{2}\right), \tag{13.18}$$

where $\Delta^* = \Delta(P+1)/h$ denotes the non-dimensional filter width. In this case,

$$g(x;\Delta) = \frac{1}{\sqrt{2\pi\Delta^2}} \exp\left(-\frac{x^2}{2\Delta^2}\right), \tag{13.19}$$

which we recognize as the standard Gaussian filter. Similarly, we can obtain the box filter by choosing

$$G(k;\Delta) = \frac{\sin(k\Delta)}{k\Delta}, \tag{13.20}$$

$$g(x;\Delta) = \frac{1}{\Delta} \mathrm{I}_{[-\frac{\Delta}{2},\frac{\Delta}{2}]}(x), \tag{13.21}$$

where I denotes the indicator function. In this case, the kernel positivity requirement, Eqn. (13.15), restricts the range of admissible filter widths or equivalently the filter width parameter γ,

$$\Delta \in \left[0, \frac{h}{P+1}\right], \quad \gamma \in \left[0, \frac{1}{(P+1)^{3/4}} \left(\frac{\Delta t_{\mathrm{CFL}}}{\Delta t}\right)^{1/2}\right]. \tag{13.22}$$

Conventionally, the filter width is fixed by requiring that $G(k^* = \pi) = 0$. However, this would lead to a very dissipative filter kernel that significantly modifies a large portion of the spectrum. Instead, we can use Eqn. (13.17) to choose Δ from Eqn. (12.19), which for $a = 2$, $b = 1/2$ gives

$$\Delta = \gamma \left(\frac{\Delta t}{\Delta t_{\mathrm{CFL}}}\right)^{1/2} \frac{h}{(P+1)^{1/4}}, \tag{13.23}$$

for some constant γ which parameterizes the family of filters. The dependence of filter

width on element size implies that the filtering operator would need to be defined separately for every single element. For large meshes, this could lead to severe memory requirements. The choice of $a = 2$ ensures that the filter width Δ depends linearly on the grid spacing h. This results in significant computational efficiency since, as shown in the next section, such dependence leads to a constant filter width $\hat{\Delta}$ in the parent space. For this same reason, we are specifically interested only in kernels that scale inversely with respect to coordinate scaling,

$$g(\phi x; \phi\Delta) = \frac{1}{\phi} g(x; \Delta). \tag{13.24}$$

13.3 Local filter formulation in 1-D

The convolution integral in Eqn. (13.13) stretches over the entire real line and a direct evaluation would require the global piecewise definition of u^δ. While such global filtering operations are performed for finite-difference methods on structured meshes, they would lead to excessive computational cost and complexity in the case of unstructured meshes. In the discussion that follows, we propose an approximate local description of the filtering operation that can be posed as a simple matrix-vector product while ensuring that the numerical solution in neighboring elements remains coupled.

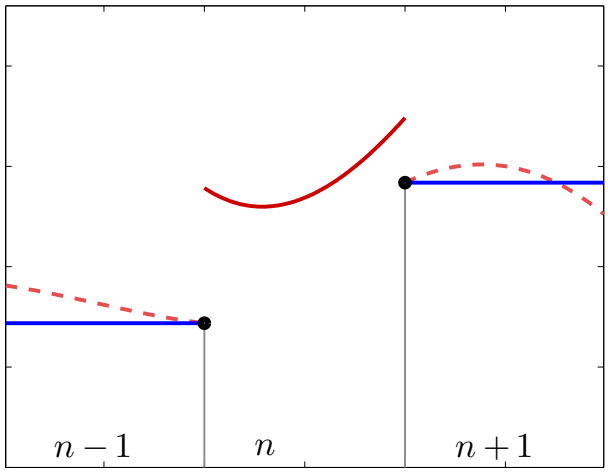
We begin by defining an operator ϕ that pads the piecewise polynomial solution in the parent space $\hat{u}^\delta$ with edge values of the neighboring elements,

$$\begin{aligned}\hat{v}_n(r) &= (\phi(\hat{u}^\delta))(r) \\ &= \hat{u}_n^\delta(r)\mathrm{I}_{(-1,1)}(r) + \hat{u}_{n-1}^\delta(1)\mathrm{I}_{(-\infty,-1]}(r) + \hat{u}_{n+1}^\delta(-1)\mathrm{I}_{[1,\infty)}(r),\end{aligned} \tag{13.25}$$

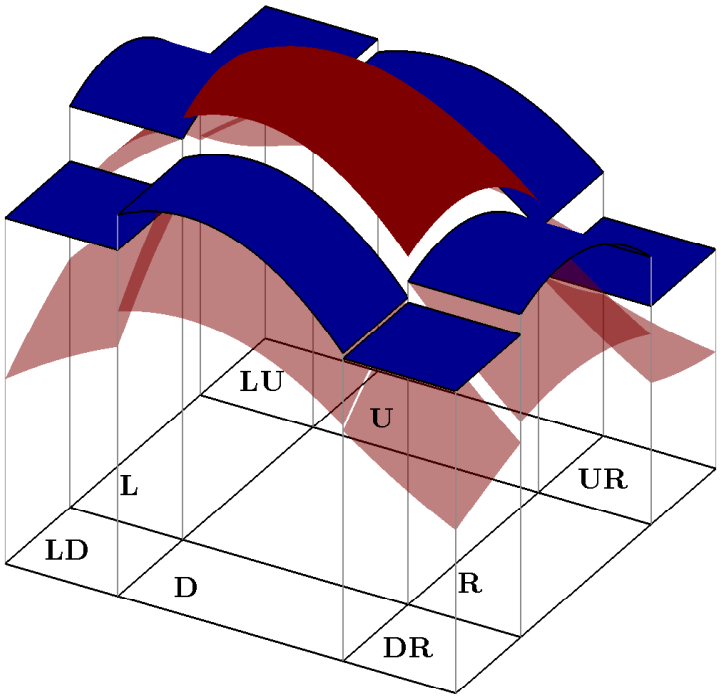
shown schematically in Fig. 13.1a. The motivation behind the padding procedure is essentially heuristic and attempts to generate a local representation of both intra and inter element discontinuities. The filtered function can now be easily expressed as

$$\bar{\hat{v}}_n(r) = \int_{\mathbb{R}} \hat{v}_n(\eta) g(x(r) - x(\eta); \Delta) J_n \mathrm{d}\eta, \tag{13.26}$$

where $J_n = h_n/2$ is the element Jacobian. This can be simplified using Eqn. (13.24) to



(a) Padded solution in 1D



(b) Padded solution in 2D

Figure 13.1: Local representation of the piecewise solution obtained by padding interface-information from neighboring elements. The dashed curves / translucent surfaces represent the original piecewise solution; the solid lines / solid surfaces denote the padded solution.

absorb the grid Jacobian and obtain a uniform filtering operator,

$$\begin{aligned}
\bar{\hat{v}}_n(r) &= \int_{-\infty}^{\infty} \hat{v}_n(\eta) g(r-\eta;\Delta J_n^{-1})\mathrm{d}\eta \\
&= \int_{-1}^{1} \hat{u}_n^{\delta}(\eta) g(r-\eta;\hat{\Delta})\mathrm{d}\eta + \hat{u}_{n-1}^{\delta}(1) \int_{-\infty}^{-1} g(r-\eta;\hat{\Delta})\mathrm{d}\eta \\
&\quad + \hat{u}_{n+1}^{\delta}(-1) \int_{1}^{\infty} g(r-\eta;\hat{\Delta})\mathrm{d}\eta \\
&= \sum_{p=1}^{P+1} \hat{u}_{n,p}^{\delta} \int_{-1}^{1} l_p(\eta) g(r-\eta;\hat{\Delta})\mathrm{d}\eta + \hat{u}_{n-1}^{\delta}(1) \int_{-\infty}^{-1} g(r-\eta;\hat{\Delta})\mathrm{d}\eta \\
&\quad + \hat{u}_{n+1}^{\delta}(-1) \int_{1}^{\infty} g(r-\eta;\hat{\Delta})\mathrm{d}\eta \\
&= \sum_{p=1}^{P+1} \hat{u}_{n,p}^{\delta} w_p(r) + \hat{u}_{n-1}^{\delta}(1) w_0(r) + \hat{u}_{n+1}^{\delta}(-1) w_{P+2}(r), \qquad (13.27)
\end{aligned}$$

where $\hat{\Delta}$ is the constant filter width in the parent domain,

$$\hat{\Delta} := \Delta J_n^{-1} = \left(\frac{\Delta t}{\Delta t_{\mathrm{CFL}}}\right)^{1/2} \frac{2\gamma}{(P+1)^{1/4}}, \qquad (13.28)$$

and the integral weights are given by

$$\begin{aligned}
w_p(r) &= \int_{-1}^{1} l_p(\eta) g(r-\eta,\hat{\Delta})\mathrm{d}\eta \ \text{ for } 1 \le p \le P+1, \\
w_0(r) &= \int_{-\infty}^{-1} g(r-\eta,\hat{\Delta})\mathrm{d}\eta = \int_{1+r}^{\infty} g(\eta,\hat{\Delta})\mathrm{d}\eta, \\
w_{P+2}(r) &= \int_{1}^{\infty} g(r-\eta,\hat{\Delta})\mathrm{d}\eta = \int_{-\infty}^{r-1} g(\eta,\hat{\Delta})\mathrm{d}\eta. \qquad (13.29)
\end{aligned}$$

Since the value of the filtered solution is required only at the solution points r_q for $q = 1, 2, \ldots P+1$, this procedure can be conveniently expressed as a matrix-vector product,

$$\begin{aligned} \bar{\mathbf{u}} &:= \left(\bar{\hat{v}}_n(\xi_r)\right)_{r=1,2,\ldots,P+1} \\ &= \mathbf{F}\,\hat{\mathbf{v}}, \end{aligned} \tag{13.30}$$

where $\mathbf{F}$ is the filter matrix of dimensions $(P+1) \times (P+3)$ such that

$$F_{q,p} = w_p(r_q) \text{ for } 0 \leq p \leq P+2, 1 \leq q \leq P, \tag{13.31}$$

and the vector of padded solution values, $\hat{\mathbf{v}}$, is given by

$$\hat{v}_p = \begin{cases} u^{\delta}_{n,p} & 1 \leq p \leq P+1, \\ u^{\delta}_{n-1}(1) & p = 0, \\ u^{\delta}_{n+1}(-1) & p = P+2 \end{cases} \tag{13.32}$$

From an implementation point of view, the integrals in Eqn. (13.29), can be computed using adaptive quadrature or a quadrature rule of sufficiently high order. For simple filter kernels, the boundary terms, w_0 and w_{P+2}, can be computed analytically. Note that such numerical integration only needs to be performed once to compute the filter matrix $\mathbf{F}$, i. e. $\mathbf{F}$ can be computed to arbitrary precision and stored thereafter.

13.4 Local filter formulation in 2-D for tensor product elements

Local filtering operators can be constructed in multiple dimensions by extending the procedure outlined in the previous section. In the case of tensor product elements, filtering can be readily performed along lines in the square parent domain. However, such an action would introduce anisotropy as the information along diagonals would be ignored and the choice of axial directions would become important. An isotropic filtering operator can be obtained by considering multi-dimensional extensions of the procedure suggested in the previous

subsection. As an example, let us define an operator ϕ^2 that pads the piecewise polynomial solution in the parent space $\hat{u}^\delta$ with the edge values of the neighboring elements,

$$\begin{aligned}\hat{v}_n(r_1,r_2) &= (\phi^2(\hat{u}^\delta))(r_1,r_2)\\ &= \hat{u}^\delta_n(r_1,r_2)\mathrm{I}_{(-1,1)\times(-1,1)}(r_1,r_2)\\ &\quad + \hat{u}^\delta_{n_L}(1,r_2)\mathrm{I}_{(-\infty,-1]\times(-1,1)} + \hat{u}^\delta_{n_R}(-1,r_2)\mathrm{I}_{[1,\infty)\times(-1,1)}\\ &\quad + \hat{u}^\delta_{n_D}(r_1,1)\mathrm{I}_{(-1,1)\times(-\infty,-1]} + \hat{u}^\delta_{n_U}(r_1,-1)\mathrm{I}_{(-1,1)\times[1,\infty)}\\ &\quad + \hat{u}^\delta_{n_{LD}}(1,1)\mathrm{I}_{(-\infty,-1]\times(-\infty,-1]} + \hat{u}^\delta_{n_{DR}}(-1,1)\mathrm{I}_{[1,\infty)\times(-\infty,-1]}\\ &\quad + \hat{u}^\delta_{n_{UR}}(-1,-1)\mathrm{I}_{[1,\infty)\times[1,\infty)} + \hat{u}^\delta_{n_{LU}}(1,-1)\mathrm{I}_{(-\infty,-1]\times[1,\infty)},\end{aligned} \tag{13.33}$$

where L,R,U,D,LD,DR,UR,LU denote the neighboring elements as illustrated in Fig. 13.1b. The convolution integral can now be simplified as

$$\begin{aligned}\bar{\hat{v}}_n(r_1,r_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{v}_n(s_1,s_2)\hat{g}_{2D}(x_1(r_1)-x_1(s_1),x_2(r_2)-x_2(s_2);\Delta)J_n\mathrm{d}s_1\mathrm{d}s_2\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{v}_n(s_1,s_2)\hat{g}(x_1(r_1)-x_1(s_1);\Delta J_n^{-1/2})\hat{g}(x_2(r_2)-x_2(s_2);\Delta J_n^{-1/2})\mathrm{d}s_1\mathrm{d}s_2.\end{aligned} \tag{13.34}$$

We now partition this double integral into the nine pieces depicted in Fig. 13.1b and substitute $\hat{v}_n$ from Eqn. (13.33). While the expressions can be directly evaluated using numerical integration, it would result in separate matrices for each element on account of the non-uniform Jacobian terms. In the case of quadrilaterals with parallel edges, the Jacobian is constant and the operation gets significantly simplified:

$$\begin{aligned}\bar{\hat{v}}_n(r_1,r_2) &= \sum_{p_1=1,p_2=1}^{P+1,P+1} \hat{u}^\delta_{n,p_1,p_2} w_{p_1}(r_1)w_{p_2}(r_2)\\ &\quad + \hat{u}^\delta_{n_L}(1,r_2)w_0(r_1)w_{p_2}(r_2) + \hat{u}^\delta_{n_R}(-1,r_2)w_{P+2}(r_1)w_{p_2}(r_2)\\ &\quad + \hat{u}^\delta_{n_D}(r_1,1)w_{p_1}(r_1)w_{P+2}(r_2) + \hat{u}^\delta_{n_U}(r_1,-1)w_{p_1}(r_1)w_0(r_2)\\ &\quad + \hat{u}^\delta_{n_{LD}}(1,1)w_0(r_1)w_0(r_2) + \hat{u}^\delta_{n_{DR}}(-1,1)w_{P+1}(r_1)w_0(r_2)\end{aligned}$$

$$+\hat{u}^{\delta}_{n_{UR}}(-1,-1)w_{P+2}(r_1)w_{P+2}(r_2)+\hat{u}^{\delta}_{n_{LU}}(1,-1)w_0(r_1)w_{P+2}(r_2), \quad (13.35)$$

where $\hat{\Delta} = \Delta J_n^{-1/2}$ is the constant filter width in the parent domain. Once again, this linear scaling results in a constant filter width in the parent space. This procedure can again be conveniently expressed as a matrix-vector product

$$\bar{\hat{u}}_{p_1,p_2} = w_i(r_{p_1})w_j(r_{p_2})\hat{v}_{i,j}, \text{ for } 1 \le p_1, p_2 \le P+1, \quad (13.36)$$

where the contraction is performed over $0 \le i, j \le P+2$, and the vector of padded solution values is given by

$$\hat{v}_{i,j} = \begin{cases} u^{\delta}_{n,i,j} & 1 \le i,j \le P+1, \\ u^{\delta}_{n_L}(1,r_2) & i=0, 1 \le j \le P+1, \\ u^{\delta}_{n_R}(-1,r_2) & i=P+2, 1 \le j \le P+1, \\ u^{\delta}_{n_D}(r_1,1) & 1 \le i \le P+1, j=0, \\ u^{\delta}_{n_U}(r_1,-1) & 1 \le i \le P+1, j=P+2, \\ u^{\delta}_{n_{LD}}(1,1) & i=0, j=0, \\ u^{\delta}_{n_{DR}}(-1,1) & i=P+2, j=0, \\ u^{\delta}_{n_{UR}}(-1,-1) & i=P+2, j=P+2, \\ u^{\delta}_{n_{LU}}(1,-1) & i=0, j=P+2 \end{cases} \quad (13.37)$$

The resultant filtering matrix $\mathbf{F}$ is of dimensions $(P+1)^2 \times (P+3)^2$.

13.5 Discontinuity detection via the concentration method

This final section briefly outlines the concentration method proposed by Gelb and Tadmor [51, 52] and further refined by Gelb and Cates [50] towards the recovery of edges in piecewise smooth functions with finitely many jump discontinuities. While prior applications of this method have mostly been concerned with image processing and post-processing of numerical solutions to differential equations, Sheshadri and Jameson [118] have recently applied it to detect shock discontinuities in transonic, inviscid as well as viscous fluid flow. In the discussion that follows, we describe a simplified version of the concentration method

which is obtained by fixing all but one of the parameters in the generalized formulation.

Consider the spectral projection of a piecewise smooth function $f : (x \in [-1,1]) \rightarrow \mathbb{R}$ onto the space of the first $P+1$ Legendre polynomials,

$$S_P(f) = \sum_{k=0}^{P} \hat{f}_k L_k(x), \tag{13.38}$$

where $L_k(x)$ is the k^{th} normalized orthogonal polynomial defined on the standard inner product and

$$\hat{f}_k = \int_{-1}^{1} f(x) L_k(x) \mathrm{d}x = \frac{-1}{k(k+1)} \int_{-1}^{1} f(x) \frac{d}{dx} \left((1-x^2) \frac{dL_k}{dx} \right) \mathrm{d}x, \tag{13.39}$$

where $(')$ denotes a derivative with respect to x. For simplicity, let f have a single jump-discontinuity at $x = c$, $[f](c) = f(c^+) - f(c^-)$. Then, integrating Eqn. (13.39) by parts and simplifying using properties of Legendre polynomials, we get

$$\hat{f}_k = \frac{-1}{k(k+1)} [f](c)(1-c^2) L_k'(c) + \mathscr{O}\left(\frac{-1}{k^2(k+1)^2} \right). \tag{13.40}$$

The first step in the concentration method is the localization of the projection in Eqn. (13.38) to the neighborhood of the jump using an appropriate concentration kernel. Towards this end, define the conjugate sum,

$$K_P * f(x) \equiv \frac{\pi \sqrt{1-x^2}}{P} \sum_{k=1}^{P} \hat{f}_k \frac{dL_k}{dx}. \tag{13.41}$$

Eqn. (13.40) can be used to show [52] that,

$$|K_P * f(x) - [f](x)| \leq \beta \frac{\log P}{P(1-x^2)^{1/4}}, \tag{13.42}$$

for $x \in \left(-1 + \frac{\beta}{P^2}, 1 - \frac{\beta}{P^2} \right)$ and some constant β. This suggests that for smooth pieces of the polynomial, the conjugate sum in Eqn. (13.41) converges to zero, and, if a jump discontinuity exists for points away from the boundaries $\{-1,1\}$ of the domain, the sum

converges to the magnitude of the jump at the rate of $\mathscr{O}\left(\frac{\log P}{P}\right)$. However, for discontinuous finite element schemes with polynomial order P, such convergence would be slow even for high orders. The separation of scales can be further increased through a non-linear enhancement,

$$P(K_P * f(x))^2 \sim \begin{cases} \frac{(\log P)^2}{P} & \text{at a smooth point,} \\ P[f]^2(x) & \text{at a discontinuity.} \end{cases} \tag{13.43}$$

This effectively exponentiates the relative convergence term by a power of two. A sharper separation can be achieved by choosing higher powers during the enhancement. Finally, for moderately large P, an accurate and reliable shock sensor can be defined as

$$\theta(x;J) = \begin{cases} 1 & P(K_P * f(x))^2 > J \\ 0 & \text{otherwise} \end{cases} \tag{13.44}$$

In general, the value of the enhanced kernel in Eqn. (13.43) increases with polynomial order, magnitude of the discontinuity, and proximity of the discontinuity from the boundaries of the bi-unit domain. This behavior has been portrayed in Table 13.1 which records the maximum value of the enhanced kernel for a centered step, $\text{sgn}(x)$, a skewed step, $\text{sgn}(x-1/2)$, a C_0 kink, $x_+ = \max(x,0)$, and a C_1 kink, $\cos(3+1.3\sin(x))$. We see that for high polynomial orders, $P \geq 5$, the concentration method clearly demarcates discontinuities in the function's value from those in the function's derivatives.

These observations certainly provide for an *a-posteriori* procedure of fixing the sensor threshold by observing the sensor values for a few initial time steps and manually selecting a suitable threshold. However, a practical choice is to select the average of the maximum values for the cases of a centered step and a C_0 kink. In other words, the sensor detects a discontinuity for normalized polynomials that are sharper than the shape midway between a centered step and a ramp.

It is important to mention that the use of the above procedure for shock detection requires appropriate normalization of the solution polynomial $u(x)$. This can be expressed as

P	C_1 kink $\cos(3+1.3\sin(x))$	C_0 kink $x_+ = \max(x,0)$	centered step $\mathrm{sgn}(x)$	skewed step $\mathrm{sgn}(x-1/2)$
5	1.7701	1.2536	3.6493	5.3244
6	1.5959	1.9925	3.5958	6.4424
7	1.4084	1.0910	4.5360	5.5975
8	1.1411	1.5527	4.6109	6.4980
9	1.0572	0.9366	5.4535	7.3697
10	1.0019	1.2722	5.6066	7.0400

Table 13.1: Maximum value of the enhanced kernel in Eqn. (13.43) over the solution points in $[-1,1]$. The functions are normalized with respect to the extrema over the point values.

$$U(x) = \frac{u(x) - u_{\min}}{u_{\max} - u_{\min} + \varepsilon}, \tag{13.45}$$

where ε represents the tolerance at machine precision. For conservation laws with linear fluxes, it suffices to take the maximum and minimum values of the solution over the entire domain. However, for non-linear problems, especially those comprising multiple shocks of different strengths, it is important to normalize the solution in each element with respect to its local extrema.

Chapter 14

Numerical experiments

This final chapter demonstrates the performance of the convolution filtering approach and the concentration based discontinuity sensor for canonical problems in 1-D and 2-D. The test cases comprise linear and non-linear scalar conservation laws as well as the Euler equations on tensor-product elements. We show that the proposed method provides non-linear stability and captures both inter and intra-element discontinuities even for exceptionally large polynomial orders. The accuracy of the filtered solution and the effect of the filter kernel are also explored.

For the numerical results that follow, the unfiltered numerical solution is computed via FR using the nodal DG correction function for the spatial derivative and the standard four-stage fourth-order Runge Kutta scheme for time integration. The interface fluxes are fully upwinded, and the time step is evaluated dynamically with a value equal to half the maximum allowable time step for linear stability. We have fixed the filter width parameter $\gamma = 1/3$, in Eqn. (13.23), and the concentration sensor threshold J, in Eqn. (13.44), to be the average of the maximum values for the cases of a centered step and a C_0 kink (see Table 13.1). These choices are essentially heuristic and have been chosen in this manner to provide a clear interpretation of the recorded results. In a practical setting, a more reliable albeit cumbersome strategy is to fine tune γ and J to a specific problem and grid by marching the numerical solution for a few time steps. However, as shown in the results that follow, these heuristic choices perform reasonably well without the need of any manual intervention.

We also perform a preliminary study of rate of convergence to ascertain the impact of filtering the solution in the entire domain without employing the discontinuity sensor. As expected, when the filter width parameter γ is relatively small, the filtered solution converges at the same rate as the unfiltered solution. As γ increases, the rate of convergence of the filtered solution decreases, until eventually, the scheme is rendered first-order accurate.

In the case of 2-D tests, we validate the consistency and performance of the isotropic filtering operator formulated in Section 13.4 for quadrilateral elements. As mentioned before, the 1-D filtering procedure can also be readily applied to tensor product elements by filtering along lines. A comparison of these two approaches shows little difference in the filtered solution. In order to identify elements with potential discontinuities, the concentration-based discontinuity sensor is evaluated along lines in all coordinate directions, for all conservative variables, in each element. We adopt a conservative approach where each conservative variable is filtered if the sensor detects a discontinuity for any of the variables.

14.1 Linear advection

We begin with the case of a linear scalar conservation law obtained by substituting $f(u) = u$ in Eqn. (11.1) defined on $\Omega = [-1, 1]$,

$$\begin{aligned} &\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \quad x \in [-1, 1],\, t > 0, \\ &u(x, 0) = u_0(x), \end{aligned} \tag{14.1}$$

where the boundary conditions are periodic. In the case of a smooth initial condition, the FR solution proceeds without filtering owing to the use of the discontinuity sensor. Hence, we are interested in $u_0(x)$ of the form [82, 77]

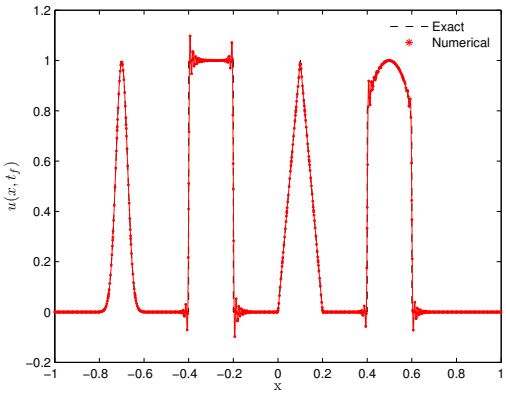
$$u_0(x) = \begin{cases} \frac{1}{6}\left(e^{-b(x+d+0.7)^2} + e^{-b(x-d+0.7)^2} + 4e^{-b(x+0.7)^2}\right), & -0.8 \le x \le -0.6 \\ 1, & -0.4 \le x \le -0.2 \\ 1 - 10|x - 0.1|, & 0.0 \le x \le 0.2 \\ \frac{1}{6}(F(x;0.5-d) + F(x;0.5+d) + 4F(x;0.5)), & 0.4 \le x \le 0.6 \\ 0, & \text{otherwise} \end{cases} \tag{14.2}$$

$$F(x;z) = \sqrt{\max((1 - 100(x-z)^2), 0)},$$

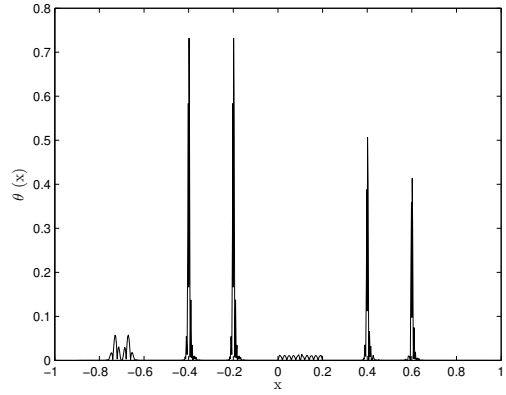
with $d = 0.005$, and $b = \frac{\log(2)}{36d^2}$. This initial condition consists of a piecewise combination of Gaussian forms, a pulse, a triangle and half ellipses. Eqns. (14.1), (14.2) are solved from $t = 0$ to 8 with and without filtering for polynomial order $P = 8$ on uniform grids having $N = 6, 12, 25, 50, 100$ elements. For a comparison between kernels, the filtered solution is obtained via both Gaussian (Eqn. (13.19)) and box (Eqn. (13.21)) filters.

Figures 14.1a-d plot the results for the case of $N = 100$. In the absence of filtering, the direct solution results in Gibbs phenomena at points of discontinuity. Since the flux is linear, these oscillations do not have a mechanism for causing instabilities and get advected subject to numerical dispersion. Figure 14.1b shows the behavior of the discontinuity sensor by plotting the variation of the enhanced concentration sum in Eqn. (13.43) applied element-wise to the unfiltered solution. The sensor accurately picks out the location of the four jump discontinuities and almost vanishes in the smooth regions of the solution. Figures 14.1c and 14.1d plot the solutions obtained by regular sensing and filtering operations via the Box and Gaussian filters respectively. Here, 'regular' implies that the operations are applied after every time-step. The action of the filter as a low-pass transfer function can be clearly observed in the filtered solution which is completely free from spurious oscillations. Additionally, owing to the sensor-selective filtering, the solution away from the discontinuities remains largely unaffected.

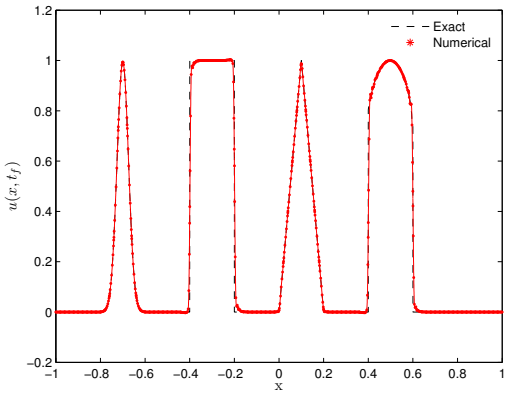
While the choice of filter kernel seems to have little visual effect, a quantitative comparison of the errors has been tabulated in Table 14.1 which includes the functional L_2



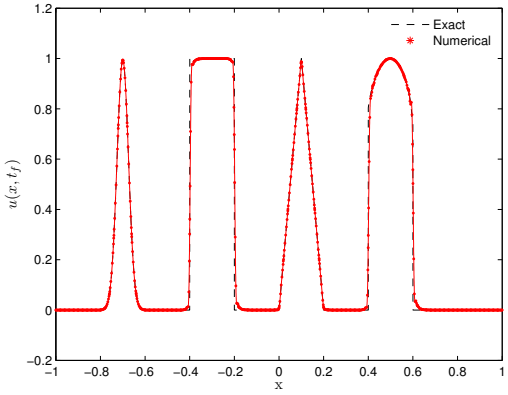
(a) Direct solution



(b) Sensor



(c) Box filter



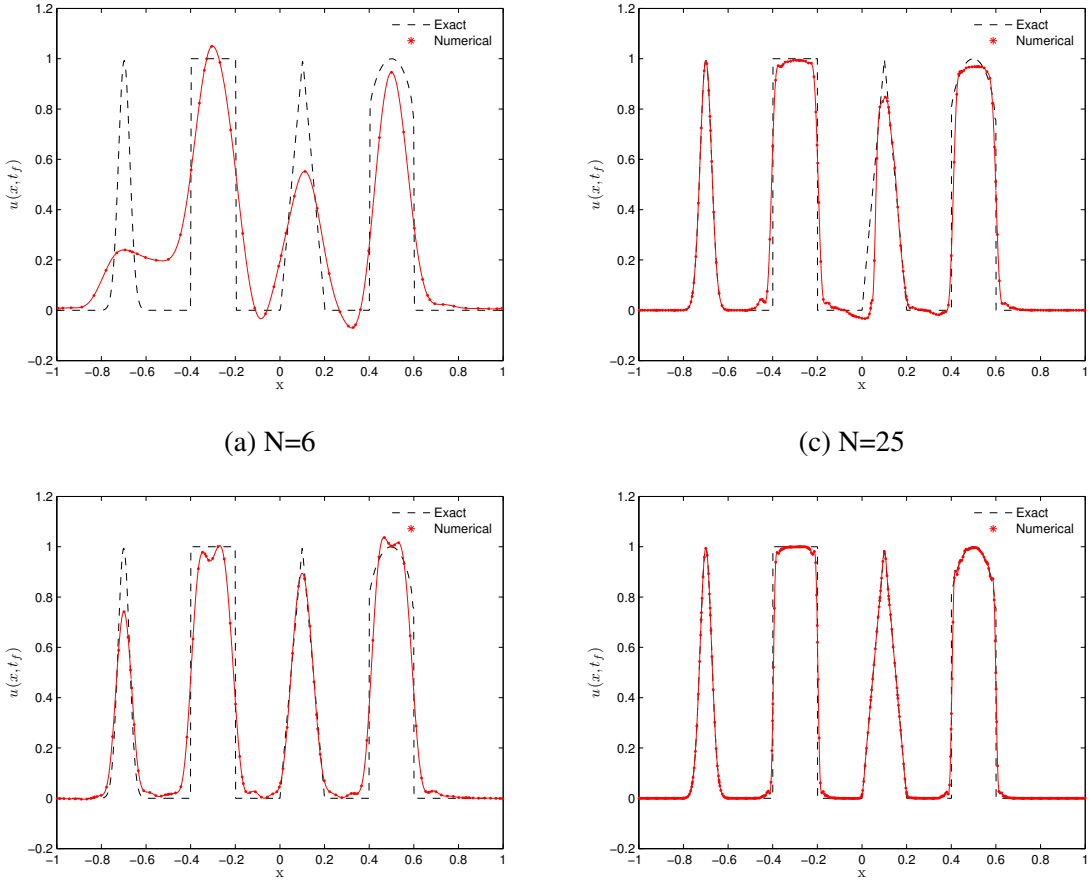
(d) Gaussian filter

Figure 14.1: Numerical solution for linear advection at $t = 8.0$ for $P = 8, N = 100$: (a) Direct solution via FR without filtering; (b) enhanced concentration sum, Eqn. (13.43), applied to the direct solution; (c) & (d) Solutions obtained with sensing and filtering operations at each time step.

Case	$\|u-u^\delta\|_{L^2(\Omega)}$	$\|u-u^\delta\|_{\ell^\infty(\Omega^\delta)}$	$\|u-u^\delta\|^\delta_{\text{O/U}}$	TV(u^δ) - TV(u)
Direct solution	4.46×10^{-2}	0.4890	9.75×10^{-2}	2.60
Box-filtered	5.38×10^{-2}	0.5812	2.93×10^{-3}	5.32×10^{-2}
Gauss-filtered	6.06×10^{-2}	0.6386	6.10×10^{-12}	1.85×10^{-3}

Table 14.1: Error metrics for the numerical solution of the linear advection problem

norm $\|u-u^\delta\|_{L^2(\Omega)}$, the discrete ℓ_∞ norm $\|u-u^\delta\|_{\ell^\infty(\Omega^\delta)}$, the discrete maximum overshoot/undershoot $\|u-u^\delta\|^\delta_{\mathrm{O/U}} = \max_{\Omega^\delta}\left(\max_{\Omega^\delta}(u^\delta)-\max(u),\ \min(u)-\min_{\Omega^\delta}(u^\delta),\ 0\right)$ and discrete total variation $\mathrm{TV}(u^\delta)$ - $\mathrm{TV}(u)$. We see that the filtering reduces the overshoot / undershoot and total variation while incurring errors due to mollification of the shocks. Between the two filters, the Gaussian filter tends to have a stronger effect for the same filter width. On account of sensor-selective filtering, there does not appear to be a significant advantage of choosing one type of filter over the other.



(a) N=6

(c) N=25

(b) N=12

(d) N=50

Figure 14.2: Grid convergence: Box-filtered solution for linear advection at $t = 8.0$ for $P = 8$ and $N = 6, 12, 25, 50$.

Figures 14.2a-d show the convergence of results with grid spacing when using the box-filter. For a very coarse grid, $N = 6$, the sharp features of the solution such as the combination of Gaussians are highly dissipated on account of the inherent numerical dissipation in the discretization scheme as well as that introduced through the filter. However, with grid refinement the filtered solution recovers rapidly and appears to have little visible error at the moderate resolution of $N = 50$. In a practical setting, there may exist a region of the flow field where the numerical solution is under-resolved and would get smoothed out by the filter due to false detection by the sensor. There are two possibilities in such a case. Either, such flow features would be a numerical artifact arising out of oscillations associated with a high-order polynomial, in which case the smoothing action would help to remove such spurious oscillations. The other possibility is when the grid is too coarse in a region where critical flow structures exist, such as a boundary layer, in which case one would need to refine the grid or change the sensor threshold J.

14.1.1 Rate of convergence for filtering without sensing

The use of a discontinuity sensor ensures that smooth regions of the solution remain unchanged. Hence, conditional on the success of the sensor, the proposed approach would preserve the order of accuracy of the underlying scheme for smooth solutions. Nevertheless, the effect of the proposed filtering operation on smooth solutions can be assessed by turning the sensor off and filtering in each element, at every time-step of the numerical solution.

Towards this end, we consider a smooth initial condition, $u_0(x) = \sin(\pi x)$, in Eqn. (14.1). The time-step is chosen to be 10 times smaller than the maximum allowable time step for linear stability, $\Delta t/\Delta t_{\text{CFL}} = 1/10$, so as to reduce errors associated with numerical integration. Eqn. (14.1) is solved from $t = 0$ to 1 with filtering in each element, at every time-step, for polynomial orders $P = 1, 2, 3, 4$ on uniform grids having $N = 8, 16$ elements. The kernel for filtering is taken to be that of the Box filter, Eqn. (13.21). The filter width parameter, γ, is varied over the entire range of possible values, Eqn. (13.22).

Figure 14.3 plots the rate of convergence of the functional L_2 norm of error $\|u -$

$u^\delta\|_{L^2(\Omega)}$ at $t = 1$ as a function of the filter width parameter, γ. The corresponding values for the unfiltered solution are also indicated for reference. Note that $t = 1$ corresponds to only half of a time period and, as proven in Part I of the dissertation, we expect to observe the short-time rate of convergence which is order $P + 1$ for fully-upwinded fluxes.

For low values of γ, the filter transfer function is almost unity and the filtering action has negligible effect. Consequently, the filtered solution converges at the same rate as the unfiltered solution. As the filter width increases, the rate of convergence smoothly decreases to first order accuracy expected from shock capturing methods. This shows that a trade-off exists between rate of convergence and stability, i. e. as the amount of artificial dissipation introduced by the filter increases, the order of accuracy of the solution decreases. For large values of γ close to the maximum limit for positivity of the filter kernel, the filtered solution, in fact, converges sublinearly, indicating that the filtered solution is highly dissipated and of little value.

It is important to note that a decrease in order of accuracy does not imply an increase in absolute error. This can be shown by comparing the L_2 norms of error for the four P values at the crossover points marked in Fig. 14.3. The corresponding error values for the grid with $N = 16$ elements are tabulated in Table 14.2. We see that, even though the rate of convergence for $P = 4$ is decreasing with γ, the actual numerical solution is still relatively accurate. For instance, at point 4 the curve for $P = 4$ has the lowest order of accuracy but still the lowest amount of error.

Point	$P = 1$	$P = 2$	$P = 3$	$P = 4$
1	9.57×10^{-3}	2.95×10^{-4}	6.43×10^{-6}	2.58×10^{-7}
2	9.57×10^{-3}	2.95×10^{-4}	6.44×10^{-6}	5.89×10^{-7}
3	9.57×10^{-3}	2.95×10^{-4}	6.48×10^{-6}	1.28×10^{-6}
4	9.57×10^{-3}	2.96×10^{-4}	1.02×10^{-5}	1.01×10^{-5}

Table 14.2: L_2 norm of error for the filtered solution at the crossover points indicated in Fig. 14.3 for the grid with $N = 16$ elements.

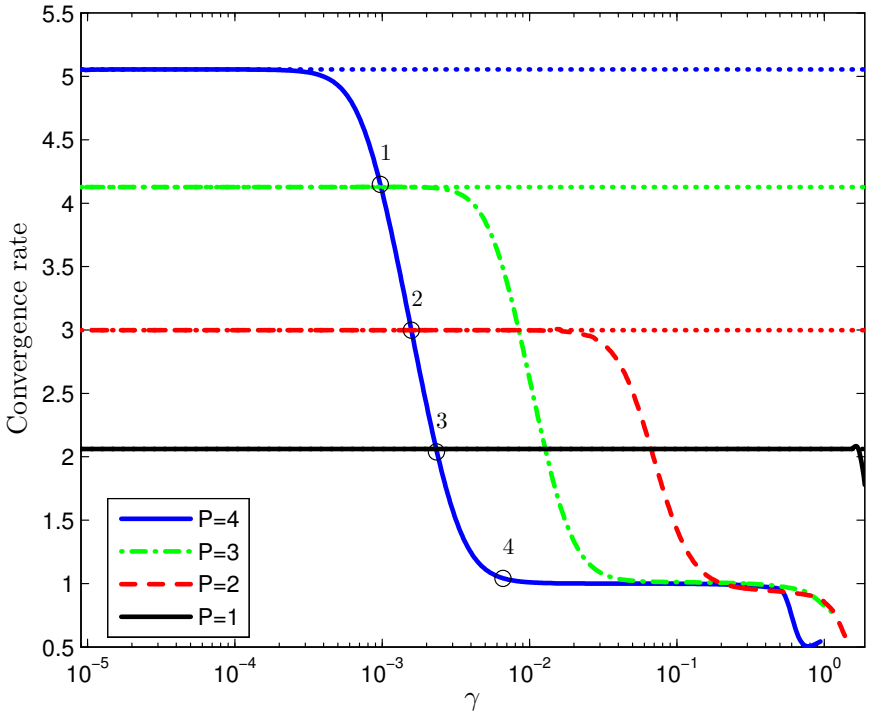


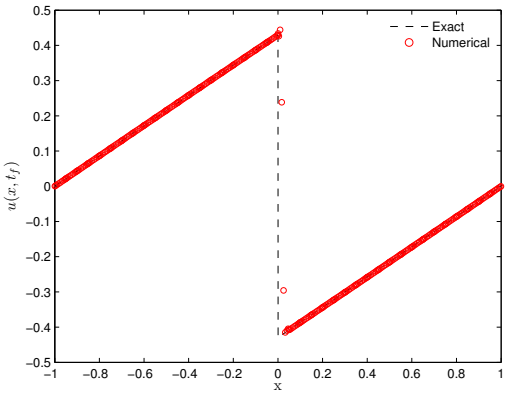
Figure 14.3: Rate of convergence for linear advection with a sinusoidal initial condition, $P = 1,2,3,4$, $N = 8,16$ at $t = 1$: The dotted lines represent the rates for the unfiltered solution. The solid and dashed curves represent the rates as a function of the filter width parameter when the solution is filtered in each element at every time step.

14.2 Inviscid Burgers equation

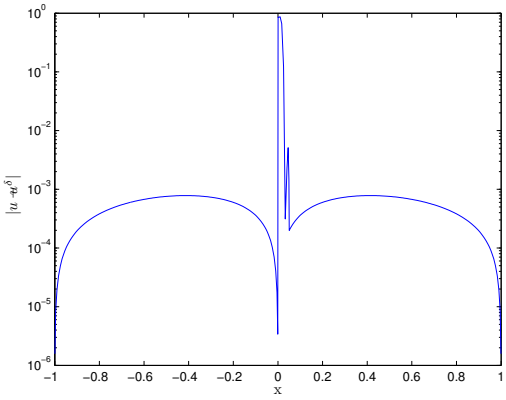
We now study the cases where the direct solution without filtering is non-linearly unstable on account of numerical instabilities. Consider the non-linear scalar conservation law obtained by substituting $f(u) = u^2$ in Eqn. (11.1) defined on $\Omega = [-1,1]$,

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{\partial u^2}{\partial x} &= 0, \quad x \in [-1,1],\, t > 0, \\ u(x,0) &= -sin(\pi x), \end{aligned} \tag{14.3}$$

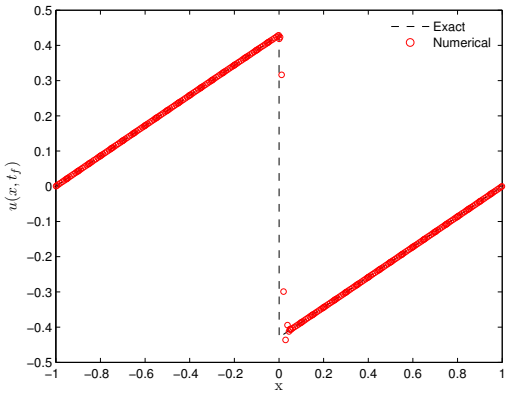
where the boundary conditions are again periodic. The analytical solution to Eqn. (14.3)



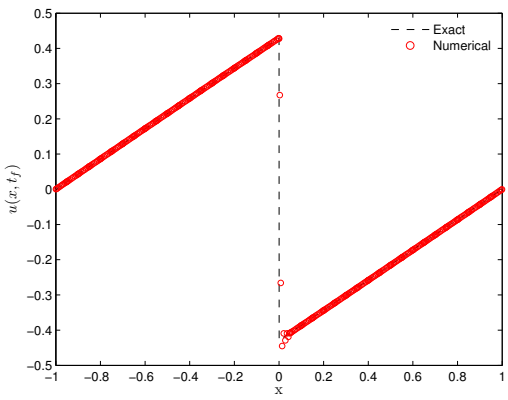
(a) Box-filtered solution, $P = 8$



(b) Point-wise error



(c) P=7



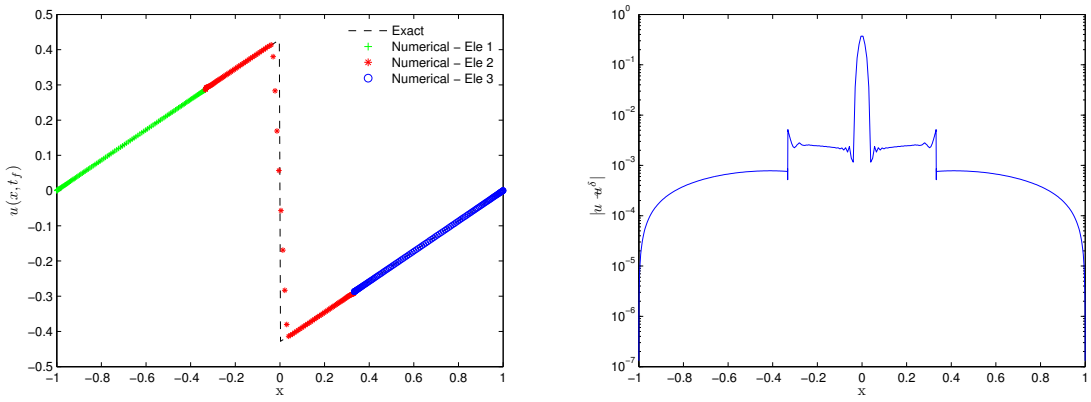
(d) P=9

Figure 14.4: (a) Numerical solution with filtering, (b) corresponding point-wise error for Burgers equation at $t = 1.0$ for $P = 8$ and $N = 40$. (c),(d) similar results for $P = 7, 9$

can be obtained via the method of characteristics. The characteristics from the two half planes eventually coalesce into a shock located at the origin whose strength decays with time. The suitability of the filtering operation can be measured through the strength and location of the shock in the numerical solution.

Figures 14.4a and b show the numerical solution with box-filtering and the associated point-wise error at $t = 1.0$ on a grid with $N = 40$ elements for $P = 8$. The filtered solution remains stable and accurately captures the shock strength. The error in smooth regions of the solution is at least three orders of magnitude smaller than near the shock. The

filtering action spreads the shock across two solution points with the mean slightly removed from the origin. Similar results are presented in Figs. 14.4c-d for $P = 7, 9$. Note that we restrict ourselves to cases with $P \geq 5$ as the performance of the chosen discontinuity sensor degrades rapidly for smaller P. An alternative sensor formulation should work seamlessly with the proposed filtering methodology.



(a) Sub-element shock capturing

(b) Point-wise error

Figure 14.5: (a) Numerical solution with filtering and (b) point-wise error for Burgers equation at $t = 1.0$ for $P = 119$ and $N = 3$.

An important property of the proposed filtering approach is the ability to capture intra-element shocks. Figures 14.5a-b show the numerical solution with filtering and associated error on a grid with $N = 3$ elements for $P = 119$. We see that both the shock location and magnitude are accurately predicted even at extremely coarse grids through such polynomial refinement. Moreover, the shock is contained completely within the interior of a single element.

14.3 Sod's shock tube

Proceeding onto a system of non-linear conservation laws, we consider the Euler equations for a calorically perfect ideal gas in 1-D,

$$\frac{\partial}{\partial t}\begin{bmatrix}\rho\\ \rho u\\ \rho E\end{bmatrix}+\frac{\partial}{\partial x}\begin{bmatrix}\rho u\\ \rho u^2+p\\ \rho H u\end{bmatrix}=0,\quad \text{in } \Omega\subset\mathbb{R},\, t>0, \tag{14.4}$$

$$p=(\gamma-1)\left(\rho E-\frac{1}{2}\rho u^2\right),$$

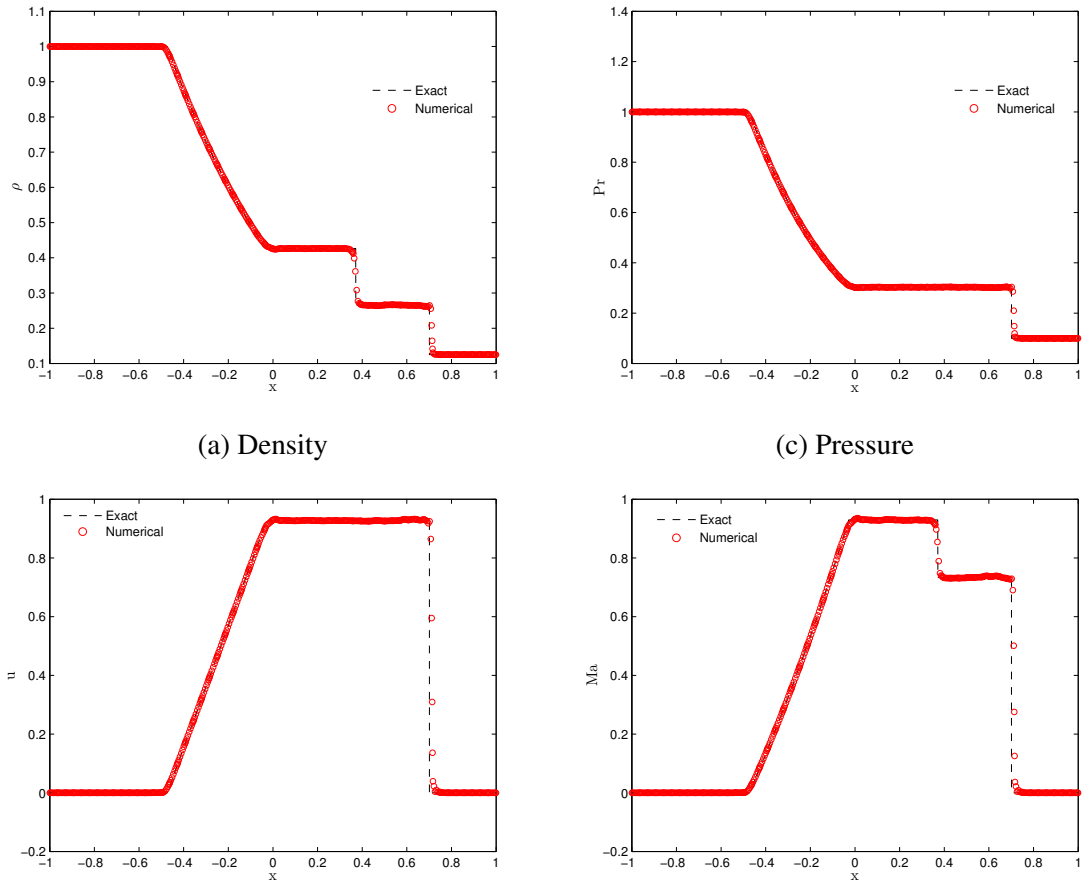
where ρ, u, E, p, H denote density, velocity, total internal energy, pressure and total enthalpy respectively, and γ denotes the ratio of specific heats. The weak form of Eqn. (14.4) naturally admits solutions with shocks depending on the initial and boundary conditions, and for high order polynomials, the direct numerical solution without filtering diverges due to numerical instabilities. In this subsection and the next, we show that the proposed filtering strategy leads to stable and accurate solutions. We adopt a conservative approach for dealing with a system of equations where an element is subjected to filtering if any of the conservative variables are diagnosed positively by the discontinuity sensor. Moreover, filtering is performed independently on each conservative variable directly.

Consider the initial and boundary conditions for Sod's shock tube problem [121] in $\Omega=[-1,1]$ for $t\leq 0.57$,

$$[\rho,u,p]^T=\begin{cases}[1,0,1]^T & x<0\\ [0.125,0,0.1]^T & x\geq 0\end{cases} \tag{14.5}$$

where $t\leq 0.57$ is chosen to ensure that none of the characteristics in the Riemann problem reach the boundaries of the domain. The exact solution can again be obtained using the method of characteristics and consists of a shock wave, a contact discontinuity and a smooth rarefaction fan.

Figure 14.6 shows the numerical solution obtained using regular sensing and box-filtering at $t=0.4$ on a grid with $N=56$ elements for polynomial order $P=8$. We see



(a) Density

(c) Pressure

(b) Velocity

(d) Mach number

Figure 14.6: Numerical solution with filtering for Sod's shock tube problem at $t = 0.4$ for $P = 8$ and $N = 56$

that the numerical solution accurately captures all the jumps and C_0 kinks in the solution with almost no overshoot/undershoot. The accuracy in smooth regions of the solution is preserved as well. While the location of the contact discontinuity is predicted exactly, the shock is spread over 4 solution points with mean slightly shifted to the right. The error metrics for each of the primitive variables are tabulated in Table 14.3 and confirm that stable, accurate and low total variation solutions can be obtained with high polynomial orders on coarse meshes.

Case	$\|u-u^\delta\|_{L^2(\Omega)}$	$\|u-u^\delta\|_{\ell^\infty(\Omega^\delta)}$	$\|u-u^\delta\|^\delta_{\text{O/U}}$	TV(u^δ) - TV(u)
ρ	1.51×10^{-2}	0.139652	5.29×10^{-6}	2.25×10^{-2}
u	8.98×10^{-2}	0.924835	5.38×10^{-3}	7.94×10^{-2}
p	1.93×10^{-2}	0.203868	5.86×10^{-6}	3.24×10^{-2}
M	7.26×10^{-2}	0.729450	6.41×10^{-3}	6.87×10^{-2}

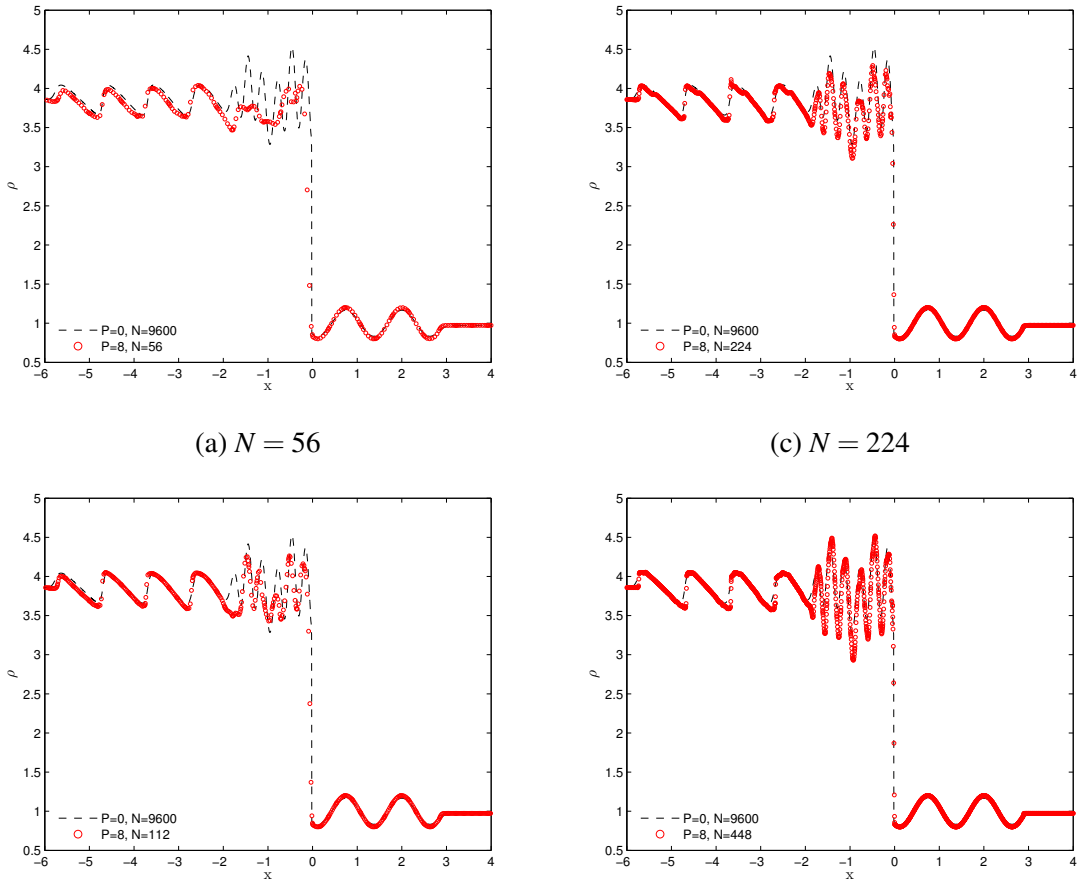
Table 14.3: Error metrics for the numerical solution of Sod's shock tube problem. Here $M = u/\sqrt{\gamma p/\rho}$

14.4 Shu-Osher shock turbulence interaction

Our final test case in 1-D consists of a flow field with sinusoidal density fluctuations moving left into a stationary Mach 3 normal shock front [82, 119]. The domain for this case is $\Omega = [-10, 10]$ and the initial condition is prescribed as

$$[\rho, u, p]^T = \begin{cases} [3.857143, -0.920279, 10.33333]^T & x < 0 \\ [1 + 0.2\sin(5x), -3.549648, 1]^T & x \geq 0 \end{cases} \tag{14.6}$$

The interaction leads to amplified high frequency components in the near wake and saw-tooth like skewed wave-forms in the far wake of the shock. We demonstrate that the proposed filtering approach provides stability without damping the density oscillations in the wake, and hence, can be employed for problems involving shock-vortex interactions as well.



(a) $N = 56$ (c) $N = 224$

(b) $N = 112$ (d) $N = 448$

Figure 14.7: Numerical solution with filtering for Shu-Osher shock turbulence interaction problem at $t = 2.0$ for $P = 8$ and $N = 56, 112, 224, 448$. The reference solution is obtained via direct solution for $P = 0$, $N = 9600$.

Figures 14.7a-d show the numerical solution obtained using regular sensing and box-filtering at $t = 2.0$ on successively refined grids with $N = 56, 112, 224, 448$ elements for $P = 8$. Since the analytical solution for this problem is not available, we use the direct solution at $t = 2.0$ for $P = 0$ on a grid with $N = 9600$ elements as the reference solution. Note that, as shown in Chapter 12, the case of $P = 0$ is naturally non-linearly stable and does not require any filtering. We see that the high-order numerical solution accurately captures the essential physical features of the flow including shock strength and location, amplitude and phase of the saw-tooth fluctuations in the far wake, and with grid refinement, even

the high frequency components in the near wake of the shock. The solution at the finest grid suggests that the high-order scheme with careful filtering can obtain accurate results with less than half the number of degrees of freedom than the first order scheme even for solutions with strong discontinuities.

14.5 Rigid body rotation

For our first 2-D test case, consider the scalar conservation law describing counter clockwise rotation in a unit square about its center,

$$\frac{\partial u}{\partial t} + \left(\frac{1}{2} - y\right)\frac{\partial u}{\partial x} + \left(x - \frac{1}{2}\right)\frac{\partial u}{\partial y} = 0, \quad (x,y) \in [0,1]^2,\ t > 0, \tag{14.7}$$
$$u(x,y,0) = u_0(x,y),$$

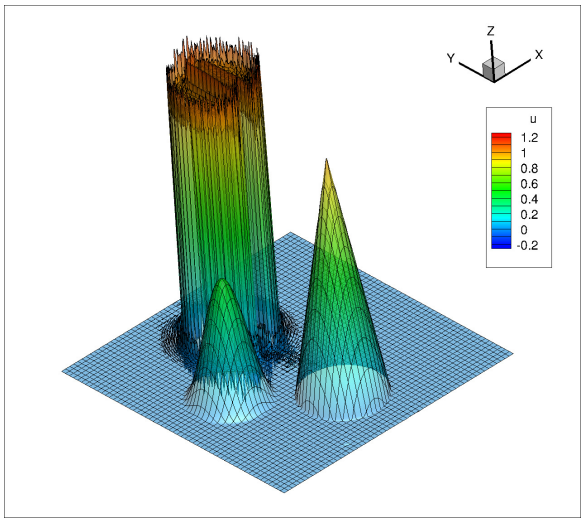
with periodic boundary conditions. Once again, in the case of a smooth initial condition, the FR solution proceeds without filtering owing to the use of the discontinuity sensor. Hence, we are interested in $u_0(x,y)$ of the form [85],

$$\begin{aligned} u_0(x,y) = {} & u_{0,1}(x,y)\, \mathrm{I}_{\{r(x,y;0.5,0.75)\leq 1\}} + u_{0,2}(x,y)\, \mathrm{I}_{\{r(x,y;0.5,0.25)\leq 1\}} \\ & + u_{0,3}(x,y)\, \mathrm{I}_{\{r(x,y;0.25,0.5)\leq 1\}}, \end{aligned} \tag{14.8}$$

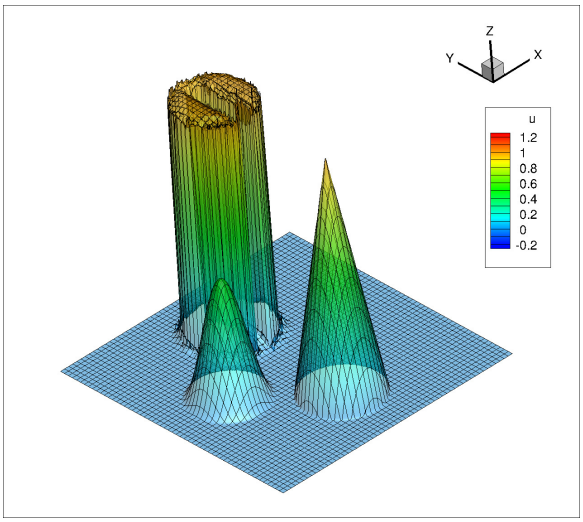
$$\begin{aligned} r(x,y;x_0,y_0) &= \sqrt{(x-0.5)^2 + (y-y_0)^2}/0.15, \\ u_{0,1}(x,y) &= \mathrm{I}_{\{|x-0.5|\geq 0.025\}\cup\{y\geq 0.85\}}, \\ u_{0,2}(x,y) &= 1 - r(x,y;0.5,0.25), \\ u_{0,3}(x,y) &= 0.25(1+\cos(\pi r(x,y;0.25,0.5))). \end{aligned}$$

which consists of a slotted cylinder, a cone and a sinusoidal bump. Eqns. (14.7), (14.8) are solved for one period from $t = 0$ to 2π with and without filtering for polynomial order $P = 8$ on a uniform grid having $N = 56 \times 56$ elements. The filtered solution is evaluated using both the isotropic filter of Section 13.4 and the direct approach of filtering along lines.

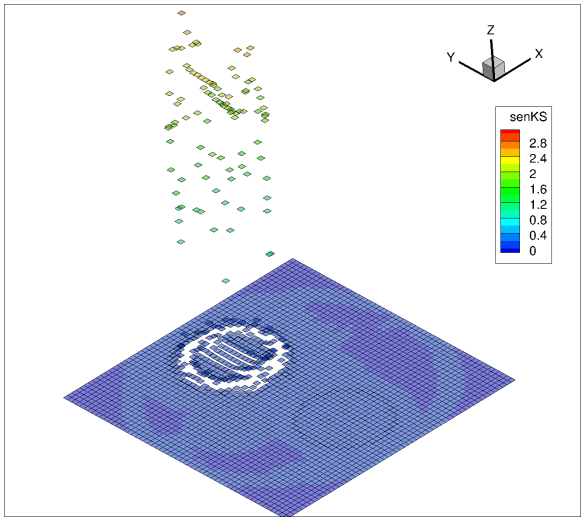
We see from Fig. 14.8a that the direct numerical solution without filtering exhibits Gibbs phenomena along the circumference of the cylinder. Figure 14.8b shows that the discontinuity sensor accurately identifies the ring of elements with jump discontinuities and almost vanishes in the smooth regions of the solution. Figures 14.8c and 14.8d plot the solutions obtained using regular sensing and filtering operations via the isotropic Box filter



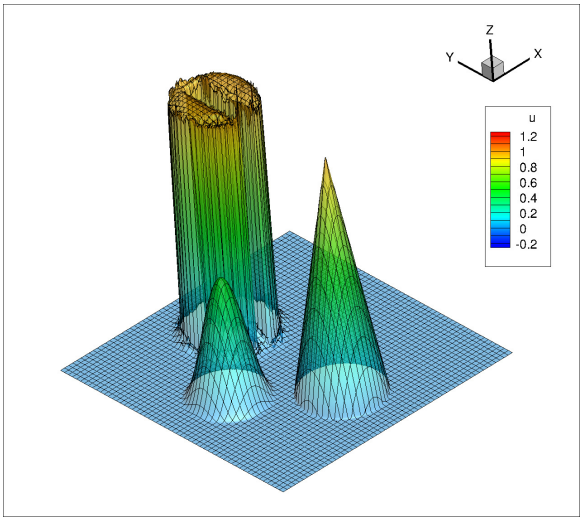
(a) Direct solution



(c) Isotropic box filter



(b) Sensor



(d) Box filter applied along lines

Figure 14.8: Numerical solution for rigid body rotation at $t = 2\pi$ for $P = 8, N = 56 \times 56$: (a) Direct solution via FR without filtering; (b) enhanced concentration sum, Eqn. (13.43), applied to the direct solution along lines; (c) & (d) Solutions obtained with sensing and filtering operations at each time step.

and the 1-D box filter applied along lines, respectively. The filtered solution is relatively free from spurious oscillations and, owing to sensor-selective filtering, the solution away from the discontinuities remains largely unaffected. A quantitative comparison of the errors

has been tabulated in Table 14.4 which includes the discrete ℓ_2 norm $\|u-u^\delta\|_{\ell^\infty(\Omega^\delta)}$, the discrete ℓ_∞ norm $\|u-u^\delta\|_{\ell^\infty(\Omega^\delta)}$ and the discrete maximum overshoot/undershoot $\|u-u^\delta\|^\delta_{\text{O/U}}$. We see that both the filtering strategies reduce error along all metrics with relatively little difference between them.

Case	$\|u-u^\delta\|_{\ell^\infty(\Omega^\delta)}$	$\|u-u^\delta\|_{\ell^\infty(\Omega^\delta)}$	$\|u-u^\delta\|^\delta_{\text{O/U}}$
Direct solution	5.95×10^{-3}	6.40×10^{-4}	6.40×10^{-4}
Isotropic box filter	1.30×10^{-4}	2.06×10^{-5}	2.06×10^{-5}
1-D box filter along lines	1.20×10^{-4}	1.57×10^{-5}	1.57×10^{-5}

Table 14.4: Error metrics for the numerical solution of the rigid body rotation problem

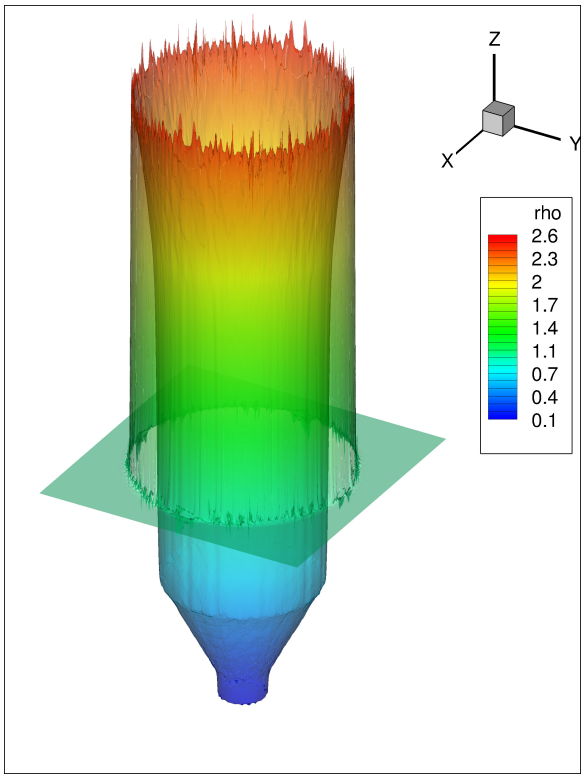
14.6 Symmetric Riemann problem

Proceeding onto a system of non-linear conservation laws, we now consider the Euler equations in 2-D for a calorically perfect ideal gas, Eqns. (4.21 - 4.22). As in 1-D, the weak form of these equations admits solutions with shocks depending on the initial and boundary conditions. For high-order polynomials, the direct numerical solution without filtering diverges due to numerical instabilities. Consider the radially symmetric Riemann problem [85] in $\Omega = \left[-\frac{1}{2},\frac{1}{2}\right]^2$,

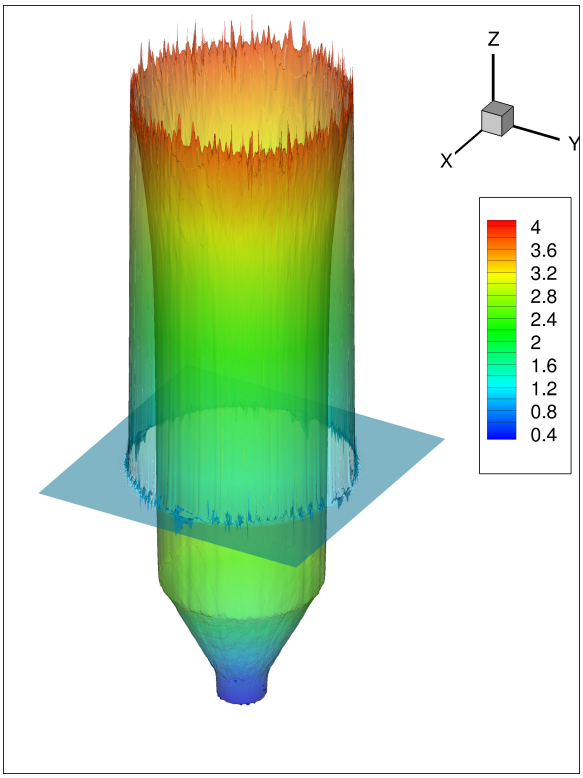
$$[\rho,u,v,p]^T = \begin{cases} [2,0,0,15]^T & \sqrt{x^2+y^2} < 0.13 \\ [1,0,0,1]^T & \sqrt{x^2+y^2} \geq 0.13 \end{cases} \tag{14.9}$$

which consists of a shock wavefront expanding radially outwards, a contact discontinuity and, most importantly, a rarefaction fan that moves radially inwards towards the origin creating a depression with a carbuncle. Thus, unlike the 1-D shock tube, the solution is not L_∞ preserving and highlights the importance of appropriate normalization of the solution for sensing discontinuities.

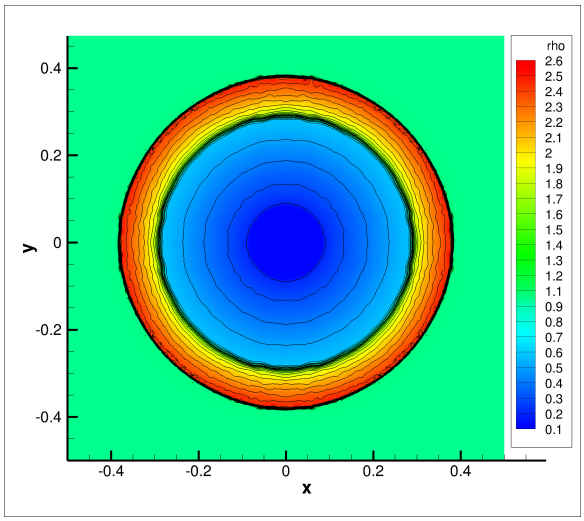
Figure 14.9 shows the numerical solution obtained using regular sensing and box-filtering at $t = 0.1$ on a grid with $N = 56\times56$ elements for polynomial order $P = 8$. We see



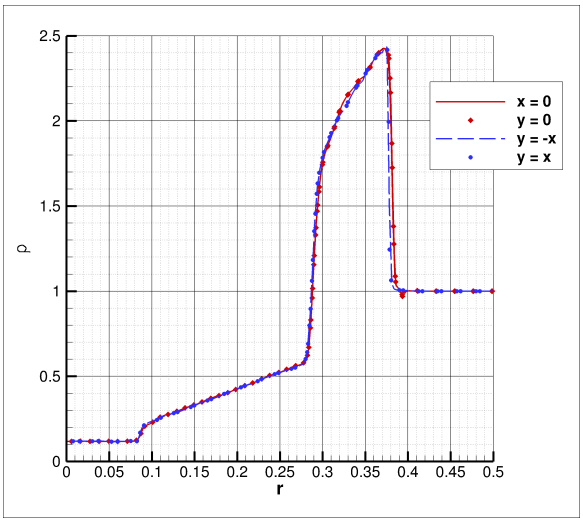
(a) Surface plot of density



(c) Surface plot of pressure



(b) Contour plot of density



(d) Cut-lines of density along coordinate and diagonal directions

Figure 14.9: Numerical solution with isotropic box filtering for radially symmetric Riemann problem at $t = 0.1$ for $P = 8$ and $N = 56 \times 56$

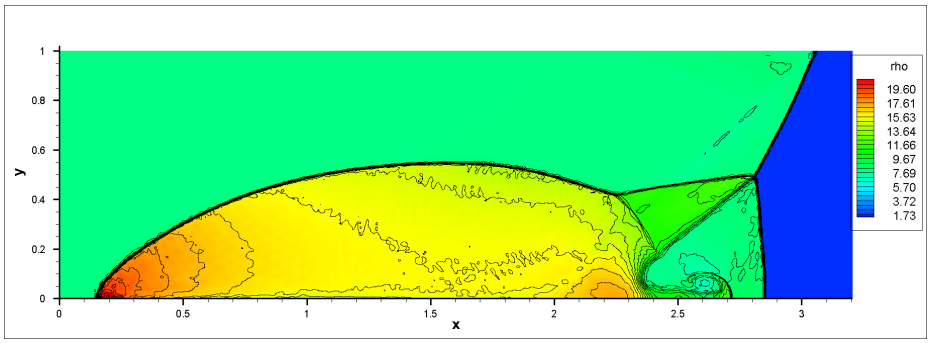
that the numerical solution accurately captures all the jumps and C_0 discontinuities in the solution including the low density depression and carbuncle in the interior of the domain. Additionally, cutlines of density in the coordinate and diagonal directions show that the numerical solution preserves radial symmetry.

14.7 Double Mach reflection

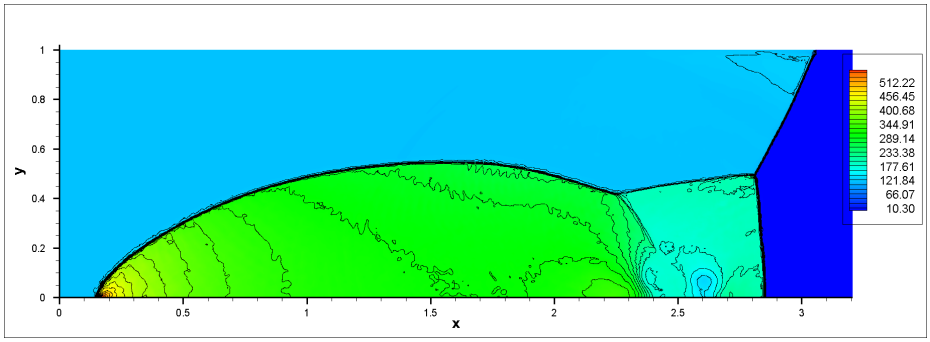
We now demonstrate the ability of the proposed sensing-filtering approach to handle exceptionally strong shocks with high polynomial orders. Towards this end, consider the double Mach reflection problem of Woodward and Colella [157] with $\Omega = [0,4] \times [0,1]$, wherein a Mach 10 shock is incident on a reflecting wall at an angle of 60°. A detailed description of the physical problem, self-similarity of the solution and the boundary conditions can be found in [157]. The complexity of treating oblique walls is avoided by tilting the domain so that the reflecting wall is aligned with the x-axis. The reflected shock is kept attached to the reflecting wall by assigning post-shock values for $\{x \in [0,1/6], y = 0\}$. A time varying boundary condition is applied on the top wall, $y = 1$, to describe the exact motion of the shock. Details regarding the implementation of the initial and boundary conditions can be found in [85].

Figure 14.10 shows the numerical solution obtained using regular sensing and box-filtering at $t = 0.2$ on a grid with $N = 56 \times 224$ elements for polynomial order $P = 8$. For this case of severely high shock strength, the threshold for the discontinuity sensor, J, was reduced to the corresponding value for a C_0 kink (see Table 13.1) in order to compute stably.

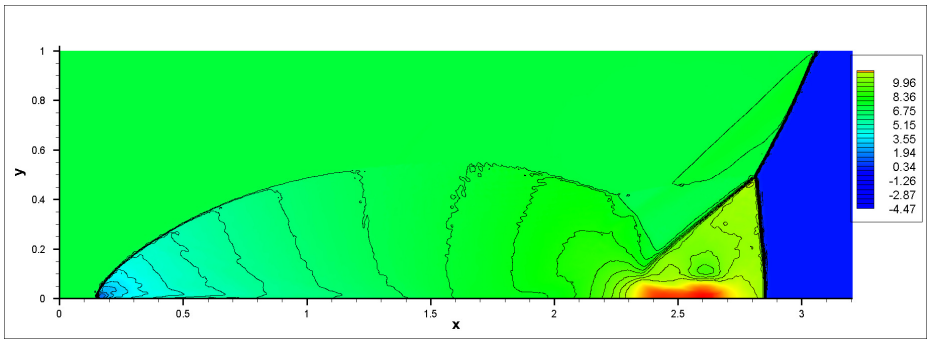
We see that, even at this coarse grid, the filtered solution consistently predicts the two Mach shocks, the contact discontinuities and the jet of dense fluid issued to the right from the base of the first contact discontinuity. However, in the region bounded by the curved reflected shock and the wall, contours of physical variables are visibly marked with high frequency noise. As predicted by Woodward and Colella, such oscillations occur since the dissipation added through the Riemann fluxes at element interfaces vanishes in the limit of zero flow velocity. Note that artificial dissipation is not added through the filter in this region since the fluctuations are too small to trigger a positive response from the



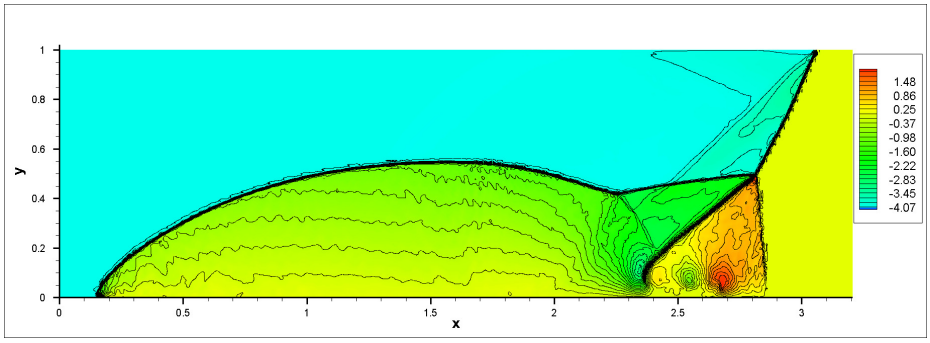
(a) Density: 30 isolines: 1.731 to 20.92



(b) Pressure: 30 isolines: 10.30 to 549.4



(c) x-velocity: 30 isolines: -4.47 to 11.28



(d) y-velocity: 30 isolines: -4.065 to 1.891

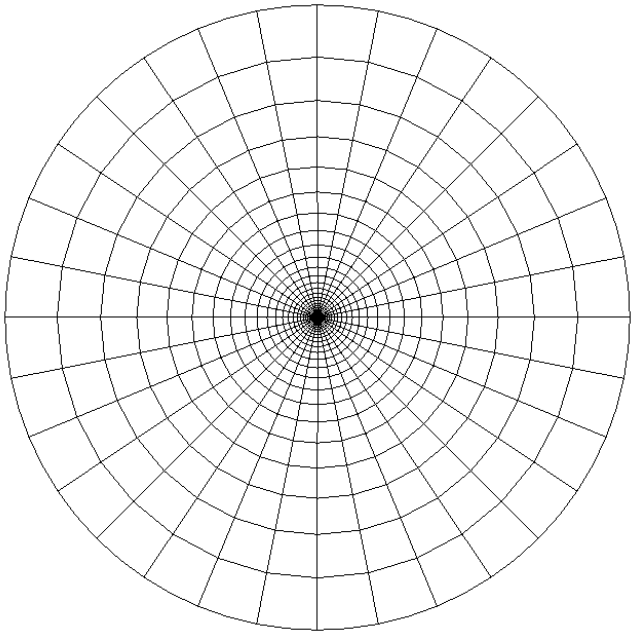
Figure 14.10: Numerical solution with isotropic box filtering for double Mach reflection problem at $t = 0.1$ for $P = 8$ and $N = 56 \times 224$.

discontinuity sensor. Finally, we can see that the solution also exhibits the 'starting error' associated with the spreading of the shock during the collocation projection of the initial condition. This is manifested most clearly in the strip originating from the upper right corner of the contour plot for v_y. Note that the flood plot reveals that this spurious contour separates levels that differ only slightly in magnitude.

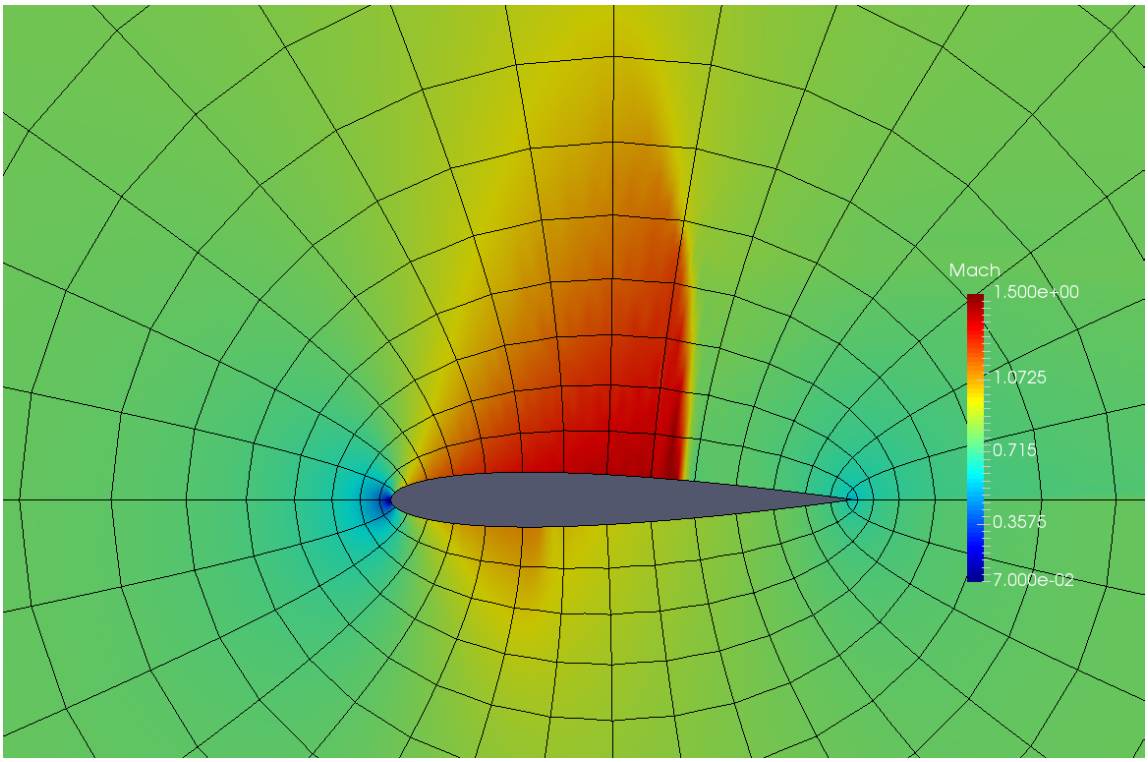
14.8 Transonic airfoil

So far in this chapter, the test cases have been concerned with time-dependent problems where the solution exhibits an unsteady or a quasi-steady state. Our final test investigates the performance of the proposed filtering framework for steady-state problems. Since the use of a discontinuity sensor ensures that smooth regions of the solution remain unchanged, we are interested in problems that exhibit shock discontinuities. As shown in the results that follow, this experiment highlights two important features of the proposed approach. First, the post-processing nature of the sensing-filtering operation leads to saturation in the equation residual which cannot be driven to machine precision despite the use of convergence acceleration techniques. This, in turn, leads to small scale fluctuations or periodicity in the integrated quantities. Second, despite the saturation, the generated pressure distribution, shock strength, shock location, and the lift and drag coefficients are very accurate. Consequently, a high-order FR scheme on a coarse grid can recover the same results as those of a second-order finite-volume scheme on a fine grid with two-orders of magnitude greater number of elements!.

In this test, we consider inviscid, Mach 0.8 flow past a NACA0012 airfoil at an angle of attack of 1.25 degrees. This problem has been studied extensively by Vassberg and Jameson [139] using well known flow solvers such as `FLO82`, `OVERFLOW` and `CFL3D` on an entire sequence of progressively refined grids. We follow their grid generation strategy to map a quasi circle to the airfoil geometry using the Karman-Trefftz conformal transformation [95]. The far-field boundary is taken to be a perfect circle at a radius of 100 chord lengths away from the leading edge. The computational domain is then populated with 32×32 quadrilateral elements in an essentially unstructured fill using Gmsh. The internal boundary of the domain, aligned with the airfoil surface, is represented by 25-node Lagrangian



(a) Mesh with 32×32 quadrilateral elements.



(b) Flood plot of Mach contours.

Figure 14.11: Numerical solution with box filtering for Mach 0.8 flow past a NACA0012 airfoil at an AoA of 1.25 degrees.

quadrilaterals whereas the interior of the domain is meshed with regular four-node quadrilaterals. A snapshot of the mesh is recorded in Fig. 14.11a. Slip-wall and characteristic boundary conditions are applied on the internal and far-field boundaries respectively. The numerical solution is obtained by marching the Euler equations, Eqns. (4.21 - 4.22), in pseudo-time using a scheme of polynomial order $P = 7$. In order to accelerate convergence, p-multigrid with a V-cycle is employed that uses all 8 available polynomial levels. The simulations are conducted using `zeFR`, a code developed at the Aerospace Computing Laboratory, which utilizes the Direct Flux Reconstruction framework for improved efficiency [109].

Figure 14.11b plots the Mach contours of the computed numerical solution after 500,000 iterations. We see that the scheme accurately recovers the strong shock on the suction side as well as the relatively weak shock on the pressure side of the airfoil. The action of the filter is to suppress instabilities arising due to Gibbs phenomenon which manifest as small disturbances in the regions close to the shocks. This can be seen clearly in Fig. 14.12 which plots the pressure distribution for each flux point located on the surface of the airfoil. Remarkably, even on such coarse grids, the scheme resolves sharp 1-flux-point shocks that are localized completely within one element. The use of a larger filter width parameter, γ, and a lower sensor threshold J would yield a smoother flow field with less fluctuations. However, the shocks would then be spread over multiple flux points.

We define the equation residual as the functional L^1 norm of the divergence of the mass flux, $\|\nabla \cdot (\rho \boldsymbol{u})\|_{L^1(\Omega)}$. Figure 14.13a plots the convergence history of this residual against the number of iterations on a logarithmic scale. We see that the solution saturates at a relatively large value despite the use of a convergence accelerator. This is on account of the choice that the sensing-filtering combination is incorporated as merely a post-processing operation applied after each time-step. The polynomial nature of the scheme has a tendency to introduce oscillations in the element containing the shock, whereas the filtering procedure inhibits them. In the absence of filtering, these oscillations lead to non-linear instabilities, as discussed in Chapter 12, leading to divergence. Clearly, either an alternative measure of residual must be utilized or the filtering procedure needs to be introduced as a limiter directly into the governing equation.

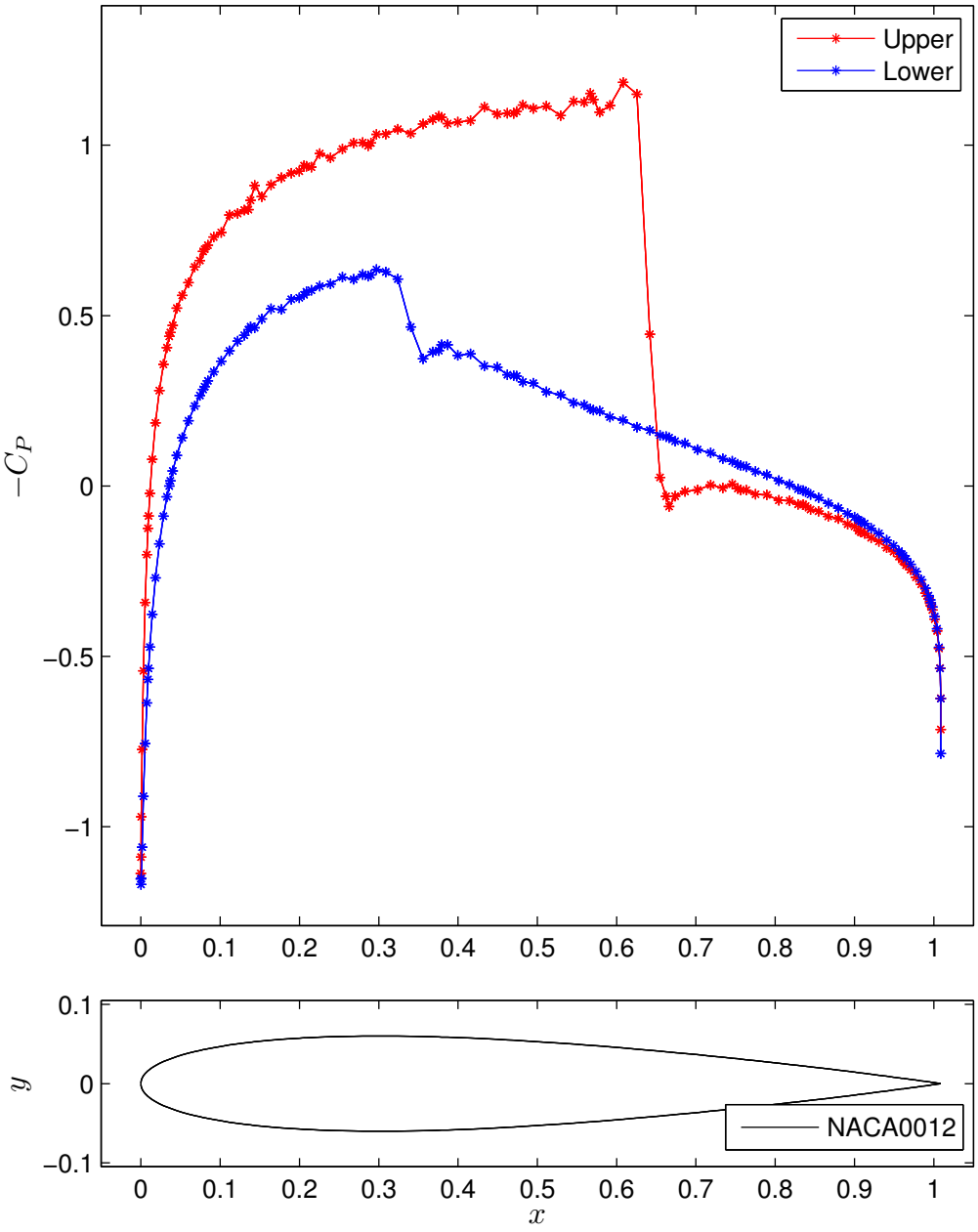
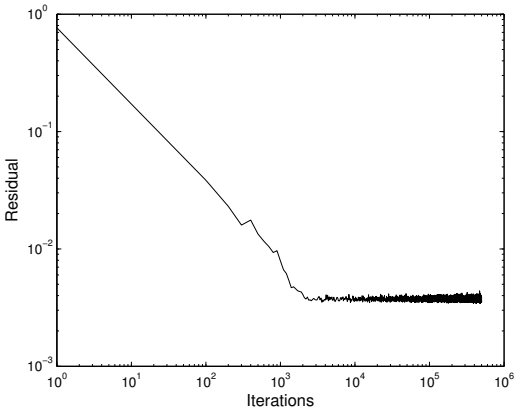
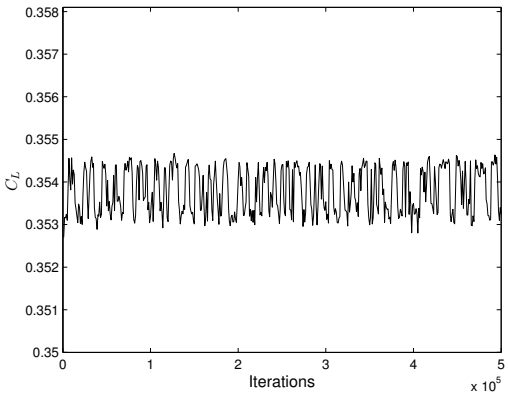


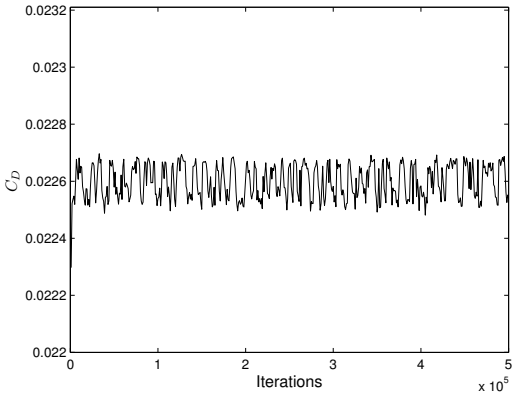
Figure 14.12: Distribution of the coefficient of pressure on the surface of the airfoil. The plotted symbols correspond to flux points.



(a) Equation residual.



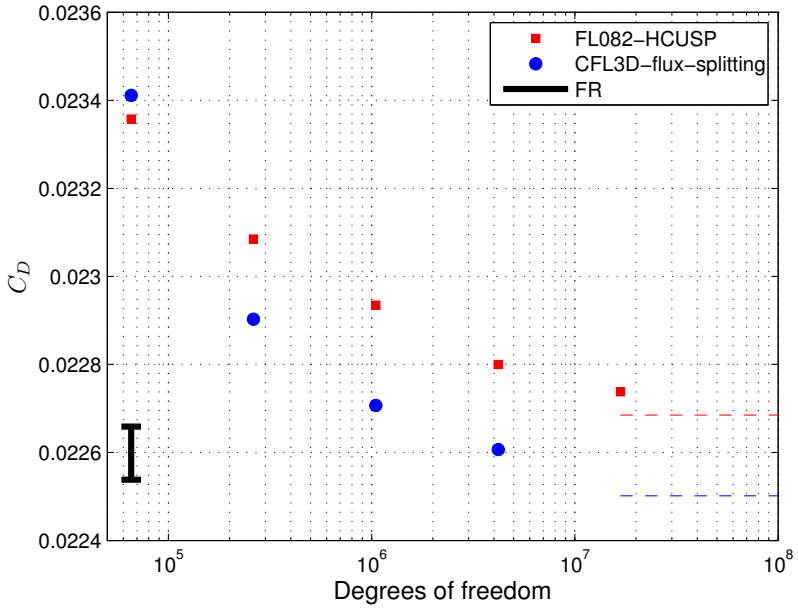
(b) Lift coefficient.



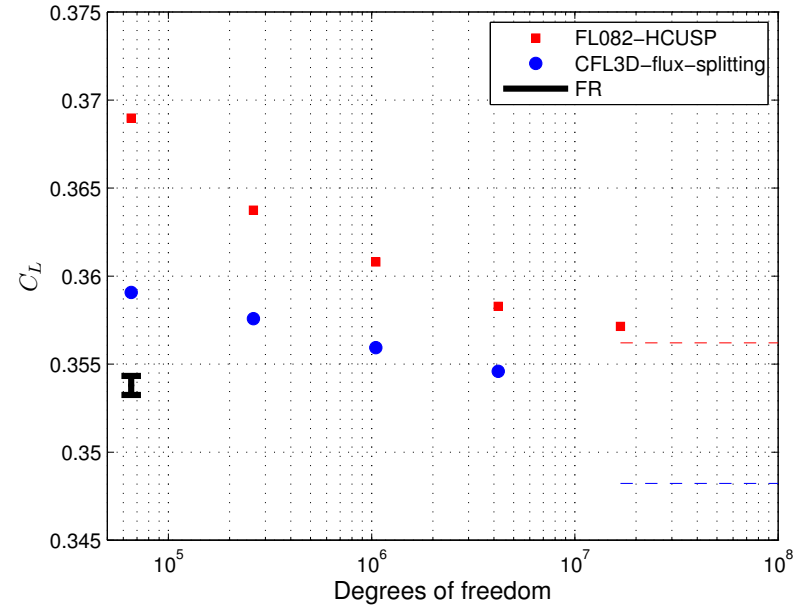
(c) Drag coefficient.

Figure 14.13: Convergence history of the residual and integrated coefficients.

Lack of convergence of the residual leads to small scale oscillations in the integrated quantities as well. This can be seen in Figs. 14.13b-c which plot the histories of the lift coefficient C_L and the drag coefficient C_D respectively. However, despite the fluctuations, when averaged in pseudo-time, both lift and drag match excellently with second-order calculations performed on high-resolution meshes. Figure 14.14 compares the results of this calculation with the grid-convergence results of Vassberg and Jameson [139] using three well established second-order computer codes. We see that the proposed sensing-filtering approach allows the FR scheme with $P = 7$ and 32^2 elements to obtain the same level of accuracy as second-order finite-volume schemes with 2048^2 elements.



(a) Drag coefficient



(b) Lift coefficient.

Figure 14.14: Comparison of integrated coefficients with the converged solutions of Vassberg and Jameson [139]. The dotted lines represent 'continuum' values and the error bar corresponds to the current FR calculation.

Appendix A

Impact of boundary conditions

Consider the advection of a scalar field, $u : \mathbb{R}_+ \mapsto \mathbb{C}$, driven by a sinusoidally forced boundary condition,

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0, \quad x > 0,\, t > 0, \tag{A.1}$$

$$u(x,0) = 0, \quad x \geq 0, \tag{A.2}$$

$$u(0,t) = \exp(i\omega t), \quad t > 0, \tag{A.3}$$

where $c \in \mathbb{R}$ is the constant wavespeed, and ω is the angular frequency of the boundary forcing. The analytical solution for this problem is obtained readily from the method of characteristics as

$$u(x,t) = \exp\left(i\omega\left(t - \frac{x}{c}\right)\right) \mathrm{I}_{\{x<ct\}}, \quad x \geq 0, \tag{A.4}$$

which shows that the disturbance injected at the origin with frequency ω is advected to the right at speed c.

In order to approximate the solution to Eqns. (A.1 - A.3) using an FR scheme of order P, the right half of the real line is partitioned into non-overlapping elements, $\mathbb{R}_+ = \bigcup_{n\geq 0} \Omega_n$ where $\Omega_n = [x_n, x_{n+1})$ with $x_0 = 0$, $x_1 = h$ and $x_n < x_{n+1}$ for $n \in \mathbb{Z}_+$. Let $\boldsymbol{x} = \frac{h}{2}(1+\boldsymbol{\xi})$ denote the $P+1$ solution points in the first element $\Omega_0 = [0,h)$. Then, for a fully upwinded

interface flux, the corrected solution polynomial in Ω_0 is given by

$$u^{\delta,c}(x,t) = u^{\delta}(x,t) + (\exp(i\omega t) - u^{\delta}(0^+,t))g_L(\xi(x)), \quad x \in \Omega_0, \tag{A.5}$$

which shows that the numerical update in the first element is completely independent from the solution values in all subsequent elements. Denoting $\boldsymbol{u}^{\boldsymbol{\delta}}$ as the vector of $P+1$ solution values on $\boldsymbol{x}$, Eqn. (A.5) leads to the semi-discrete nodal update equation

$$\frac{d\boldsymbol{u}^{\boldsymbol{\delta}}}{dt} + c\frac{2}{h}\boldsymbol{C_0}\boldsymbol{u}_{\boldsymbol{n}}^{\boldsymbol{\delta}} = -c\frac{2}{h}\exp(i\omega t)\boldsymbol{g}_{L,\xi}, \quad t > 0, \tag{A.6}$$

$$\boldsymbol{u}^{\boldsymbol{\delta}}(0) = \boldsymbol{0}, \tag{A.7}$$

$$\boldsymbol{C_0} = \boldsymbol{D} - \boldsymbol{g}_{L,\xi}\boldsymbol{\ell}_-^T, \tag{A.8}$$

where $\boldsymbol{D}$ is the polynomial differentiation matrix, $\boldsymbol{g}_{L,\xi}$ is the vector of derivatives of the left-boundary correction function in parent space, and $\boldsymbol{\ell}_-$ is the vector of Lagrange function evaluations on the left boundary. Explicit forms of these quantities are recorded in Chapter 2. Eqns. (A.6 - A.8) represent a linear dynamical system with an input term. It is therefore expedient to diagonalize the differentiation operator,

$$\boldsymbol{C_0} = \boldsymbol{W}\boldsymbol{M}\boldsymbol{W}^{-1}, \tag{A.9}$$

where $\boldsymbol{M}$ is the diagonal matrix of eigenvalues, $m_p \in \mathbb{C}$ for $p = 1,2,\ldots,P+1$, and $\boldsymbol{W} \in \mathbb{C}^{P+1\times P+1}$ is the dense matrix containing eigenvectors of the operator, $\boldsymbol{w}_p$. The analytical solution to Eqns. (A.6 - A.8) can now be written using convolution,

$$\begin{aligned}
\boldsymbol{u}^{\boldsymbol{\delta}}(t) &= \int_0^t \exp\left(-c\frac{2}{h}(t-\tau)\boldsymbol{C_0}\right)\left(-c\frac{2}{h}\exp(i\omega\tau)\boldsymbol{g}_{L,\xi}\right)\mathrm{d}\tau \\
&= -c\frac{2}{h}\int_0^t \exp\left(-c\frac{2}{h}(t-\tau)\boldsymbol{W}\boldsymbol{M}\boldsymbol{W}^{-1}\right)\boldsymbol{g}_{L,\xi}\exp(i\omega\tau)\mathrm{d}\tau \\
&= -c\frac{2}{h}\boldsymbol{W}\left[\int_0^t \exp\left(-c\frac{2}{h}(t-\tau)\boldsymbol{M}\right)\exp(i\omega\tau)\mathrm{d}\tau\right]\boldsymbol{W}^{-1}\boldsymbol{g}_{L,\xi}
\end{aligned}$$

$$= \exp(i\omega t)\boldsymbol{W}\mathscr{M}\boldsymbol{W}^{-1}\boldsymbol{g}_{L,\xi}, \tag{A.10}$$

where $\mathscr{M}$ is a diagonal matrix such that

$$\mathscr{M}_{pp} = \frac{1}{m_p + i\hat{\omega}}\left(\exp\left(-c\frac{2}{h}(m_p + i\hat{\omega})t\right) - 1\right), \quad p = 1,\ldots,P+1, \tag{A.11}$$

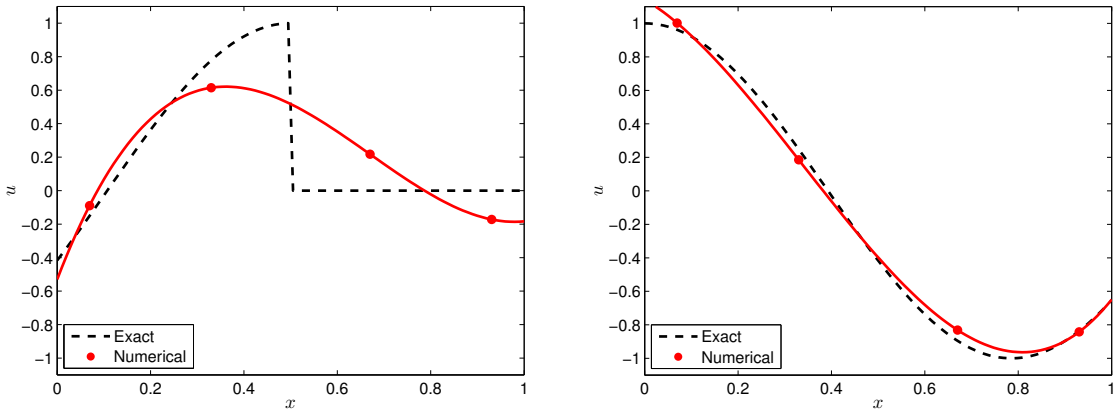
$$\hat{\omega} = \frac{\omega h}{2c}, \tag{A.12}$$

and $\hat{\omega}$ is the characteristic normalized frequency for the numerical problem. Eqn. (A.10) shows that the numerical solution is a combination of $P+1$ eigenmodes,

$$\boldsymbol{u}^{\delta}(t) = \exp(i\omega t)\sum_{p=1}^{P+1}\mathscr{M}_{pp}G_p\boldsymbol{w}_p, \tag{A.13}$$

where $\boldsymbol{G} = \boldsymbol{W}^{-1}\boldsymbol{g}_{L,\xi}$, and the condition for stability in the von Neumann sense is given by

$$m_p^{\mathrm{Re}} \geq 0 \text{ for } p = 1,2,\ldots,P+1. \tag{A.14}$$



(a) $t = \frac{1}{2}\frac{h}{c}$ (b) $t = 10\frac{2\pi}{\omega}$

Figure A.1: Real part of the numerical and exact solutions in the first element for the fully upwinded DG scheme of order $P = 3$ on Gauss-Legendre points.

Figure A.1 compares the real part of the numerical and exact solutions in Ω_0 at two different times for $\omega = 4$, $c = 1$ and $h = 1$ using an upwinded DG scheme of order $P = 3$. The

plot at $t = \frac{h}{2c}$ shows the start-up error in the propagation of the initial characteristic which manifests as a discontinuity due to the null initial datum. The other plot, at $t = 10T$ for $T = \frac{2\pi}{\omega}$, shows the quasi-steady behavior of the solution when the effect of the initial condition has been completely flushed out.

Using Eqns. (A.4) and (A.13), we can now evaluate the vector of error at the solution points,

$$\begin{aligned} \boldsymbol{e}(t) &:= \boldsymbol{u}^{\boldsymbol{\delta}}(t) - \boldsymbol{u}(t) \\ &= \exp(i\omega t) \left[\sum_{p=1}^{P+1} \mathscr{M}_{pp} G_p \boldsymbol{w}_p - \exp\left(-i\frac{\omega}{c}\boldsymbol{x}\right) \mathrm{I}_{\{x<ct\}} \right] \\ &= \exp(i\omega t) \sum_{p=1}^{P+1} \left[\mathscr{M}_{pp} G_p - \beta_p \mathrm{I}_{\{x<ct\}} \right] \boldsymbol{w}_p, \end{aligned} \tag{A.15}$$

where $\boldsymbol{\beta}(\hat{\omega})$ is the vector of expansion coefficients of the sinusoid when expressed in the basis of the eigenvectors,

$$\exp\left(-i\hat{\omega}(\mathbf{1} + \boldsymbol{\xi})\right) = \boldsymbol{W}\boldsymbol{\beta}. \tag{A.16}$$

Comparing with Eqns. (2.19) and (2.17), we see that these expansion coefficients are identical to the modal weights for the initial value problem (IVP) when $kc = -\omega$. However, there are two important differences. First, $\boldsymbol{\beta}$ in the current boundary-driven setting has no relation with the numerical solution. Second, the eigenvectors now correspond to the differentiation operator $\boldsymbol{C_0}$ instead of $\boldsymbol{Q}$ for the IVP, which for $\alpha = 0$ is given by $\boldsymbol{Q} = \boldsymbol{C_0} + \exp(-ikh)\boldsymbol{g}_{L,\xi}\boldsymbol{\ell}_+^T$. The distribution of these expansion coefficients is recorded in Fig. A.2 which plots $|\beta|^2$ against $\hat{\omega}$ for the 3 eigenmodes of the fully-upwinded DG scheme of order $P = 3$ using Gauss-Legendre solution points.

Figure A.3 plots the Euclidean norm of error, $\|\boldsymbol{e}(t)\|_2$ against the number of time periods for the same setting as for Fig. A.1. We see that there is an initial surge of error associated with the propagation of the first characteristic. Once this wavefront exits the element, the error plunges and settles to a steady amplitude.

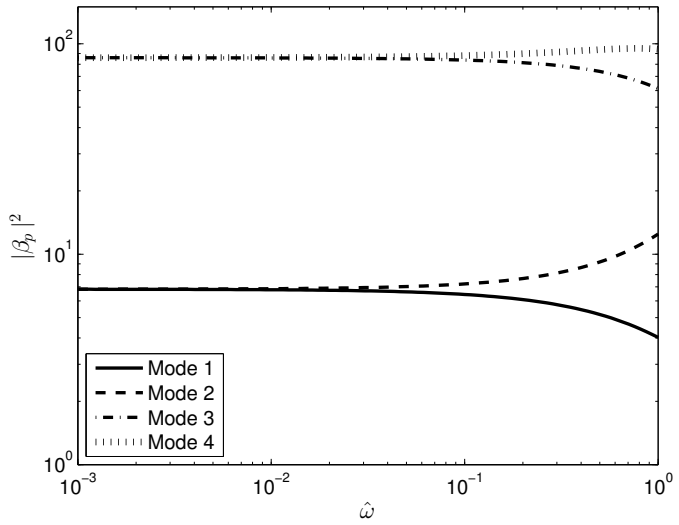


Figure A.2: $|\beta|^2$ for the 4 eigenmodes of the DG scheme of order $P = 3$ using Gauss-Legendre solution points.

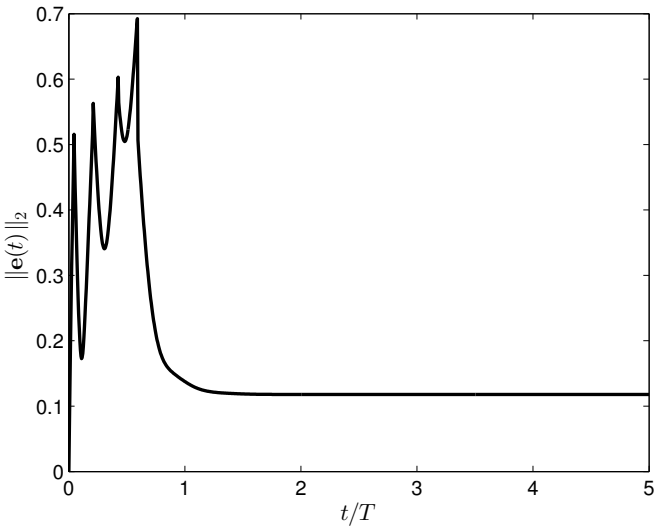


Figure A.3: Evolution of the norm of error in the first element for the fully upwinded DG scheme of order $P = 3$ on Gauss-Legendre points. Here, $T = \frac{2\pi}{\omega}$ is the time period of oscillations.

In order to proceed further with analysis, we restrict our attention to the quasi-steady solution acquired after the initial condition has been flushed out. Towards this end, note that full-upwinding ensures that $m_p^{\mathrm{Re}} > 0$ for $p = 1, 2, \ldots, P+1$. Then, for a finite h and $t \gg 0$,

$$\begin{aligned} \boldsymbol{e}(t) &\to \exp(i\omega t) \sum_{p=1}^{P+1} \left[\mathscr{M}_{pp} G_p - \beta_p\right] \boldsymbol{w}_p \\ &\to -\exp(i\omega t) \sum_{p=1}^{P+1} \left[\frac{G_p}{m_p + i\hat{\omega}} + \beta_p\right] \boldsymbol{w}_p, \end{aligned} \tag{A.17}$$

which shows that the quasi-steady error varies sinusoidally with frequency ω in time and the amplitude of oscillation in the pth mode is given by

$$A_p = \frac{G_p}{m_p + i\hat{\omega}} + \beta_p, \quad p = 1, \ldots, P+1. \tag{A.18}$$

Figure A.4 plots the magnitude of A_p against the normalized frequency for the $P+1$ modes in the case of DG for $P = 2$. We see that error vanishes with $\hat{\omega}$ showing that the FR formulation for this boundary condition is indeed consistent. Let $\boldsymbol{A} := (A_p)_{p=1,\ldots,P+1}$. Then the condition for consistency,

$$\boldsymbol{A} \to 0, \quad \text{as } \hat{\omega} \to 0, \tag{A.19}$$

can be written equivalently as

$$\boldsymbol{C_0 1} = -\boldsymbol{g}_{L,\xi}. \tag{A.20}$$

by noting that

$$\begin{aligned} \boldsymbol{W A} &= \boldsymbol{W} \left(\boldsymbol{M} + i\hat{\omega}\boldsymbol{I}\right)^{-1} \boldsymbol{G} + \boldsymbol{W}\boldsymbol{\beta} \\ &\to \boldsymbol{W M}^{-1} \boldsymbol{W}^{-1} \boldsymbol{g}_{L,\xi} + \boldsymbol{1} \\ &= \boldsymbol{C_0}^{-1} \boldsymbol{g}_{L,\xi} + \boldsymbol{1}. \end{aligned} \tag{A.21}$$

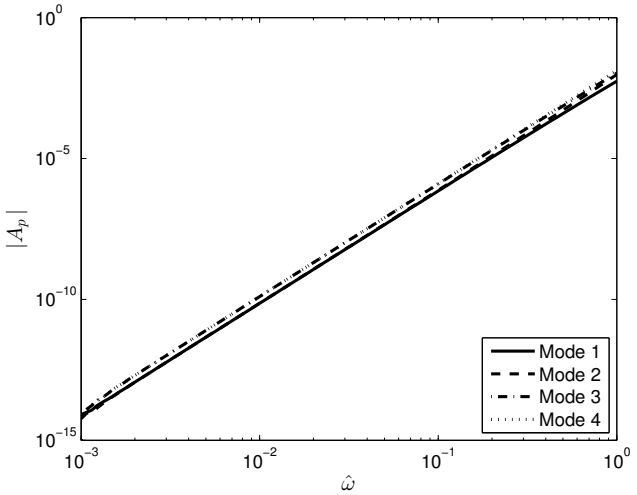


Figure A.4: Quasi-steady amplitude of error $|A_p|$ for the 4 numerical modes for the fully upwinded DG scheme of order $P = 3$.

We can also determine the rate of convergence of this amplitude with refinement in $\hat{\omega}$. Note that $\hat{\omega}$ varies proportionally to the grid-spacing h, so that such refinement is identical to an increase in the grid resolution. Alternatively, $\hat{\omega}$ can be reduced by decreasing the physical frequency ω or increasing the advection speed c. An eigensolution estimate for the rate of convergence is summarized in the following result.

Theorem A.1. *The quasi-steady amplitude of error in the numerical solution to Eqns. (A.6 - A.8) using a fully upwinded FR scheme of order P converges as*

$$\|\boldsymbol{e}(t)\|_2 \leq C \left(\frac{\omega h}{c}\right)^{\sigma}, \tag{A.22}$$

for $t \gg 0$, where C is some constant and σ is such that

$$\boldsymbol{C_0^j s^j} = \mathbf{1}, \quad j = 0, \ldots, \sigma - 1, \tag{A.23}$$

$$\left(s^j\right)_p = \frac{(1+\xi_p)^j}{j!}, \quad p = 1, \ldots, P+1. \tag{A.24}$$

Proof. Begin by expanding A_p in Eqn. (A.18) as

$$\begin{aligned} A_p &= G_p(m_p + i\hat{\omega})^{-1} + \beta_p \\ &= \frac{G_p}{m_p}\left(1 + i\frac{\hat{\omega}}{m_p}\right)^{-1} + \beta_p \\ &= G_p \sum_{j=0}^{\infty} \frac{(-i\hat{\omega})^j}{m_p^{j+1}} + \beta_p, \end{aligned}$$

so that vector of modal amplitudes can be written as

$$\begin{aligned} \boldsymbol{A} &= \sum_{j=0}^{\infty} (-i\hat{\omega})^j \boldsymbol{M}^{-(j+1)} \boldsymbol{G} + \boldsymbol{\beta} \\ &= \sum_{j=0}^{\infty} (-i\hat{\omega})^j \boldsymbol{M}^{-(j+1)} \boldsymbol{W}^{-1} \boldsymbol{g}_{L,\xi} + \boldsymbol{W}^{-1} \exp\left(-i\hat{\omega}(\mathbf{1}+\boldsymbol{\xi})\right) \\ &= \boldsymbol{W}^{-1} \left[\sum_{j=0}^{\infty} (-i\hat{\omega})^j \boldsymbol{C}_{\mathbf{0}}^{-(j+1)} \boldsymbol{g}_{L,\xi} + \exp\left(-i\hat{\omega}(\mathbf{1}+\boldsymbol{\xi})\right) \right] \\ &= \boldsymbol{W}^{-1} \sum_{j=0}^{\infty} (-i\hat{\omega})^j \left[\boldsymbol{C}_{\mathbf{0}}^{-(j+1)} \boldsymbol{g}_{L,\xi} + \frac{1}{j!} (\mathbf{1}+\boldsymbol{\xi})^j \right] \\ &= \boldsymbol{W}^{-1} \sum_{j=0}^{\infty} (-i\hat{\omega})^j \left[\frac{1}{j!} (\mathbf{1}+\boldsymbol{\xi})^j - \boldsymbol{C}_{\mathbf{0}}^{-j} \mathbf{1} \right], \end{aligned} \tag{A.25}$$

where we have used the condition of consistency, Eqn. (A.20) in the last step. The proof is completed by substituting Eqn. (A.25) in Eqn. (A.17), i.e. for $t \gg 0$,

$$\begin{aligned} \|\boldsymbol{e}(t)\|_2 &= \| - \exp(i\omega t) \boldsymbol{W} \boldsymbol{A} \|_2 \\ &= \| \boldsymbol{W} \boldsymbol{A} \|_2 \\ &= \left\| \sum_{j=0}^{\infty} (-i\hat{\omega})^j \left[\frac{1}{j!} (\mathbf{1}+\boldsymbol{\xi})^j - \boldsymbol{C}_{\mathbf{0}}^{-j} \mathbf{1} \right] \right\|_2 \\ &\leq \hat{\omega}^{\sigma} \sum_{j=0}^{\infty} |\hat{\omega}|^j \left\| \frac{1}{(j+\sigma)!} (\mathbf{1}+\boldsymbol{\xi})^{j+\sigma} - \boldsymbol{C}_{\mathbf{0}}^{-(j+\sigma)} \mathbf{1} \right\|_2 . \end{aligned}$$

□

Theorem A.1 provides a simple and direct mechanism for evaluating the rate of convergence of any upwinded FR scheme for this boundary-driven problem. Figure A.5 plots this rate, σ, as a function of the ESFR parameter C for $P = 2,3,4,5$. We see that the rate is identical to $P+1$ for all choices of correction functions. This observation has important ramifications for the convergence of the FR formulation. In particular, as in the case of steady-state hyperbolic problems, the long-time super-convergence derived in Chapter 4 for time-dependent problems is not relevant for boundary-driven solutions. In other words, for most practical applications, we expect the solution to converge at the short-time rate associated with projection into the polynomial space of the numerical solution.

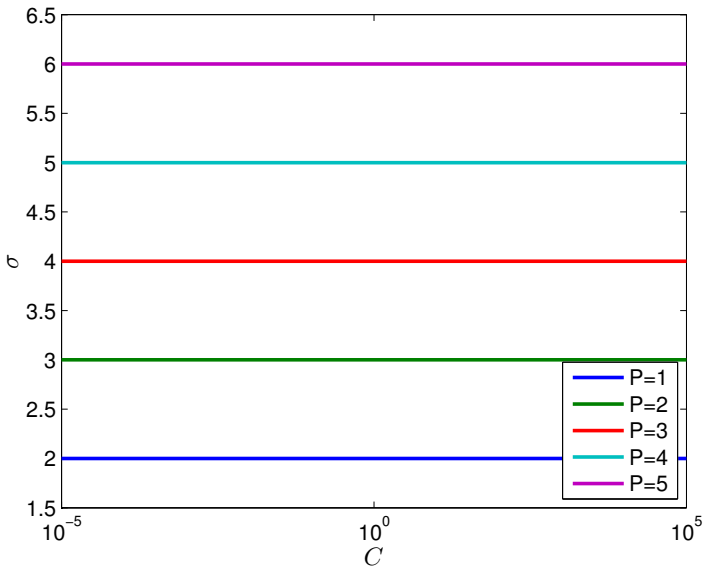


Figure A.5: Rate of convergence against the ESFR parameter C for $P = 1,\ldots,5$.

Appendix B

Group velocity and mesh reflection

Most of the theoretical developments in this dissertation have been concerned with the behavior of the numerical solution when the problem is well resolved. More precisely, the exposition has often utilized series expansions applicable only in the limit of asymptotic grid resolution. While this has generated valuable estimates for the rate of convergence, it fails to capture important non-asymptotic phenomena relevant to practical simulations which are almost never fully resolved. A characterization of the dispersion and dissipation errors for the entire range of resolvable wavenumbers was introduced in Chapter 2. Thereafter, spectrally optimal schemes were numerically derived in Chapter 5 where we integrated wave-propagation error over the entire range of wavenumbers. In this supplementary chapter, we show that FR formulation can potentially generate spurious numerical artifacts called q-waves [140, 131] in coarse regions of the computational domain. These numerical waves travel in the direction opposite to that prescribed by the underlying physics and can arise whenever the numerical solution is not adequately resolved [42] such as a sharp initial condition, a wave traveling through stretched elements, shock waves, boundary conditions [104] etc.

Returning to the model problem in Part I, consider the advection of a scalar $u : \mathbb{R} \mapsto \mathbb{C}$,

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R},\, t > 0, \tag{B.1}$$

$$u(x,0) = u_0(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right)\exp(-ik_0 x), \tag{B.2}$$

where the initial condition comprises a sinusoid of wavenumber k_0 filtered by a Gaussian of standard deviation $\sigma > 0$. Since u_0 is absolutely integrable, $u_0(x) \in L^1(\mathbb{R})$, it's Fourier transform exists and is obtained readily as

$$U_0(k) = (\mathscr{F} u_0)(k) = \sigma \exp\left(-\frac{\sigma^2}{2}(k-k0)^2\right), \tag{B.3}$$

which shows that, if $\sigma \gg 0$, then $U_0(k)$ is centered at k_0 and decays rapidly on either side. Hence, this problem deals with the advection of a packet of modulated waves. The analytical solution is non-dispersive, $u(x,t) = u_0(x-t)$, i.e. all component waves propagate at unit speed to the right. However, as shown in Eqn. (2.20), the numerical solution u^δ for an FR scheme of order P consists of $P+1$ modes all of which are dispersive. For instance, on a uniform mesh with element size h, the solution in the nth element is given by

$$\boldsymbol{u}_{\boldsymbol{n}}^{\boldsymbol{\delta}}(t) = \int_{\mathbb{R}} U_0(k) \exp(ikx_n) \sum_{p=1}^{P+1} \exp(-i\omega_p(k)t)\beta_p \boldsymbol{w}_{\boldsymbol{p}} \mathrm{d}k, \tag{B.4}$$

where $\omega_p(k) = k\lambda_p(k)$ is the effective angular velocity for the pth mode, and λ_p, β_p and $\boldsymbol{w}_{\boldsymbol{p}}$ are as described in Chapter 2. Substituting Eqn. (B.3) and expanding ω_p in a Taylor series about k_0, we see that

$$\begin{aligned}
\boldsymbol{u}_{\boldsymbol{n}}^{\boldsymbol{\delta}}(t) &= \sigma \sum_{p=1}^{P+1} \int_{\mathbb{R}} \exp\Bigg[-\frac{\sigma^2}{2}(k-k0)^2 + ikx_n - i\omega_p(k_0)t \\
&\qquad\qquad - i\left(\omega_p'(k_0)(k-k_0) + \mathscr{O}(k-k_0)^2\right)t \Bigg] \mathrm{d}k\, \beta_p \boldsymbol{w}_{\boldsymbol{p}} \\
&= \sigma \sum_{p=1}^{P+1} \int_{\mathbb{R}} \exp\left[ikx_n - i\omega_p(k_0)t - i\omega_p'(k_0)(k-k_0)t\right] \mathrm{d}k\, \beta_p \boldsymbol{w}_{\boldsymbol{p}} + \mathscr{O}(\sigma\exp(-\sigma^2)) \\
&\simeq \sigma \sum_{p=1}^{P+1} \exp\left(i(k_0 x_n - \omega_p(k_0)t)\right) \int_{\mathbb{R}} \exp\left[i(k-k_0)\left(x_n - \omega_p'(k_0)t\right)\right] \mathrm{d}k\, \beta_p \boldsymbol{w}_{\boldsymbol{p}},
\end{aligned} \tag{B.5}$$

when $\sigma \gg 0$. Hence, each of the $P+1$ modes consists of a packet of wavenumbers where

the envelope propagates at the speed

$$v_{g_p}(k_0) := \omega_p'(k_0) = \frac{\partial(k\lambda_p)}{\partial k}, \quad p = 1, \ldots, P+1, \tag{B.6}$$

referred to as the group velocity [131, 114] of the packet. Within the envelope, the disturbance propagates at the phase speed $\lambda_p(k_0)$. In other words, the consistent measure of dispersion for a wave-packet is given by the group velocity and not the phase speed.

The existence of multiple eigenmodes complicates the illustration of group velocity for a given FR scheme. In order to simplify matters, consider the energy-weighted quantity

$$\hat{v}_g(k_0) := \sum_{p=1}^{P+1} v_{g_p}(k_0)\hat{\beta}_p,$$
$$\hat{\beta}_p := \frac{|\beta_p|^2}{\sum_{p=1}^{P+1} |\beta_p|^2},$$

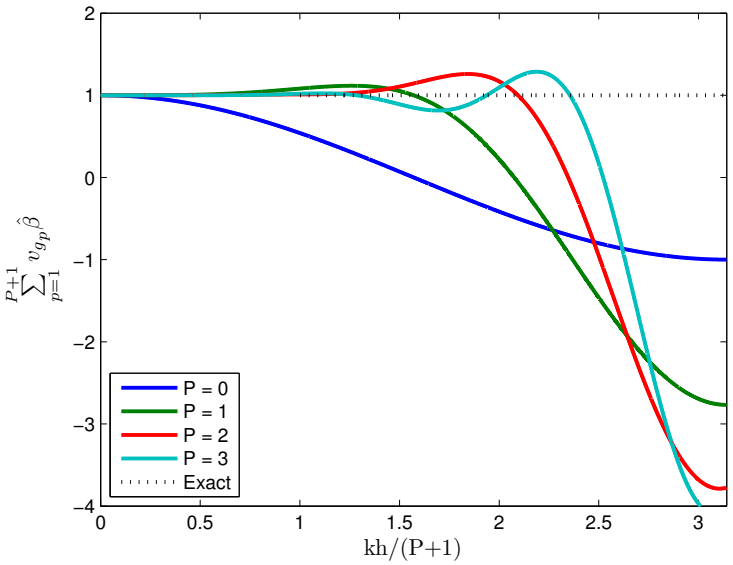
where the group velocity for mode p has been weighted by the fraction of energy it receives from a sinusoidal initial condition of wavenumber k_0 (c.f. Chapter 2). Figure B.1 plots $\hat{v}_g(k_0)$ for $P = 0, \ldots, 3$ in the case of fully upwinded and centered DG schemes. Several observations are now in line.

1. In the asymptotic limit, $kh \to 0$, the numerical group velocity follows the analytical expectation, i.e. dispersion vanishes with sufficient resolution.

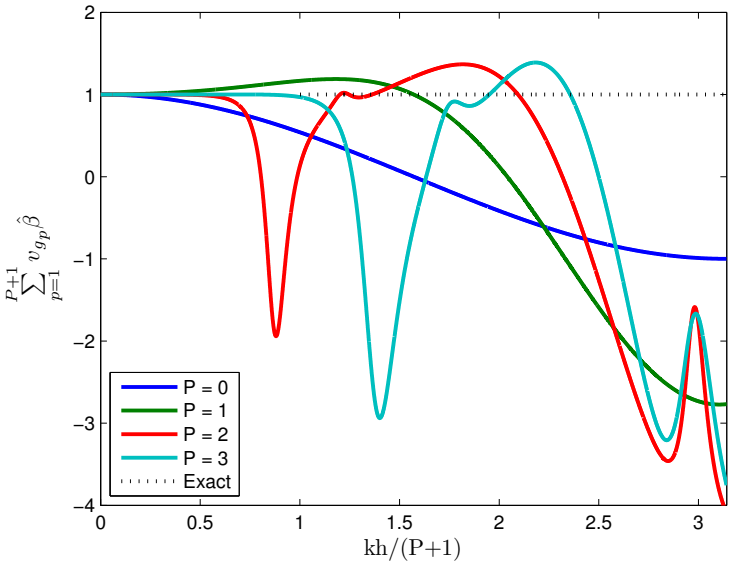
2. Numerical group velocity can indeed be negative, resulting in q-waves that move to the left. These spurious wavefronts can incorrectly trigger flow instabilities [112].

3. Just like in finite-difference schemes [88], the maximum speed of such q-waves increases with the order of the scheme. Consequently, q-waves arising out of reflections would travel faster when P is higher.

4. q-waves can, in turn, be reflected back from inflow boundaries turning into p-waves [104] which are low-wavenumber waves that travel to the right. These p-waves would get reflected back again forming a feedback loop [140].



(a) Fully upwind interface flux, $\alpha = 0$



(b) Centered interface flux, $\alpha = \frac{1}{2}$

Figure B.1: Energy weighted group velocity, $\hat{v}_g$, for the DG scheme of orders $P = 0, \ldots, 3$ using Gauss-Legendre solution points.

5. For upwinded schemes, such q-waves are rapidly dissipated and for full-upwinding, cannot pollute the solution upstream. This is expected since any element for a fully-upwinded scheme is completely independent of all elements downstream. However, upwinded schemes pay for this benefit through numerical dissipation.

6. For centered schemes, such q-waves propagate without any numerical dissipation. Moreover, for higher polynomial order, the group velocity can be negative for moderate wavenumbers as well! Clearly, fully centered FR schemes are not a feasible option for time-dependent problems.

With these analytical expectations in place, we now perform a simple numerical experiment to demonstrate the presence of q-waves. Towards this end, let Eqn. (B.1) hold on the bi-unit interval $\Omega = [-1,1]$,

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \quad x \in (-1,1),\, t > 0, \tag{B.7}$$

$$u(x,0) = \exp\left(-\frac{(x+0.5)^2}{2\sigma^2}\right), \quad x \in [-1,1], \tag{B.8}$$

$$u(-1,t) = 0, \quad t > 0, \tag{B.9}$$

$$u(1,t) = u(1^-,t), \quad t > 0, \tag{B.10}$$

where a null Dirichlet condition is supplied on the left boundary and an outflow condition is imposed on the right boundary. The standard deviation σ is fixed as 0.01. The computational domain consists of a piecewise uniform grid where the element size is given by

$$h = \begin{cases} \frac{1}{50}, & x < 0, \\ \frac{1}{5}, & x > 0. \end{cases} \tag{B.11}$$

Correspondingly, the Gaussian is well resolved in the left half of the domain which allocates 10 elements per two standard deviations, but poorly resolved in the right half which can only afford 1 element per two standard deviations. The correction functions are chosen to recover the DG scheme of order $P = 2$ and the solution is time-marched on Gauss-Legendre points using the fourth order four-stage RK scheme with a fixed time step of

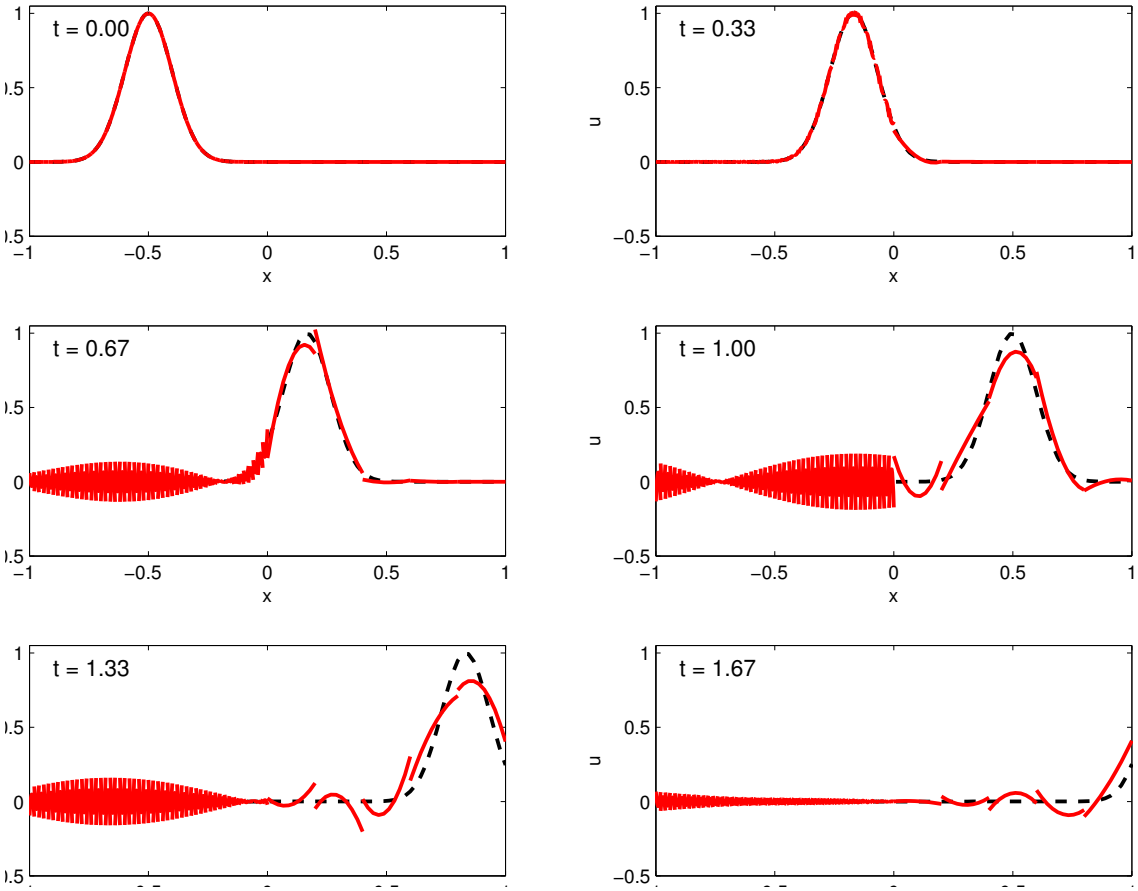


Figure B.2: Snapshots of the numerical solution for the centered DG scheme of order $P = 2$. The dotted line represents the exact solution.

$\Delta t = 0.00235$. We are interested in the behavior of the numerical solution as it crosses the interface between the two halves of the grid.

Figure B.2 provides snapshots of the numerical and exact solutions when the interface fluxes are centered, $\alpha = \frac{1}{2}$. We see that the Gaussian is advected accurately on the fine grid with negligible dispersion. However, on reaching the coarse grid, part of the Gaussian energy is reflected upstream through q-waves. These are the wavenumber packets for which the group velocity is negative. The Dirichlet condition on the left boundary absorbs much of these wavefronts and the rest is reflected back into the domain as p-waves.

The generation of such spurious waves can be completely inhibited by upwinding the

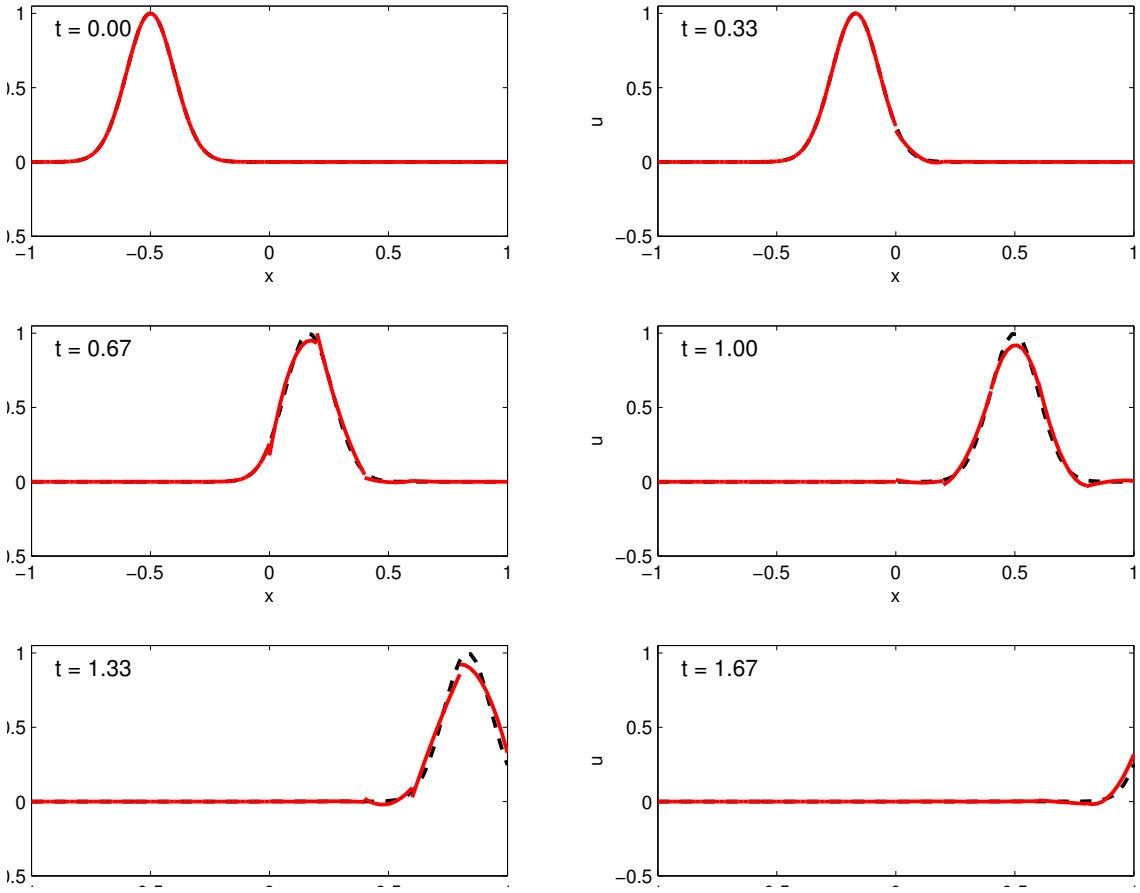


Figure B.3: Snapshots of the numerical solution for the fully-upwinded, $\alpha = 0$, DG scheme of order $P = 2$. The dotted line represents the exact solution.

interface fluxes. This can be observed from Fig. B.3 which plots snapshots for the same scheme but now with $\alpha = 0$. The only effect in this case is the added numerical dissipation on the coarse grid.

Alternatively, the amplitude of these waves can be suppressed by increasing the polynomial order which effectively increases the resolution of the simulation. This can be observed from Fig. B.4 which plots snapshots for the centered DG scheme but now with $P = 3$.

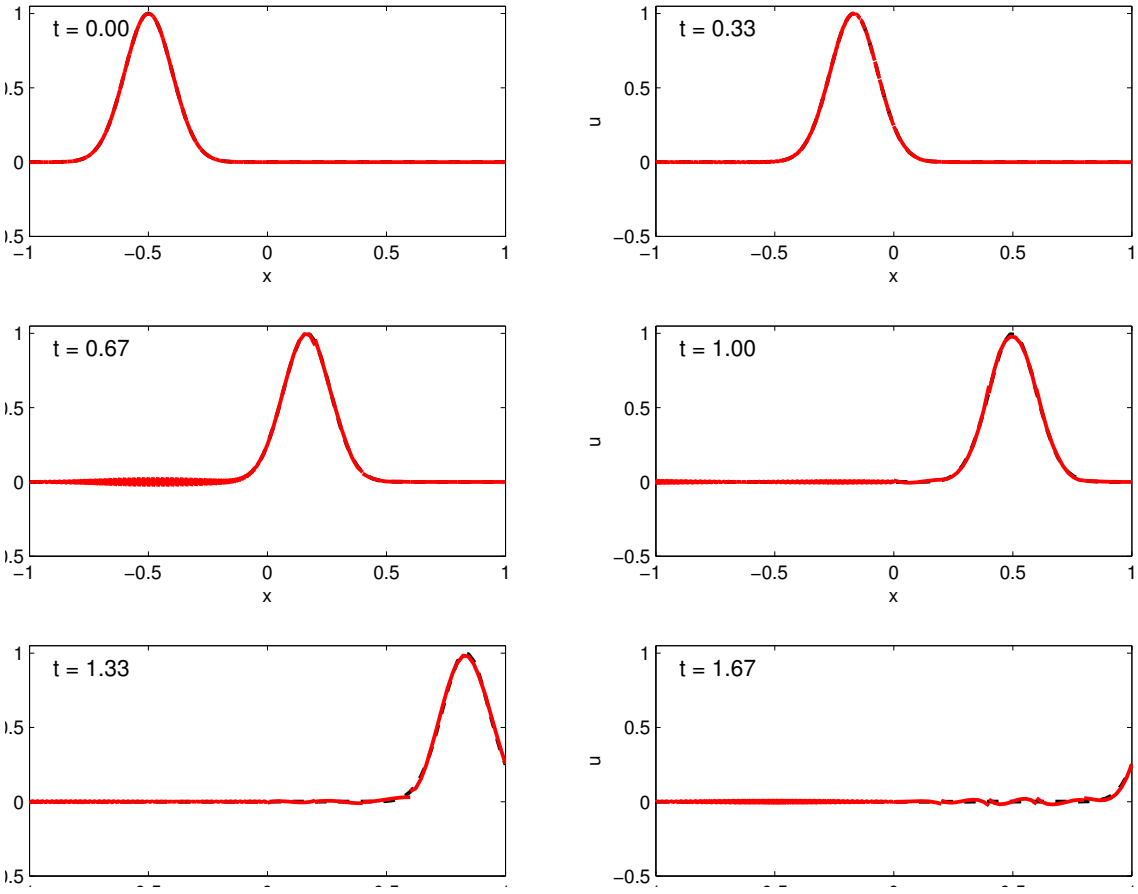


Figure B.4: Snapshots of the numerical solution for the centered DG scheme of order $P = 3$. The dotted line represents the exact solution.

Bibliography

[1] Yves Adam. Highly accurate compact implicit methods and boundary conditions. *Journal of Computational Physics*, 24(1):10–22, 1977.

[2] NA Adams and S Stolz. A subgrid-scale deconvolution approach for shock capturing. *Journal of Computational Physics*, 178(2):391–426, 2002.

[3] Slimane Adjerid, Karen D Devine, Joseph E Flaherty, and Lilia Krivodonova. A posteriori error estimation for discontinuous galerkin solutions of hyperbolic problems. *Computer methods in applied mechanics and engineering*, 191(11):1097–1112, 2002.

[4] M Ainsworth, P Monk, and W Muniz. Dispersive and dissipative properties of discontinuous Galerkin finite element methods for the second-order wave equation. *Journal of Scientific Computing*, 27(1-3):5–40, 2006.

[5] Mark Ainsworth. Discrete dispersion relation for hp-version finite element approximation at high wave number. *SIAM Journal on Numerical Analysis*, 42(2):553–575, 2004.

[6] Mark Ainsworth. Dispersive and dissipative behaviour of high order discontinuous Galerkin finite element methods. *Journal of Computational Physics*, 198(1):106–130, 2004.

[7] Douglas N Arnold, Franco Brezzi, Bernardo Cockburn, and L Donatella Marini. Unified analysis of discontinuous galerkin methods for elliptic problems. *SIAM journal on numerical analysis*, 39(5):1749–1779, 2002.

[8] Kartikey Asthana and Antony Jameson. High-order flux reconstruction schemes with minimal dispersion and dissipation. *Journal of Scientific Computing*, pages 1–32, 2014.

[9] Kartikey Asthana, Jerry Watkins, and Antony Jameson. On consistency, convergence and rate of convergence of flux reconstruction for time-dependent problems. *SIAM Journal of Numerical Analysis*, Under review, submitted July 2015.

[10] Harold L Atkins and Brian Helenbrook. Super-convergence of discontinuous galerkin method applied to the navier–stokes equations. *AIAA Paper*, 3787:2009, 2009.

[11] Garrett E Barter and David L Darmofal. Shock capturing with higher-order, pde-based artificial viscosity. *AIAA paper*, 3823:2007, 2007.

[12] Francesco Bassi and Stefano Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible navier–stokes equations. *Journal of computational physics*, 131(2):267–279, 1997.

[13] Roland Becker, Peter Hansbo, and Mats G Larson. Energy norm a posteriori error estimation for discontinuous galerkin methods. *Computer Methods in Applied Mechanics and Engineering*, 192(5):723–733, 2003.

[14] C Béraudo, A Fortin, T Coupez, Y Demay, B Vergnes, and JF Agassant. A finite element method for computing the flow of multi-mode viscoelastic fluids: comparison with experiments. *Journal of non-newtonian fluid mechanics*, 75(1):1–23, 1998.

[15] Christine Bernardi and Yvon Maday. Polynomial interpolation results in sobolev spaces. *Journal of computational and applied mathematics*, 43(1):53–80, 1992.

[16] Rupak Biswas, Karen D Devine, and Joseph E Flaherty. Parallel, adaptive finite element methods for conservation laws. *Applied Numerical Mathematics*, 14(1):255–283, 1994.

[17] Hugh M Blackburn and S Schmidt. Spectral element filtering techniques for large eddy simulation with dynamic estimation. *Journal of Computational Physics*, 186(2):610–629, 2003.

[18] Christophe Bogey and Christophe Bailly. A family of low dispersive and low dissipative explicit schemes for flow and noise computations. *Journal of Computational Physics*, 194(1):194–214, 2004.

[19] Christophe Bogey, Nicolas De Cacqueray, and Christophe Bailly. A shock-capturing methodology based on adaptative spatial filtering for high-order non-linear computations. *Journal of Computational Physics*, 228(5):1447–1465, 2009.

[20] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[21] J.R. Bull and A. Jameson. Simulation of the compressible Taylor Green vortex using high-order flux reconstruction schemes. In *AIAA Aviation 2014, 16–20 June, Atlanta, Georgia*, June 2014. unpublished.

[22] A Burbeau, P Sagaut, and Ch-H Bruneau. A problem-independent limiter for high-order runge–kutta discontinuous galerkin methods. *Journal of Computational Physics*, 169(1):111–150, 2001.

[23] Wei Cai, David Gottlieb, and Chi-Wang Shu. Essentially nonoscillatory spectral fourier methods for shock wave calculations. *Mathematics of Computation*, 52(186):389–410, 1989.

[24] C Canuto and A Quarteroni. Approximation results for orthogonal polynomials in sobolev spaces. *Mathematics of Computation*, 38(157):67–86, 1982.

[25] Mark H Carpenter, Travis C Fisher, Eric J Nielsen, and Steven H Frankel. Entropy stable spectral collocation schemes for the navier–stokes equations: Discontinuous interfaces. *SIAM Journal on Scientific Computing*, 36(5):B835–B867, 2014.

[26] Mark H Carpenter and Christopher A Kennedy. Fourth-order 2n-storage runge-kutta schemes. *Nasa tm*, 109112:871–885, 1994.

[27] Corentin Carton de Wiart, Koen Hillewaert, Matthieu Duponcheel, and Grégoire Winckelmans. Assessment of a discontinuous galerkin method for the simulation of vortical flows at high reynolds number. *International Journal for Numerical Methods in Fluids*, 74(7):469–493, 2014.

[28] P Castonguay, DM Williams, PE Vincent, and A Jameson. Energy stable flux reconstruction schemes for advection–diffusion problems. *Computer Methods in Applied Mechanics and Engineering*, 267:400–417, 2013.

[29] P. Castonguay, D.M. Williams, P.E. Vincent, M. R. López-Morales, and A. Jameson. On the development of a high-order, multi-gpu enabled, compressible viscous flow solver for mixed unstructured grids. In *AIAA Computational Fluid Dynamics Conference, AIAA-2011-3229*, 2011.

[30] Patrice Castonguay. *High-order energy stable flux reconstruction schemes for fluid flow simulations on unstructured grids*. PhD thesis, Citeseer, 2012.

[31] Patrice Castonguay, PE Vincent, and Antony Jameson. Application of high-order energy stable flux reconstruction schemes to the euler equations. In *49th AIAA Aerospace Sciences Meeting*, volume 686, 2011.

[32] Patrice Castonguay, Peter E Vincent, and Antony Jameson. A new class of high-order energy stable flux reconstruction schemes for triangular elements. *Journal of Scientific Computing*, 51(1):224–256, 2012.

[33] Yingda Cheng, Irene M Gamba, Armando Majorana, and Chi-Wang Shu. Discontinuous galerkin solver for boltzmann-poisson transients. *Journal of Computational Electronics*, 7(3):119–123, 2008.

[34] Yingda Cheng and Chi-Wang Shu. A discontinuous galerkin finite element method for directly solving the hamilton–jacobi equations. *Journal of Computational Physics*, 223(1):398–415, 2007.

[35] Yingda Cheng and Chi-Wang Shu. Superconvergence and time evolution of discontinuous galerkin finite element solutions. *Journal of Computational Physics*, 227(22):9612–9627, 2008.

[36] Yingda Cheng and Chi-Wang Shu. Superconvergence of discontinuous galerkin and local discontinuous galerkin schemes for linear hyperbolic and convection-diffusion equations in one space dimension. *SIAM Journal on Numerical Analysis*, 47(6):4044–4072, 2010.

[37] Bernardo Cockburn, Claes Johnson, C-W Shu, Eitan Tadmor, and Alfio Quarteroni. Advanced numerical approximation of nonlinear hyperbolic equations: lectures given at the 2nd session of the centro internazionale matematico estivo. 1998.

[38] Bernardo Cockburn, Guido Kanschat, Ilaria Perugia, and Dominik Schötzau. Superconvergence of the local discontinuous galerkin method for elliptic problems on cartesian grids. *SIAM Journal on Numerical Analysis*, 39(1):264–285, 2001.

[39] Bernardo Cockburn, George E Karniadakis, and Chi-Wang Shu. *The development of discontinuous Galerkin methods*. Springer, 2000.

[40] Bernardo Cockburn, Mitchell Luskin, Chi-Wang Shu, and Endre Süli. *Post-processing of Galerkin methods for hyperbolic problems*. Springer, 2000.

[41] Thomas Frederick Coleman, Mary Ann Branch, and Andrew Grace. *Optimization toolbox for Use with matlab: User's guide*. Mathworks Incorporated, 1999.

[42] Tim Colonius and Sanjiva K Lele. Computational aeroacoustics: progress on nonlinear problems of sound generation. *Progress in Aerospace sciences*, 40(6):345–416, 2004.

[43] Andrew W Cook and William H Cabot. Hyperviscosity for shock-turbulence interactions. *Journal of Computational Physics*, 203(2):379–385, 2005.

[44] James DeBonis. Solutions of the taylor-green vortex problem using high-resolution explicit finite difference methods. *AIAA paper*, 382:2013, 2013.

[45] Richard S Falk. Analysis of finite element methods for linear hyperbolic problems. In *Discontinuous Galerkin Methods*, pages 103–112. Springer, 2000.

[46] Krzysztof J Fidkowski and David L Darmofal. Review of output-based error estimation and mesh adaptation in computational fluid dynamics. *AIAA journal*, 49(4):673–694, 2011.

[47] Krzysztof J Fidkowski, Todd A Oliver, James Lu, and David L Darmofal. p-multigrid solution of high-order discontinuous galerkin discretizations of the compressible navier–stokes equations. *Journal of Computational Physics*, 207(1):92–113, 2005.

[48] Travis C Fisher and Mark H Carpenter. High-order entropy stable finite difference schemes for nonlinear conservation laws: Finite domains. *Journal of Computational Physics*, 252:518–557, 2013.

[49] Haiyang Gao and ZJ Wang. A conservative correction procedure via reconstruction formulation with the chain-rule divergence evaluation. *Journal of Computational Physics*, 232(1):7–13, 2013.

[50] Anne Gelb and Dennis Cates. Detection of edges in spectral data iiirefinement of the concentration method. *Journal of Scientific Computing*, 36(1):1–43, 2008.

[51] Anne Gelb and Eitan Tadmor. Detection of edges in spectral data. *Applied and computational harmonic analysis*, 7(1):101–135, 1999.

[52] Anne Gelb and Eitan Tadmor. Detection of edges in spectral data ii. nonlinear enhancement. *SIAM Journal on Numerical Analysis*, 38(4):1389–1408, 2000.

[53] Sergei Konstantinovich Godunov. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik*, 89(3):271–306, 1959.

[54] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, 2013.

[55] David Gottlieb and Jan S Hesthaven. Spectral methods for hyperbolic problems. *Journal of Computational and Applied Mathematics*, 128(1):83–131, 2001.

[56] David Gottlieb, Steven A Orszag, and CAMBRIDGE HYDRODYNAMICS INC MA. *Numerical analysis of spectral methods*, volume 2. SIAM, 1977.

[57] David Gottlieb and Chi-Wang Shu. On the gibbs phenomenon and its resolution. *SIAM review*, 39(4):644–668, 1997.

[58] Zigo Haras and Shlomo Ta'asan. Finite difference schemes for long-time integration. *Journal of Computational Physics*, 114(2):265–279, 1994.

[59] Ralf Hartmann and Paul Houston. Adaptive discontinuous galerkin finite element methods for the compressible euler equations. *Journal of Computational Physics*, 183(2):508–532, 2002.

[60] Jan Hesthaven and Robert Kirby. Filtering in legendre spectral methods. *Mathematics of Computation*, 77(263):1425–1452, 2008.

[61] Jan S Hesthaven, Sigal Gottlieb, and David Gottlieb. *Spectral methods for time-dependent problems*, volume 21. Cambridge University Press, 2007.

[62] Jan S Hesthaven and Tim Warburton. *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*, volume 54. Springer Science & Business Media, 2007.

[63] Jan S Hesthaven and Timothy Warburton. Nodal high-order methods on unstructured grids: I. time-domain solution of maxwell's equations. *Journal of Computational Physics*, 181(1):186–221, 2002.

[64] Richard S Hirsh. Higher order accurate difference solutions of fluid mechanics problems by a compact differencing technique. *Journal of computational physics*, 19(1):90–109, 1975.

[65] Fang Q Hu and Harold L Atkins. Eigensolution analysis of the discontinuous Galerkin method with nonuniform grids: I. one space dimension. *Journal of Computational Physics*, 182(2):516–545, 2002.

[66] Fang Q Hu and Harold L Atkins. Two-dimensional wave analysis of the discontinuous Galerkin method with non-uniform grids and boundary conditions. In *Proceedings of the 8th AIAA/CEAS Aeroacoustics Conference*, 2002.

[67] Fang Q Hu, MY Hussaini, and Patrick Rasetarinera. An analysis of the discontinuous Galerkin method for wave propagation problems. *Journal of Computational Physics*, 151(2):921–946, 1999.

[68] Thomas JR Hughes. *The finite element method: linear static and dynamic finite element analysis*. Courier Corporation, 2012.

[69] HT Huynh. A flux reconstruction approach to high-order schemes including discontinuous Galerkin methods. *AIAA paper*, 4079:2007, 2007.

[70] HT Huynh. High-order methods including discontinuous galerkin by reconstructions on triangular meshes. *AIAA Paper*, 44:2011, 2011.

[71] Hung T Huynh. A reconstruction approach to high-order schemes including discontinuous Galerkin for diffusion. *AIAA paper*, 403:2009, 2009.

[72] Antony Jameson. Airfoils admitting non-unique solutions of the euler equations. *AIAA paper*, 91(1625):1–13, 1991.

[73] Antony Jameson. Time dependent calculations using multigrid, with applications to unsteady flows past airfoils and wings. *AIAA paper*, 1596:1991, 1991.

[74] Antony Jameson and Kui Ou. 50 years of transonic aircraft design. *Progress in Aerospace Sciences*, 47(5):308–318, 2011.

[75] Antony Jameson, Wolfgang Schmidt, Eli Turkel, et al. Numerical solutions of the euler equations by finite volume methods using runge-kutta time-stepping schemes. *AIAA paper*, 1259:1981, 1981.

[76] Antony Jameson, Peter E Vincent, and Patrice Castonguay. On the non-linear stability of flux reconstruction schemes. *Journal of Scientific Computing*, 50(2):434–445, 2012.

[77] Guang-Shan Jiang and Chi-Wang Shu. Efficient implementation of weighted eno schemes. Technical report, DTIC Document, 1995.

[78] Claes Johnson and Juhani Pitkäranta. An analysis of the discontinuous galerkin method for a scalar hyperbolic equation. *Mathematics of computation*, 46(173):1–26, 1986.

[79] Ohannes A Karakashian and Frederic Pascal. A posteriori error estimates for a discontinuous galerkin approximation of second-order elliptic problems. *SIAM Journal on Numerical Analysis*, 41(6):2374–2399, 2003.

[80] Soshi Kawai and Sanjiva K Lele. Localized artificial diffusivity scheme for discontinuity capturing on curvilinear meshes. *Journal of Computational Physics*, 227(22):9498–9526, 2008.

[81] David A Kopriva and John H Kolias. A conservative staggered-grid Chebyshev multidomain method for compressible flow. Technical report, DTIC Document, 1995.

[82] Lilia Krivodonova. Limiters for high-order discontinuous galerkin methods. *Journal of Computational Physics*, 226(1):879–896, 2007.

[83] 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):323–338, 2004.

[84] Ethan J Kubatko, Benjamin A Yeager, and David I Ketcheson. Optimal strong-stability-preserving runge–kutta time discretizations for discontinuous galerkin methods. *Journal of Scientific Computing*, 60(2):313–344, 2014.

[85] Dmitri Kuzmin. Hierarchical slope limiting in explicit and implicit discontinuous galerkin methods. *Journal of Computational Physics*, 257:1140–1162, 2014.

[86] Peter D Lax. Hyperbolic systems of conservation laws ii. *Communications on Pure and Applied Mathematics*, 10(4):537–566, 1957.

[87] Sanjiva K Lele. Direct numerical simulation of compressible free shear flows. *AIAA paper*, 374:1989, 1989.

[88] Sanjiva K Lele. Compact finite difference schemes with spectral-like resolution. *Journal of Computational Physics*, 103(1):16–42, 1992.

[89] P Lesaint and Pierre-Arnaud Raviart. On a finite element method for solving the neutron transport equation. *Mathematical aspects of finite elements in partial differential equations*, (33):89–123, 1974.

[90] Chunlei Liang, Antony Jameson, and ZJ Wang. Spectral difference method for compressible flow on unstructured grids with mixed elements. *Journal of Computational Physics*, 228(8):2847–2858, 2009.

[91] Yen Liu, Marcel Vinokur, and ZJ Wang. Spectral difference method for unstructured grids I: basic formulation. *Journal of Computational Physics*, 216(2):780–801, 2006.

[92] M López-Morales, Jonathan Bull, Jacob Crabill, Thomas D Economon, David Manosalvas, Joshua Romero, Abhishek Sheshadri, JE Watkins, David Williams, Francisco Palacios, et al. Verification and validation of hifiles: a high-order les unstructured solver on multi-gpu platforms. In *32nd AIAA applied aerodynamics conference, Atlanta, Georgia, USA*, pages 16–20, 2014.

[93] Manuel R López-Morales, Kartikey Asthana, and Antony Jameson. Flux reconstruction schemes with corrected fluxes continuous in *m* derivatives. *in preparation*, 2015.

[94] A Meister, S Ortleb, and Th Sonar. Application of spectral filtering to discontinuous galerkin methods on triangulations. *Numerical Methods for Partial Differential Equations*, 28(6):1840–1868, 2012.

[95] Louis Melville Milne-Thomson. *Theoretical hydrodynamics*. Courier Corporation, 1968.

[96] Hanieh Mirzaee, Jennifer K Ryan, and Robert M Kirby. Smoothness-increasing accuracy-conserving (siac) filters for discontinuous galerkin solutions: Application to structured tetrahedral meshes. *Journal of Scientific Computing*, 58(3):690–704, 2014.

[97] Parviz Moin and Krishnan Mahesh. Direct numerical simulation: a tool in turbulence research. *Annual review of fluid mechanics*, 30(1):539–578, 1998.

[98] RC Moura, SJ Sherwin, and J Peiro. Linear dispersion-diffusion analysis and its application to under-resolved turbulence simulations using discontinuous galerkin spectral/hp methods. *Journal of Computational Physics*, 298:695–710, 2015.

[99] Cristian R Nastase and Dimitri J Mavriplis. High-order discontinuous galerkin methods using an hp-multigrid approach. *Journal of Computational Physics*, 213(1):330–357, 2006.

[100] Ngoc Cuong Nguyen, Jaume Peraire, and Bernardo Cockburn. An implicit high-order hybridizable discontinuous galerkin method for linear convection–diffusion equations. *Journal of Computational Physics*, 228(9):3232–3254, 2009.

[101] Francisco Palacios, Michael R Colonno, Aniket C Aranake, Alejandro Campos, Sean R Copeland, Thomas D Economon, Amrita K Lonkar, Trent W Lukaczyk, Thomas WR Taylor, and Juan J Alonso. Stanford university unstructured (su2): An open-source integrated computational environment for multi-physics simulation and design. *AIAA Paper*, 287:2013, 2013.

[102] P-O Persson and Jaime Peraire. Newton-gmres preconditioning for discontinuous galerkin discretizations of the navier-stokes equations. *SIAM Journal on Scientific Computing*, 30(6):2709–2733, 2008.

[103] Per-Olof Persson and Jaime Peraire. Sub-cell shock capturing for discontinuous galerkin methods. *AIAA paper*, 112:2006, 2006.

[104] T J&amp Poinsot and SK Lelef. Boundary conditions for direct simulations of compressible viscous flows. *Journal of computational physics*, 101(1):104–129, 1992.

[105] Patrick Rasetarinera and MY Hussaini. An efficient implicit discontinuous spectral galerkin method. *Journal of Computational Physics*, 172(2):718–738, 2001.

[106] Wm H Reed and TR Hill. Triangular mesh methods for the neutron transport equation. *Los Alamos Report LA-UR-73-479*, 1973.

[107] Jean-François Remacle, Joseph E Flaherty, and Mark S Shephard. An adaptive discontinuous galerkin technique with an orthogonal basis applied to compressible flow problems. *SIAM review*, 45(1):53–72, 2003.

[108] Béatrice Rivière. *Discontinuous Galerkin methods for solving elliptic and parabolic equations: theory and implementation*. Society for Industrial and Applied Mathematics, 2008.

[109] J Romero, K Asthana, and A Jameson. A simplified formulation of the flux reconstruction method. *Journal of Scientific Computing*, pages 1–24, 2015.

[110] Joshua Romero and Antony Jameson. Extension of the flux reconstruction method to triangular elements using collapsed-edge quadrilaterals. In *54th AIAA Aerospace Sciences Meeting*, page 1825, 2016.

[111] P Sagaut and R Grohens. Discrete filters for large eddy simulation. *International Journal for Numerical Methods in Fluids*, 31(8):1195–1220, 1999.

[112] T. Sengupta and T. Poinsot. *Instabilities of Flows: With and Without Heat Transfer and Chemical Reaction*. CISM International Centre for Mechanical Sciences. Springer Vienna, 2010.

[113] T. Sengupta and T.K. Sengupta. *High Accuracy Computing Methods: Fluid Flows and Wave Phenomena*. Cambridge University Press, 2013.

[114] Tapan K Sengupta, SK Sircar, and Anurag Dipankar. High accuracy schemes for dns and acoustics. *Journal of Scientific Computing*, 26(2):151–193, 2006.

[115] TK Sengupta and S Bhaumik. Onset of turbulence from the receptivity stage of fluid flows. *Physical review letters*, 107(15):154501, 2011.

[116] TK Sengupta, Y Bhumkar, and V Lakshmanan. Design and analysis of a new filter for les and des. *Computers & Structures*, 87(11):735–750, 2009.

[117] Spencer Sherwin. Dispersion analysis of the continuous and discontinuous Galerkin formulations. In *Discontinuous Galerkin Methods*, pages 425–431. Springer, 2000.

[118] Abhishek Sheshadri and Antony Jameson. Shock detection and capturing methods for high order discontinuous-galerkin finite element methods. In *32nd AIAA Applied Aerodynamics Conference*, volume 2688, 2014.

[119] Chi-Wang Shu and Stanley Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *Journal of Computational Physics*, 77(2):439–471, 1988.

[120] Chi-Wang Shu and Stanley Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes, ii. In *Upwind and High-Resolution Schemes*, pages 328–374. Springer, 1989.

[121] Gary A Sod. A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. *Journal of computational physics*, 27(1):1–31, 1978.

[122] T Strouboulis and JT Oden. A posteriori error estimation of finite element approximations in fluid mechanics. *Computer methods in applied mechanics and engineering*, 78(2):201–242, 1990.

[123] E. Süli and D.F. Mayers. *An Introduction to Numerical Analysis*. Cambridge University Press, 2003.

[124] Endre Süli. A posteriori error analysis and adaptivity for finite element approximations of hyperbolic problems. In *An introduction to recent developments in theory and numerics for conservation laws*, pages 123–194. Springer, 1999.

[125] Magnus Svärd and Siddhartha Mishra. Shock capturing artificial dissipation for high-order finite difference schemes. *Journal of Scientific Computing*, 39(3):454–484, 2009.

[126] Eitan Tadmor. The numerical viscosity of entropy stable schemes for systems of conservation laws. i. *Mathematics of Computation*, 49(179):91–103, 1987.

[127] Eitan Tadmor. Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Acta Numerica*, 12:451–512, 2003.

[128] Eitan Tadmor, MJ Baines, and KW Morton. Super viscosity and spectral approximations of nonlinear conservation laws. 1993.

[129] Christopher KW Tam and Jay C Webb. Dispersion-relation-preserving finite difference schemes for computational acoustics. *Journal of Computational Physics*, 107(2):262–281, 1993.

[130] Lony L Thompson and Peter M Pinsky. Complex wavenumber fourier analysis of the p-version finite element method. *Computational Mechanics*, 13(4):255–275, 1994.

[131] Lloyd N Trefethen. Group velocity in finite difference schemes. *SIAM review*, 24(2):113–136, 1982.

[132] Kris Van den Abeele. Development of high-order accurate schemes for unstructured grids. *Phd thesis in Vrije Universiteit Brussel*, 2009.

[133] JJW Van der Vegt and H Van der Ven. Discontinuous galerkin finite element method with anisotropic local grid refinement for inviscid compressible flows. *Journal of Computational Physics*, 141(1):46–77, 1998.

[134] JJW Van der Vegt and H Van der Ven. Space–time discontinuous galerkin finite element method with dynamic grid motion for inviscid compressible flows: I. general formulation. *Journal of Computational Physics*, 182(2):546–585, 2002.

[135] Harmen van der Ven. A family of large eddy simulation (les) filters with nonuniform filter widths. *Physics of Fluids (1994-present)*, 7(5):1171–1172, 1995.

[136] 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.

[137] Hervé Vandeven. Family of spectral filters for discontinuous problems. *Journal of Scientific Computing*, 6(2):159–192, 1991.

[138] Oleg V Vasilyev, Thomas S Lund, and Parviz Moin. A general class of commutative filters for les in complex geometries. *Journal of Computational Physics*, 146(1):82–104, 1998.

[139] John C Vassberg and Antony Jameson. In pursuit of grid convergence for two-dimensional euler solutions. *Journal of Aircraft*, 47(4):1152–1166, 2010.

[140] Robert Vichnevetsky and John B Bowles. *Fourier analysis of numerical approximations of hyperbolic equations*, volume 5. Siam, 1982.

[141] Peter E Vincent, Patrice Castonguay, and Antony Jameson. Insights from von neumann analysis of high-order flux reconstruction schemes. *Journal of Computational Physics*, 230(22):8134–8154, 2011.

[142] Peter E Vincent, Patrice Castonguay, and Antony Jameson. A new class of high-order energy stable flux reconstruction schemes. *Journal of Scientific Computing*, 47(1):50–72, 2011.

[143] Peter E Vincent, Antony M Farrington, Freddie D Witherden, and Antony Jameson. An extended range of stable-symmetric-conservative flux reconstruction correction functions. *Computer Methods in Applied Mechanics and Engineering*, 296:248–272, 2015.

[144] Miguel R Visbal and Datta V Gaitonde. Very high-order spatially implicit schemes for computational acoustics on curvilinear meshes. *Journal of Computational Acoustics*, 9(04):1259–1286, 2001.

[145] Miguel R Visbal and Datta V Gaitonde. On the use of higher-order finite-difference schemes on curvilinear and deforming meshes. *Journal of Computational Physics*, 181(1):155–185, 2002.

[146] ZJ Wang, Krzysztof Fidkowski, Rémi Abgrall, Francesco Bassi, Doru Caraeni, Andrew Cary, Herman Deconinck, Ralf Hartmann, Koen Hillewaert, HT Huynh, et al.

High-order cfd methods: current status and perspective. *International Journal for Numerical Methods in Fluids*, 72(8):811–845, 2013.

[147] ZJ Wang and Haiyang Gao. A unifying lifting collocation penalty formulation including the discontinuous Galerkin, spectral volume/difference methods for conservation laws on mixed grids. *Journal of Computational Physics*, 228(21):8161–8186, 2009.

[148] Frank M White and Isla Corfield. *Viscous fluid flow*, volume 3. McGraw-Hill New York, 2006.

[149] Gerald Beresford Whitham. *Linear and nonlinear waves*, volume 42. John Wiley & Sons, 2011.

[150] D. M. Williams. *Energy Stable High-Order Methods for Simulating Unsteady, Viscous, Compressible Flows on Unstructured Grids*. PhD thesis, Stanford University, 2013.

[151] David M Williams, Patrice Castonguay, Peter E Vincent, and Antony Jameson. An extension of energy stable flux reconstruction to unsteady, non-linear, viscous problems on mixed grids. In *20th AIAA Computational Fluid Dynamics Conference*, pages 27–30, 2011.

[152] DM Williams, P Castonguay, PE Vincent, and A Jameson. Energy stable flux reconstruction schemes for advection-diffusion problems on triangles. *Journal of Computational Physics*, 2013.

[153] DM Williams, Patrice Castonguay, Peter E Vincent, and Antony Jameson. Energy stable flux reconstruction schemes for advection–diffusion problems on triangles. *Journal of Computational Physics*, 250:53–76, 2013.

[154] DM Williams and A Jameson. Energy stable flux reconstruction schemes for advection–diffusion problems on tetrahedra. *Journal of Scientific Computing*, pages 1–39, 2013.

[155] DM Williams and A Jameson. Energy stable flux reconstruction schemes for advection–diffusion problems on tetrahedra. *Journal of Scientific Computing*, 59(3):721–759, 2014.

[156] Freddie D Witherden, Brian C Vermeire, and Peter E Vincent. Heterogeneous computing on mixed unstructured grids with pyfr. *Computers & Fluids*, 120:173–186, 2015.

[157] Paul Woodward and Phillip Colella. The numerical simulation of two-dimensional fluid flow with strong shocks. *Journal of computational physics*, 54(1):115–173, 1984.

[158] Lin Yin, Amit Acharya, Nahil Sobh, Robert B Haber, and Daniel A Tortorelli. A space-time discontinuous galerkin method for elastodynamic analysis. In *Discontinuous Galerkin Methods*, pages 459–464. Springer, 2000.

[159] Xiaolin Zhong. High-order finite-difference schemes for numerical simulation of hypersonic boundary-layer transition. *Journal of Computational Physics*, 144(2):662–709, 1998.

[160] Jun Zhu, Xinghui Zhong, Chi-Wang Shu, and Jianxian Qiu. Runge–kutta discontinuous galerkin method using a new type of weno limiters on unstructured meshes. *Journal of Computational Physics*, 248:200–220, 2013.