RICE UNIVERSITY

Discontinuous Galerkin and Finite Difference Methods for the Acoustic Equations with Smooth Coefficients

by

Mario Bencomo

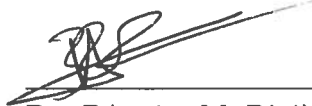
A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Master of Arts

Approved, Thesis Committee:

Dr. William W. Symes, Chair
Noah Harding Professor of Computational and Applied Mathematics and Professor of Earth Science

Dr. Tim Warburton
Full Professor of Computational and Applied Mathematics



Dr. Béatrice M. Rivière
Full Professor of Computational and Applied Mathematics

Houston, Texas

April, 2015

ABSTRACT

Discontinuous Galerkin and Finite Difference Methods for the Acoustic Equations with Smooth Coefficients

by

Mario Bencomo

This thesis analyzes the computational efficiency of two types of numerical methods: finite difference (FD) and discontinuous Galerkin (DG) methods, in the context of 2D acoustic equations in pressure-velocity form with smooth coefficients. The acoustic equations model propagation of sound waves in elastic fluids, and are of particular interest to the field of seismic imaging. The ubiquity of smooth trends in real data, and thus in the acoustic coefficients, validates the importance of this novel study. Previous work, from the discontinuous coefficient case of a two-layered media, demonstrates the efficiency of DG over FD methods but does not provide insight for the smooth coefficient case. Floating point operation (FLOPs) counts are compared, relative to a prescribed accuracy, for standard 2-2 and 2-4 staggered grid FD methods, and a myriad of standard DG implementations. This comparison is done in a serial framework, where FD code is implemented in C while DG code is written in Matlab. Results show FD methods considerably outperform DG methods in FLOP count. More interestingly, implementations of quadrature based DG with mesh refinement (for lower velocity zones) yield the best results in the case of highly variable media, relative to other DG methods.

Acknowledgement

Firstly, I would like to thank the committee members for their commitment and their aid throughout this thesis work. In particular, Dr. Warburton has been my primary source for all DG related matters. My advisor, Dr. Symes, has been the tremendous guiding force all through this research project. I am grateful for your support.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1450681.

Contents

Illustrations

Tables

Chapter 1

Introduction

This thesis analyzes the computational efficiency of two types of numerical methods, finite difference (FD) and discontinuous Galerkin (DG) methods, in the context of the 2D acoustic equations in velocity-pressure form with smooth coefficients. The key question addressed is as follows: How does computational efficiency of FD and DG methods compare in the case where coefficients of the acoustic equations vary smoothly? I provide results for some canonical smooth media examples and analyze their computational cost relative to accuracy. A better understanding of the computational behavior of FD and DG methods, applied to acoustics, will have implications in seismic imaging and oil prospecting.

1.1 Motivation

Seismic Surveys

Seismic surveying is a common prospecting practice throughout the oil industry, consisting of generating and measuring seismic waves as they propagate through a medium of interest from which relevant geophysical information may be recovered. Seismic waves are induced by impulsive energy sources, such as: projectiles, impactors

(i.e., weight drops), explosives, and electrical impulse sources. The choice of source is of course dependent on application and particular source considerations. For example, air guns are the source of choice in marine surveying while the impact of a sledge-hammer hitting the ground would suffice for shallow seismic surveying; see (Sheriff and Geldart, 1995) and Burger et al. (2006) for more on sources in applications to seismic surveying. Fig. 1.1 shows the canonical setup for marine seismic surveying. A sound wave source (air guns releasing compressed air in this application) induces seismic waves that propagate down to the ocean floor, interact with the subsea rock formations, and later recorded by acoustic receivers on streamers. The data is then analyzed in order to obtain information that will help identify and localize oil and gas reserves. Several streamers are towed across the water during data acquisition, with lengths reaching up to 5 kilometers spanning over 500 meters wide, roughly a 30 by 3 city block area (FishSAFE, 2013). Clearly the dimensions involved in seismic surveying alone convey value of the data acquired, and more importantly highlight the necessity optimal data analysis.

Inversion theory provides the mathematical framework for the analysis of seismic data and ultimately the ability to extract parameters pertaining to physical phenomena. Furthermore, the forward problem plays a crucial role in the accuracy and development of inversion methods. The problem of waveform inversion, at least that of full waveform inversion, is typically formulated as a least squares problem with a nonlinear objective function. Iterative optimization techniques used to solve the

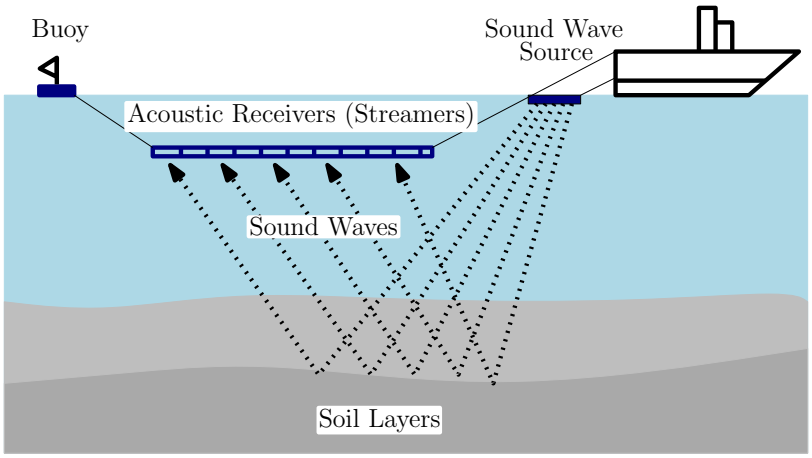


Figure 1.1 : Marine seismic surveying setup, adapted from FishSAFE (2013).

least squares problem require the solution of several forward problems per iteration. Hence, the "quality of the results of any inverse method," and thus the recovered seismic image, "depends heavily on the realism of the forward modeling" (Gauthier, Virieux, and Tarantola, 1986).

The Forward Problem: Acoustic Equations and Smooth Coefficients

The forward problem, in the context of seismic imaging, comprises of modeling the phenomena of seismic wave propagation through the Earth's subsurface. The acoustic equations, given by Eq. (2.1), describe the first order perturbation of the conservation laws of mass and momentum obeyed by an elastic fluid close to equilibrium. See Gurtin (1981), pp. 122-137, for a complete discussion of elastic fluids and a derivation of the acoustic equations. The acoustic equations model the propagation of compressional waves (or P-waves) in elastic fluids, however a more realistic phys-

ical representative of Earth's subsurface formations is an elastic material in which the dynamics of this system is given by the elasticity equations. The core difference between elastic fluids and elastic materials lies in the nature of their constitutive equations; stress is a function of the relative deformation for elastic materials, where in the case of elastic fluids (and in general inviscid fluids) stress is given by a pressure (Gurtin, 1981, §17). Consequently shear waves (or S-waves) arise in elasto-dynamics and are indeed observed in seismic applications. Not surprisingly, higher fidelity in the physics gives rise to complexity in the model; there are 21 parameters associated with the stress tensor for anisotropic linear elasticity. Currently the discrepancy between accuracy gain and computational cost of the elasticity equations is considerable for the large scale nature of seismic surveying, thus making the acoustic equations the traditional choice for seismic wave modeling and inversion in the oil industry.

In this study I consider the acoustic equations with smooth coefficients (i.e., density and bulk modulus will vary smoothly in space), a case relevant for seismic imaging. Layered media arise naturally in geological formations, corresponding to discontinuities in coefficients. Equally relevant are gradual variations in the media due to large time scale influences such as sediment accumulation. Consider Fig. 1.2, adapted from Sheriff and Geldart (1995), a plot of velocity (speed of sound) versus depth for Gulf Coast sands and shales and offshore Venezuela. Clearly "smooth" trends are observed for the various data motivating the use of smooth coefficients.

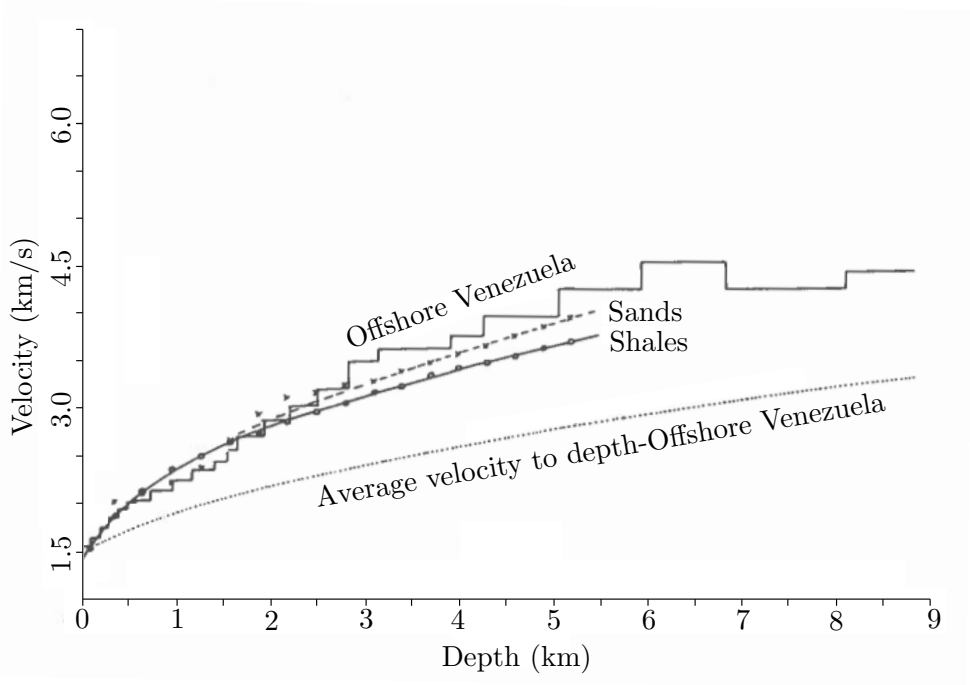


Figure 1.2 : Velocity-depth profiles for Gulf Coast sands and shales, and offshore Venezuela (Sheriff and Geldart, 1995, pp. 120).

1.2 Literature Review

Finite Difference Methods

Finite difference (FD) methods are classical numerical methods for solving a variety of differential equations. An extensive overview of FD methods for seismic wave simulations is given by Moczo et al. (2007). Standard FD methods consist of discretizing the domain of interest into uniformly spaced grid points, where partial derivatives are approximated by finite differences of the function evaluated at specified grid points, also referred to as the stencil of the FD method.

Staggered grid FD methods were introduced by Madariaga (1976) for propagation

of P-SV elastic waves in fault dynamics and by Virieux (1984, 1986) in the context of P-SV and SH elastic waves in heterogeneous media. SV and SH denote the vertical and horizontal polarization of shear waves, or S-waves, respectively. These methods resolve stability issues present in conventional grid FD for elastic waves in media with large Poisson's ratio, and provide a unified scheme for acoustic-elastic coupled systems, of importance to marine-seismic surveying. Levander (1988) later generalized staggered grid FD to higher order methods, in particular the $\mathcal{O}(\Delta t^2, h^4)$ or 2-4 order staggered grid scheme (second-order and fourth-order accuracy in time and space respectively) for P-SV wave propagation. Along with preserving stability properties of staggered grid methods, the 2-4 scheme reduces the number of nodes needed in memory storage to one fourth of the nodes required for second-order in space P-SV staggered grid methods, while maintaining computation times comparable to other fourth-order in space methods.

The finite difference methods considered in this thesis are the acoustic variant of the 2-2 and 2-4 order staggered grid methods proposed by Virieux (1986) and Levander (1988). These numerical schemes can be viewed as a two-part discretization process: first a discretization in space, for a semi-discrete scheme, and then in time resulting in a fully-discretized scheme. This procedure of discretization is known as the method of lines (MOL) and provides a methodology for analyzing stability and convergence of fully discretized schemes. Stability criteria of these numerical methods yield relationships between the spatial and temporal discretization that must be satis-

fied for stability of the numerical method. In particular, Virieux (1986) and Levander (1988) have shown the following stability criterion for the 2-2 and 2-4 staggered grid FD methods, respectively:

$$\Delta t < \frac{1}{\sqrt{2}V_P}h \qquad \text{(2-2 FD)} \tag{1.1}$$

$$\Delta t < \frac{0.606}{V_P}h \qquad \text{(2-4 FD)} \tag{1.2}$$

where V_P is the compressional velocity. Equations 1.1 and 1.2 were derived in the context of 2-D isotropic elasticity, assuming $\Delta x = \Delta z = h$. Note that these stability limits are independent of the shear velocity V_S, and hence independent of Poisson's ratio, unlike some standard FD methods (Virieux, 1986; Levander, 1988). Stability limits will effectively impose a constraint on how large the time step Δt can be taken relative to the spatial discretization h, something to keep in mind when considering computational cost. Clearly, in terms of computational cost, methods with larger stability regions are more attractive, though the higher accuracy of a method can offset these associated costs.

Another numerical property that will impact the accuracy and overall computational cost of FD methods is grid dispersion, also analyzed in Virieux (1986) and Levander (1988) for the 2-2 and 2-4 staggered grid FD methods respectively. Consider a plane wave with given wave vector $\boldsymbol{k}$ and frequency ω. The relationship of ω with respect to $\boldsymbol{k}$ as the plane wave propagates through the medium is known as

the dispersion relation. Grid dispersion is the discrepancy in the dispersion relation between the continuum equations and discrete system given by the numerical scheme. Moreover, mismatch in the dispersion relation corresponds to errors in arrival times for waves at given frequencies, thus potentially degrading substantially the accuracy of the numerical method. Grid dispersion can be minimized however by refining the spatial discretization h, more precisely increasing the number of grid points per wavelength. Sampling ratios of 10 and 5 grid points per wavelength, for the shortest wavelength, are used as a rule of thumb for the 2-2 and 2-4 staggered grid FD methods respectively (Virieux, 1986; Levander, 1988).

Overall, the numerical behavior of FD methods is well understood and well cataloged throughout the applied mathematics community and applications thereof. The strengths of FD methods however rely heavily on the simplicity of the domain geometries and smoothness of medium parameters. The necessity for dealing with discontinuities in material parameters and the rise of high performance computing within the oil industry has lead to applications of numerical methods other than FD methods, namely finite elements and variants such as spectral elements and discontinuous Galerkin. Of course this thesis work is primarily concerned with the case of smooth coefficients, as alluded in the title. Nevertheless I will briefly detour into modeling wave propagation under discontinuous medium parameters in order to recount how the discontinuous Galerkin method has gained traction in the field of seismic modeling.

Discontinuous Galerkin Methods

The discontinuous Galerkin (DG) method comprises of approximating the solution of a PDE via a Galerkin approximation where discontinuities are allowed, though penalized in some manner. In more precise words, the PDE solution is approximated by a numerical solution such that, when the numerical solution is applied to a specified weak formulation, it yields a residual orthogonal with respect to a chosen finite dimensional space. The DG method was first introduced by Lesaint and Raviart (1974) for the neutron transport problem and since then has been extended to a variety of applications: fluid transport in porous media, Navier-Stokes flow, and wave propagation to name a few. I refer to Cockburn (2003) for a comprehensive overview of DG methods and applications. DG methods have since gained popularity due to their geometric flexibility and mesh and polynomial order adaptivity, also known as *hp* adaptivity. Moreover, these methods can yield explicit schemes at each time-step (after inverting a block diagonal matrix), thus resulting in tractable algorithms for time-dependent hyperbolic problems such as wave propagation.

There is almost a zoology of DG schemes for the modeling of wave propagation that results from availability of choice, for example: the choice of approximation spaces, choice of basis functions, how to handle jump discontinuities across elements, choice of elements, etc.. Penalty DG methods are amongst the more popular schemes for acoustic and elastic wave propagation problems in their second order form, sometimes referred to as the displacement formulation. The following papers cover some of the

analysis in numerical stability, grid dispersion, and error estimates of these methods: De Basabe and Sen (2010), De Basabe et al. (2008), Grote et al. (2006), Riviere and Wheeler (2003). This thesis work will be concerned with the first order form of the acoustic wave equation, also known as the velocity-stress or velocity-pressure formulation. I chose the first order formulation primarily for the fact that the FD schemes selected in this comparison are based on this formulation. In particular, I will be considering the Runge-Kutta DG scheme with upwind flux for the acoustic equations in velocity-pressure form.

The implementation details and the development of the DG method used in this work is based off a nodal approach, where the approximation spaces are spanned by multivariate Lagrange polynomials (Hesthaven and Warburton, 2007). The article on time domain solutions to Maxwell's equations via nodal high-order methods on unstructured nodes by Hesthaven and Warburton (2002) contains methodology and analysis that carries over to my thesis. In particular, Hesthaven and Warburton (2002) prove and demonstrate that the global error of the semi-discrete solution grows at most linearly in time and can be minimized through *hp* refinement:

$$\| \boldsymbol{q}(t) - \boldsymbol{q}_N(t) \|_\Omega \leq \mathcal{O}(h^{N+1}) + t\mathcal{O}(h^N), \tag{1.3}$$

assuming that the true solution is smooth enough, where $\boldsymbol{q}$ and $\boldsymbol{q}_N$ denote the true and numerical solutions respectively and N is the maximal order of the polynomials in the approximation space.

The low-storage fourth order Runge-Kutta method used for time discretization in this thesis work was first proposed by Carpenter and Kennedy (1994). Runge-Kutta DG (RK-DG) methods have since then gained popularity primarily attributed to having a time-limiting step size proportional to the mesh size h divided by the maximum propagation speed v_{max},

$$\Delta t \leq \frac{\mathrm{CFL}}{v_{max}} \min_{\Omega} h, \tag{1.4}$$

as mentioned by Hesthaven and Warburton (2002), with maximum velocity v_{max} in the medium. Note that for a 2D triangular mesh, h consists of the diameter of the largest circle inscribed at a triangle. There are two main drawbacks of RK-DG methods that are current research topics and worth mentioning: the time step size is dependent on the globally smallest h rather than a local h which can be of issue for h-adaptive methods. Second, the CFL constant turns out to be dependent on the polynomial order N of the DG scheme, in particular CFL $= \mathcal{O}(\frac{1}{N^2})$, which hampers the convergence and computational benefits of higher order methods. A time-space DG method is proposed by Monk and Richter (2005) allowing for a more efficient "local" time-space stability criterion appropriate for h adaptive algorithms. The spatial order dependency of the CFL condition in Eq.(1.4) is addressed by Warburton and Hagstrom (2008), acquiring a CFL condition independent of polynomial order. Despite these time-limiting step size disadvantages, the RK-DG with upwind flux method is still a competitive candidate for the simulation of wave propagation as

shown in many of the research studies to be cited shortly, and is thus in part the subject of this thesis.

As discussed previously for FD methods, grid dispersion is an important numerical phenomenon that plays a crucial role in the accuracy of wave propagation. Aside from errors in the phase velocity, that is grid dispersion, DG schemes considered here suffer from dissipation errors, where amplitudes of waves appear to be over attenuated due to numerical errors. This dissipation error is inherent to the choice of flux; in particular one can show that the total energy of the semi-discrete system decays with time due to the dissipative nature of the upwind flux, as oppose to the true conservative nature of hyperbolic systems. De Basabe and Sen (2007) provide analysis and numerical results of grid dispersion for some common finite element and spectral element methods in application to acoustic and elastic wave propagation, along with some comparisons to FD methods. Ainsworth (2004) studies the dissipation and dispersion errors under hp refinement of the DG method applied to a linear advection equation. Specifically, Ainsworth (2004) shows that the polynomial order N can be chosen such that the dispersion error decays super-exponentially if $2N + 1 \approx chk$ for a given mesh size h and wavenumber k, for some constant $c > 1$. Work by Hu et al. (1999) provide an interesting analysis of anisotropy in dispersion and dissipation errors on quadrilateral and triangular uniform meshes for DG methods applied to 2D wave propagation problems. More importantly, the 1D analysis by Hu et al. (1999) will serve as a guideline to spatially discretize the domain with respect to a resolvable wavenumber;

specifics will be revealed in the results section.

DG schemes, similar to the RK-DG method implemented in this work, have recently gained popularity in the field of seismic modeling. In particular, current work focuses on the analysis, applications, and extensions of DG schemes to the context of discontinuous media. Work by Wang (2009) and Wang et al. (2010), on the comparison of DG and FD methods for time domain acoustics, reveals the efficiency of the DG method over staggered FD methods for the case of complex piecewise constant media. Interface errors over the discontinuity reduce the convergence rate of FD methods to first order, while a DG scheme with an appropriately aligned mesh results in a sub-optimal second order method making DG a more efficient method for complicated models.

Zhebel et al. (2013) perform a study on the parallel scalability of FD and finite element methods, including mass lumped finite elements and DG, for 3D acoustic wave propagation in piecewise constant media with a dipping interface on an Intel Sandy Bridge dual 8-core machine and Intel's 61-core Xeon Phi. Overall the DG method demonstrated larger speed up on Sandy Bridge as the number of cores was increased and the problem size is kept constant, partly due to the fact that DG involves more net FLOPs (floating point operations) relative to other methods. Interestingly enough, for Intel's Xeon Phi, FD and DG methods showed similar strong scalability performance, for an "optimal" choice of FD domain subdivisions. Lastly, convergence results by Zhebel et al. (2013) demonstrate again the superior accuracy

of finite element type methods with mesh alignment to that of FD methods. Another interesting study was conducted by Simonaho et al. (2012), where DG simulations of acoustic wave propagation were compared to real data. Author's simulated and acquired 3D experimental data for pulse propagation and scattering from a cylinder in air. Results show that simulated data matches measurement time-series well up to 4.5 ms for the pulse propagation study. Moreover, amplitude spectrum of the simulated data closely resembles that of real data for frequencies less than 2kHz. For the scattering cylinder case, simulated data contained all of the representative qualitative characteristics present in the real data, that is interference patterns from reflections and diffractions due to the cylinder, for 2D spatial slices at given time-shots.

A lot of the work on DG applied to wave propagation cited above, assumes that the medium parameters are piecewise constant for implementation purposes. Atkins and Shu (1998) and Hesthaven and Warburton (2007) propose efficient quadrature-free DG implementation strategies in the case of piecewise constant media on triangular meshes. Aside from reduction in code complexity and overhead computational cost, these quadrature-free implementations result in lower memory costs associated with storing DG operators relative to their quadrature based counterparts. On the other hand, quadrature based implementations have the upper hand in terms of accuracy, since medium parameters as variable functions in space within mesh elements are better representatives than piecewise constant coefficients when computing element-wise integrals in the definition of DG operators such ass mass and stiffness matrices.

For example, Ober et al. (2010) employ quadrature based DG implementations in the context of acoustic seismic inversion. Collis et al. (2010) provide a comparison between the quadrature-free and quadrature based DG implementations for elastic wave propagation in variable media. Their results show that the numerical solutions for piecewise constant and variable models do not converge to the same limit as the polynomial order is increased. However, Collis et al. (2010) does not provide any insight into the performance and efficiency of quadrature-free and quadrature based implementations, a topic of importance that will be discussed in this thesis work.

1.3 Claim

Few studies have been made comparing the computational efficiency of FD and DG methods in the context of the acoustic equations. Wang (2009) provides a comparison for 2D acoustics with discontinuous coefficients. He demonstrates the efficiency of a standard DG method, with curvilinear elements, over the 2-4 staggered grid FD method for a 2D dome model, results attributed to the inadequacy of FD methods to resolve complex geometric interfaces. Collis et al. (2010), Ober et al. (2010) apply DG to acoustic and elastic wave propagation in variable media, but do not offer an elucidated comparison between quadrature-free and quadrature based DG implementations relative to FD solvers.

Recent computational studies of FD and DG methods consider top of the line hardware, software, and programming implementations in order to achieve efficient

solvers that will ultimately minimize runtimes (Zhebel et al., 2013; Zhou, 2014). It is well observed that DG methods suffer from a higher computational costs in terms of FLOP counts, however the total runtime can be significantly mitigated as consequence of the method's higher accuracy, parallel scalability, and computational efficiency (Zhebel et al., 2013; Wang, 2009; Wang et al., 2010). On the other hand, Zhou (2014) has experimented with optimizing FD stencil code on multi-core programming by improving vectorization and minimizing memory traffic, in particular for the constant density acoustic wave equation, yielding 20% to 30% of peak FLOPs per second depending on compiler and FD approximation order. The end goal of modeling software within seismic imaging is to provide fast accurate solvers to be used in inversion algorithms; clearly, runtimes are a direct and highly relevant metric of the speed of these solvers. However, measuring runtimes is dependent on the hardware, software, implementation, etc., a measure that is at times not very portable. I consider instead a more fundamental manner in comparing FD and DG, that is through counting FLOPs, which will only depend on the algorithm, and it's implementation to some degree, of the numerical method. The relationship between runtimes and FLOP counts can of course be complicated, especially when comparing between two very different numerical methods like FD and DG. FLOP counting will nonetheless offer a point of reference and insight into the computational costs and efficiency of these methods.

In my work, I offer a baseline comparison of FD and DG methods in the context

of the acoustic wave equation in first order form with smoothly varying coefficients, a regime in which interface errors are not of issue. Moreover, both quadrature-free and quadrature based implementations will be studied in this analysis for the sake of completeness. In particular, I compare computational cost for a prescribed accuracy on several smooth models.

1.4 Agenda

The remainder of the thesis will consist of three more chapters. Chapter 2 will develop in some detail the finite difference and discontinuous Galerkin methods considered here, in particular highlighting considered implementation strategies. Numerical simulations and results will be discussed in Chapter 3. Finally, Chapter 4 will summarize and interpret the results and give some concluding remarks.

Chapter 2

Methods

2.1 Introduction

This Chapter derives to some detail the finite difference and discontinuous Galerkin methods considered in this study. The staggered grid finite difference methods applied to the acoustic equations are derived for the general 2-2k order case in §2.3. Several aspects of the discontinuous Galerkin method are discussed in §2.4, in particular the derivation of the semi-discrete scheme, choice of flux, time discretization, and implementation options for handling variable media.

2.2 Model Problem

The 2D acoustic equations, in velocity-pressure form, are given by the following system of first order PDEs:

$$\rho(\boldsymbol{x})\frac{\partial \boldsymbol{v}}{\partial t}(\boldsymbol{x},t) + \nabla p(\boldsymbol{x},t) = \boldsymbol{f}(\boldsymbol{x},t), \tag{2.1a}$$

$$\beta(\boldsymbol{x})\frac{\partial p}{\partial t}(\boldsymbol{x},t) + \nabla \cdot \boldsymbol{v}(\boldsymbol{x},t) = g(\boldsymbol{x},t), \tag{2.1b}$$

for $\boldsymbol{x} = [x, y]^T \in \Omega \subset \mathbb{R}^2$ and $0 \leq t \leq T$, for some final time $T > 0$. Here ρ and $\beta = 1/\kappa$ represent density and compressibility (i.e., the reciprocal of the bulk modulus κ) respectively. Furthermore, $\boldsymbol{v} = [v_x, v_y]^T$ is the velocity vector, p is the pressure, and $\boldsymbol{f} = [f_x, f_y]^T$ and g are source terms. It will be assumed that ρ and β are smooth functions in the spatial variables (x, y). Homogeneous initial and boundary conditions will be considered.

2.3 Finite Difference Methods

The finite difference method consists of approximating derivatives via linear combinations of the field in question at given points. Coefficients of these linear combinations are uniquely determined, as prescribed by related Taylor expansions, by the choice of field points and the maximum order of approximation. A conventional finite difference scheme applied to Eq.(2.1) would use field points over a regular grid, the same grid points for pressure and velocities. In the staggered grid variant, field points for the pressure and velocity fields are over regular grids such that grid points do not overlap with each other, i.e., pressure and velocity field points do not coincide. Fig.(2.1) depicts the relative location of grid points for the pressure and velocity fields. Note that the pressure field points reside in what is dubbed as the *primary grid*, where indexing pertains to integers, while velocity grid points lie on *shifted grids* where fractional indexing is allowed for a particular index.

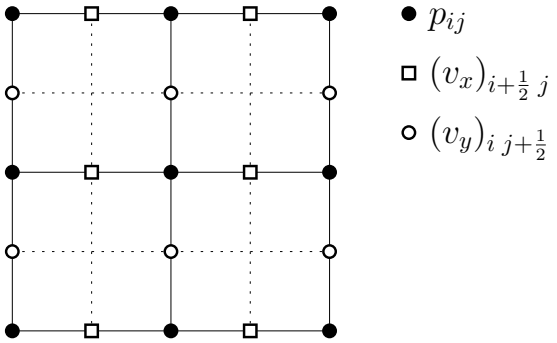


Figure 2.1 : Staggered grid points for 2D acoustics.

Space Discretization

Let $D_x^{h,(k)} f(x_0)$ denote the $2k^{th}$ order centered finite difference approximation of $\frac{\partial f}{\partial x}$ at x_0, of step size h as follows:

$$D_x^{h,(k)} f(x_0) := \frac{1}{h} \sum_{n=1}^{k} a_n^{(k)} \Big\{ f\left(x_0 + \left(n - \tfrac{1}{2}\right) h\right) - f\left(x_0 - \left(n - \tfrac{1}{2}\right) h\right) \Big\}.$$

The coefficients $a_n^{(k)}$ can be derived by a straightforward Taylor approximation of f centered at the various points $x_0 \pm (n - \frac{1}{2})h$. The following table includes coefficients $a_n^{(k)}$ for $k = 1, 2$:

	$n = 1$	$n = 2$
$k = 1$	1	-
$k = 2$	1/24	$-9/8$

Table 2.1 : Table of coefficients for centered finite difference formulas of order $k = 1, 2$.

I now give the staggered grid semi-discrete scheme for acoustics of arbitrary $2k^{th}$ order:

$$(\rho)_{i+\frac{1}{2}\,j}\frac{d}{dt}(v_x)_{i+\frac{1}{2}\,j} = -D_x^{h,(k)}(p)_{i+\frac{1}{2}\,j} - (f_x)_{i+\frac{1}{2}\,j}, \tag{2.2}$$

$$(\rho)_{i\,j+\frac{1}{2}}\frac{d}{dt}(v_y)_{i\,j+\frac{1}{2}} = -D_y^{h,(k)}(p)_{i\,j+\frac{1}{2}} - (f_y)_{i\,j+\frac{1}{2}}, \tag{2.3}$$

$$(\beta)_{ij}\frac{d}{dt}(p)_{ij} = -D_x^{h,(k)}(v_x)_{ij} - D_y^{h,(k)}(v_y)_{ij} + (g)_{ij}, \tag{2.4}$$

where $p_{ij} = p(ih, jh)$. Note that $D_x^{h,(k)}(p)_{i+\frac{1}{2}\,j}$ will consist of linear combinations of $(p)_{i\pm(n-1)\,j}$, that is the pressure field at primal grid points. Similarly, one can show that any arithmetic resulting from discrete operators in Eq.(2.2) will involve fields in their respective grid.

Time Discretization

System from Eq.(2.2) is discretized in time by applying a centered difference approximation of the time derivative, again in a staggered manned for the pressure and velocity fields. The general $2-2k$ staggered grid method takes the following form:

$$(v_x)^{n+1}_{i+\frac{1}{2}\,j} = (v_x)^{n}_{i+\frac{1}{2}\,j} + \Delta t\frac{1}{(\rho)_{i+\frac{1}{2}\,j}}\Big\{ - D_x^{h,(k)}(p)^{n+\frac{1}{2}}_{i+\frac{1}{2}\,j} + (f_x)^{n+\frac{1}{2}}_{i+\frac{1}{2}\,j}\Big\} \tag{2.5}$$

$$(v_y)^{n+1}_{i\,j+\frac{1}{2}} = (v_y)^{n}_{i\,j+\frac{1}{2}} + \Delta t\frac{1}{(\rho)_{i\,j+\frac{1}{2}}}\Big\{ - D_y^{h,(k)}(p)^{n+\frac{1}{2}}_{i\,j+\frac{1}{2}} + (f_y)^{n+\frac{1}{2}}_{i\,j+\frac{1}{2}}\Big\} \tag{2.6}$$

$$(p)^{n+\frac{1}{2}}_{ij} = (p)^{n-\frac{1}{2}}_{ij} + \Delta t\frac{1}{(\beta)_{ij}}\Big\{ - D_x^{h,(k)}(v_x)^{n}_{ij} - D_y^{h,(k)}(v_y)^{n}_{ij} + (g)^{n}_{ij}\Big\}, \tag{2.7}$$

where $p^{n+\frac{1}{2}}_{ij} = p(ih, jh, (n+\frac{1}{2})\Delta t)$.

2.4 Discontinuous Galerkin Method

In this section I derive the fully discrete DG method applied to the first-order system given by Eq.(2.1): a DG in space discretization and a Runge-Kutta method in time. I will use a nodal approach, as described by Hesthaven and Warburton (2007), with Lagrangian polynomials as basis functions. The derivation provided here is developed for the general case where the model parameters ρ and κ vary within each element in the partitioned domain. In particular, I will employ standard strategies from literature to deal with variable coefficients within DG elements in an efficient manner.

Approximation Space and Basis Functions

Let $\mathcal{T}_h$ denote a triangulation of Ω into K triangles τ^k, i.e., $\mathcal{T}_h = \{\tau^k\}_{k=1}^K$. I define the finite dimensional approximation space U_h as follows:

$$U_h := \{v_h : v_h|_\tau \in \mathbb{P}^N(\tau), \forall \tau \in \mathcal{T}_h\}, \tag{2.8}$$

where $\mathbb{P}^N(\tau)$ is the space of 2D polynomials of degree up to N restricted to some triangle τ in the triangulation $\mathcal{T}_h$. Thus, U_h is set of piecewise polynomials of order up to $N \in \mathbb{N}$ associated with $\mathcal{T}_h$.

There are $N^* := \frac{1}{2}(N+1)(N+2)$ degrees of freedom associated for each triangle $\tau \in \mathcal{T}_h$; given $\tau \in \mathcal{T}_h$ and a $u \in U_h$, restricted to τ, $u|_\tau$ can be decomposed into N^* linearly independent basis functions. A nodal DG method comprises of spanning the approximation space U_h by nodal basis functions, or equivalently Lagrange polyno-

mials. Suppose $\{\ell_j^k\}_{j=1}^{N^*}$ is such a nodal basis corresponding to τ^k. In order to specify this basis one would be required to identify a set of N^* distinct points $\{x_j^k\}_{j=1}^{N^*} \subset \tau^k$ (referred to as nodes) such that $\ell_j^k(x_i^k) = \delta_{ij}$. It is observed that the relationship between nodes and nodal basis functions is unique, assuming that the nodes are not degenerate, yielding a connection between degrees of freedom, basis functions, and nodes. Hence the number of degrees of freedom is equal to the number of distinct nodes one must provide in order to specify a particular set of nodal basis functions. In this thesis I choose the *α-optimized nodal set*, proposed by Hesthaven and Warburton (2007). These nodes are computed using a warp and blend technique on the equidistant conforming node set for some reference equilateral triangle. The space U_h is thus spanned after specifying the nodal sets $\{x_j^k\}_j^{N^*}$ for each triangle τ^k,

$$U_h = \bigoplus_{k=1}^{K} \text{span}\{\ell_j^k(\boldsymbol{x})\}_{j=1}^{N^*}, \qquad \text{where } \ell_j^k(\boldsymbol{x}_i^k) = \delta_{ij},\ \forall j = 1, ..., N^*.$$

Space Discretization

Assume Ω is triangulated into a partition $\mathcal{T}_h$, and let φ be some test function with the following properties: supp $\{\varphi\} \subseteq \tau$ and $\varphi \in C^\infty(\Omega)$, for some $\tau \in \mathcal{T}_h$. Multiplying the x-component of Eq.(2.1a) by φ and integrating over Ω gives,

$$\int_\tau \rho \frac{\partial v_x}{\partial t} \varphi \, d\boldsymbol{x} + \int_\tau \frac{\partial p}{\partial x} \varphi \, d\boldsymbol{x} = \int_\tau f_x \varphi \, d\boldsymbol{x}.$$

Integrating by parts and replacing p in the boundary integral with p^* yields

$$\int_\tau \rho \frac{\partial v_x}{\partial t} \varphi \, d\boldsymbol{x} - \int_\tau p \frac{\partial \varphi}{\partial x} \, d\boldsymbol{x} + \int_{\partial\tau} \hat{n}_x p^* \varphi \, d\sigma = \int_\tau f_x \varphi \, d\boldsymbol{x},$$

also referred to as the *weak formulation*, where $\hat{n}_x$ denotes the x-component of the outward unit vector $\hat{\boldsymbol{n}}$, normal to $\partial\tau$. The p^* term is referred to as the *numerical flux* of p and plays a vital role in the stability of the method. I will postpone further details of p^* to §2.4. Applying integration by parts yet again gives the *strong formulation*:

$$\int_\tau \rho \frac{\partial v_x}{\partial t} \varphi \, d\boldsymbol{x} + \int_\tau \frac{\partial p}{\partial x} \varphi \, d\boldsymbol{x} + \int_{\partial\tau} \hat{n}_x (p^* - p) \varphi \, d\sigma = \int_\tau f_x \varphi \, d\boldsymbol{x}.$$

Repeating this process for equations in Eq.(2.1) gives the following system:

$$\int_\tau \rho \frac{\partial v_x}{\partial t} \varphi \, d\boldsymbol{x} + \int_\tau \frac{\partial p}{\partial x} \varphi \, d\boldsymbol{x} + \int_{\partial\tau} \hat{n}_x (p^* - p) \varphi \, d\sigma = \int_\tau f_x \varphi \, d\boldsymbol{x}, \tag{2.9a}$$

$$\int_\tau \rho \frac{\partial v_y}{\partial t} \varphi \, d\boldsymbol{x} + \int_\tau \frac{\partial p}{\partial y} \varphi \, d\boldsymbol{x} + \int_{\partial\tau} \hat{n}_y (p^* - p) \varphi \, d\sigma = \int_\tau f_y \varphi \, d\boldsymbol{x}, \tag{2.9b}$$

$$\int_\tau \beta \frac{\partial p}{\partial t} \varphi \, d\boldsymbol{x} + \int_\tau (\nabla \cdot \boldsymbol{v}) \, \varphi \, d\boldsymbol{x} + \int_{\partial\tau} \varphi (\boldsymbol{v}^* - \boldsymbol{v}) \cdot \hat{\boldsymbol{n}} \, d\sigma = \int_\tau g \varphi \, d\boldsymbol{x}, \tag{2.9c}$$

referred to as the *strong formulation* for the acoustic equations, with corresponding numerical flux $\boldsymbol{v}^*$ for $\boldsymbol{v}$.

The semi-discrete problem consists of finding $p_h, v_{x,h}, v_{y,h} \in U_h$ such that Eq.(2.9) is satisfied for all test functions $\varphi \in U_h$. With the choice of nodal basis functions $\{\ell_j\}_{j=1}^{N^*}$, along with a given nodal set $\{\boldsymbol{x}_j\}_{j=1}^{N^*} \subset \tau$, the numerical approximations to

the velocity and pressure fields can be written as linear combinations of the nodal basis:

$$v_{x,h}(\boldsymbol{x},t)|_\tau = \sum_{j=1}^{N^*}(v_{x,h})_j(t)\ell_j(\boldsymbol{x}) \tag{2.10a}$$

$$v_{y,h}(\boldsymbol{x},t)|_\tau = \sum_{j=1}^{N^*}(v_{y,h})_j(t)\ell_j(\boldsymbol{x}) \tag{2.10b}$$

$$p_h(\boldsymbol{x},t)|_\tau = \sum_{j=1}^{N^*}(p_h)_j(t)\ell_j(\boldsymbol{x}) \tag{2.10c}$$

for each $\tau \in \mathcal{T}_h$, where $(v_{x,h})_j, (v_{y,h})_j, (p_h)_j$ are unknowns that serve as the coefficients in the nodal basis expansion. Hence integral terms over τ from Eq.(2.9a), and similarly for Eq.(2.9b) and Eq.(2.9c), result in computing weighted inner products between basis functions:

$$\int_\tau \rho\frac{\partial v_{x,h}}{\partial t}\ell_i \, d\boldsymbol{x} = \sum_{j=1}^{N^*}\frac{d}{dt}(v_{x,h})_j \int_\tau \rho\ell_i\ell_j \, d\boldsymbol{x}, \quad \text{and} \quad \int_\tau \frac{\partial p}{\partial x}\ell_i \, d\boldsymbol{x} = \sum_{j=1}^{N^*}(p_h)_j \int_\tau \ell_i\frac{\partial \ell_j}{\partial x} \, d\boldsymbol{x}$$

where $\varphi = \ell_i$ for $i = 1, ..., N^*$, in the general case where ρ and β may be varying within τ.

Now consider the surface integral in Eq.(2.9a) and note that the boundary $\partial\tau$ can be subdivided into integrals over edges $e \in \partial\tau$. Thus,

$$\int_{\partial\tau} \hat{n}_x(p_h^* - p_h)\ell_i \, d\sigma = \sum_{e\in\partial\tau}\int_e \hat{n}_x^{(e)}\left((p_h^{(e)})^* - p_h^{(e)}\right)\ell_i \, d\sigma, \tag{2.11}$$

where the superscript (e) is used to emphasize the dependency on of the fields and

$\hat{n}_x$ with respect to edge $e \in \partial\tau$. The α-optimized nodal set will have $N+1$ subsets of points $\{x_{m_j}\}_{j=1}^{N+1}$ located at each edge e by construction. Consequently, $v_{x,h}, v_{y,h}$, and p_h are uniquely defined on (e) by the $N+1$ nodes along with a subset of $N+1$ nodal basis functions for to the respective edge. In other words, a subset of the total element information is required to fully describe terms evaluated at edges, unlike the case for modal DG where all of the modes are needed. For example, $v_{x,h}$ at edge e, denoted by $v_{x,h}^{(e)}$, is given by

$$v_{x,h}(\boldsymbol{x},t)|_e = \sum_{j=1}^{N+1} (v_{x,h}^{(e)})_{m_j}(t)\ell_{m_j}^{(e)}(\boldsymbol{x}), \tag{2.12}$$

where $(v_{x,h}^{(e)})_{m_j}$ denote the $N+1$ subset of nodal coefficients associated with edge e, with corresponding nodal set $\{\boldsymbol{x}_{m_j}^{(e)}\}_{j=1}^{N+1}$ and basis functions $\{\ell_{m_j}^{(e)}\}_{j=1}^{N+1}$.

Conglomerating the results above for the semi-discrete form of Eq.(2.9a) yields the following:

$$\begin{aligned}\sum_{j=1}^{N^*} \frac{d}{dt}(v_{x,h})_j \int_\tau \rho\ell_i\ell_j \, d\boldsymbol{x} \;\; + \;\; & \sum_{j=1}^{N^*}(p_h)_j \int_\tau \ell_i \frac{\partial \ell_j}{\partial x} \, d\boldsymbol{x} \\ & + \sum_{e\in\partial\tau}\sum_{j=1}^{N+1} \hat{n}_x \left((p_h^{(e)})^* - p_h^{(e)}\right)_{m_j} \int_e \ell_i \ell_{m_j}^{(e)} \, d\sigma = \int_\tau f_x \ell_i \, d\boldsymbol{x}\end{aligned} \tag{2.13}$$

for all $i = 1, ..., N^*$. For a given $\tau \in \mathcal{T}_h$, I define the coefficient vectors in the following manner,

$$
\begin{aligned}
\mathbf{v}_{x,h} &:= [(v_{x,h})_1, (v_{x,h})_2, \ldots, (v_{x,h})_{N^*}]^T, && \mathbf{v}_{x,h} \in \mathbb{R}^{N^*}, \\
\mathbf{v}_{y,h} &:= [(v_{y,h})_1, (v_{y,h})_2, \ldots, (v_{y,h})_{N^*}]^T, && \mathbf{v}_{y,h} \in \mathbb{R}^{N^*}, \\
\mathbf{p}_h &:= [(p_h)_1, (p_h)_2, \ldots, (p_h)_{N^*}]^T, && \mathbf{p}_h \in \mathbb{R}^{N^*}, \\
\mathbf{v}_{x,h}^{(e)} &:= [(v_{x,h}^{(e)})_{m_1}, (v_{x,h}^{(e)})_{m_2}, \ldots, (v_{x,h}^{(e)})_{m_{N+1}}]^T, && \mathbf{v}_{x,h}^{(e)} \in \mathbb{R}^{(N+1)}, \\
\mathbf{v}_{y,h}^{(e)} &:= [(v_{y,h}^{(e)})_{m_1}, (v_{y,h}^{(e)})_{m_2}, \ldots, (v_{y,h}^{(e)})_{m_{N+1}}]^T, && \mathbf{v}_{y,h}^{(e)} \in \mathbb{R}^{(N+1)}, \\
\mathbf{p}_h^{(e)} &:= [(p_h)_{m_1}^{(e)}, (p_h)_{m_2}^{(e)}, \ldots, (p_h)_{m_{N+1}}^{(e)}]^T, && \mathbf{p}_h^{(e)} \in \mathbb{R}^{(N+1)}, \\
\mathbf{v}_{n,h}^{(e)} &:= \hat{n}_x \mathbf{v}_{x,h}^{(e)} + \hat{n}_y \mathbf{v}_{y,h}^{(e)}.
\end{aligned}
$$

With this notation I write the matrix form of Eq.(2.13),

$$
M[\rho] \frac{d}{dt} \mathbf{v}_{x,h} + S^x \mathbf{p}_h + \sum_{e \in \partial\tau} \hat{n}_x M^{(e)} \left((\mathbf{p}_h^{(e)})^* - \mathbf{p}_h^{(e)} \right) = \mathbf{f}^x
$$

where the local weighted mass matrices $M[\rho]$ and mass matrix $M^{(e)}$, x-stiffness matrix S^x, and right-hand-side vector $\mathbf{f}^x$ are given below:

$$
\begin{aligned}
M[\rho]_{ij} &:= \int_\tau \rho \ell_i \ell_j \, d\boldsymbol{x}, && M \in \mathbb{R}^{N^* \times N^*}, \\
M_{ij}^{(e)} &:= \int_e \ell_i^{(e)} \ell_{m_j}^{(e)} \, d\sigma, && M^{(e)} \in \mathbb{R}^{N^* \times (N+1)}, \\
S_{ij}^x &:= \int_\tau \ell_i \frac{\partial \ell_j}{\partial x} \, d\boldsymbol{x}, && S^x \in \mathbb{R}^{N^* \times N^*}, \\
\mathbf{f}_i^x &:= \int_\tau f_x \ell_i \, d\boldsymbol{x}, && \mathbf{f}^x \in \mathbb{R}^{N^*}.
\end{aligned}
$$

A similar approach is taken for Eq.(2.9b) and Eq.(2.9c):

$$M[\rho]\frac{d}{dt}\mathbf{v}_{y,h} + S^y\mathbf{p}_h + \sum_{e\in\partial\tau}\hat{n}_y M^{(e)}\left((\mathbf{p}_h^{(e)})^* - \mathbf{p}_h^{(e)}\right) = \mathbf{f}^y,$$

$$M[\beta]\frac{d}{dt}\mathbf{p}_h + S^x\mathbf{v}_{x,h} + S^y\mathbf{v}_{y,h} + \sum_{e\in\partial\tau} M^{(e)}\left((\mathbf{v}_{n,h}^{(e)})^* - \mathbf{v}_{n,h}^{(e)}\right) = \mathbf{g},$$

where the local weighted mass matrix $M[\beta]$, y-stiffness matrix S^y, and the right-hand-side vectors $\mathbf{f}^y, \mathbf{g}$ are given below:

$$\begin{aligned}
M[\beta]_{i,j} &:= \int_\tau \beta\ell_i\ell_j\, d\boldsymbol{x}, \quad M[\beta] \in \mathbb{R}^{N^*\times N^*},\\
S^y_{i,j} &:= \int_\tau \ell_i\frac{\partial\ell_j}{\partial y}\, d\boldsymbol{x}, \qquad S^y \in \mathbb{R}^{N^*\times N^*},\\
\mathbf{f}^y_i &:= \int_\tau f_y\ell_i\, d\boldsymbol{x}, \qquad \mathbf{f}^y \in \mathbb{R}^{N^*},\\
\mathbf{g}_i &:= \int_\tau g\ell_i\, d\boldsymbol{x}, \qquad \mathbf{g} \in \mathbb{R}^{N^*}.
\end{aligned}$$

Finally the explicit semi-discrete DG scheme is given in the following system of ODE's:

$$\frac{d}{dt}\mathbf{v}_{x,h} = -D^x[\tfrac{1}{\rho}]\mathbf{p}_h - \sum_{e\in\partial\tau}\hat{n}_x L^{(e)}[\tfrac{1}{\rho}]\left((\mathbf{p}_h^{(e)})^* - \mathbf{p}_h^{(e)}\right) + F^x[\tfrac{1}{\rho}] \tag{2.14a}$$

$$\frac{d}{dt}\mathbf{v}_{y,h} = -D^y[\tfrac{1}{\rho}]\mathbf{p}_h - \sum_{e\in\partial\tau}\hat{n}_y L^{(e)}[\tfrac{1}{\rho}]\left((\mathbf{p}_h^{(e)})^* - \mathbf{p}_h^{(e)}\right) + F^y[\tfrac{1}{\rho}] \tag{2.14b}$$

$$\frac{d}{dt}\mathbf{p}_h = -D^x[\tfrac{1}{\beta}]\mathbf{v}_{x,h} - D^y[\tfrac{1}{\beta}]\mathbf{v}_{y,h} - \sum_{e\in\partial\tau} L^{(e)}[\tfrac{1}{\beta}]\left((\mathbf{v}_{n,h}^{(e)})^* - \mathbf{v}_{n,h}^{(e)}\right) + G[\tfrac{1}{\beta}] \tag{2.14c}$$

where

$$D^{\alpha}[\tfrac{1}{\omega}] := M[\omega]^{-1}S^{\alpha},$$

$$L^{(e)}[\tfrac{1}{\omega}] := M[\omega]^{-1}M^{(e)},$$

$$F^{\alpha}[\tfrac{1}{\omega}] := M[\omega]^{-1}\mathbf{f}^{\alpha},$$

$$G[\tfrac{1}{\omega}] := M[\omega]^{-1}\mathbf{g},$$

are the DG operators weighted by some general function ω, where ω in this context is either ρ or β, and $\alpha \in \{x, y\}$.

Handling Variable Coefficients

The derivation of the semi-discrete scheme given by Eq.(2.14) was carried out under the general assumption of varying acoustic coefficients ρ and β within DG elements, hence resulting in weighted operators $D^{\alpha}[\frac{1}{\omega}], L^{(e)}[\frac{1}{\omega}]$. The efficiency and accuracy in computing these weighted operators will no doubt affect the computation time of the overall DG method. Common DG schemes will assume that physical coefficients can be considered constant within elements, leading to a piecewise constant representation of the medium. An alternative would be to carry out the actual integration via quadrature rules, leading to an increase in accuracy though at some cost. For completeness I have implemented both variations of the DG scheme given in Eq.(2.14) based on quadrature-free and quadrature based computations of weighted operators which I briefly discuss subsequently.

The **first approach** consists of assuming that ρ and β are constant within elements $\tau \in \mathcal{T}_h$. In particular for smoothly varying media, as done in finite volume methods (LeVeque, 2002), I will take a homogenization approach and use the average of the coefficients over τ as the representative constant,

$$\rho\big|_\tau \approx \bar{\rho}_\tau \equiv \frac{1}{|\tau|}\int_\tau \rho(\boldsymbol{x})\, d\boldsymbol{x}, \tag{2.15a}$$

$$\beta\big|_\tau \approx \bar{\beta}_\tau \equiv \frac{1}{|\tau|}\int_\tau \beta(\boldsymbol{x})\, \boldsymbol{x}, \tag{2.15b}$$

denoting the averages of ρ and β over τ. With this approach, local operators from Eq.(2.14) take a simplified form:

$$D^\alpha[\tfrac{1}{\omega}] = \frac{1}{\bar{\omega}_\tau} M^{-1} S^\alpha,$$

$$L^{(e)}[\tfrac{1}{\omega}] = \frac{1}{\bar{\omega}_\tau} M^{-1} M^{(e)},$$

$$F^\alpha[\tfrac{1}{\omega}] = \frac{1}{\bar{\omega}_\tau} M^{-1} \mathbf{f}^\alpha,$$

$$G[\tfrac{1}{\omega}] = \frac{1}{\bar{\omega}_\tau} M^{-1} \mathbf{g},$$

where $M \equiv M[1]$ is the standard unweighted mass matrix on τ.

Standard mass matrices M and $M^{(e)}$, and stiffness matrices S^α for $\alpha \in \{x, y\}$, consist of evaluating inner products between basis functions of the following form:

$$\int_\tau \ell_i \ell_j \, d\boldsymbol{x}, \quad \int_e \ell_i \ell_{m_j} \, d\sigma, \quad \int_\tau \ell_i \frac{\partial \ell_j}{\partial \alpha} \, d\boldsymbol{x}.$$

These integrals can be related via an affine map $\Psi : \tau \to \hat{\tau}$, with $(x, y) \mapsto (\hat{x}, \hat{y})$, to integrals on some reference triangle $\hat{\tau}$ and reference edge $\hat{e}$, i.e.,

$$J(\tau) \int_{\hat{\tau}} \hat{\ell}_i \hat{\ell}_j \, d\hat{\boldsymbol{x}}, \quad J(e) \int_{\hat{e}} \hat{\ell}_i \hat{\ell}_{m_j} \, d\hat{\sigma}, \quad J(\tau) \int_{\hat{\tau}} \hat{\ell}_i \left(\frac{\partial \hat{x}}{\partial \alpha} \frac{\partial}{\partial \hat{x}} + \frac{\partial \hat{y}}{\partial \alpha} \frac{\partial}{\partial \hat{y}} \right) \hat{\ell}_j \, d\hat{\boldsymbol{x}}, \tag{2.16}$$

where $J(\cdot)$ is the Jacobian of transformation between $\tau \to \hat{\tau}$ or $e \to \hat{e}$, and $\{\hat{\ell}_i\}_{i=1}^{N^*}$ are the basis functions for the space $\mathbb{P}^N(\hat{\tau})$. Consequently, it suffices to store scalar transformation factors $J(\tau), J(e)$, and scalar geometric factors $\frac{\partial \hat{x}}{\partial \alpha}, \frac{\partial \hat{y}}{\partial \alpha}$ for each τ and one copy of reference operators resulting from inner products on reference elements given in Eq.(2.16). Computing the element-wise operators will thus consist of linear combinations of the reference operators as hinted by Eq.(2.16), as oppose to having to compute integrals for each element. Hesthaven and Warburton (2007) give a detailed account of the approach used in this thesis for computing and assembling in an efficient manner, along with a quadrature free implementation, of the reference mass and stiffness matrices and transformation factors.

The homogenization step in the method sketched above, in which the medium is effectively approximated by a carefully chosen piecewise constant analog, is instrumental to an efficient quadrature-free implementation of the DG method. The accuracy of this approach thus rests in part with how well the wave response of the

homogenized medium approximates wave propagation in the true medium. Two limiting factors to this homogenization approach include: the size of elements τ, and the variability and complexity of the medium. Intuitively, higher accuracy in complicated media can be achieved by naively decreasing the size of elements τ, indeed the only option there is if one is to maintain a quadrature-free implementation.

A natural alternative to the first approach is a quadrature based implementation, and thus the **second approach** considered in this thesis. Quadrature rules accurate up to polynomials of order $2N + N_\omega$, where N_ω is extra precision imposed on the numerical integration to account for the weighting factor ω, are used to compute the weighted integrals

$$\int_\tau \omega \ell_i \ell_j \, d\boldsymbol{x}.$$

This quadrature based approach is thus more accurate since it does not suffer from approximation errors, aside from numerical integration, that are of issue for the first approach. As is common, higher accuracy comes at a price, in particular having to store each of the weighted operators for each element. For example, suppose there are K elements, then the amount of memory needed to store all of the weighted operators is

$$\begin{aligned} K \times \Big(sizeof(D^x) + sizeof(D^y) + 3 * sizeof(L^{(e)}) \Big) \\ &= K \times N^* \times \Big(2N^* + 3(N+1) \Big) \\ &= \mathcal{O}(KN^4) \end{aligned}$$

modulus $sizeof(float)$. This is contrasted with

$$\begin{aligned} K \times \Big(sizeof(J(\tau)) + 3 * sizeof(J(e)) + 2 * sizeof(geo.fac.) \Big) & \\ + sizeof(M) + 2 * sizeof(S^{\alpha}) & \\ = 6K + 3N^* \times N^* & \\ = \mathcal{O}(K + N^4) & \end{aligned}$$

for the quadrature-free approach.

Numerical Fluxes

The numerical flux terms p^* and $\boldsymbol{v}^*$ in Eq.(2.9), and consequently in the semi-discrete DG scheme Eq.(2.14), can be interpreted as the means through which information between elements is propagated. These flux terms more importantly play a crucial role in the numerical stability of the method. In particular, with regards to hyperbolic problems where energy is conserved, the flux terms essentially restrict the discrete energy analog from growing over time through dissipative or conservative mechanisms. Consequently flux terms also have an impact in the methods accuracy, e.g., convergence rates, dissipation and dispersion properties. The choice of flux is not unique and continuous to be a topic of interest; see Cockburn (2003) for an accessible overview of DG methods including an excellent discussion on numerical fluxes. In this subsection I will give an overview of the derivation, stemming from *Riemann solvers*, of the flux terms used for the DG method considered in this thesis. A complete exposé

of this derivation, and a more in depth discussion on Riemann solvers can be found in LeVeque (2002), within the context of finite volume methods.

The acoustic equations in Eq.(2.1) can be reformulated in matrix-vector notation as follows:

$$Q\frac{\partial \boldsymbol{q}}{\partial t} + \mathcal{A}_x\frac{\partial \boldsymbol{q}}{\partial x} + \mathcal{A}_y\frac{\partial \boldsymbol{q}}{\partial y} = \tilde{\boldsymbol{s}}$$

where

$$Q = \begin{bmatrix} \rho & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & \beta \end{bmatrix}, \quad \mathcal{A}_x = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mathcal{A}_y = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

and

$$\boldsymbol{q} = \begin{bmatrix} v_x \\ v_y \\ p \end{bmatrix}, \quad \tilde{\boldsymbol{s}} = \begin{bmatrix} f_x \\ f_y \\ g \end{bmatrix}.$$

Define the operator $\Pi = \hat{n}_x\mathcal{A}_x + \hat{n}_y\mathcal{A}_y$, and note that

$$Q^{-1}\Pi\boldsymbol{q} = \begin{bmatrix} \frac{1}{\rho}\hat{n}_x p \\ \frac{1}{\rho}\hat{n}_y p \\ \frac{1}{\beta}\hat{\boldsymbol{n}}\cdot\boldsymbol{v} \end{bmatrix}$$

corresponds to the integrands in the surface integrals of Eq.(2.9), i.e., the flux terms. One can show that the eigenvalues of $Q^{-1}\Pi$ are given by $\{-c, 0, c\}$, where $c = 1/\sqrt{\rho\beta}$ is the acoustic wave velocity.

The Riemann solver provides a solution to the case where an acoustic wave travels between two media with constant and differing coefficients ρ and β. A jump discontinuity in the wave field occurs when the incident wave travels through the two media, from which two smooth wave fields, a reflected and transmitted wave, originate and propagate with appropriate velocities. The intermediate wave field that exhibits the jump discontinuity at the interface of the boundary between the two media is what will be used as the flux term for the DG formulation.

Let Q^- and Q^+ denote the coefficient matrices for the two media and similarly let $\boldsymbol{q}^-$ and $\boldsymbol{q}^+$ denote the respective states. The Riemann jump conditions state that intermediate states $\boldsymbol{q}^*$ and $\boldsymbol{q}^{**}$ must satisfy the following relations between $\boldsymbol{q}^+$ and $\boldsymbol{q}^-$,

$$c^- Q^-(\boldsymbol{q}^* - \boldsymbol{q}^-) + (\Pi \boldsymbol{q})^* - (\Pi \boldsymbol{q})^- = 0$$

$$(\Pi \boldsymbol{q})^{**} - (\Pi \boldsymbol{q})^* = 0$$

$$-c^+ Q^+(\boldsymbol{q}^{**} - \boldsymbol{q}^+) + (\Pi \boldsymbol{q})^{**} - (\Pi \boldsymbol{q})^+ = 0$$

With some algebraic manipulation one can derive the following formulas for the intermediate states p^* and $\boldsymbol{v}^*$:

$$p^* = \frac{Z^- p^+ + Z^+ p^-}{Z^- + Z^+} - \frac{Z^- Z^+}{Z^- + Z^+} \hat{\boldsymbol{n}} \cdot (\boldsymbol{v}^+ - \boldsymbol{v}^-) \tag{2.17a}$$

$$\boldsymbol{v}^* = \frac{Z^- \boldsymbol{v}^- + Z^+ \boldsymbol{v}^+}{Z^- + Z^+} - \frac{1}{Z^- + Z^+} \hat{\boldsymbol{n}} (p^+ - p^-) \tag{2.17b}$$

where $Z = \sqrt{\rho/\beta}$ is the acoustic impedance.

Consider a triangulation $\mathcal{T}$ of the domain Ω, and let $\tau \in \mathcal{T}$. For some field q defined on Ω, the '$-$' and '$+$' superscripts refer to the inner and outer trace, respectively, of q at $\partial\tau$ with respect to τ. In other words,

$$q^+(\boldsymbol{x}) = \lim_{\epsilon \to 0} q(\boldsymbol{x} + \epsilon \hat{\boldsymbol{n}}) \quad \text{and} \quad q^-(\boldsymbol{x}) = \lim_{\epsilon \to 0} q(\boldsymbol{x} - \epsilon \hat{\boldsymbol{n}})$$

for $\boldsymbol{x} \in \partial\tau$, and $\hat{\boldsymbol{n}}$ is the unit vector normal and outward to $\partial\tau$. The numerical flux terms $(p^* - p^-)$ and $(\boldsymbol{v}^* - \boldsymbol{v}^-)$ that appear in the surface integrals in Eq.(2.9) are given as follows, when taking p^* and $\boldsymbol{v}^*$ to be the intermediate states from the Riemann jump conditions in Eq.(2.17):

$$p^* - p^- = \frac{Z^-}{2\langle\langle Z \rangle\rangle} \Big(\alpha Z^+ \hat{\boldsymbol{n}} \cdot [\![\boldsymbol{v}]\!] - [\![p]\!] \Big) \tag{2.18a}$$

$$\boldsymbol{v}^* - \boldsymbol{v}^- = \frac{Y^-}{2\langle\langle Y \rangle\rangle} \Big(\alpha Y^+ \hat{\boldsymbol{n}} [\![p]\!] - [\![\boldsymbol{v}]\!] \Big) \tag{2.18b}$$

with

$$\langle\langle q \rangle\rangle = \frac{q^- + q^+}{2} \quad \text{and} \quad [\![q]\!] = q^- - q^+$$

and acoustic conductance $Y = 1/Z$. The auxiliary variable α controls the amount of dissipation of the numerical flux and takes values in $[0, 1]$; e.g., $\alpha = 0$ and $\alpha = 1$ correspond to a non-dissipative central flux and a dissipative upwind flux respectively.

In the case where τ has an edge on $\partial\Omega$, the exterior values are extended by taking

$$p^+ = -p^- \quad \text{and} \quad \boldsymbol{v}^+ = \boldsymbol{v}^-,$$

for the free surface boundary condition $p = 0$ at $\partial\Omega$.

Perfectly Matched Layer Implementation

The physical region modeled in seismology applications corresponds to a small window view of the Earth, essentially an unbounded medium. The simplest manner in which to mimic an unbounded medium numerically would be to extend the computational domain such that the boundary effects (depending on what boundary conditions are considered) will not reach the region of interest at the end of the simulation time. Though simple to implement, this naive computational domain extension trick is rather expensive and inefficient, i.e., one ends up computing fields at a considerable amount of unnecessary points. The perfectly matched layer (PML) is a popular technique for simulating wave propagation through an unbounded domain. The method of PMLs was first developed in the context of electromagnetic wave propagation in free space by Berenger (1994). The idea consist of appending a lossy layer around your physical domain with the following properties:

1. there should be no reflected waves between the PML and the physical domain,

2. and the wave should decay exponentially within the PML.

Computationally, the PML method amounts to solving for auxiliary fields and auxiliary PDEs or ODEs, specifics depending on construction and implementation. I have implemented the PML construction derived by Abarbanel and Gottlieb (1998), where the PML equations take form of an inhomogeneous damped acoustic wave equation in first order form augmented by a set of auxiliary ODEs:

$$\rho\frac{\partial v_x}{\partial t} + \frac{\partial p}{\partial x} = f_x + \rho\eta_x(P_x - 2v_x) \tag{2.19a}$$

$$\rho\frac{\partial v_y}{\partial t} + \frac{\partial p}{\partial y} = f_y + \rho\eta_y(P_y - 2v_y) \tag{2.19b}$$

$$\beta\frac{\partial p}{\partial t} + \nabla\cdot\boldsymbol{v} = g - \beta\frac{\partial \eta_x}{\partial x}Q_x - \beta\frac{\partial \eta_y}{\partial y}Q_y \tag{2.19c}$$

$$\frac{\partial P_x}{\partial t} + \eta_x v_x = 0 \tag{2.20a}$$

$$\frac{\partial P_y}{\partial t} + \eta_y v_y = 0 \tag{2.20b}$$

$$\beta\frac{\partial Q_x}{\partial t} + \beta\eta_x Q_x = v_x \tag{2.20c}$$

$$\beta\frac{\partial Q_y}{\partial t} + \beta\eta_y Q_y = v_y \tag{2.20d}$$

where P_x, P_y, Q_x and Q_y are the auxiliary fields and η_x and η_y are the PML coefficients given by the following equations (for a domain setup similar to Fig.(2.2)):

$$\eta_\alpha(x,y) = \begin{cases} \eta_{\alpha_{max}} \left[\dfrac{\frac{L_\alpha}{2} + \alpha}{d_\alpha} \right]^2 & , \quad \alpha \in [-d_\alpha - \frac{L_\alpha}{2}, -\frac{L_\alpha}{2}] \\ 0 & , \quad \alpha \in (-\frac{L_\alpha}{2}, \frac{L_\alpha}{2}) \\ \eta_{\alpha_{max}} \left[\dfrac{\frac{L_\alpha}{2} - \alpha}{d_\alpha} \right]^2 & , \quad \alpha \in [\frac{L_\alpha}{2}, \frac{L_\alpha}{2} + d_\alpha] \end{cases} \tag{2.21}$$

for $\alpha \in \{x, y\}$. PML constants $\eta_{\alpha_{max}}$ and d_α denote the PML amplitude and layer widths.

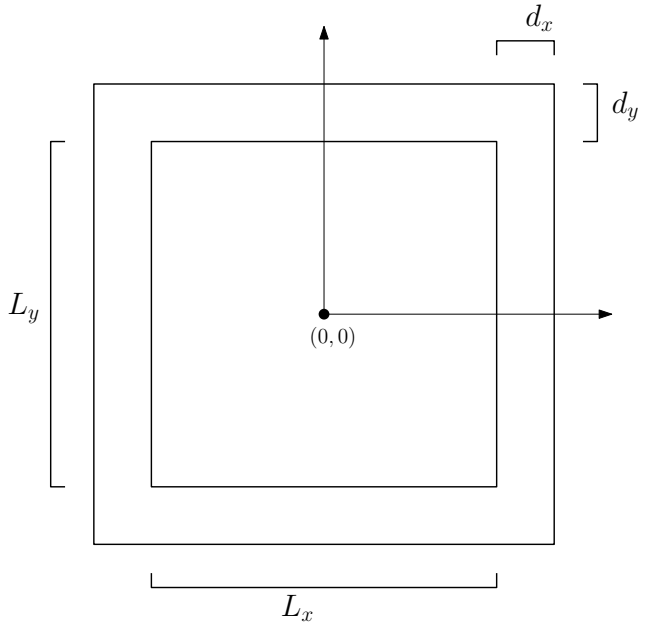


Figure 2.2 : Schematic of physical domain and PML.

Time Discretization

A member of the Runge-Kutta (RK) discretization family, the low-storage five-stage fourth-order explicit RK method, is used to discretize in time the system of ODE's in Eq.(2.14).

Chapter 3

Numerical Experiments and Results

3.1 Introduction

This chapter elaborates on the setup and results of numerical experiments carried out. FD and DG methods are tested for several different smooth models: homogeneous, linear-in-depth, and negative-lens velocity models with constant density. Lastly for the sake of completeness, a medium with a dip discontinuity and a mixture of linear-in-depth and negative lens velocity models is tested. In application, the velocity and density coefficients are given at spatial points corresponding to a uniform grid, from which one would have to interpolate for coefficients at points required by the solver. This process is idealized in that coefficients are given by some analytical formula and hence can be computed at any desired point without worrying about interpolation error.

The DG code was implemented in Matlab and is based off of code from Hesthaven and Warburton (2007), while the FD solver is taken from IWAVE (Symes et al., 2009) which is written in C and C++. Computations carried out on a personal laptop with the following specifications: 2.9GHz dual-core Intel Core i7 processor (Turbo Boost up to 3.6GHz) with 4MB L3 cache, and 8GB of 1600MHz DDR3 memory. I emphasize

that there was no attempt to parallelize the DG code since it was written in Matlab, out of simplicity, which will of course limit any type of comparison between FD and DG. Further discussion on the choices of implementation and its effect on this comparison will be discussed in the last chapter. Nonetheless, FLOP counts will be used to determine in a more software/hardware independent manner a metric for computational costs per method given a prescribed accuracy. In this thesis FLOPs are counted in an algorithmic manner taking into consideration PML regions, since they contribute in a significant manner the total number of operations computed.

Throughout the numerical experiments discussed below I will be considering a source function that is a Ricker wavelet in time and some smooth approximation to the delta function in space, see Appendix-A for more detail. The choice of Ricker in time, with a central frequency of 10Hz, is common in computational seismology for seismic imaging. Moreover, an isotropic point source serves as an "approximation" to explosive sources, a reasonable heuristic given the length scale of the survey region versus the volume where the explosion takes place. In practice the numerical representation of singularities is of course nontrivial, in particular within the context of finite difference methods. Though methodology for discretizing singular sources does exist (e.g., Petersson and Sjögreen (2010)), I have decided in replacing $\delta(\boldsymbol{x})$ with a smooth non-singular alternative of bounded support, thus relieving this study of complications due to discretizing singularities.

I point out that my numerically tractive source term has unintentionally resulted

in complicating the analytics of the problem. In other words writing an analytical solution to the acoustic equations Eq.(2.1) is at best difficult even in the simplest case of a homogeneous media for the type of source considered here. Customary practice dictates the use of highly discretized numerical solutions as the "true solution" when analytical ones are not easily available. A highly discretized 2-4 FD solution will be considered as the "true solution" in the subsequent test cases.

3.2 Defining Errors

The end goal of seismic simulators, at least in the context of seismic imaging, is to produce accurate seismic traces given information about the medium. Seismic traces are simply time series of the pressure field p evaluated at some spatial points $\boldsymbol{x}_r$ related to the location of receivers in the application. Hence, it will suffice to consider $p(\boldsymbol{x}_r, t)$ for some $\boldsymbol{x}_r \in \Omega$ and $t \in [0, T]$ while computing errors.

As alluded in the introduction of this chapter, errors will be measured relative to a highly discretized numerical solution. In particular, the FD 2-4 method with grid size of $dx = dy = 0.5\ m$ is used as the "true solution" when measuring error for all subsequent numerical experiments. Given a numerical solution p_h, the relative error at a specified spatial point $\boldsymbol{x}_r$ is denoted by $E_h(\boldsymbol{x}_r)$;

$$E_h(\boldsymbol{x}_r) = \frac{\|p_h(\boldsymbol{x}_r, \cdot) - p(\boldsymbol{x}_r, \cdot)\|}{\|p(\boldsymbol{x}_r, \cdot)\|},$$

where p will denote the "true solution" and $\|\cdot\|$ is as defined by

$$\|p_h(\boldsymbol{x}_r, \cdot)\| = \sqrt{\sum_{n=0}^{N} |p_h(\boldsymbol{x}_r, t_i)|^2}. \tag{3.1}$$

The following are accuracy conditions imposed on all preceding FLOP and computation time comparisons:

$$\underset{\boldsymbol{x}_r}{\text{RMS}}\, E_h(\boldsymbol{x}_r) < 5\% \tag{3.2}$$

$$\max_{\boldsymbol{x}_r} E_h(\boldsymbol{x}_r) < 6\% \tag{3.3}$$

I emphasize that this choice of accuracy conditions is somewhat arbitrary though justifiable and derived from engineering practices. The RMS error provides a rough average of the error while the max error criterion limits the maximum worst observed variability of the error. Mesh/grid size h, along with time steps, are chosen such that the accuracy criterion is met while attempting to minimize runtime.

3.3 Convergence Rates

Prior to carrying out numerical experiments for the comparison of FD and DG methods, a numerical convergence rate analysis is considered as a means of validating the correctness of the implemented methods. Let p denote the true solution to the acoustic equations, for some specified boundary conditions and domain Ω. The numerical solution will be denoted by p_h, where the subscript h emphasizes the dependency of the solution p_h on the mesh/grid size. Suppose that p_h was computed via a numerical

method with a convergence rate of $R > 0$, i.e.,

$$p_h = p + Ch^R + \mathcal{O}(h^{R+1}).$$

Taking the ratio of the difference of numerical solutions for $h, h/2, h/4$ can yield an approximation to the convergence rate R as follows:

$$\begin{aligned} \frac{p_h - p_{h/2}}{p_{h/2} - p_{h/4}} &= \frac{Ch^R - C(h/2)^R + \mathcal{O}(h^{R+1})}{C(h/2)^R - C(h/4)^R + \mathcal{O}(h^{R+1})} \\ &= \frac{1 - 2^{-R} + \mathcal{O}(h)}{2^{-R} - 4^{-R} + \mathcal{O}(h)} \\ &= 2^R + \mathcal{O}(h) \end{aligned}$$

$$\implies R \approx \ log_2 \left(\frac{p_h - p_{h/2}}{p_{h/2} - p_{h/4}} \right). \tag{3.4}$$

The convergence rates are computed by taking the ℓ^2-norm in time of the differences in Eq.(3.4), for a specified receiver location $\boldsymbol{x}_r$:

$$R(\boldsymbol{x}_r) = log_2 \left(\frac{\|p_h(\boldsymbol{x}_r, \cdot) - p_{h/2}(\boldsymbol{x}_r, \cdot)\|}{\|p_{h/2}(\boldsymbol{x}_r, \cdot) - p_{h/4}(\boldsymbol{x}_r, \cdot)\|} \right).$$

The numerical pressure fields for this convergence analysis are computed by solving the acoustic equations in a homogeneous density and velocity medium*: $\rho =$

*Velocity as a property of the medium, denoted by c, is related to κ (or β) and ρ by $(\kappa/\rho)^{1/2} = (\kappa\beta)^{-1/2} = c$. Computations with regards to the DG method will indeed involve compressibility β, however it is more intuitive and qualitative to show the velocity of the medium, as done throughout this chapter.

2.3 g/cm^3, and $c = 3\ km/s$. Physical domain Ω is shown in Fig. (3.1a), along with receiver locations $\boldsymbol{x}_r$ and source configuration. A PML boundary is also shown Fig.(3.1a), effectively simulating an unbounded medium†. A sample of the structured triangular meshes considered here is shown in Fig.(3.1b).

Numerical convergence rates $R(\boldsymbol{x}_r)$ are plotted in Fig.(3.2) for FD and DG methods respectively. Special care was taken to suppress second order errors stemming from the time derivative methods considered for the FD schemes by choosing a small enough time step.

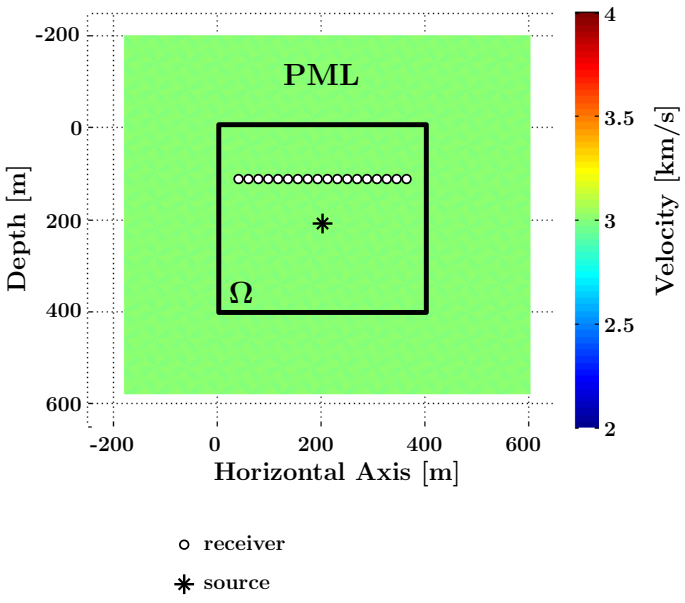
	FD 2-2	FD 2-4	DG N=2	DG N=4
RMS $R(\boldsymbol{x}_r)$	2.076	3.844	1.398	3.971

Table 3.1 : RMS with respect to $\boldsymbol{x}_r$ for convergence rates $R(\boldsymbol{x}_r)$.

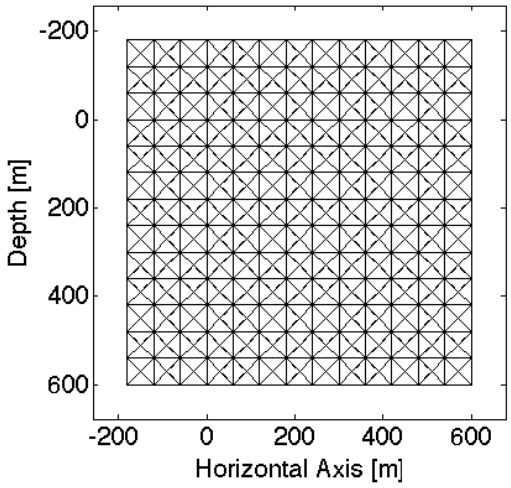
3.4 Homogenous Test Case

The first test case considered here is that of the homogeneous medium, with $\rho = 2.3\ g/cm^3$ and $c = 3\ km/s$. Physical setup of the numerical experiment in shown in Fig.(3.3a). Note that the physical region consists of the rectangular section $\Omega = [0\ m, 400\ m] \times [0\ m, 1200\ m]$, where regions to the left of 0 m and right of 400 m with respect to the horizontal axis correspond to PML regions, absorbing any waves propagating outside of Ω. Furthermore, free surface boundary conditions are applied for boundaries at depths 0 m and 1200 m. The setup depicted in Fig. (3.3a) was

†PML widths and configurations will be model and method dependent.

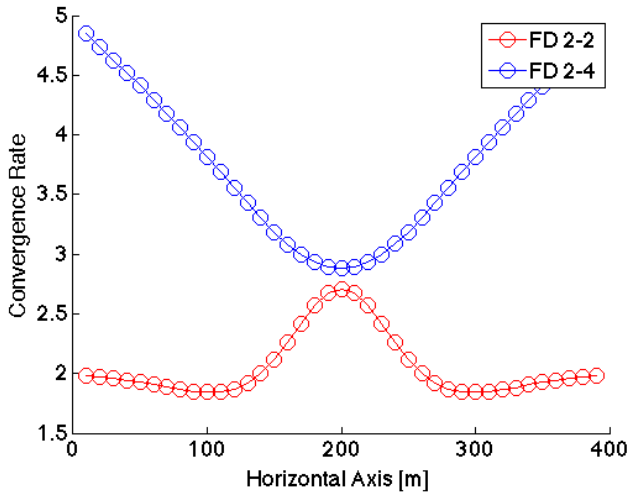


(a) Experimental setup.

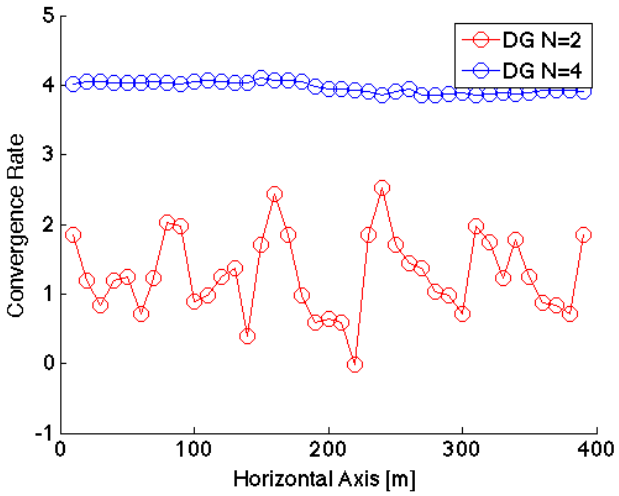


(b) Sample mesh on domain.

Figure 3.1 : Setup and sample structured mesh for convergence rate test.



(a) $R(\boldsymbol{x}_r)$ for FD methods.



(b) $R(\boldsymbol{x}_r)$ for DG methods.

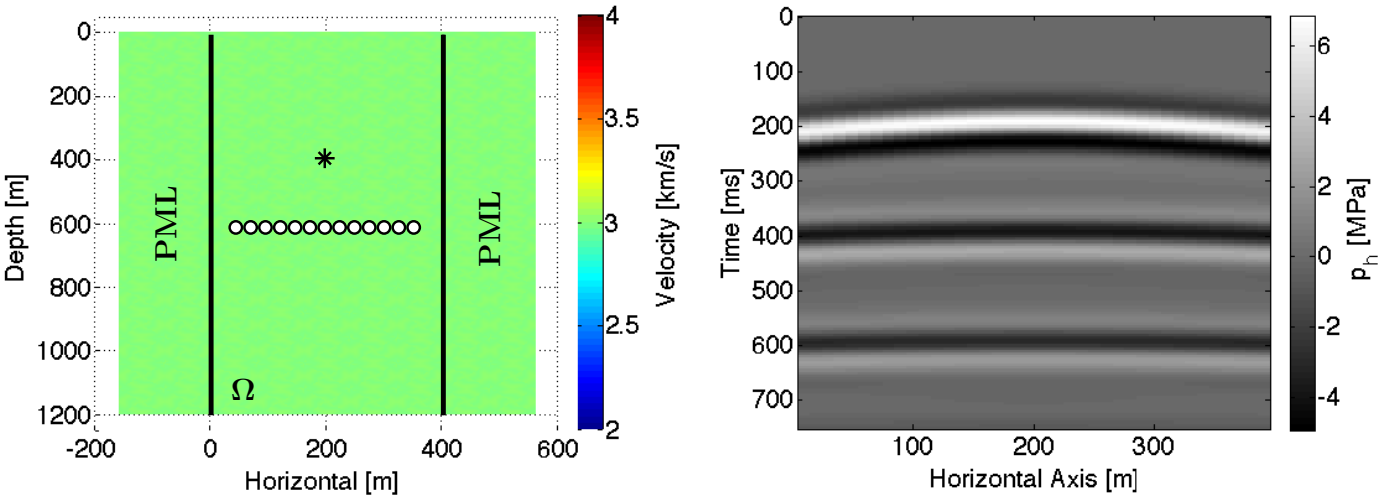
Figure 3.2 : Convergence rates for numerical methods at points $\boldsymbol{x}_r$.

constructed in order for receivers to measure the propagation of the initial wave over several wavelengths. In fact, due to the constant medium velocity of 3 km/s and a Ricker wavelet with central frequency of 10 Hz, it follows that the receivers at the depth of 600 m will measured up to 5 wavelengths ($5 \times 300 = 1500\ m$) of wave propagation.

Relative errors as functions of horizontal displacement are plotted for both FD and DG methods in Fig.(3.4). Simulations were carried out as to yield computed p_h that satisfy accuracy conditions Eq.(3.2) and Eq.(3.3). Table 3.2 provides a summary overview of the performances for FD and DG methods. Grid points per wavelength (GPW) is computed as $c_{min}/(f_{peak}h)$ for FD methods, while $c_{min}/(f_{peak}hN)$ gives an approximate analog for DG methods. Tables in Appendix B include more information about the discretization and computed relative errors for this and all test cases. Overall, FD methods yield GFLOP counts that are 150 to 25 times smaller than that of the DG methods. Further discussion and comparison of computational efficiency between DG and FD is reserved to the conclusion chapter.

	dt [ms]	GPW	GFLOPs
FD 2-2	0.9546	50	0.55
FD 2-4	1.6122	25	0.12
DG N=2	1.2032	15	14.28
DG N=4	1.2058	15	17.99

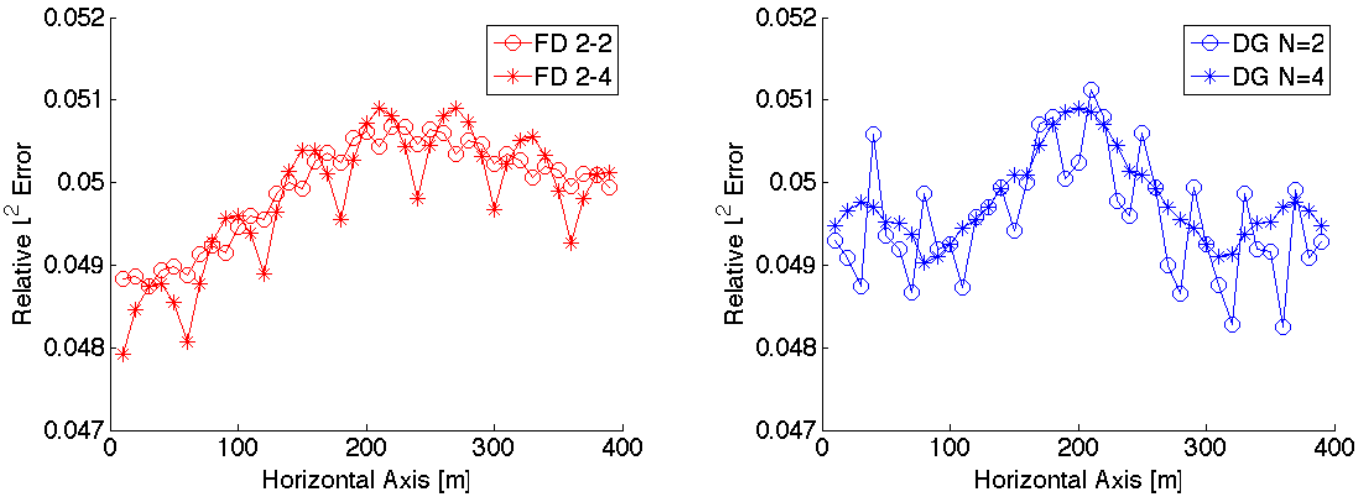
Table 3.2 : Results for homogeneous test case.



(a) Experimental setup.

(b) Traces of highly discretized numerical solution.

Figure 3.3 : Experimental setup and traces for homogeneous test case.



(a) Relative errors for FD.

(b) Relative error for DG.

Figure 3.4 : Relative errors for homogeneous model.

3.5 Linear-in-Depth Velocity Test Case

The following test case compares the performance of FD and DG methods on a slowly varying medium: consttant density and linearly varying velocity in depth medium, as shown in Fig.(3.5a). The receiver-source setup used here is similar to that of the homogeneous test case. Seismic traces for the highly discretized FD 2-4 method are shown in Fig.(3.5b).

Several combinations of different DG implementations are considered for this test case: mainly quadrature-free and quadrature based DG implementations with and without mesh refinement options. Mesh refinement is applied in regions of lower velocity where waves, as they propagate through such regions, their wavelength decrease dictated by the relationship between frequency f, velocity c and wavelength λ:

$$c = \lambda f. \tag{3.5}$$

The stability and accuracy of FD and DG methods in fact depend on how finely discretized the method is with respect to all of the quantitates present in Eq.(3.5). In particular regions of lower velocity will need to be discretized at a finer level than other regions. This study will follow that heuristic and apply mesh refinement for such regions in variable media.

The use of quadrature based versus quadrature-free DG will of course be important to this and the following test case where coefficients β and ρ vary within

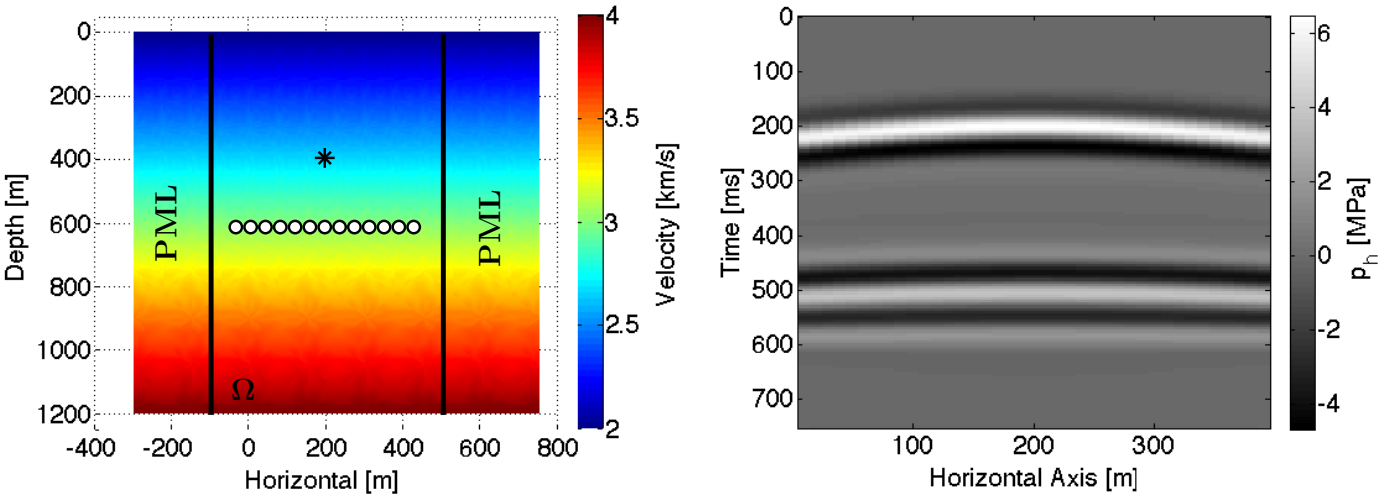
triangular elements. Moreover, the impact in overall accuracy and computation cost will notably depend on the mesh and, if any, mesh refinement. Exemplary meshes, with and without refinement, are shown in Fig.(3.6) and Fig.(3.7) along with their corresponding piecewise constant approximations to the variable velocity model in the case of quadrature-free DG.

Relative errors and their variability across receiver locations is shown in Fig.(3.8) for all methods, including combinations of DG implementations dealing with variable media. Among the DG methods, and their variable implementations, it is observed that lower oder methods and implementations with mesh refinement seem to exhibit higher variations in errors for the most part. Note that Table 3.3 includes runtimes for DG methods, this is done in order to make comparisons between quadrature-free and quadrature based methods, as well as mesh refinement options.The following are some observations summarized in Table 3.3:

- FD GFLOP count is again roughly between 192 to 9 times smaller than DG;
- the DG method with fastest runtimes has a GFLOP counter of roughly 175 times that of the 2-4 FD method;
- quadrature-free DG resulted in faster runtimes;
- mesh refinement provide some decrease in runtimes;
- among DG methods, the $N = 4$ quadrature-free (with or without mesh refinement) implementation yielded the smaller runtimes.

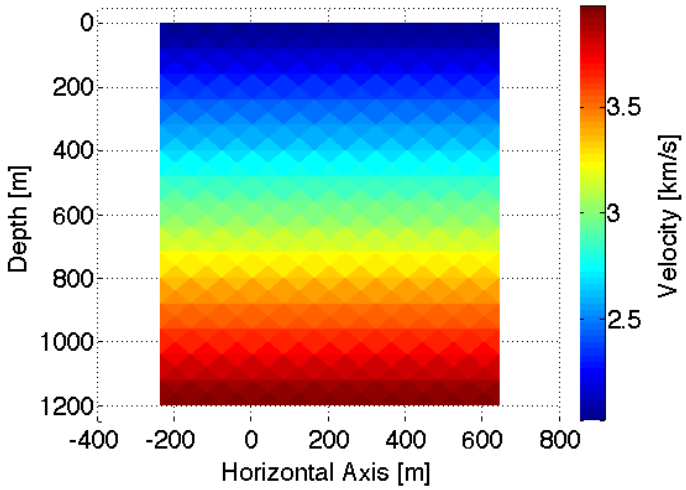
	dt [ms]	GPW	GFLOPs		Runtime [s]
FD 2-2	0.4679	33.3	1.1358		-
FD 2-4	1.5563	16.6	0.1261		-
no mesh ref.					
DG N=2, Q-free	1.2232	10	16.16		1.85 E+3
DG N=2, with Q	1.2232	10	11.82		3.79 E+3
DG N=4, Q-free	1.3327	10	22.03		1.17 E+3
DG N=4, with Q	0.9281	6.6	11.86		1.35 E+3
mesh ref.					
DG N=2, Q-free	1.1631	10	13.52		1.77 E+3
DG N=2, with Q	1.2032	10	9.66		3.08 E+3
DG N=4, Q-free	1.2271	10.6	24.22		1.17 E+3
DG N=4, with Q	0.8686	8	11.56		1.39 E+3

Table 3.3 : Results for linear-in-depth velocity test case.

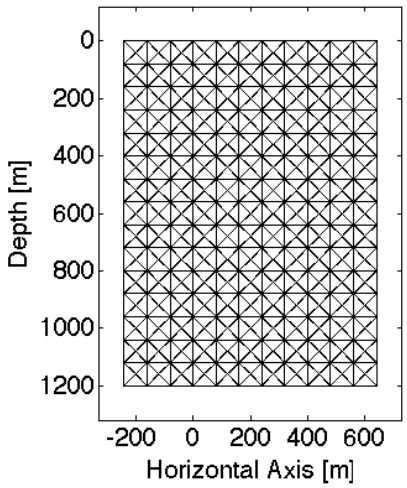


(a) Velocity model and experimental setup. (b) Seismic traces of highly discretized numerical solution.

Figure 3.5 : Experimental setup and traces for linear-in-depth velocity test case.

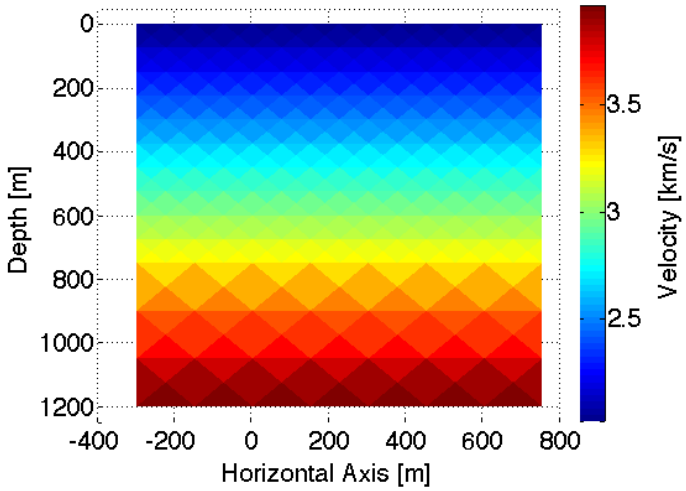


(a) Piecewise constant approximation.

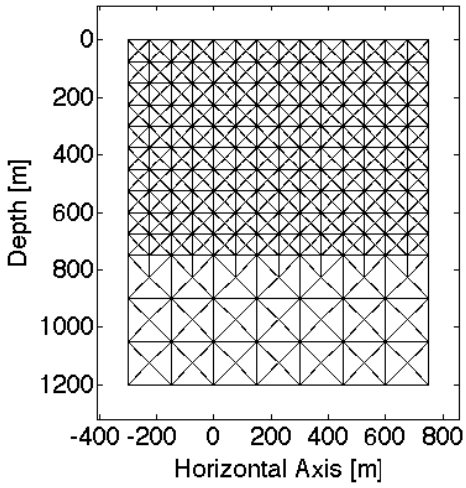


(b) Corresponding structured mesh.

Figure 3.6 : Piecewise approximation of velocity model and corresponding mesh, with no mesh refinement.

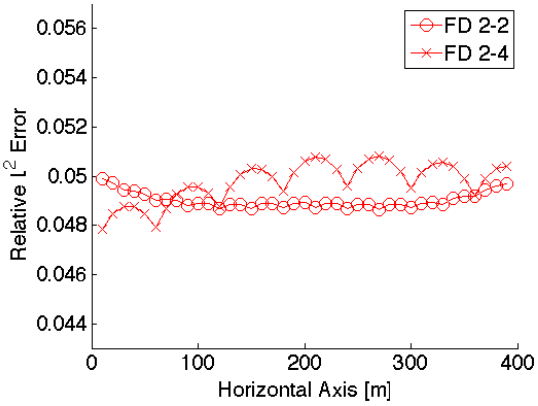


(a) Piecewise constant approximation.

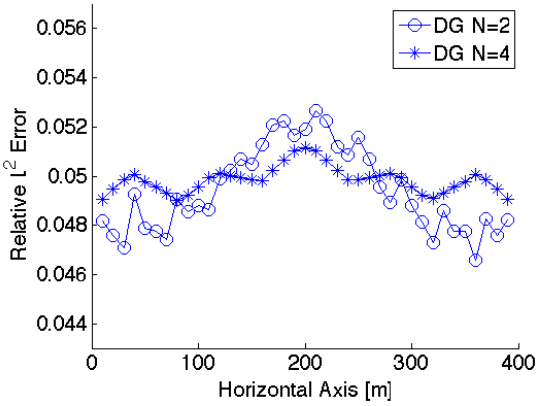


(b) Corresponding structured mesh, with refinement.

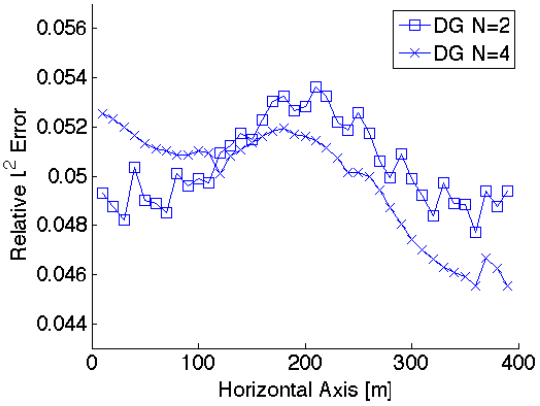
Figure 3.7 : Piecewise approximation of linear-in-depth velocity model and corresponding mesh, with mesh refinement.



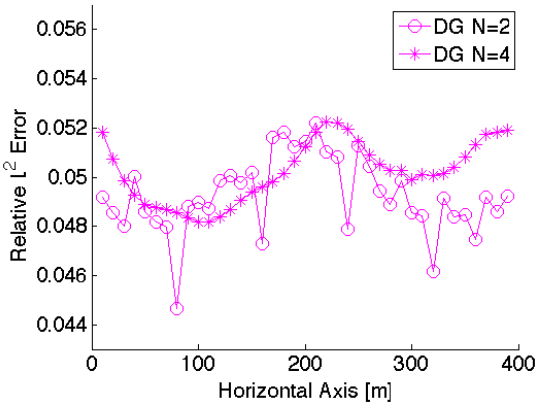
(a) FD errors



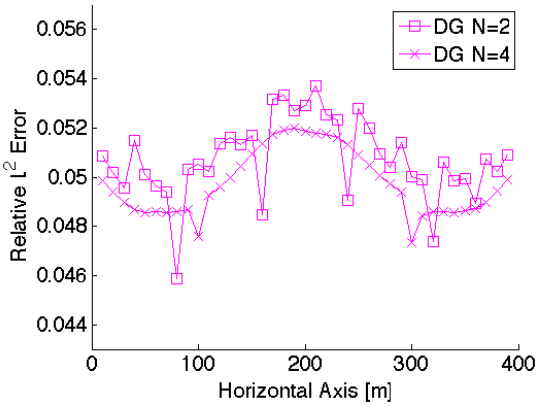
(b) quadrature-free DG



(c) quadrature based DG



(d) quadrature-free DG, with mesh refinement



(e) quadrature based DG, with mesh refinement

Figure 3.8 : Relative errors for linear-in-depth velocity model.

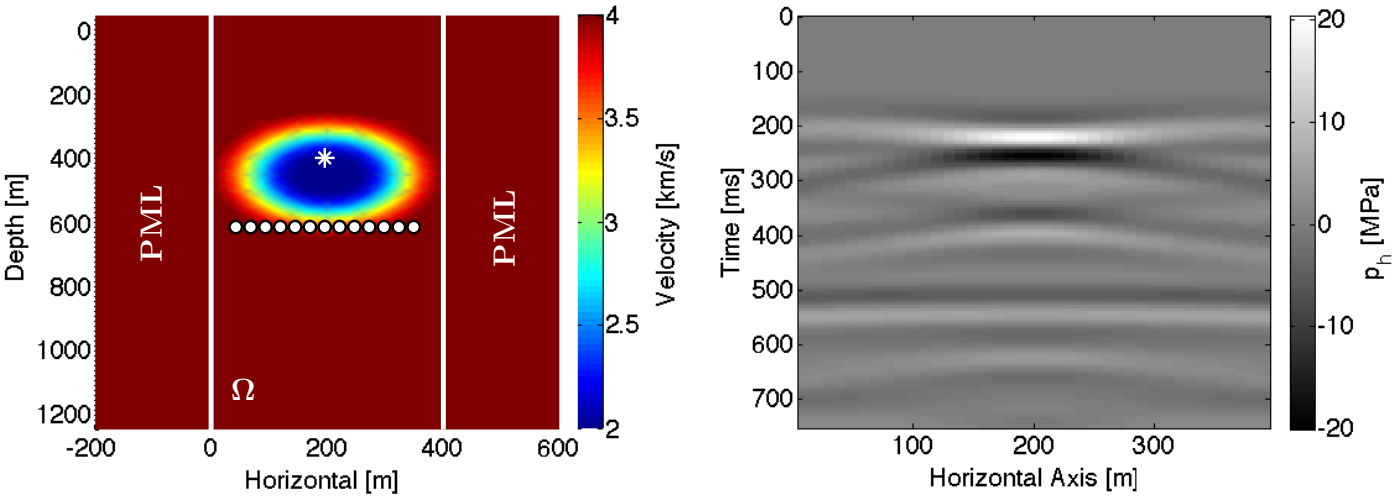
3.6 Negative-Lens Test Case

The final test case studied here is that of the negative-lens velocity model with constant density. Fig.(3.9a) depicts the receiver-source configuration along with the velocity model, with computed seismic traces shown in Fig.(3.9b). Moreover, Fig.(3.10) and Fig.(3.11) illustrate some examples of piecewise approximations of the medium along with their meshes, with and without mesh refinement. Errors are plotted in Fig. (3.12) while Table 3.4 summarize performance of the different methods. The following are some note worthy remarks of the results accumulated for this test case:

- FD GLFOP counts are over 1219 to 5.7 times smaller than that of DG methods;
- the best performing DG scheme has a GFLOP count of 189 times the GFLOP count of 2-4 FD;
- among DG methods, the $N = 4$ with quadrature and mesh refinement result in the faster runtimes;
- overall, quadrature based implementations resulted in the use of larger mesh size h and a decrease in runtime;
- mesh refinement have yielded significant reductions in runtime.

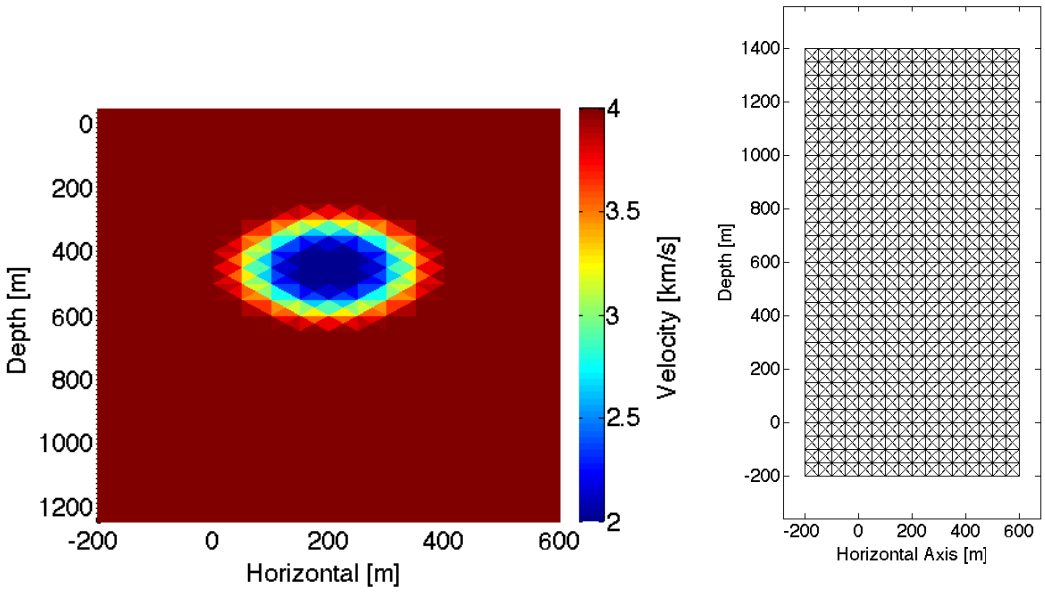
	dt [ms]	GPW	GFLOPs		Runtime [sec]
FD 2-2	0.8379	33.3	0.6296		-
FD 2-4	1.5645	13.3	0.0820		-
no mesh ref.					
DG N=2, Q-free	1.0027	10	19.72		2.16 E+3
DG N=2, with Q	0.9625	6.6	7.72		2.51 E+3
DG N=4, Q-free	0.6545	16	99.92		2.58 E+3
DG N=4, with Q	1.1994	10	19.99		1.54 E+3
mesh ref.					
DG N=2, Q-free	0.9826	10	7.44		1.33 E+3
DG N=2, with Q	0.8523	8	3.61		1.69 E+3
DG N=4, Q-free	0.6545	16	32.19		1.57 E+3
DG N=4, with Q	1.2048	10.6	8.29		1.09 E+3

Table 3.4 : Results for negative-lens test case.



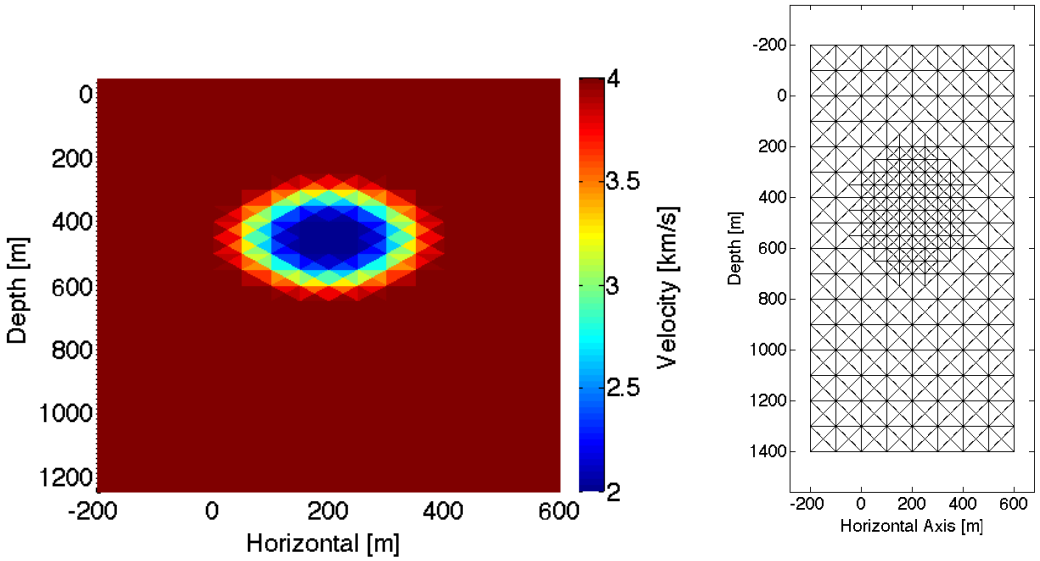
(a) Velocity model and experimental setup. (b) Seismic traces of highly discretized numerical solution.

Figure 3.9 : Experimental setup and traces for negative-lens test case.



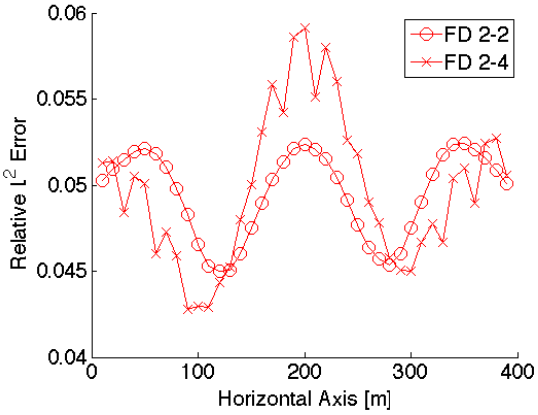
(a) Piecewise constant approximation. (b) Corresponding structured mesh.

Figure 3.10 : Piecewise approximation of negative-lens velocity model and corresponding mesh, with no mesh refinement.

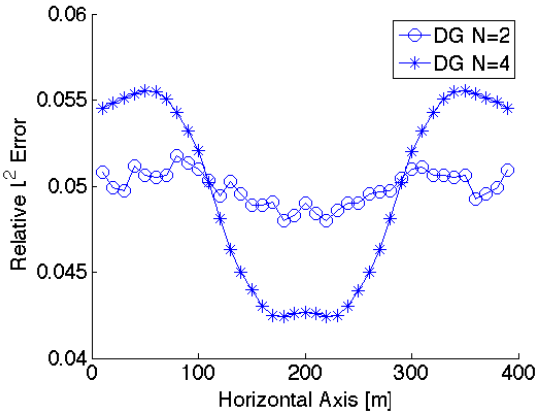


(a) Piecewise constant approximation. (b) Corresponding structured mesh, with refinement.

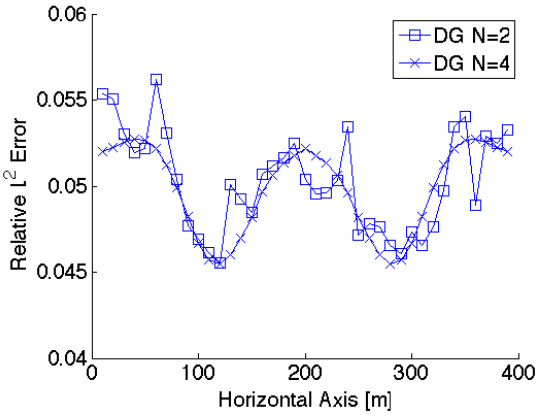
Figure 3.11 : Piecewise approximation of negative-lens velocity model and corresponding mesh, with mesh refinement.



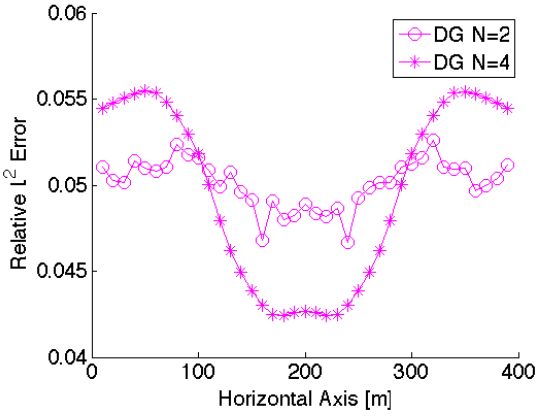
(a) FD errors



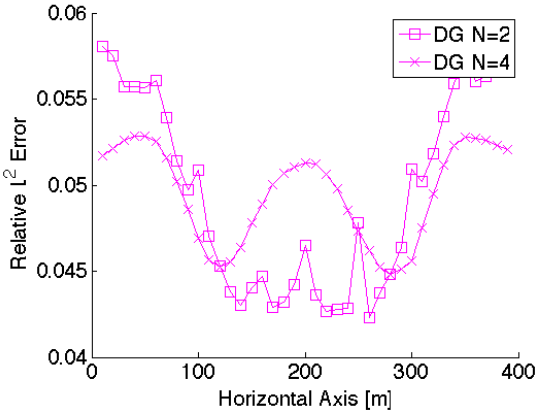
(b) quadrature-free DG



(c) quadrature based DG



(d) quadrature-free DG, with mesh refinement



(e) quadrature based DG, with mesh refinement

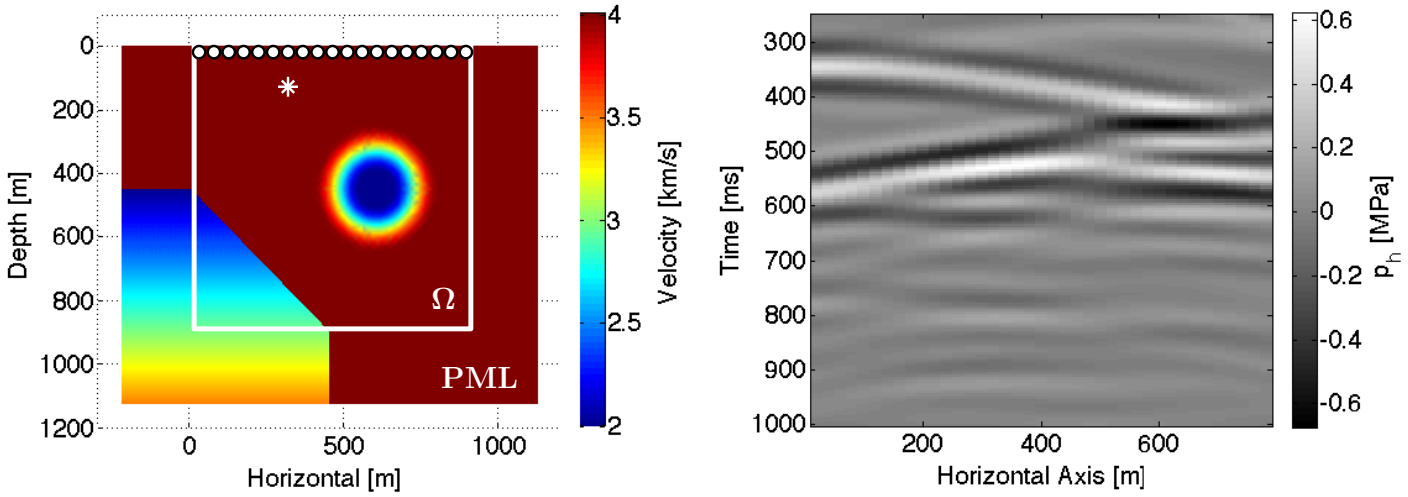
Figure 3.12 : Relative errors for negative-lens velocity model.

3.7 Mixed Test Case

The last test case considered is that of a mixed model with a dip discontinuity, incorporating both the linear-in-depth and lens models, see Fig.(3.13a). The pressure wave recorded resulting from the direct source, also known as the direct wave, is ignored when computing error for this particular test case by truncating the first 250 ms of the trace data. Essentially the direct wave data is discarded allowing for a comparison that is more sensitive to the reflections and refractions induced by the medium. The traces in Fig. (3.13b) illustrates these recorded reflections from the dip discontinuity, refracted waves from the negative lens, and multiple reflections between the free surface and layered dip discontinuity. Only one DG implementation is used for this test case, based on results from the previous test cases: polynomial order $N = 4$, quadrature based, with mesh refinement. Relative errors are plotted in Fig.(3.14), and overall performance is summarized in Table 3.5. Overall, DG GFLOP count is roughly 33 to 18 times greater than that of the FD schemes.

	dt [ms]	GPW	GFLOPs
FD 2-2	0.74246	33.3	1.4308
FD 2-4	1.1300	25	0.7793
DG	1.0375	14.2	25.68

Table 3.5 : Results for mixed test case.



(a) Velocity model and experimental setup.

(b) Seismic traces of highly discretized FD solution.

Figure 3.13 : Experimental setup and traces for mixed model test case.

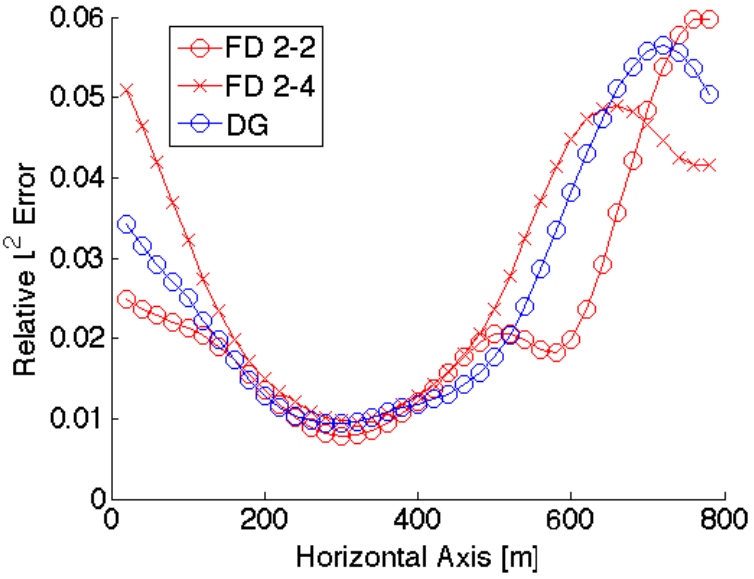


Figure 3.14 : Relative errors for mixed velocity model.

Chapter 4

Conclusion

The purpose of this thesis work is to provide a first attempt at a comparison between FD and DG methods in the context of seismic imaging, particularly in acoustics with smoothly varying medium. Recent works demonstrate the efficiency and advantages of DG over FD methods when considering discontinuous media, that is piecewise constant media (Wang (2009), Wang et al. (2010), Zhebel et al. (2013)). It has been DG's geometric flexibility, allowing for mesh alignment with discontinuous interfaces, and high parallelism that have propelled DG as a strong competitor in seismic imaging. On the other hand, gradual variations in physical parameters are indeed physical and observed in nature as well as relevant to the inversion process. My thesis incorporates established methodology for dealing with variable media within DG via efficient quadrature-free, and accurate quadrature based, implementations. Furthermore the use of mesh refinement is also considered in order to efficiently tackle the difficulties encountered in variable media, i.e., low velocity regions that require higher discretization.

Results reported here are of course preliminary, in that the DG code was implemented in Matlab while the FD code was written in C. Moreover, no attempt was made to parallelize the code at the moment, though this will be a natural extension

for future directions. All methods were subject to accuracy conditions Eq.(3.2) and Eq.(3.3), where the RMS and maximum relative error must be roughly less than 5% and 6% respectively, from which GFLOP counts were recorded to provide some measurement of computational cost and efficiency. Table 4.1 condenses the results from Chapter 3 into GFLOP ratios for the DG methods with best runtimes, relative to FD GFLOP counts for the test cases discussed. Results over the various test cases show that FD vastly outperforms DG. This was to be expected especially for the smooth media test cases. However not all hope is lost, as the mixed medium test case demonstrates a 33 fold difference in GFLOPS between DG and FD, the smallest difference observed amongst DG implementations with the best running times. This suggests that DG may have a chance to compete with FD under complex media with perhaps localized low velocity zones and where the geometry of the discontinuity is sufficiently relevant, under an HPC implementation framework.

Computational runtime is without a doubt one of the most important forms of measuring computational cost, especially for an industry that deals with enormous amounts of data such as the energy industry. The relation between the amount of work done by a computer, measured in FLOPs here, and computational runtime can be dubious since it depends highly on so many factors such as software, hardware, algorithmic design, and their codependence. Despite this, I argue that some simplified *back-of-the-envelope* calculations can provide some insight into the computational runtime and thus validate the use of FLOPs in this study. Zhou (2014) showed that

$20\% \sim 30\%$ of peak performance can be achieved for similar FD methods via vectorization and cache optimization on a Sandy Bridge Xeon E5-2660 processor. Using Zhou's results, I can provide a comparison of the runtimes for FD and DG methods based on their FLOP counts. Let T_{FD} and T_{DG} denote the computational running time of some FD and DG method, and let X and Y denote the total GFLOP count and peak performance rate respectively. Thus, assuming a 20% peak performance for the FD method, the running time T_{FD} can be estimated to be

$$T_{FD} = \frac{X}{0.2Y}.$$

Consider the case where the DG method at hand has a peak performance of $\epsilon \in (0, 1]$ and a total GFLOP count of $33X$, corresponding to the mixed test case. The ratio between DG and FD runtime is given as follows:

$$T_{DG}/T_{FD} = \frac{33X}{\epsilon Y} \Big/ \frac{X}{0.2Y} = \frac{6.6}{\epsilon} \leq 6.6.$$

In other words, the GFLOP discrepancy between FD and DG is too great for DG to overcome by shear computational efficiency, considering smooth media keeping in mind the choice of accuracy defined here. Despite FD's dominance in these test cases, I would argue that results shown here offer some insight into the application of DG in general when considering variable media, discontinuous or not.

	hom.	linear	lens	mixed
GFLOP ratio	150	175	101	33

Table 4.1 : Approximate GFLOP ratios between best of DG over FD, for each test case.

It was observed in the linear-in-depth velocity model that mesh refinement offered little alleviation in computational cost, while in the negative-lens test case refinement offered a clear reduction in run time as well as GFLOPs. The nature of the medium, that is a slowly varying medium versus another medium with a more localized anomaly, indeed played a role in how suitable mesh refinement can be. I personally attribute the lack of benefit in mesh refinement to my crude refinement algorithm, where elements are effectively refined from h to $h/2$. Perhaps an unstructured mesh with a more incremental refining algorithm could have produced better results.

More importantly, the use of quadrature for computing DG operators leads to overall less GFLOPs than the quadrature-free implementation. Chapter 2 hinted at the low memory storage properties of quadrature-free DG where reference operators, along with geometric factors, are required throughout calculations. This of course came at the price of having to partially recompute DG operators on the fly for each time step, hence the higher FLOP count as shown for all test cases. The efficiency of quadrature-free DG stems from the fact that it is faster to carrying out computations as oppose to making memory calls; it is faster to partially recompute operators than call them from memory. All in all, the efficiency of quadrature-free DG hinges on the

assumption that the physical medium is well approximated by some mesh-dependent piecewise constant analogue. Results show overall that smaller, or equal*, sized mesh elements are required in order maintain desired accuracy. Note that for the linear-in-depth velocity test case, quadrature-free implementations did in fact achieve reduced runtimes relative to their quadrature based counterparts. On the other hand, the negative-lens test case reveals the benefit of quadrature based implementation especially for higher order methods, due to the localized nature of the lens and its higher variability within mesh elements.

This work, along with complementary research on FD versus DG in discontinuous media (Wang, 2009; Wang et al., 2010), highlight the strengths and drawbacks of these two numerical methods in application to seismic modeling. Mainly, mesh and media discontinuity alignments, in conjunction with the discontinuous nature of DG solutions, characterize the efficiency and accuracy of DG solvers in cases where geometry of reflectors play a dominant role in the accuracy of the modeling problem, while FD methods thrive in smooth coefficient media. Seismic modeling problems model wave propagation in regions that contain both discontinuities and relatively smooth variations. For example, in marine-seismic modeling common discontinuities are the seafloor and potentially salt domes, or other geophysical phenomena that produce abrupt changes in media parameters. The "localized" nature of these discontinuities has prompted research in hybrid FD/Hermitian-FEM methods (Rylander and Bonde-

*As was the case for DG, N=2, where mesh sizes for both quadrature-free and quadrature based DG were the same, independent of mesh adaptivity for the linear-in-depth velocity test case.

son, 2000; Chen et al., 2014). The idea consists on applying FEM to a smaller region containing the interface of discontinuity while FD or Hermitian interpolation is used throughout the rest of the domain in order to capitalize on the strengths of each method. The crux of the matter is the issue of coalescing in time and space the FD regions of structured grids and FEM domains of unstructured meshes. Nonetheless, current implementations and applications in solving Maxwell's equations have shown promising results (Rylander and Bondeson, 2000; Chen et al., 2014).

Appendix A

A.1 Source Function

A uniform choice of source function is used throughout the numerical experiments described in Chaper 3. In particular, the righthand side terms $\boldsymbol{f}(\boldsymbol{x},t)$, $g(\boldsymbol{x},t)$ from Eq.(2.1) take the following form:

$$\boldsymbol{f}(\boldsymbol{x},t)=0, \quad \text{and} \quad g(\boldsymbol{x},t)=\chi(\boldsymbol{x})\Psi(t) \tag{A.1}$$

where $\Psi(t)=\Psi(t;f_{peak})$ is a Ricker wavelet with central frequency f_{peak},

$$\Psi(t;f_{peak})=\left(1-2(\pi t f_{peak})^2\right)e^{-(\pi t f_{peak})^2}, \tag{A.2}$$

and $\chi(\boldsymbol{x})=\chi(\boldsymbol{x};\boldsymbol{x}_0,\delta\boldsymbol{x})$ is a cosine bump function centered at $\boldsymbol{x}_0$ and width $\delta\boldsymbol{x}$. Fig.(A.1) gives a 1D example of $\chi(x;0,0.5)$. Note that the spatial support of $\chi(\boldsymbol{x})$ is compact, in fact $supp\ \chi(x;0,0.5)=[-0.25,0.25]$.

The following are some of uniform choices for the source term made throughout Chapter 2:

$$f_{peak}=10\ Hz, \quad \delta\boldsymbol{x}=[50\ m, 50\ m].$$

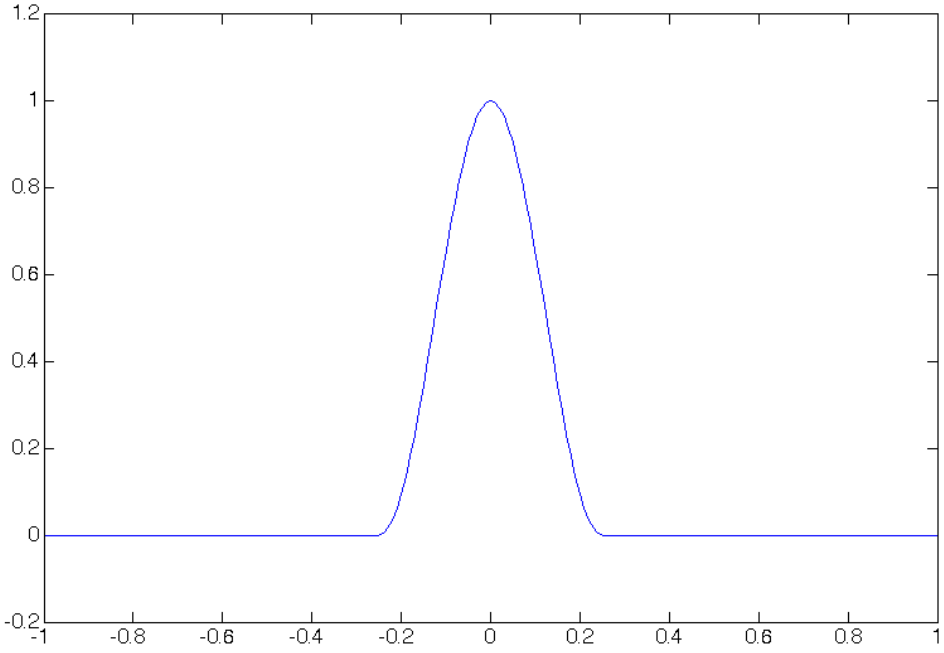


Figure A.1 : 1D example of cosine bump function; $\boldsymbol{x}_0 = 0, \delta\boldsymbol{x} = 0.5$

Appendix B

B.1 Auxiliary Result Tables

The following tables include supplementary information for the various test cases considered in Chapter 3. Maximum and root mean square (RMS) relative errors are computed. Discretization parameters for the "true solution", i.e., highly discretized FD 2-4 solution, are also included.

	dt [ms]	h [m]	max[%]	RMS [%]
FD 2-4	0.0589	0.5	-	-
FD 2-2	0.9546	6	5.07	4.99
FD 2-4	1.6122	12	5.09	4.98
DG N=2	1.2032	40	5.11	4.96
DG N=4	1.2058	80	5.09	4.98

Table B.1 : Results for homogeneous test case.

	dt [ms]	h [m]	max[%]	RMS [%]
FD 2-4	0.0442	0.5	-	-
FD 2-2	0.7425	6	5.98	2.71
FD 2-4	1.1300	8	5.10	3.20
DG	1.0375	112.5 : 56.25	5.65	3.08

Table B.2 : Results for mixed test case.

	dt [ms]	h [m]	max[%]	RMS [%]
FD 2-4	0.0442	0.5	-	-
FD 2-2	0.4679	6	4.99	4.90
FD 2-4	1.5563	12	5.08	4.98
no mesh ref.				
DG N=2, Q-free	1.2232	40	5.27	4.94
DG N=2, with Q	1.2232	40	5.36	5.05
DG N=4, Q-free	1.3327	80	5.12	4.99
DG N=4, with Q	0.9281	120	5.25	4.98
mesh ref.				
DG N=2, Q-free	1.1631	80 : 40	5.22	4.93
DG N=2, with Q	1.2032	80 : 40	5.37	5.08
DG N=4, Q-free	1.2271	150 : 75	5.23	5.02
DG N=4, with Q	0.8686	200 : 100	5.20	4.98

Table B.3 : Results for linear-in-depth velocity test case.

	dt [ms]	h [m]	max[%]	RMS [%]
FD 2-4	0.0442	0.5	-	-
FD 2-2	0.8379	6	5.24	4.96
FD 2-4	1.5645	15	5.91	5.00
no mesh ref.				
DG N=2, Q-free	1.0027	40	5.18	4.99
DG N=2, with Q	0.9625	60	5.62	5.04
DG N=4, Q-free	0.6545	50	5.56	5.01
DG N=4, with Q	1.1994	80	5.28	4.99
mesh ref.				
DG N=2, Q-free	0.9826	80 : 40	5.26	5.01
DG N=2, with Q	0.8523	100 : 50	5.81	4.99
DG N=4, Q-free	0.6545	100 : 50	5.55	5.00
DG N=4, with Q	1.2048	150 : 75	5.29	4.96

Table B.4 : Results for negative-lens test case.

Bibliography

Abarbanel, S. and Gottlieb, D. (1998). On the construction and analysis of absorbing layers in cem. *Applied Numerical Mathematics*, 27(4):331–340.

Ainsworth, M. (2004). Dispersive and dissipative behaviour of high order discontinuous galerkin finite element methods. *Journal of Computational Physics*, 198(1):106–130.

Atkins, H. L. and Shu, C.-W. (1998). Quadrature-free implementation of discontinuous galerkin method for hyperbolic equations. *AIAA journal*, 36(5):775–782.

Berenger, J.-P. (1994). A perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics*, 114(2):185–200.

Burger, H. R., Sheehan, A. F., and Jones, C. H. (2006). *Introduction to applied geophysics: Exploring the shallow subsurface*. WW Norton.

Carpenter, M. H. and Kennedy, C. A. (1994). Fourth-order 2n-storage runge-kutta schemes. *Nasa tm*, 109112.

Chen, X. R., Appelö, D., and Hagstrom, T. (2014). A hybrid hermite-discontinuous galerkin method for hyperbolic systems with application to maxwell's equations. *J. Comput. Phys.*, 257:501–520.

Cockburn, B. (2003). Discontinuous galerkin methods. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 83(11):731–754.

Collis, S. S., Ober, C. C., van Bloemen Waanders, B. G., et al. (2010). Unstructured discontinuous galerkin for seismic inversion. In *SEG Technical Program Expanded Abstracts 2010*, pages 2767–2772. Society of Exploration Geophysicists.

De Basabe, J. D. and Sen, M. K. (2007). Grid dispersion and stability criteria of some common finite-element methods for acoustic and elastic wave equations. *Geophysics*, 72(6):T81–T95.

De Basabe, J. D. and Sen, M. K. (2010). Stability of the high-order finite elements for acoustic or elastic wave propagation with high-order time stepping. *Geophysical Journal International*, 181(1):577–590.

De Basabe, J. D., Sen, M. K., and Wheeler, M. F. (2008). The interior penalty discontinuous galerkin method for elastic wave propagation: grid dispersion. *Geophysical Journal International*, 175(1):83–93.

FishSAFE (2013). Seismic surveys. `http://fishsafe.eu/en/offshore-structures/seismic-surveys.aspx`.

Gauthier, O., Virieux, J., and Tarantola, A. (1986). Two-dimensional nonlinear inversion of seismic waveforms; numerical results. *Geophysics*, 51(7):1387–1403.

Grote, M. J., Schneebeli, A., and Schötzau, D. (2006). Discontinuous galerkin finite element method for the wave equation. *SIAM Journal on Numerical Analysis*, 44(6):2408–2431.

Gurtin, M. E. (1981). *An introduction to continuum mechanics*, volume 158. Academic Pr.

Hesthaven, J. S. and Warburton, T. (2002). Nodal high-order methods on unstructured grids: I. time-domain solution of maxwell's equations. *Journal of Computational Physics*, 181(1):186–221.

Hesthaven, J. S. and Warburton, T. (2007). *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*, volume 54. Springer.

Hu, F. Q., Hussaini, M., and Rasetarinera, P. (1999). An analysis of the discontinuous galerkin method for wave propagation problems. *Journal of Computational Physics*, 151(2):921–946.

Lesaint, P. and Raviart, P.-A. (1974). *On a finite element method for solving the neutron transport equation*. Univ. Paris VI, Labo. Analyse Numérique.

Levander, A. (1988). Fourth-order finite-difference pw seismograms. *Geophysics*, 53(11):1425.

LeVeque, R. J. (2002). *Finite volume methods for hyperbolic problems*, volume 31. Cambridge university press.

Madariaga, R. (1976). Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America*, 66(3):639–666.

Moczo, P., Robertsson, J., and Eisner, L. (2007). The finite-difference time-domain method for modeling of seismic wave propagation. *Advances in Geophysics*, 48:421–516.

Monk, P. and Richter, G. (2005). A discontinuous galerkin method for linear symmetric hyperbolic systems in inhomogeneous media. *Journal of Scientific Computing*, 22(1):443–477.

Ober, C. C., Smith, T. M., Collis, S. S., Overfelt, J. R., and Schwaiger, H. (2010). Elastic wave propagation in variable media using a discontinuous galerkin method. Technical report, Sandia National Laboratories.

Petersson, N. A. and Sjögreen, B. (2010). Stable grid refinement and singular source discretization for seismic wave simulations. *Commun. Comput. Phys.*, 8(5):1074–1110.

Riviere, B. and Wheeler, M. F. (2003). Discontinuous finite element methods for acoustic and elastic wave problems. *Contemporary Mathematics*, 329:271–282.

Rylander, T. and Bondeson, A. (2000). Stable fem-fdtd hybrid method for maxwell's equations. *Computer Physics Communications*, 125(1):75–82.

Sheriff, R. E. and Geldart, L. P. (1995). *Exploration seismology / R.E. Sheriff, L.P. Geldart.* Cambridge ; New York : Cambridge University Press, 1995.

Simonaho, S.-P., Lhivaara, T., and Huttunen, T. (2012). Modeling of acoustic wave propagation in time-domain using the discontinuous galerkin method - a comparison with measurements. *Applied Acoustics*, 73(2):173 – 183.

Symes, W., Terentyev, I. S., Vdovina, T., et al. (2009). Getting it right without knowing the answer: Quality control in a large seismic modeling project. 28(1):2602–2606.

Virieux, J. (1984). Wave propagation in heterogeneous media: velocity-stress finite-difference method. *Geophysics*, 49(11):1933–1957.

Virieux, J. (1986). P-SV wave propagation in heterogeneous media; velocity-stress finite-difference method. *Geophysics*, 51(4):889–901.

Wang, X. (2009). Discontinuous Galerkin Time Domain Methods for Acoustics and Comparison with Finite Difference Time Domain Methods. Master's thesis, Rice University, Houston, Texas.

Wang, X., Symes, W. W., and Warburton, T. (2010). Comparison of discontinuous galerkin and finite difference methods for time domain acoustics. In *SEG Technical Program Expanded Abstracts*, volume 29, pages 3060–3065.

Warburton, T. and Hagstrom, T. (2008). Taming the cfl number for discontinuous galerkin methods on structured meshes. *SIAM Journal on Numerical Analysis*, 46(6):3151–3180.

Zhebel, E., Minisini, S., Kononov, A., and Mulder, W. (2013). Performance and scalability of finite-difference and finite-element wave-propagation modeling on intel's xeon phi. In *SEG Technical Program Expanded Abstracts*, pages 3386–3390.

Zhou, M. (2014). Wave Equation Based Stencil Optimizations on Multi-core CPU. Master's thesis, Rice University, Houston, Texas.