Applications of Reverse-time Migration

by

Robert Marlin Ferner

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science
in
Geophysics

Department of Physics
University of Alberta

Abstract

In a reflection seismic experiment, a controlled source function injects energy into the Earth. This causes the subsurface to undergo elastic deformation referred to as seismic wave propagation. Contrasts in the elastic properties of the propagation medium create interfaces which scatter the seismic wave-field. Scattered source energy is recorded at the Earth's surface as seismic data, these recordings indirectly contain information of subsurface scattering locations. Seismic imaging or migration is an inverse scattering problem which aims to produce structural images of the subsurface from seismic data.

Formulations of finite-difference solutions to the elastic wave-equation are well documented in geophysical literature. These finite-difference solutions model seismic wave propagation within a medium on a discrete grid. A major application of this is the seismic imaging technique known as reverse-time migration. This thesis will outline practical applications of elastic staggered-grid finite-difference modelling and the reverse-time migration algorithm.

Acknowledgements

I would like to thank my supervisor, Professor Mauricio Sacchi for his patience and support during my master's research at the University of Alberta. His insights into how scientific research should be conducted, and ability to articulate complex problems in a simple and understandable manner are skills that I wish to one day emulate. In addition, I would like to thank all the other professors whom I had the pleasure of working with at the University of Alberta, in particular Jeff Gu, Doug Schmitt, Martyn Unsworth, Murray Gingras and Claire Currie. The passion you all possess for teaching inspires me to not only expand my wealth of knowledge but share it with others.

I would like to acknowledge the sponsors of the Signal Analysis and Imaging Group within the University of Alberta for contributing to the funding of this research. In addition, the Madagascar open-source software package freely available from http://www.reproducibility.org, and Compute Canada for the use of computational resources.

I would like to also thank my colleagues and friends at the University of Alberta for the many enjoyable times and conversations shared together. I have learned many valuable lessons from relationships with people during my master's program.

To my family, thank you for the continued love and support you have provided throughout my time at the University of Alberta. Your are the most influential people in my life and play a large role in my personal achievements and successes. And to Amy, thank you for always being there for me and believing in me.

Contents

Appendices

List of Tables

List of Figures

CHAPTER 1

Introduction

1.1 Reflection seismology

Reflection seismology is a technique used in exploration geophysics. The method utilizes an active source such as an air gun (marine) or a Vibroseis truck or explosives (land) to propagate energy into the Earth. This energy results in the elastic deformation of the subsurface referred to as seismic wave propagation. The seismic wave-field is described as a temporal and spacial varying function governed by the source function and the elastic properties of the propagation medium.

In heterogeneous media, interfaces exist between materials with different elastic properties. These interfaces cause part of an incident seismic wave to scatter or reflect off the interface, and part to transmit/refract through the interface. In a reflection seismic experiment, source energy that is reflected from subsurface interfaces is recorded by an array of receivers at the Earth's surface known as geophones (land) or hydrophones (marine). These recordings are referred to as seismic data, and indirectly contain information about the locations of scattering events. With an approximate description of how seismic wave-fields propagate, seismic data can be used to produce structural images of subsurface interfaces via inverse scattering theory. This will be discussed further in sections 1.2 and 1.3.

A reflection experiment may record seismic data for several thousand individual sources emanating from different spatial locations. The separate recording of each source is known as a shot-profile. Sources are spatially distributed such that the wave-field illuminates an entire region of interest. Receiver arrays can be positioned such that the data is two-dimensional (one spatial dimension and time) or three-dimensional (two spatial dimensions and time). This thesis will contain examples using two-dimensional data, however it is possible to expand the outlined methods to three-dimensions.

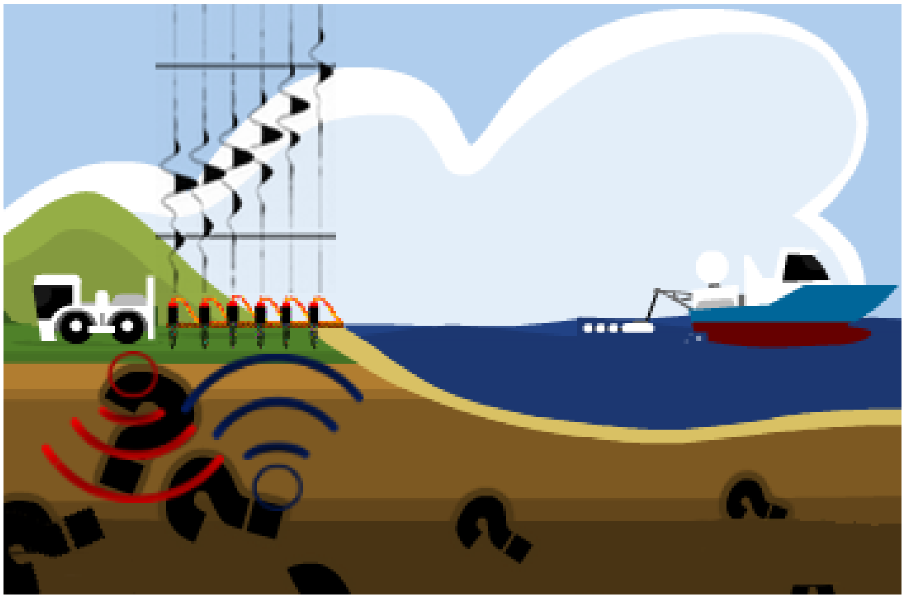


Figure 1.1: Cartoon illustrating a reflection experiment.

1.2 Staggered-grid elastic finite-difference modelling

Seismic wave-fields propagate through the subsurface via elastic deformation. The wave-field consists of two wave modes, compressional waves (P waves) and shear waves (S waves). Compressional waves propagate as volumetric deformation, and the shear waves as shape or shearing deformations. Our physical understanding of seismic wave propagation can therefore be described via generalized Hooke's law for elastic solids. Generalized Hooke's law is a constitutive equation which relates the second-order stress and strain tensors (τ_{ij},ϵ_{kl}) via the fourth-order stiffness tensor c_{ijkl}. Assuming the propagation medium has rotationally invariant material properties (isotropic elastic solid), the stiffness tensor can be composed of two parameters λ and μ. These are known as the Lamé parameters and along with density ρ, describe the elastic properties of the isotropic material. A first order hyperbolic system of partial differential equations can be derived from the basic equations of motion and generalized Hooke's Law to describe seismic wave propagation within isotopic elastic materials. This is known as the isotropic elastodynamic system of equations, an accurate approximation of seismic wave propagation capable of describing both compressional and shearing wave modes.

For simple scenarios such as homogenous media, analytic solutions to the elastodynamic system can be derived to model elastic wave propagation (Sherwood, 1958). However, for complex heterogeneous media determining closed-form solutions is not possible. In such cases, finite-difference methods can be used to generate explicit numerical solutions by discretizing the derivatives of the differential equations and updating wave-field components for a specified time interval. Conventional finite-difference methods such as those of Boore (1972) and Kelly et al. (1976) define the components of the elastic wave-field and model parameters at every position on a discrete mesh. However, for realistic exploration scale problems these algorithms become memory and computationally expensive. The staggered-grid finite-difference method of Virieux (1984, 1986) spatially and temporally offsets the nodes at which the components of the wave-field are defined, resulting in an improvement in memory and computational efficiency.

Finite-difference modelling allows the extrapolation of elastic wave-fields through time within a discrete model. This makes possible the ability to numerically model synthetic seismic data containing both compressional waves and shear waves. Proper implementation of boundary conditions and source functions along with stable model parametrization, ensures numerical solutions are in good agreement with analytically calculated solutions.

1.3 Reverse-time migration

Seismic data is composed of incident source energy scattered from subsurface interfaces back to receivers at the surface. Therefore, the seismic data indirectly contains information about contrasts in the elastic properties of the propagation medium. Assuming weak scattering, seismic data can be formulated in terms of the Born approximation to the Lippmann-Schwinger equation (Beylkin and Burridge, 1990). This formulation of seismic data contains a Green's function describing a forward modelling operator for a background elastic model, and a variable corresponding to perturbations within the background elastic model. Seismic imaging or migration is the method of producing images of subsurface interfaces from recorded seismic data. Under the weak scattering assumption, the migrated image (model perturbations) can be described as the adjoint of the forward modelling operator applied to the seismic data (Schuster et al., 2002). Physically this can be understood as repositioning events in the seismic data to there scattering locations in the subsurface. Three main categories of migration algorithms exist, each differing in the method of forward/adjoint modelling operators:

1. Ray tracing based methods such as Kirchhoff (Wiggins, 1984) and Gaussian beam migration (Hill, 1990).
2. One-way wave-equation extrapolation methods such as Gazdag, PSPI (Gazdag, 1978; Gazdag and Sguazzero, 1984) and split-step (Stoffa et al., 1990) commonly referred to

as wave-field extrapolation methods (WEM).

3. Finite-difference two-way wave-equation extrapolation methods known as reverse-time migration (RTM) (Baysal et al., 1983; Whitmore, 1983; McMechan, 1983).

Shot-profile migration produces a separate partial image for each individual shot recording. Each partial image only resolves the subsurface region illuminated by a lone source. Summing all of the partial images produces the final migrated image, resolving the region illuminated by all sources. The work in this thesis will focus on shot-profile reverse-time migration, using staggered-grid finite-difference methods for wave-field extrapolation. Two-way wave-equation extrapolation allows imaging of steeply dipping structures, accommodates complex variations in the elastic model, and models all types of waveforms including turning waves (Yao, 2013). This makes reverse-time migration the most accurate contemporary migration method.

The reverse-time migration algorithm can be broken up into two stages, a wave-field extrapolation and an imaging condition. The wave-field extrapolation maps positions of source and receiver energy within a smooth background velocity model. The cross-correlation imaging condition of Claerbout (1971) then generates images of contrasts in the elastic properties of the subsurface. This is done under the assumption that extrapolated source and receiver energy exist spatially and temporally in the same position where incident source energy scattering occurred. Assuming only compressional wave modes exist, the deformation becomes purely volumetric allowing the seismic wave-field to be described as a scalar field. This results in a scalar imaging algorithm which produces a single image for only compressional wave scattering. However,when both compressional and shear wave modes exist a vector imaging condition must be implemented, producing two images (Yan and Sava, 2008). One images corresponds compressional wave scattering (PP scattering) and one to converted wave scattering (PS scattering).

In a conventional migration two wave-field extrapolations are performed for each shot-profile, one for the source wave-field and one for the receiver wave-field. These wave-field extrapolations make up the bulk of the algorithms cost as they are done via explicit finite-differences. This makes the cost of the migration proportional to the number of migrated shot-profiles. In efforts to increase efficiency, multiple shot-profiles can be blended together to form super-shots prior to wave-field extrapolation. This process known as shot encoding reduces the computational cost of the migration by the number of combined shots. However, the final image is degraded by crosstalk interferences; artifacts generated from the cross-correlation of unrelated wave-field components within blended wave-field extrapolations (Romero et al., 2000). Thus, a trade off exists for blended migrations between computational cost and image quality governed by the number of shot-profiles combined to form a super-shot. Research in improving the image quality of blended migrations has become of increasing interest. Many different encoding schemes have been investigated to minimize the aforementioned crosstalk artifacts (Godwin et al., 2011). Though it remains unclear if blended migration is capable of outperforming migration of a selectively decimated shot-record.

1.4 Overview of thesis work

Chapter 2: We will derive the elastodynamic system of equations via generalized Hooke's law for elastic solids. This first order hyperbolic system of partial differential equations describes two-dimensional isotropic elastic deformation. The chapter will outline the staggered-grid finite-difference algorithm of Levander (1988) which numerically solves the elastodynamic system. The algorithm uses fourth-order space second-order time $O(\Delta t^2, h^4)$ central finite-difference operators to approximate the equations derivatives. Implementation of the algorithms finite-difference stencils, source functions, and boundary conditions will be described in detail. In addition, the stability and grid dispersion criterion of the algorithm will be investigated. It will be shown that numerical solutions produced by the finite-difference algorithm are in agreement with analytic solutions to the elastodynamic system.

Chapter 3: Under the weak scattering assumption we will formulate recorded seismic data via the Born approximation to the Lippmann-Schwinger equation. This formulation describes data using forward modelling operators for a background elastic model (Green's functions) and a term corresponding to perturbations within the background model. From this formulation of the seismic data we will arrive at the cross-correlation imaging condition of Claerbout (1971) as the adjoint modelling operator (adjoint Green's function) applied to the seismic data. This imaging condition along with elastic staggered-grid finite-difference modelling will compose a reverse-time migration algorithm. Implementation of the algorithm to a synthetic data set containing only compressional wave modes will outline the two step procedure of reverse-time migration. A more complex data set with both compressional and shear waves will outline the algorithms ability to image using both wave modes by separating the seismic wave-field into scalar and vector potentials (Yan and Sava, 2008).

Chapter 4: This chapter will review the concepts of phase encoding, blended migration and shot-decimated migration. In addition, investigate the crosstalk artifacts produced from blended migration. The chapter will also give an overview of the maximum likelihood estimator of location (M-estimator) technique, and introduce several different M-estimators (Huber, 1964). We will show how M-estimators can attenuate crosstalk artifacts in blended migration. A comparison of image quality between blended migration and shot-decimated migration is given and the viability of the two methods discussed.

1.5 Contribution of thesis work

Chapter 2: Conducted a detailed review and classic derivation of the two-dimensional isotropic elastodynamic system of equations. Composed sfea2dfd, the $O(\Delta t^2, h^4)$ two-dimensional elastic staggered-grid finite-difference algorithm described in detail. The algorithm was written in standard C using the Madagascar development platform, freely available from http://www.reproducibility.org. The Open-MP (Open Multi-Processing) API was used to implement shared memory parallelization of sfea2dfd resulting in a significant speed-up of the algorithm (refer to Appendix B.1).

Chapter 3: Modelled synthetic seismic data sets (acoustic/elastic) using a faulted elastic model. The synthetic data was then used to produce structural images via reverse-time migration from a smooth background model. The sfea2dfd algorithm was used to perform the wave-field extrapolations of the reverse-time migrations. Several different imaging conditions were composed using standard C (sfxcorr_acoustic, sfxcorr_acoustic_sourcenorm, sfxcorr_elastic) and implemented to produce the migration results. Python scripting within the Madagascar environment was used to carry out the synthetic data generation and reverse-time migration work flows.

Chapter 4: Modelled a conventional synthetic acoustic data set and a blended synthetic acoustic data set for 4, 8, and 16 shots per super shot using the Marmousi model. Migrated the conventional and blended data along with selectively decimated subsets of the conventional data set. Quantitatively analyzed the ability of different M-estimators to attenuate crosstalk artifacts within the blended images via relative L2 norm differences with the conventional migration. In addition, analyzed the relative L2 norm differences produced by selectively decimating the shot record prior to migration. This work was carried out on the Compute Canada symmetric multiprocessor system Hungabee, Python scripting within the Madagascar environment was used to carry out the synthetic data generation and reverse-time migration work flows.

CHAPTER 2

Elastic staggered-grid finite-difference

2.1 Introduction

Partial differential equations describe our physical understanding of many natural dynamic processes. These include fluid flow, heat transfer, electromagnetic field behaviour, elastic deformation and many others. However, in complex scenarios determining analytic solutions to these equations is often not possible. In such cases, finite-difference methods can be used to generate explicit numerical solutions by discretizing the derivatives of the differential equations. One application of this is modelling elastic deformation within complex heterogeneous media. This is of specific interest in seismology, the investigation of elastic wave propagation within the Earth's subsurface.

The finite-difference algorithms of Boore (1972) and Kelly et al. (1976) model numerical solutions to the two-dimensional isotropic elastodynamic equations. These algorithms operate on a conventional grid, where the components of the system are defined at every node of a discrete mesh. Consequentially, for large mesh sizes these algorithms become computationally and memory intensive, making realistic scale problems expensive to solve. The methods of Virieux (1984, 1986), similar to that of Yee et al. (1966), spatially and temporally offsets the nodes which components of the wave-field are defined. This staggered-grid implementation of the algorithm results in a reduction in memory and computational cost.

This chapter will derive the elastodynamic system of equations describing isotropic elastic deformation in two dimensions. In addition, will outline a staggered-grid finite-difference algorithm that produces realistic numerical solutions for this system. This will include implementation the of algorithms differencing operators, source excitation, boundary conditions, and a comparison of numerical solutions to analytic predictions for a simple single layer model.

2.2 The elastodynamic system of equations

2.2.1 Hooke's Law

The investigation of wave propagation phenomenon is largely comprised of elastic theory. In general elastic theory is considered a component of continuum mechanics; a branch of classical mechanics where motion and deformation of bodies is studied neglecting the discrete nature of matter (Pujol, 2003). Under the assumption of continuum, parameters influencing wave propagation through a medium can be described as averages over many molecules of matter. We will investigate Hooke's Law, the linear relationship between force and displacement of a spring, or stress and strain for an elastic medium as our starting point in describing elastic wave propagation. Hooke's Law for a rigid scalar system is shown by 2.1.

$$f_r = -kl \tag{2.1}$$

For a scalar spring under extension from a suspended mass, assuming equilibrium a force must exist opposing the gravitational force. This is known as the restoring force f_r and is proportional to the springs stiffness k and displacement from a relaxed state l. The stiffness is a positive scalar value that describes the linear relationship between restoring force and displacement of the mass.

If the system is displaced from equilibrium, the net force on the mass is no longer equal zero. This results in motion of the suspended mass due to the restoring force. This system is described by the homogenous, second order, linear differential equation shown in 2.2, with u being the displacement from equilibrium.

$$\frac{\partial u^2}{\partial t^2} + \omega^2 u = 0 \qquad \text{where,} \qquad \omega = \sqrt{\frac{k}{m}} \tag{2.2}$$

From inspection it can be determined the general solution is exponential and takes the closed-form of 2.3 where B and C are arbitrary co-efficient.

$$u(t) = Be^{i\omega t} + Ce^{-i\omega t} \tag{2.3}$$

2.3 can be rewritten via Euler's formula $e^{\pm i\omega t} = \cos(\omega t) + i\sin(\omega t)$ as a simple harmonic oscillator in the form of a phase shifted cosine 2.4. Here ω is the angular frequency of the oscillation, A the maximum amplitude, and δ the phase rotation or starting position of the system. It can be observed that the angular frequency or speed of the oscillations is governed by the material properties of the system k and m.

$$u(t) = A\cos{(\omega t - \delta)} \qquad \text{where,} \qquad A = \sqrt{(B+C)^2 + i(B-C)^2} \tag{2.4}$$

The example above assumes the mass and spring behave as rigid bodies. Mechanics of rigid bodies consider the ideal case where relative positions of masses within the system are fixed. However, when dealing with continuous elastic bodies relative positions of masses within the system are dynamic and alter if deformation occurs (Taylor, 2005). This concept is essential in describing the elastic deformation of continuous bodies.

2.2.2 Stress and strain

For a continuous three-dimensional system two prevalent forces exist, volume forces and surface forces. Volume forces $\mathbf{f_v}$ are forces proportional to the volume of a body which they act upon, where as surface forces $\mathbf{f_s}$ are forces proportional to the area of a surface which they act upon. Assuming gravity is the only volume force acting on a system, the volume forces will be ignored as exploration scale problems are small enough to assume fluctuations in gravity negligible. Surface forces oriented normal to a surface can be expressed as $\mathbf{f_s} = \tau\hat{\mathbf{n}}A$ where τ is the magnitude, $\hat{\mathbf{n}}$ the outward normal vector to the surface and A the surface area. Negative magnitudes denote a force applied opposite to the normal representing a compressive force, while a positive magnitude denotes a force inline with the normal representing a tensile force. Since surface forces are proportional to the area of the surface they act upon it becomes natural to express them as the ratio of force per area or stress. This is shown in equation 2.5 as the stress vector $\mathbf{T}(\hat{\mathbf{n}})$, expressed as the finite-limit of the force per area ratio. It can be seen that the stress vector is dependent on the orientation of the surface normal, and can be represented with units of pressure.

$$\mathbf{T}(\hat{\mathbf{n}}) = \lim_{A\to 0}\frac{\mathbf{f_s}}{A} = \frac{\partial \mathbf{f_s}}{\partial A} \tag{2.5}$$

For a three-dimensional system, stress becomes a second order tensor $\mathbf{T}_i(\hat{\mathbf{n}}_j) = \tau_{ij}\hat{n}_j$ with nine scalar components τ_{ij}. Each stress vector $\mathbf{T}_i(\hat{\mathbf{n}}_j)$ contains three components describing the magnitudes of forces with normals to the dimensional planes $(j = 1, 2, 3)$ and three mutually perpendicular stress vectors are required to describe stress normal to any surface in the system $(i = 1, 2, 3)$.

$$\mathbf{T}_i(\hat{\mathbf{n}}_j) = \begin{bmatrix} \mathbf{T}_i(\hat{\mathbf{n}}_j) \\ \mathbf{T}_i(\hat{\mathbf{n}}_j) \\ \mathbf{T}_i(\hat{\mathbf{n}}_j) \end{bmatrix} = \begin{bmatrix} \tau_{11}\hat{n}_1 \; \tau_{12}\hat{n}_2 \; \tau_{13}\hat{n}_3 \\ \tau_{21}\hat{n}_1 \; \tau_{22}\hat{n}_2 \; \tau_{23}\hat{n}_3 \\ \tau_{31}\hat{n}_1 \; \tau_{32}\hat{n}_2 \; \tau_{33}\hat{n}_3 \end{bmatrix} = \begin{bmatrix} \tau_{11} \; \tau_{12} \; \tau_{13} \\ \tau_{21} \; \tau_{22} \; \tau_{23} \\ \tau_{31} \; \tau_{32} \; \tau_{33} \end{bmatrix} \begin{bmatrix} \hat{n}_1 \\ \hat{n}_2 \\ \hat{n}_3 \end{bmatrix} \tag{2.6}$$

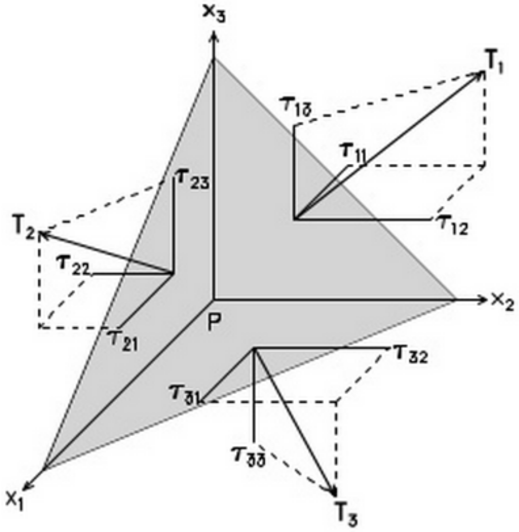


Figure 2.1: The stress tetrahedron illustrating the components of the stress tensor τ_{ij} and the stress vectors $\mathbf{T}_i(\hat{\mathbf{n}}_j) = \mathbf{T}_i$. (From Pujol (2003))

The diagonal components $\tau_{ij} = \delta_{ij}\tau_{ij}$ are normal stresses acting along the principle planes of stress $(i = j)$. The off diagonal components $\tau_{ij} = (1 - \delta_{ij})\tau_{ij}$ are the shear stresses and act perpendicular to principle planes of stress. For an elastic body at equilibrium the sum of moments must equal zero, resulting in symmetry about the stress tensor $(\tau_{ij} = \tau_{ji})$.

The application of stress to a body results in deformation, the alteration of the body's dimensions. A deformation is said to be elastic if the body returns to it's original dimensions after an applied stress vanishes. Deformation expressed as a ratio over the initial dimensions of a body is referred to as strain. We will use rate of displacement within a body as the measure of strain. This is shown below for one dimension in equation 2.8 where u is a displacement from its original position x, and ϵ the strain.

$$\epsilon = \frac{\partial u}{\partial x} \tag{2.7}$$

For a three dimensional system strain is expressed as a second order derivative tensor with nine scalar components ϵ_{kl}. However, the strain tensor must be symmetric $(\epsilon_{kl} = \epsilon_{lk})$ such that rigid rotations are not considered in the definition of strain (Taylor, 2005). Thus, the strain tensor must only be the symmetric portion of the derivative tensor. This results in strain being defined as the average of two spacial derivatives shown below in equation 2.8. Nine components are required to describe the position of a point in the deformed body $k = 1, 2, 3$ relative to its position in the original body $l = 1, 2, 3$.

$$\epsilon_{kl} = \frac{1}{2}\left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k}\right) \tag{2.8}$$

The elements of ϵ_{kl} relate to changes in lengths and angles, and the trace of ϵ_{kl} relates to changes in volume of a deformed body (Pujol, 2003). The diagonal elements of the stain tensor $\epsilon_{kl} = \delta_{kl}\epsilon_{kl}$ correspond to normal strains describing length changes along the coordinate axis. The off diagonal elements $\epsilon_{kl} = (1 - \delta_{kl})\epsilon_{kl}$ correspond to shear strains, describing changes in angles.

2.2.3 The Lamé parameters

For small three dimensional elastic deformations the relationship between stress and strain tensors is assumed to be linear. This linear relationship is referred to as generalized Hooke's law for elastic solids shown by equation 2.9. The variable c_{ijkl} is a fourth order tensor with 81 components corresponding to the stiffness or elastic properties of a medium.

$$\tau_{ij} = c_{ijkl}\epsilon_{kl} \tag{2.9}$$

However, because the stress and strain tensors are symmetric for linearly elastic solids, c_{ijkl} is also symmetric ($c_{ijkl} = c_{lkji}$) and reduces to 36 linearly independent components. For elastic solids with rotationally invariant material properties (isotropic elastic solids), c_{ijkl} further reduces to 6 linearly independent components shown by equation 2.10. The two variables λ and μ are referred to as the Lamé parameters.

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{2.10}$$

This allows the generalized Hooke's law for isotropic elastic solids to be written in terms of the lamé parameters shown by equation 2.11. Due to the elastic medium having no preferred direction, the deformation will occur along the applied stresses direction. The repeated indicies denotes diagonal elements of the strain tensor.

$$\tau_{ij} = \lambda\delta_{ij}\epsilon_{kk} + 2\mu\epsilon_{ij} \tag{2.11}$$

It can be observed that the parameter μ solely governs the relationship between the shear stresses and strains, and is therefore referred to as the shear modulus. The shear modulus describes the rigidity of an elastic body, or the ability of the body to resist shape/shearing deformation. The parameter λ governs the relationship between uni-axial and lateral strain to uni-axial stress. Therefore, λ describes the incompressibility of an elastic body, the ability of the body to resist volumetric deformation. However, it can be seen that both of the lamé parameters are required to describe the axial stress-strain relationship. The physical explanation of this given by Goodway (2001) is that in order for a volumetric deformation to occur so must a shape deformation, thereby invoking resistance as a function of both λ and μ. For a heterogeneous elastic body the lamé parameters vary spatially, making them functions of spatial position within the body, $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ where $\mathbf{x} = (x_1, x_2, x_3)$.

2.2.4 Isotropic elastodynamic equations

Although we discuss stress and strain in three-dimensions for simplicity we will express the isotropic elastodynamic system of equations in two dimensions using Cartesian co-ordinates $\mathbf{x} = (x, z)$, where x is the horizontal direction and z the vertical direction. The Lagrangian form for the general equation of motion in two dimensions can be expressed as equations 2.12, where ρ is density of the elastic medium, (u_x, u_z) horizontal and vertical displacements, (τ_{xx}, τ_{zz}) the horizontal and vertical normal stresses and τ_{xz} the shear stress (Aki and Richards, 2002).

$$\rho\frac{\partial^2 u_x}{\partial t^2} = \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xz}}{\partial z} \qquad \rho\frac{\partial^2 u_z}{\partial t^2} = \frac{\partial \tau_{zz}}{\partial z} + \frac{\partial \tau_{xz}}{\partial x} \tag{2.12}$$

Equation 2.11 for two dimensions can be expanded to describe the stresses shown below as equations 2.13. Strain is represented as rate of displacement, ρ the density, and λ and μ the lamé parameters of the medium. The equations of 2.12 and 2.13 form a second-order hyperbolic system of partial differential equations. Five equations are needed to solve for the five unknowns $u_x, u_z, \tau_{xx}, \tau_{zz}, \tau_{xz}$.

$$\begin{gathered}\tau_{xx} = (\lambda + 2\mu)\frac{\partial u_x}{\partial x} + \lambda\frac{\partial u_z}{\partial z} \qquad \tau_{zz} = (\lambda + 2\mu)\frac{\partial u_z}{\partial z} + \lambda\frac{\partial u_x}{\partial x} \\ \tau_{xz} = \mu\left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}\right)\end{gathered} \tag{2.13}$$

The second-order hyperbolic system of displacement (u_x, u_z) and stress $(\tau_{xx}, \tau_{zz}, \tau_{xz})$ can be transformed into the first-order hyperbolic system of particle velocity (v_x, v_z) and stress by differentiating the stress equations with respect to time, (Madariaga, 1976). This is shown below as the system of equations 2.14.

$$
\begin{gathered}
\frac{\partial v_x}{\partial t} = \frac{1}{\rho}\left(\frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xz}}{\partial z}\right) \qquad \frac{\partial v_z}{\partial t} = \frac{1}{\rho}\left(\frac{\partial \tau_{zz}}{\partial z} + \frac{\partial \tau_{xz}}{\partial x}\right) \\
\frac{\partial \tau_{xx}}{\partial t} = (\lambda + 2\mu)\frac{\partial v_x}{\partial x} + \lambda\frac{\partial v_z}{\partial z} \qquad \frac{\partial \tau_{zz}}{\partial t} = (\lambda + 2\mu)\frac{\partial v_z}{\partial z} + \lambda\frac{\partial v_x}{\partial x} \\
\frac{\partial \tau_{xz}}{\partial t} = \mu\left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right)
\end{gathered} \tag{2.14}
$$

The material properties determine the speed of P-SV wave propagation in the medium, similar to how material properties of the spring system governs the angular frequency of oscillations. Two types of wave modes exist in elastic wave propagation, compressional waves and shear waves. The compressional wave modes propagate at speed V_p as volumetric perturbations of the medium. The compressional wave speed depends on the lamé parameters that govern the axial stress-strain relationship $(\lambda + 2\mu)$ and the density of the medium ρ. The shear wave modes propagate at speed V_s as shape or shearing perturbations of the medium, with shear wave speed being a function of the shear modulus μ and the density of the medium ρ. Setting the shear modulus μ to zero reduces the system of equations to the acoustic case described in detail by Virieux (1984). This makes the modelling of solid/liquid interfaces straight forward.

$$
V_p = \sqrt{\frac{\lambda + 2\mu}{\rho}} \qquad V_s = \sqrt{\frac{\mu}{\rho}} \tag{2.15}
$$

In the example in section 2.1.1 we find the analytic solution to the ordinary differential equation describing a suspended mass displaced from equilibrium. However, we will solve this system of partial differential equations numerically as an boundary value problem, with the initial condition of equilibrium at time $t = 0$.

2.3 Explicit solution via staggered-grid finite-difference

2.3.1 Staggered-grid discretization of the elastodynamic equations

Explicit solutions to the elastodynamic equations can be numerically modelled on a discrete mesh. The elastic parameters of a continuous medium are quantized into cells or nodes, whereby each node is the smallest possible differentiable unit of parametrization. Virieux

(1986) outlines a numerical solution to the first-order hyperbolic elastic system of equations shown above in equation 2.14, using second-order accurate centered finite-difference operators in time and space $O(\Delta t^2, h^2)$ to discretize the derivatives (refer to Appendix A.1). The method implements a classic staggered-grid formulation similar to Yee et al. (1966), where computation of different wave-field components are staggered spatially and/or temporally by a half node index. This means that not all elastic parameters and components of the wave-field are defined at every node. Thus, the P-SV Madariaga-Virieux staggered-grid scheme offers an economic improvement for memory requirements and computational cost in comparison to the non staggered-grid formulations of Boore (1972) and Kelly et al. (1976).

One downside of explicit finite-difference schemes is the computational expense and large memory requirements needed for realistic exploration scale simulations. Dablain (1986) demonstrates the ability of higher-order finite difference approximations to reduce the memory requirements of wave equation finite-difference schemes, allowing for larger models and higher frequency solutions. Levander (1988) outlines a second-order accurate time, and fourth-order accurate space $O(\Delta t^2, h^4)$ formulation to the Madariaga-Virieux finite-difference scheme (refer to Appendix A.2). This scheme utilizes a nine-point Cartesian stencils to implement the finite-difference operators. The Levander (1988) $O(\Delta t^2, h^4)$ scheme provides an optimal trade-off between computational efficiency and numerical accuracy (Hustedt et al., 2004).

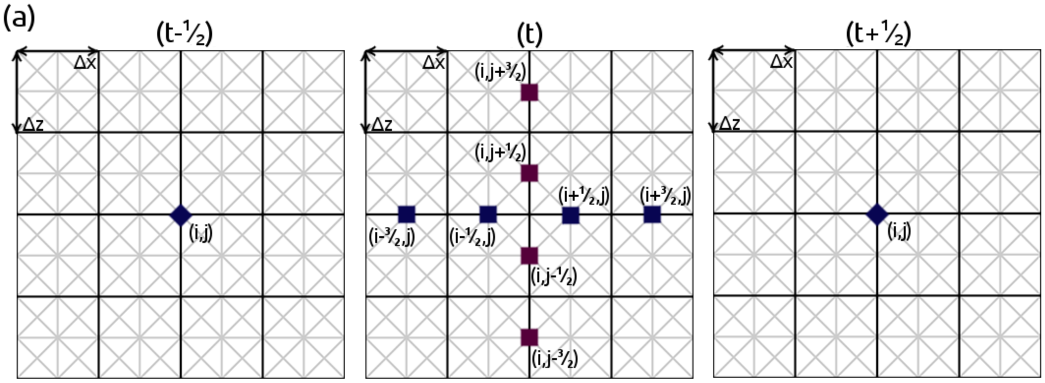
The nine-point Cartesian finite-difference stencils can be formulated by replacing the derivatives from elastodynamic equations from 2.14 with central finite-difference approximations to the first derivative. The time derivatives should be substituted by the second-order approximation $O(\Delta t^2)$ from Appendix A.1 and the spatial derivatives with the fourth-order approximation $O(h^4)$ from Appendix A.2. This procedure is outlined below in equation 2.16 for the horizontal particle velocity component v_x, using continuous notation and assuming $\Delta x = \Delta z = h$.

$$
\begin{aligned}
\frac{v_x^{(t+\frac{1}{2}\Delta t)}(x,z) - v_x^{(t-\frac{1}{2}\Delta t)}(x,z)}{\Delta t} &= \qquad (2.16) \\
\frac{1}{\rho}\left(\frac{-C_2\left(\tau_{xx}^{(t)}(x+\frac{3}{2}h,z) - \tau_{xx}^{(t)}(x-\frac{3}{2}h,z)\right) + C_1\left(\tau_{xx}^{(t)}(x+\frac{1}{2}h,z) - \tau_{xx}^{(t)}(x-\frac{1}{2}h,z)\right)}{h}\right) \\
+\frac{1}{\rho}\left(\frac{-C_2\left(\tau_{xz}^{(t)}(x,z+\frac{3}{2}h) - \tau_{xz}^{(t)}(x,z-\frac{3}{2}h)\right) + C_1\left(\tau_{xz}^{(t)}(x,z+\frac{1}{2}h) - \tau_{xz}^{(t)}(x,z-\frac{1}{2}h)\right)}{h}\right)
\end{aligned}
$$

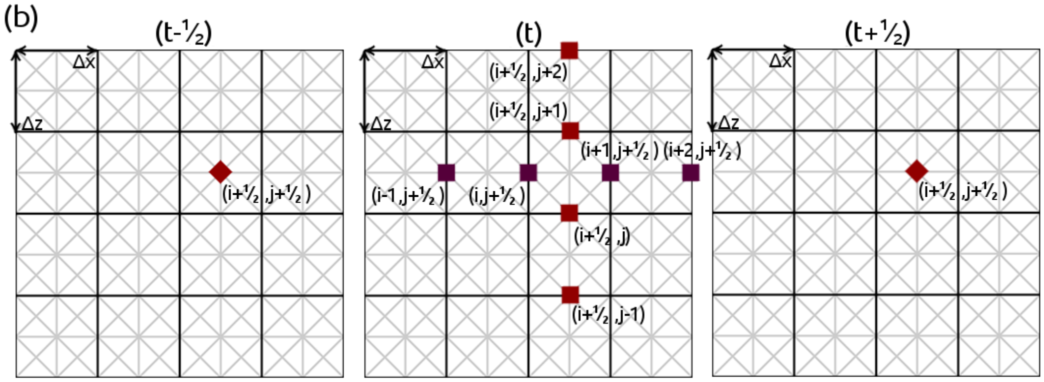
However, when mapping the stencils onto a discrete grid four stencils are required to compute the five independent wave-field components. This is due to wave-field components being defined at discrete positions, staggered spatially and/or temporally by a half node index. For the explicit numerical equations 2.17-2.21, i denotes the index along the horizontal axis, j the index along the vertical axis, and t the time index. $C_1 = \frac{9}{8}$ and $C_2 = \frac{1}{24}$ are the finite-difference co-efficients for the fourth-order approximation to the first derivative. The node size is defined by Δx and Δz, and the time step by Δt. The numerical material properties are delineated as $B = \frac{1}{\rho}$ and $(L, M, L2M) = (\lambda, \mu, \lambda + 2\mu)$, the numerical particle velocities $(\mathbf{U}, \mathbf{V}) = (v_x, v_z)$, and the numerical stresses $(\mathbf{X}, \mathbf{Z}, \mathbf{S}) = (\tau_{xx}, \tau_{zz}, \tau_{xz})$. It can be observed that numerical particle velocities are defined at half nodes in time $(t + \frac{1}{2}), (t - \frac{1}{2})$. The horizontal particle velocities $\mathbf{U}$ are defined spatially at full nodes along both axis (i, j), and the vertical particle velocites $\mathbf{V}$ are defined spatially at half nodes along both axis $(i + \frac{1}{2}, j + \frac{1}{2})$. Similarly, the numerical stresses are defined at full nodes in time $(t), (t + 1)$, with normal stresses $(\mathbf{X}, \mathbf{Z})$ defined spatially at half nodes along the horizontal axis and full nodes along the vertical axis $(i + \frac{1}{2}, j)$. Lastly, shear stresses $\mathbf{S}$ are defined spatially at half nodes along the vertical axis and full nodes along the horizontal axis $(i, j + \frac{1}{2})$. The arrangement of the staggered-grid intertwines the stencils such that the finite-difference operators only index nodes where relevant wave-field components are defined.

Figure 2.2 illustrates how the nine-point Cartesian stencils uses the numerical stresses at time (t), and previously computed particle velocities at time $(t - \frac{1}{2})$, to update particle velocities at time $(t + \frac{1}{2})$. Similarly, Figure 2.3 illustrates how numerical particle velocities at time $(t + \frac{1}{2})$ and previously computed stresses at time (t) update numerical stresses at time $(t + 1)$. Staggered-grid finite-difference is commonly referred as a leapfrog scheme as the method alternates between updating numerical particle velocities (Figure 2.2) and stresses (Figure 2.3). Iterating the update equations recursively through time for all nodes in a discrete model, numerically models the elastic wave propagation.

It can be observed that $B = \frac{1}{\rho}$ is defined at the same nodes as the particle velocity components $(\mathbf{U}, \mathbf{V}) = (v_x, v_z)$, $(L, M) = (\lambda, \mu)$ at the same nodes as the normal stresses $(\mathbf{X}, \mathbf{Z}) = (\tau_{xx}, \tau_{zz})$, and $M = \mu$ at the same nodes as the shear stress component $\mathbf{S} = \tau_{xz}$. However the input or reference model is only defined at a full node interval, therefore the material properties required at a half-node must be calculated. Mittet (2002) outlines an averaging method for interpolating the material properties to the required half-node interval. A modified version of this method is shown below as equations 2.22-2.25.

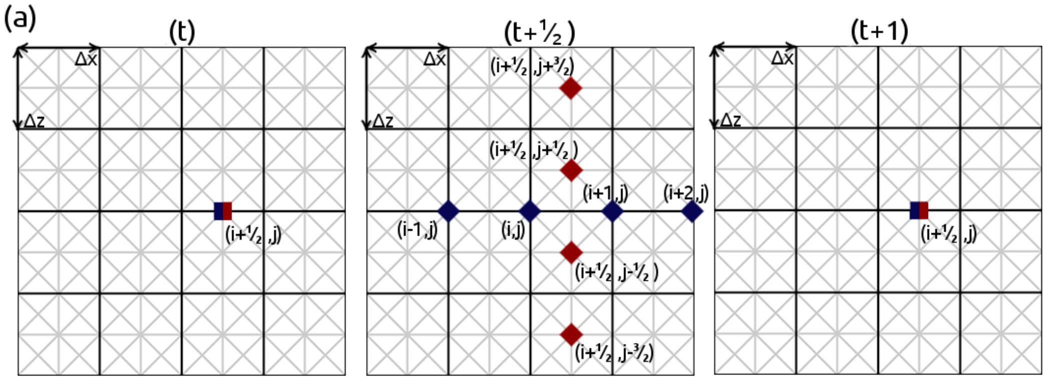


$$
\begin{aligned}
&\mathbf{U}^{t+\frac{1}{2}}_{(i,j)} = \mathbf{U}^{t-\frac{1}{2}}_{(i,j)} \\
&+B_{(i,j)}\frac{\Delta t}{\Delta x}\left(-C_2\left(\mathbf{X}^t_{(i+\frac{3}{2},j)} - \mathbf{X}^t_{(i-\frac{3}{2},j)}\right) + C_1\left(\mathbf{X}^t_{(i+\frac{1}{2},j)} - \mathbf{X}^t_{(i-\frac{1}{2},j)}\right)\right) \\
&+B_{(i,j)}\frac{\Delta t}{\Delta z}\left(-C_2\left(\mathbf{S}^t_{(i,j+\frac{3}{2})} - \mathbf{S}^t_{(i,j-\frac{3}{2})}\right) + C_1\left(\mathbf{S}^t_{(i,j+\frac{1}{2})} - \mathbf{S}^t_{(i,j-\frac{1}{2})}\right)\right)
\end{aligned}
\tag{2.17}
$$



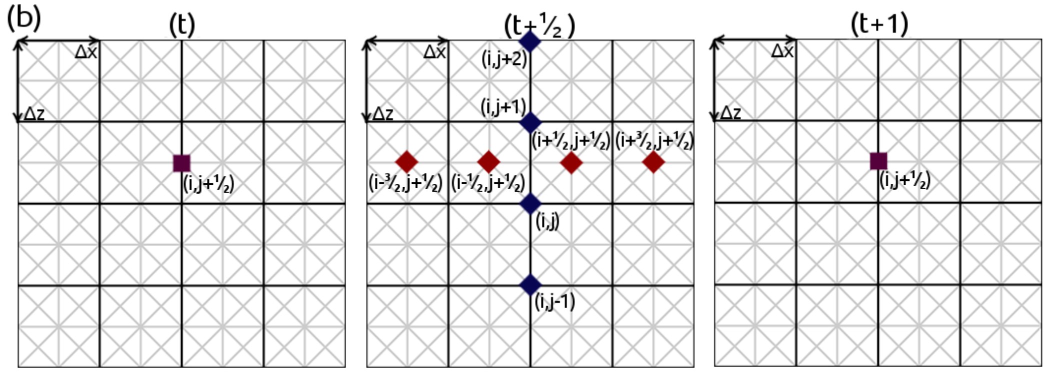
$$
\begin{aligned}
&\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})} = \mathbf{V}^{t-\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})} \\
&+B_{(i+\frac{1}{2},j+\frac{1}{2})}\frac{\Delta t}{\Delta x}\left(-C_2\left(\mathbf{S}^t_{(i+2,j+\frac{1}{2})} - \mathbf{S}^t_{(i-1,j+\frac{1}{2})}\right) + C_1\left(\mathbf{S}^t_{(i+1,j+\frac{1}{2})} - \mathbf{S}^t_{(i,j+\frac{1}{2})}\right)\right) \\
&+B_{(i+\frac{1}{2},j+\frac{1}{2})}\frac{\Delta t}{\Delta z}\left(-C_2\left(\mathbf{Z}^t_{(i+\frac{1}{2},j+2)} - \mathbf{Z}^t_{(i+\frac{1}{2},j-1)}\right) + C_1\left(\mathbf{Z}^t_{(i+\frac{1}{2},j+1} - \mathbf{Z}^t_{(i+\frac{1}{2},j)}\right)\right)
\end{aligned}
\tag{2.18}
$$

Figure 2.2: Nine-point Cartesian particle velocity stencils and corresponding discretized equations. Diamonds are the numerical particle velocities and squares the numerical stresses. Blue indicates the horizontal components $(\mathbf{U},\mathbf{X}) = (v_x, \tau_{xx})$, red the vertical components $(\mathbf{V},\mathbf{Z}) = (v_z, \tau_{zz})$, and purple the shear stress $\mathbf{S} = \tau_{xz}$. (a) Horizontal particle velocity stencil and corresponding equation. (b) Vertical particle velocity stencil and corresponding equation.



$$\begin{aligned}\mathbf{X}^{t+1}_{(i+\frac{1}{2},j)} &= \mathbf{X}^{t}_{(i+\frac{1}{2},j)}\\ &+L_{(i+\frac{1}{2},j)}\frac{\Delta t}{\Delta z}\left(-C_2\left(\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{3}{2})}-\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j-\frac{3}{2})}\right)+C_1\left(\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})}-\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j-\frac{1}{2})}\right)\right)\\ &+L2M_{(i+\frac{1}{2},j)}\frac{\Delta t}{\Delta x}\left(-C_2\left(\mathbf{U}^{t+\frac{1}{2}}_{(i+2,j)}-\mathbf{U}^{t+\frac{1}{2}}_{(i-1,j)}\right)+C_1\left(\mathbf{U}^{t+\frac{1}{2}}_{(i+1,j)}-\mathbf{U}^{t+\frac{1}{2}}_{(i,j)}\right)\right)\end{aligned}\tag{2.19}$$

$$\begin{aligned}\mathbf{Z}^{t+1}_{(i+\frac{1}{2},j)} &= \mathbf{Z}^{t}_{(i+\frac{1}{2},j)}\\ &+L2M_{(i+\frac{1}{2},j)}\frac{\Delta t}{\Delta z}\left(-C_2\left(\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{3}{2})}-\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j-\frac{3}{2})}\right)+C_1\left(\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})}-\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j-\frac{1}{2})}\right)\right)\\ &+L_{(i+\frac{1}{2},j)}\frac{\Delta t}{\Delta x}\left(-C_2\left(\mathbf{U}^{t+\frac{1}{2}}_{(i+2,j)}-\mathbf{U}^{t+\frac{1}{2}}_{(i-1,j)}\right)+C_1\left(\mathbf{U}^{t+\frac{1}{2}}_{(i+1,j)}-\mathbf{U}^{t+\frac{1}{2}}_{(i,j)}\right)\right)\end{aligned}\tag{2.20}$$



$$\begin{aligned}\mathbf{S}^{t+1}_{(i,j+\frac{1}{2})} &= \mathbf{S}^{t}_{(i,j+\frac{1}{2})}\\ &+M_{(i,j+\frac{1}{2})}\frac{\Delta t}{\Delta x}\left(-C_2\left(\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{3}{2},j+\frac{1}{2})}-\mathbf{V}^{t+\frac{1}{2}}_{(i-\frac{3}{2},j+\frac{1}{2})}\right)+C_1\left(\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})}-\mathbf{V}^{t+\frac{1}{2}}_{(i-\frac{1}{2},j+\frac{1}{2})}\right)\right)\\ &+M_{(i,j+\frac{1}{2})}\frac{\Delta t}{\Delta z}\left(-C_2\left(\mathbf{U}^{t+\frac{1}{2}}_{(i,j+2)}-\mathbf{U}^{t+\frac{1}{2}}_{(i,j-1)}\right)+C_1\left(\mathbf{U}^{t+\frac{1}{2}}_{(i,j+1)}-\mathbf{U}^{t+\frac{1}{2}}_{(i,j)}\right)\right)\end{aligned}\tag{2.21}$$

Figure 2.3: Nine-point Cartesian stress stencils and corresponding discretized equations. (a) Normal stress stencil and corresponding equations. (b) Shear stress stencil and corresponding equation. The half blue half red square is due to the normal stresses $(\mathbf{X}, \mathbf{Z}) = (\tau_{xx}, \tau_{zz})$ being defined at the same nodes.

$$B_{(i+\frac{1}{2},j+\frac{1}{2})} = \frac{B_{(i+1,j+1)} + B_{(i-1,j-1)} + B_{(i+1,j-1)} + B_{(i-1,j+1)}}{4} \tag{2.22}$$

$$M_{(i,j+\frac{1}{2})} = \begin{cases} \text{if } M_{(i,j)} \text{ and } M_{(i,j+1)} \neq 0 \cdots \frac{M_{(i,j)}+M_{(i,j+1)}}{2} \\ \text{if } M_{(i,j)} \text{ or } M_{(i,j+1)} = 0 \cdots 0 \end{cases} \tag{2.23}$$

$$M_{(i+\frac{1}{2},j)} = \begin{cases} \text{if } M_{(i,j)} \text{ and } M_{(i+1,j)} \neq 0 \cdots \frac{M_{(i,j)}+M_{(i+1,j)}}{2} \\ \text{if } M_{(i,j)} \text{ or } M_{(i+1,j)} = 0 \cdots 0 \end{cases} \tag{2.24}$$

$$L_{(i+\frac{1}{2},j)} = \frac{L_{(i,j)} + L_{(i+1,j)}}{2} \tag{2.25}$$

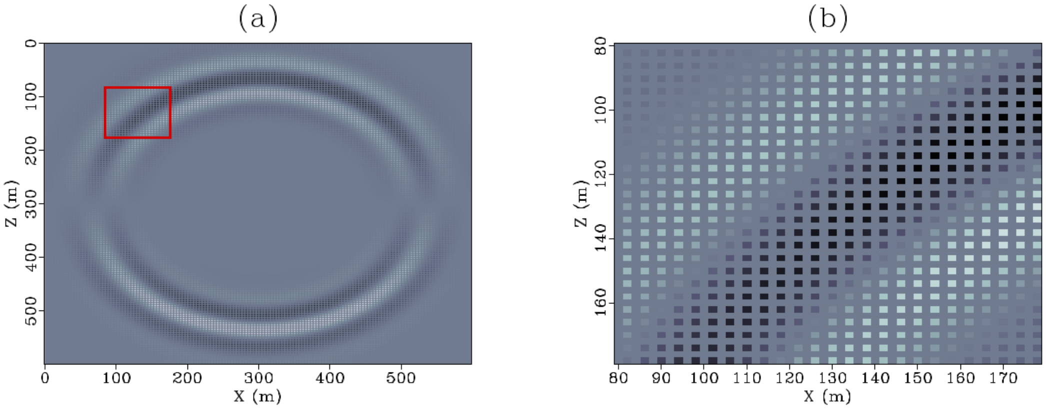


Figure 2.4: Coarse-grid representation of the wave-field. (a) Vertical particle velocity **V**. (b) Magnified view of region within red box, notice the checker board appearence.

Due to the wave-field components spatial staggering, the propagation has a checker board like appearance. This is known as the coarse-grid representation of the wave-field. The coarse grid representation has a node separation of Δx Δz and shown below in Figure 2.4. Interpolation of the wave-field components to a node separation of $\frac{\Delta x}{2}$ $\frac{\Delta z}{2}$, via nearest neighbour interpolation infills the node points where a wave-field component is not defined. This is known as the fine-grid representation of the wave-field shown below in Figure 2.5. The interpolation of the wave-field to the fine-grid is cosmetic, as interpolated nodes are not indexed by the update stencils.

The P-SV Madariaga-Virieux staggered-grid scheme offers a robust explicit numerical solution to the elastic two-way wave equation. The scheme is stable for all values of Poisson's

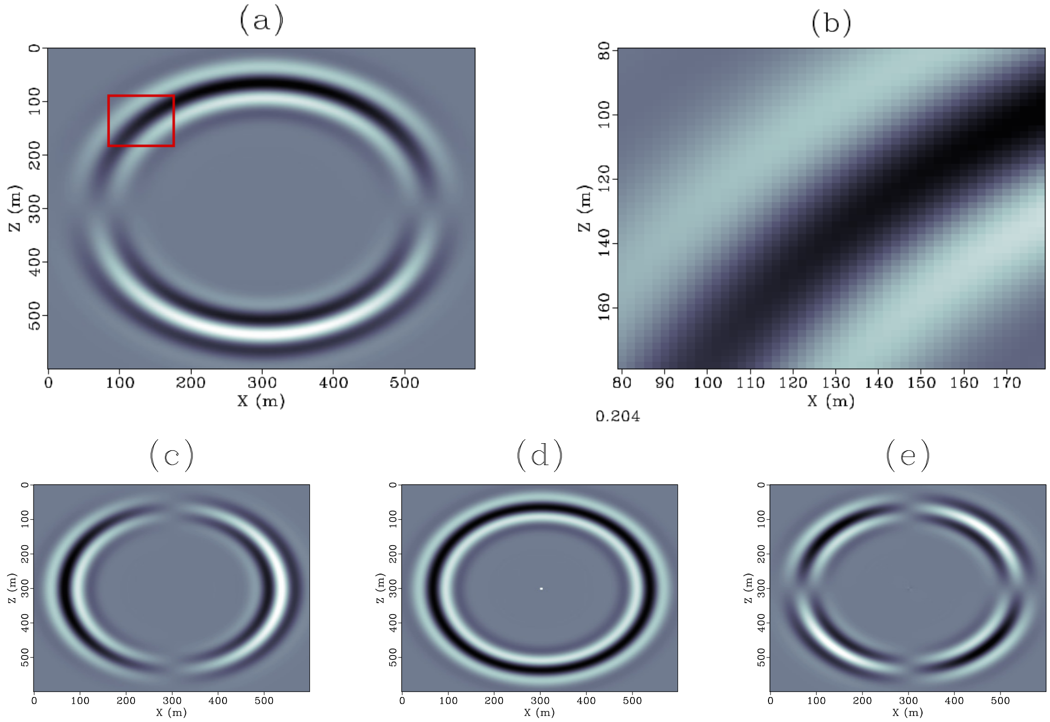


Figure 2.5: Wave-field components after nearest neighbour interpolation to fine-grid. (a) Vertical particle velocity $\mathbf{V}$. (b) Magnified view of region within red box. (c) Horizontal particle velocity $\mathbf{U}$. (d) Normal stresses shown as pressure $\frac{1}{2}(\mathbf{X}+\mathbf{Z})$. (e) Shear stress $\mathbf{S}$.

ratio. This makes possible modelling of solid liquid interfaces necessary in marine exploration problems. In addition, it is simple to place sources and receivers at desired locations within a discrete model. This makes it an ideal scheme for modelling complex acquisition geometries, such as those implemented using simultaneous sources or ocean bottom cables/nodes.

2.3.2 Source implementation

For a medium to be in a state of equilibrium at time $t = 0$, numerical stresses and particle velocities must equal zero everywhere on the grid. Therefore, numerical excitation of the discrete mesh is required to initiate wave propagation. A sampling function defines at which nodes excitation will occur, and a time evolving source function drives wave-propagation imbedded at the specified nodal points. The sampling function equals one at a defined source position and zero elsewhere. The temporal source function and its derivatives must be continuous, to ensure accuracy of the numerical solution. Gaussian functions and the first and second order derivatives of Gaussian functions are common source function, in addition to Ricker and Ormsby wavelets. The source function can be input into any component of the wave-field. Coutant et al. (1995) and Graves (1996) outline how different source mechanisms can be generated via implementation of the source as varying combinations of wave-field components. Due to the numerical normal stresses being defined at the the same nodal points, implementation of an impulsive/explosive source as normal stresses $(\mathbf{X}, \mathbf{Z}) = (\tau_{xx}, \tau_{zz})$ is a simple and effective source excitation method reducing the grid asymmetry (Coutant et al., 1995). Source implementation is referred to as a "hard" source, if the nodal points where the source is defined are not affected by the update equations. This causes artifacts as these nodal points act as wave-field scatterers. To eliminate "hard" source artifacts the source function can be set to vanish after a set period of time, after which the source defined nodal points are set by the update stencils. Assuming the propagated wave-field does not interact with the source defined nodal points prior to the source vanishing, no artifacts will be generated (Schneider et al., 1998). Source implementation is described below by equations 2.26 and 2.27, where κ is the sampling function, and $s(t)$ the source function and t_s the source duration.

$$\kappa_{(i+\frac{1}{2},j)}\mathbf{X}^{t}_{(i+\frac{1}{2},j)} = \begin{cases} s(t) \text{ if } t \leq t_s \\ 2.19 \text{ if } t \geq t_s \end{cases} \tag{2.26}$$

$$\kappa_{(i+\frac{1}{2},j)}\mathbf{Z}^{t}_{(i+\frac{1}{2},j)} = \begin{cases} s(t) \text{ if } t \leq t_s \\ 2.20 \text{ if } t \geq t_s \end{cases} \tag{2.27}$$

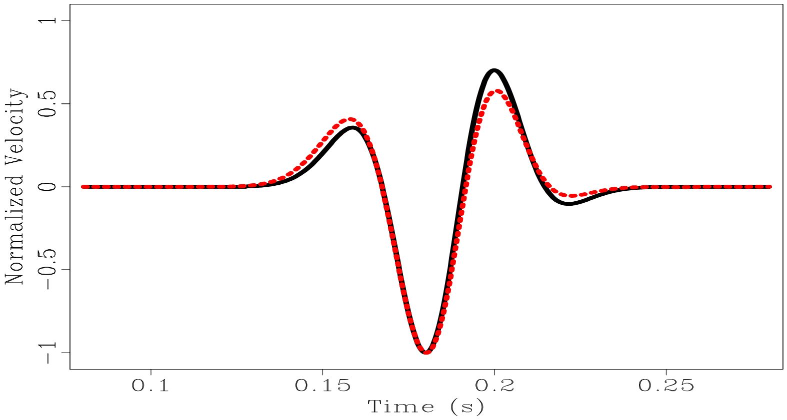


Figure 2.6: Analytic solution (black-continuous) and vertical component numerical solution (red-dash) for 75° phase rotated $17.5\,Hz$ central frequency Ricker wavelet propagated as pressure (τ_{xx}, τ_{zz}). Wave-field recorded $200\,m$ directly below source node in homogenous propagation medium $[V_p = 1500\,\frac{m}{s},\ \rho = 1000\,\frac{kg}{m^3}]$

Coutant et al. (1995) compares numerically propagated source functions to analytic solutions. Travel time, shape, and amplitude discrepancies can be evaluated to determine if the source is correctly represented in the propagation. Figure 2.6 illustrates the comparison of numerical and analytic solutions for propagation of a phase rotated Ricker wavelet in a homogenous acoustic medium. The numerical solution was produced via convolution of the finite-difference source with a Dirac delta function. The numerical and analytic solutions must be scaled before comparison as the source is implemented as normal stress $(\mathbf{X}, \mathbf{Z}) = (\tau_{xx}, \tau_{zz})$, but recorded as particle velocity $\mathbf{V} = v_z$. Agreement between the numerical and analytic solution confirms the source function is correctly represented in the numerical wave-field.

2.3.3 Free-surface boundary condition

Implementation of a free-surface boundary condition models the solid/air interface observed the earth's surface. Due to the harsh acoustic impedance contrast between elastic solids and the earths atmosphere, the solid/air interface can be approximated as a solid/vacuum interface. The free-surface boundary condition for two-dimensional P-SV wave propagation requires the vertical stresses τ_{zz} and shearing stresses τ_{xz} to be equal zero along the free-surface (Levander, 1988; Mittet, 2002). This is shown as equation 2.28 and 2.29 below for a free-surface at $z = 0$.

$$\tau_{zz}\big|_{z=0} = 0 \tag{2.28}$$

$$\tau_{xz}\big|_{z=0} = 0 \tag{2.29}$$

All wave-field components should equal zero above the free-surface boundary, as propagation is not possible within a vacuum. This can be achieved by setting all material properties above the free-surface equal to zero. For the $O(\Delta t^2, h^4)$ P-SV staggered-grid finite-difference scheme described in subsection 2.3.1, the numerical shear stresses $\mathbf{S}$ and vertical stresses $\mathbf{Z}$ are staggered along the z-axis by a half index. To satisfy 2.28 the numerical vertical stresses must be set to zero along the free-surface located at $j = 0$ ($\mathbf{Z}_{(i,0)} = 0$). It can be seen in Figure 2.3 that the numerical shear stresses are defined at a half z-axis index, and therefore will be undefined along the free-surface nodes at $j = 0$; however, are required to update the numerical horizontal particle velocities along the free-surface. To satisfy equation 2.29 when updating the numerical horizontal particle velocities along the free-surface, the numerical shear stresses must be mirrored above the free-surface $\mathbf{S}_{(i,-\frac{1}{2})} = -\mathbf{S}_{(i,\frac{1}{2})}$ and $\mathbf{S}_{(i,-\frac{3}{2})} = -\mathbf{S}_{(i,\frac{3}{2})}$ producing a fictitious layer. This ensures that the numerical shear stresses do not effect the update equations along the free surface, and are effectively zero. Similar false layer free-surface boundary conditions are mentioned by Levander (1988), Graves (1996), Jiang (2012) and Marcinkovich and Olsen (2003), and can be easily implemented by reducing equation 2.17 to equation 2.30 along the free-surface along with setting the numerical vertical stresses to zero.

$$\begin{aligned}\mathbf{U}^{t+\frac{1}{2}}_{(i,j)} &= \mathbf{U}^{t-\frac{1}{2}}_{(i,j)} \\ &+ B_{(i,j)}\frac{\Delta t}{\Delta x}\left(-C_2\left(\mathbf{X}^t_{(i+\frac{3}{2},j)} - \mathbf{X}^t_{(i-\frac{3}{2},j)}\right) + C_1\left(\mathbf{X}^t_{(i+\frac{1}{2},j)} - \mathbf{X}^t_{(i-\frac{1}{2},j)}\right)\right)\end{aligned} \tag{2.30}$$

2.3.4 Perfectly matched layer boundary condition

The propagation of the wave-field components into unbounded model edges will generate undesirable boundary reflections. Artificial boundary conditions are required to simulate the outgoing nature of the wave-field as it propagates into the model edges, emulating an infinite region of propagation. A perfectly matched layer (PML) is a lossy interface between two half spaces that does not reflect plane wave energy for all frequencies and angles of incidence. Chew and Liu (1996) demonstrate that a perfectly matched layer can be designed for the elastodynamic equations, and propose its use as a boundary condition in numerical elastic wave-field modelling. Collino and Tsogka (2001) introduce a perfectly matched layer boundary condition for the Madariaga-Virieux P-SV finite-difference scheme. The method splits each of the wave-field components into directional derivative auxiliary equations, parallel and perpendicular to the boundary. A dampening parameter is then applied to the auxiliary equation perpendicular to the boundary as it propagates into the layer. The method resembles that of a damped harmonic oscillator, in that a dampening parameter is applied to decay the wave-field as it propagates within the layer. The $O(\Delta t^2, h^4)$ explicit numerical equations for propagation within the perfectly matched layer are shown below in equations 2.31-2.45, where D^x and D^z are the horizontal and vertical dampening parameters. The superscript above the wave-field components denotes the directional derivative auxiliary. Outside of the perfectly matched layer the dampening parameters equal zero, reducing the system of equations to the one described by 2.17-2.21.

$$\mathbf{U}^{t+\frac{1}{2}}_{(i,j)} = (\mathbf{U}^x)^{t+\frac{1}{2}}_{(i,j)} + (\mathbf{U}^z)^{t+\frac{1}{2}}_{(i,j)} \tag{2.31}$$

$$\frac{(\mathbf{U}^x)^{t+\frac{1}{2}}_{(i,j)} - (\mathbf{U}^x)^{t-\frac{1}{2}}_{(i,j)}}{\Delta t} + (D^x)_{(i,j)} \frac{(\mathbf{U}^x)^{t+\frac{1}{2}}_{(i,j)} - (\mathbf{U}^x)^{t-\frac{1}{2}}_{(i,j)}}{2} =$$

$$B_{(i,j)} \frac{\left(-C_2 \left(\mathbf{X}^t_{(i+\frac{3}{2},j)} - \mathbf{X}^t_{(i-\frac{3}{2},j)}\right) + C_1 \left(\mathbf{X}^t_{(i+\frac{1}{2},j)} - \mathbf{X}^t_{(i-\frac{1}{2},j)}\right)\right)}{\Delta x} \tag{2.32}$$

$$\frac{(\mathbf{U}^z)^{t+\frac{1}{2}}_{(i,j)} - (\mathbf{U}^z)^{t-\frac{1}{2}}_{(i,j)}}{\Delta t} + (D^z)_{(i,j)} \frac{(\mathbf{U}^z)^{t+\frac{1}{2}}_{(i,j)} - (\mathbf{U}^z)^{t-\frac{1}{2}}_{(i,j)}}{2} =$$

$$B_{(i,j)} \frac{\left(-C_2 \left(\mathbf{S}^t_{(i,j+\frac{3}{2})} - \mathbf{S}^t_{(i,j-\frac{3}{2})}\right) + C_1 \left(\mathbf{S}^t_{(i,j+\frac{1}{2})} - \mathbf{S}^t_{(i,j-\frac{1}{2})}\right)\right)}{\Delta z} \tag{2.33}$$

$$\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})} = (\mathbf{V}^x)^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})} + (\mathbf{V}^z)^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})} \tag{2.34}$$

$$\frac{(\mathbf{V}^x)^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})} - (\mathbf{V}^x)^{t-\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})}}{\Delta t} + (D^x)_{(i+\frac{1}{2},j+\frac{1}{2})} \frac{(\mathbf{V}^x)^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})} + (\mathbf{V}^x)^{t-\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})}}{2} =$$

$$B_{(i+\frac{1}{2},j+\frac{1}{2})} \frac{\left(-C_2 \left(\mathbf{S}^t_{(i+2,j+\frac{1}{2})} - \mathbf{S}^t_{(i-1,j+\frac{1}{2})}\right) + C_1 \left(\mathbf{S}^t_{(i+1,j+\frac{1}{2})} - \mathbf{S}^t_{(i,j+\frac{1}{2})}\right)\right)}{\Delta x} \tag{2.35}$$

$$\frac{(\mathbf{V}^z)^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})} - (\mathbf{V}^z)^{t-\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})}}{\Delta t} + (D^z)_{(i+\frac{1}{2},j+\frac{1}{2})} \frac{(\mathbf{V}^z)^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})} + (\mathbf{V}^z)^{t-\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})}}{2} =$$

$$B_{(i+\frac{1}{2},j+\frac{1}{2})} \frac{\left(-C_2 \left(\mathbf{Z}^t_{(i+\frac{1}{2},j+2)} - \mathbf{Z}^t_{(i+\frac{1}{2},j-1)}\right) + C_1 \left(\mathbf{Z}^t_{(i+\frac{1}{2},j+1} - \mathbf{Z}^t_{(i+\frac{1}{2},j)}\right)\right)}{\Delta z} \tag{2.36}$$

$$\mathbf{X}^{t}_{(i+\frac{1}{2},j)} = (\mathbf{X}^{x})^{t}_{(i+\frac{1}{2},j)} + (\mathbf{X}^{z})^{t}_{(i+\frac{1}{2},j)} \tag{2.37}$$

$$\frac{(\mathbf{X}^{x})^{t+1}_{(i+\frac{1}{2},j)} - (\mathbf{X}^{x})^{t}_{(i+\frac{1}{2},j)}}{\Delta t} + (D^{x})_{(i+\frac{1}{2},j)} \frac{(\mathbf{X}^{x})^{t+1}_{(i+\frac{1}{2},j)} + (\mathbf{X}^{x})^{t}_{(i+\frac{1}{2},j)}}{2} = L2M_{(i+\frac{1}{2},j)} \frac{\left(-C_2\left(\mathbf{U}^{t+\frac{1}{2}}_{(i+2,j)} - \mathbf{U}^{t+\frac{1}{2}}_{(i-1,j)}\right) + C_1\left(\mathbf{U}^{t+\frac{1}{2}}_{(i+1,j)} - \mathbf{U}^{t+\frac{1}{2}}_{(i,j)}\right)\right)}{\Delta x} \tag{2.38}$$

$$\frac{(\mathbf{X}^{z})^{t+1}_{(i+\frac{1}{2},j)} - (\mathbf{X}^{z})^{t}_{(i+\frac{1}{2},j)}}{\Delta t} + (D^{z})_{(i+\frac{1}{2},j)} \frac{(\mathbf{X}^{z})^{t+1}_{(i+\frac{1}{2},j)} + (\mathbf{X}^{z})^{t}_{(i+\frac{1}{2},j)}}{2} = L_{(i+\frac{1}{2},j)} \frac{\left(-C_2\left(\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{3}{2})} - \mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j-\frac{3}{2})}\right) + C_1\left(\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})} - \mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j-\frac{1}{2})}\right)\right)}{\Delta z} \tag{2.39}$$

$$\mathbf{Z}^{t}_{(i+\frac{1}{2},j)} = (\mathbf{Z}^{x})^{t}_{(i+\frac{1}{2},j)} + (\mathbf{Z}^{z})^{t}_{(i+\frac{1}{2},j)} \tag{2.40}$$

$$\frac{(\mathbf{Z}^{x})^{t+1}_{(i+\frac{1}{2},j)} - (\mathbf{Z}^{x})^{t}_{(i+\frac{1}{2},j)}}{\Delta t} + (D^{x})_{(i+\frac{1}{2},j)} \frac{(\mathbf{Z}^{x})^{t+1}_{(i+\frac{1}{2},j)} + (\mathbf{Z}^{x})^{t}_{(i+\frac{1}{2},j)}}{2} = L_{(i+\frac{1}{2},j)} \frac{\left(-C_2\left(\mathbf{U}^{t+\frac{1}{2}}_{(i+2,j)} - \mathbf{U}^{t+\frac{1}{2}}_{(i-1,j)}\right) + C_1\left(\mathbf{U}^{t+\frac{1}{2}}_{(i+1,j)} - \mathbf{U}^{t+\frac{1}{2}}_{(i,j)}\right)\right)}{\Delta x} \tag{2.41}$$

$$\frac{(\mathbf{Z}^{z})^{t+1}_{(i+\frac{1}{2},j)} - (\mathbf{Z}^{z})^{t}_{(i+\frac{1}{2},j)}}{2} + (D^{z})_{(i+\frac{1}{2},j)} \frac{(\mathbf{Z}^{z})^{t+1}_{(i+\frac{1}{2},j)} + (\mathbf{Z}^{z})^{t}_{(i+\frac{1}{2},j)}}{2} = L2M_{(i+\frac{1}{2},j)} \frac{\left(-C_2\left(\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{3}{2})} - \mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j-\frac{3}{2})}\right) + C_1\left(\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})} - \mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j-\frac{1}{2})}\right)\right)}{\Delta z} \tag{2.42}$$

$$\mathbf{S}^{t}_{(i,j+\frac{1}{2})} = (\mathbf{S}^{x})^{t}_{(i,j+\frac{1}{2})} + (\mathbf{S}^{z})^{t}_{(i,j+\frac{1}{2})} \tag{2.43}$$

$$\frac{(\mathbf{S}^{x})^{t+1}_{(i,j+\frac{1}{2})} - (\mathbf{S}^{x})^{t}_{(i,j+\frac{1}{2})}}{\Delta t} + (D^{x})_{(i,j+\frac{1}{2})} \frac{(\mathbf{S}^{x})^{t+1}_{(i,j+\frac{1}{2})} - (\mathbf{S}^{x})^{t}_{(i,j+\frac{1}{2})}}{2} = M_{(i,j+\frac{1}{2})} \frac{\left(-C_2\left(\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{3}{2},j+\frac{1}{2})} - \mathbf{V}^{t+\frac{1}{2}}_{(i-\frac{3}{2},j+\frac{1}{2})}\right) + C_1\left(\mathbf{V}^{t+\frac{1}{2}}_{(i+\frac{1}{2},j+\frac{1}{2})} - \mathbf{V}^{t+\frac{1}{2}}_{(i-\frac{1}{2},j+\frac{1}{2})}\right)\right)}{\Delta x} \tag{2.44}$$

$$\frac{(\mathbf{S}^{z})^{t+1}_{(i,j+\frac{1}{2})} - (\mathbf{S}^{z})^{t}_{(i,j+\frac{1}{2})}}{\Delta t} + (D^{z})_{(i,j+\frac{1}{2})} \frac{(\mathbf{S}^{z})^{t+1}_{(i,j+\frac{1}{2})} + (\mathbf{S}^{z})^{t}_{(i,j+\frac{1}{2})}}{2} = M_{(i,j+\frac{1}{2})} \frac{\left(-C_2\left(\mathbf{U}^{t+\frac{1}{2}}_{(i,j+2)} - \mathbf{U}^{t+\frac{1}{2}}_{(i,j-1)}\right) + C_1\left(\mathbf{U}^{t+\frac{1}{2}}_{(i,j+1)} - \mathbf{U}^{t+\frac{1}{2}}_{(i,j)}\right)\right)}{\Delta z} \tag{2.45}$$

The horizontal and vertical dampening parameters must be computed such that no incident energy will generate reflections from propagation within the perfectly matched layer. Collino and Tsogka (2001) derive model dependent dampening parameters within the lossy boundary layer to sufficiently attenuate the wave-field. The dampening parameters are dependent on the speed of wave propagation V_p, the node width of the PML N, and a theoretical reflection co-efficient R. The theoretical reflection co-efficient can be selected based off the thickness of the layer ($N = 5 \rightarrow R = 0.01, N = 10 \rightarrow R = 0.001, N = 20 \rightarrow R = 0.0001$). Equations 2.46-2.47, delinate computation of the dampening parameters.

$$D^x(x) = D^{x_{max}} \left(\frac{x}{N\Delta x}\right)^2 \quad \text{where,} \quad D^{x_{max}} = log\left(\frac{1}{R}\right)\frac{3V_p}{2N\Delta x} \tag{2.46}$$

$$D^z(z) = D^{z_{max}} \left(\frac{z}{N\Delta z}\right)^2 \quad \text{where,} \quad D^{z_{max}} = log\left(\frac{1}{R}\right)\frac{3V_p}{2N\Delta z} \tag{2.47}$$

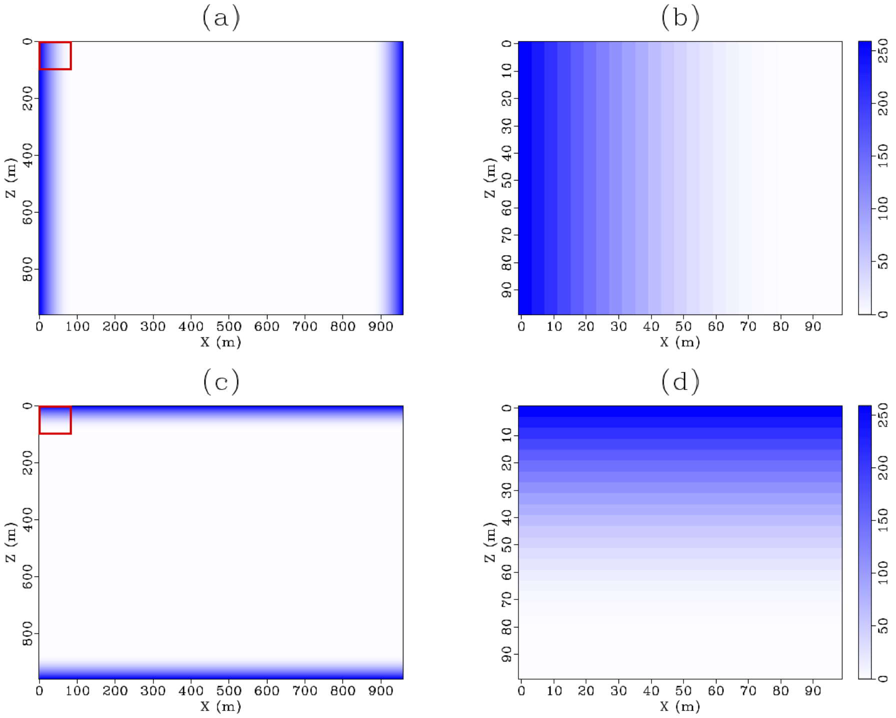


Figure 2.7: Perfect matching layer compressional dampening parameters for a homogenous acoustic medium [$N = 20$ $R = 0.0001$ $V_p = 1500$ $\frac{m}{s}$ $\Delta x = \Delta z = 4m$]. (a) Horizontal boundary $D^x(x)$. (b) Zoomed in view of red box from (a). (c) Vertical boundary $D^z(z)$. (d) Zoomed in view of red box from (c).

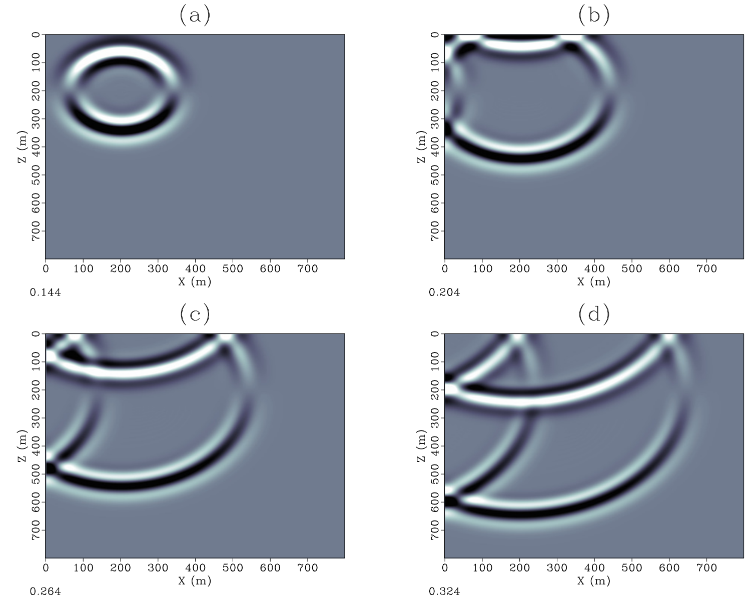


Figure 2.8: Snapshots of vertical particle velocity components for explosive source located at $(x, z) = (200, 200)m$ with no boundary conditions. Snapshots at times (a) 0.144s. (b) 0.204s. (c) 0.264s. (d) 0.324s.

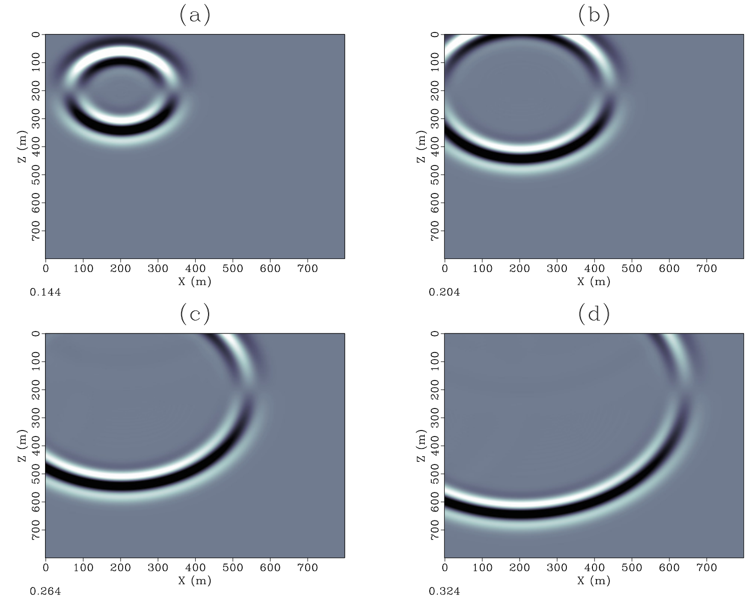


Figure 2.9: Snapshots of vertical particle velocity components for explosive source located at $(x, z) = (200, 200)m$ with perfectly match layer boundary shown in Figure 2.7. Snapshots at times (a) 0.144s. (b) 0.204s. (c) 0.264s. (d) 0.324s.

Figure 2.7 depicts equations 2.46-2.47. It can be observed that the boundary decays quadratically from $D^{x_{max}}$ and $D^{z_{max}}$ located at the edge of the PML, and equals zero outside of the boundary. In the corners of the boundary both D^x and D^z are applied to the wave-field components, this can be observed in the zoomed regions. Figure 2.8 shows wave-field propagation for a explosive source in homogenous acoustic medium $V_p = 1500\frac{m}{s}$ with no boundary conditions implemented, note the undesirable boundary reflections generated from unbounded model edges. Figure 2.9 is the same propagation as that of 2.8 but with implementation of perfectly match layer boundary conditions, note that the boundary condition attenuates the wave-field emulating an infinite region of propagation.

2.3.5 Stability criterion and grid dispersion relation

For the Levander staggered-grid finite-difference scheme the nine-point Cartesian stencils update the wave-field components at each time step. Due to the discrete approximation of continuous derivatives, restrictions must be placed on the parametrization of the numerical scheme such that the finite-difference algorithm be stable. The stability criterion is a restriction placed upon the parametrization of the time increment Δt, forcing the difference between the exact and numerical solutions to remain bounded as time index t increases (Mitchell, 1969). It can be physically understood as the largest time interval such that the update equations are uniquely determined by the nine-point stencils (Aki and Richards, 1980). This requires the stencils be spatially and temporally adequate, encompassing the region of deformational influence governing the update of a node. The value is dependent on the the size of node discretization $\Delta x = \Delta z = h$, the speed of wave propagation V_p and the finite-difference co-coefficients for the fourth-order approximation to the first derivative C_1 and C_2. The stability criterion for the Levander (1988) $O(\Delta t^2, h^4)$ staggered-grid finite-difference scheme is shown by equation 2.48. For heterogeneous models, the finite-difference algorithm will be stable provided the stability criterion is satisfied for all nodes within the model.

$$\Delta t < \frac{h}{\sqrt{2V_p\left(C_1 + C_2\right)}} \quad \text{or,} \quad \Delta t < 0.606\frac{h}{V_p} \tag{2.48}$$

The elastodynamic system of equations 2.14 is non-dispersive. That is the velocity of elastic waves are frequency independent and propagate at the compressional velocity V_p for P wave modes and at the shearing velocity V_s for S wave modes. However, discrete modelling of the continuous system introduces dispersive behaviour in the numerical propagation. This phenomenon known as grid dispersion produces a variation of wave propagation velocity with frequency, resulting higher frequencies of a signal to be delayed relative to the lower frequencies (Kelly et al., 1976). Dispersive behaviour in propagated source functions can be observed as a delay/broadening of the signal with the development of an oscillatory tail (Alford et al., 1974). The numerical grid dispersion manifests through time, causing the dispersive behaviour to become more prominent as propagation distance increases. A quantitative understanding of the cause of numerical grid dispersion, will allow parametrization of the finite-difference algorithm to satisfactorily limit its effects.

Taking the Fourier transform along the temporal and spatial dimensions of the wave equation results in its frequency-wave number representation. In this form the dispersion relation of the equation can be expressed, relating the wave number and frequency of a wave. This is shown for a single dimension by equation 2.49 where ω is the angular frequency, v the speed of propagation and k the wave number.

$$\omega(k) = v(k)k \tag{2.49}$$

Dispersion relations in this form are useful as they allow for the inspection of the phase and group velocity. Equation 2.50 below defines the phase velocity c_p and group velocity c_g. The continuous wave equation dispersion relation is linear ($v(k)$ a constant), such that the group velocity and phase velocity are equal and the equation non-dispersive.

$$c_p = \frac{\omega}{k} \quad \text{and,} \quad c_g = \frac{\partial \omega}{\partial k} \tag{2.50}$$

However, the explicit finite-difference solutions to the wave equation have nonlinear dispersion relations due to the use of discrete differencing operators. Trefethen (1982) shows the dispersion relations for different finite-difference schemes can be determined by inserting the plane wave solution $e^{i(\omega t + kx)}$ into a differencing formula and cancelling common factors. This can be done for a simple one-dimensional first-order wave equation, to gain a better understanding of nonlinear dispersion relation behaviour. This is shown below by equation 2.51-2.52 using second-order finite-difference operators $O(\Delta t^2, h^2)$ to approximate the derivatives for the hyperbolic equation $\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}$.

$$\frac{e^{i(\omega(t+\Delta t)+kx)} - e^{i(\omega(t-\Delta t)+kx)}}{\Delta t} = \frac{e^{i(\omega t+k(x+h))} - e^{i(\omega t-+k(x-h))}}{h} \tag{2.51}$$

Expanding the exponentials and cancelling common factors results in 2.52. Application of Euler's formula further simplifies the dispersion relation to 2.53.

$$\left(e^{i\omega\Delta t} - e^{-i\omega\Delta t}\right) = \frac{\Delta t}{h}\left(e^{ikh} - e^{-ikh}\right) \tag{2.52}$$

$$sin(\omega \Delta t) = \frac{\Delta t}{h} sin(kh) \tag{2.53}$$

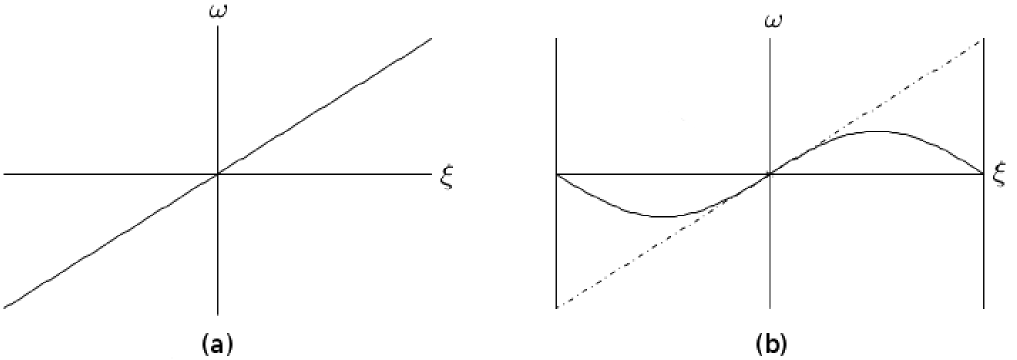


Figure 2.10: (a) Linear dispersion relation, for the continuous hyperbolic equation $\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}$. (b) Nonlinear dispersion relation from equation 2.53 for a fixed $\frac{\Delta t}{h}$. Dispersion relations depicted within the region $[\frac{-\pi}{h}, \frac{\pi}{h}] \times [\frac{-\pi}{\Delta t}, \frac{\pi}{\Delta t}]$ (Modified from Trefethen (1996)).

It can be observed the dispersion relation for one-dimensional finite-difference scheme is 2π periodic in kh and $\omega \Delta t$. This allows them to be described in the region $[\frac{-\pi}{h}, \frac{\pi}{h}] \times [\frac{-\pi}{\Delta t}, \frac{\pi}{\Delta t}]$, with all other values being an alias of a frequency in this domain. In addition, it is shown that these dispersion relations are sinusoidal and behave linearly near $(\omega, k) = (0, 0)$. In order to honour the non dispersive nature of the continuous wave equation, the size of node discretization h must be selected such that all numerically propagated wave numbers lie within the linear region of the dispersion relation near $(\omega, k) = (0, 0)$. This requires that the Nyquist wave number be chosen to be larger than the maximum propagated wave number.

In two dimensions, wave number becomes a vector with a horizontal and vertical component. This results in the dispersion relation to become a function of propagation angle. In addition, for the elastodynamic system of equations compressional and shearing velocities govern the propagation. Thus, separate dispersion relations must be determined for P and S wave modes. Levander (1988) derives the two-dimensional dispersion relations for the nine-point Cartesian finite-difference scheme outlined in equations 2.17-2.21. It is shown that the grid spacing must be selected such that the Nyquist wave number be five times that of the maximum propagated wave number to satisfactorily limit the dispersion effects for both wave modes at all angles. Figure 2.11 depicts the dispersive nature of the finite-difference algorithm if the model parametrization is not selected correctly.

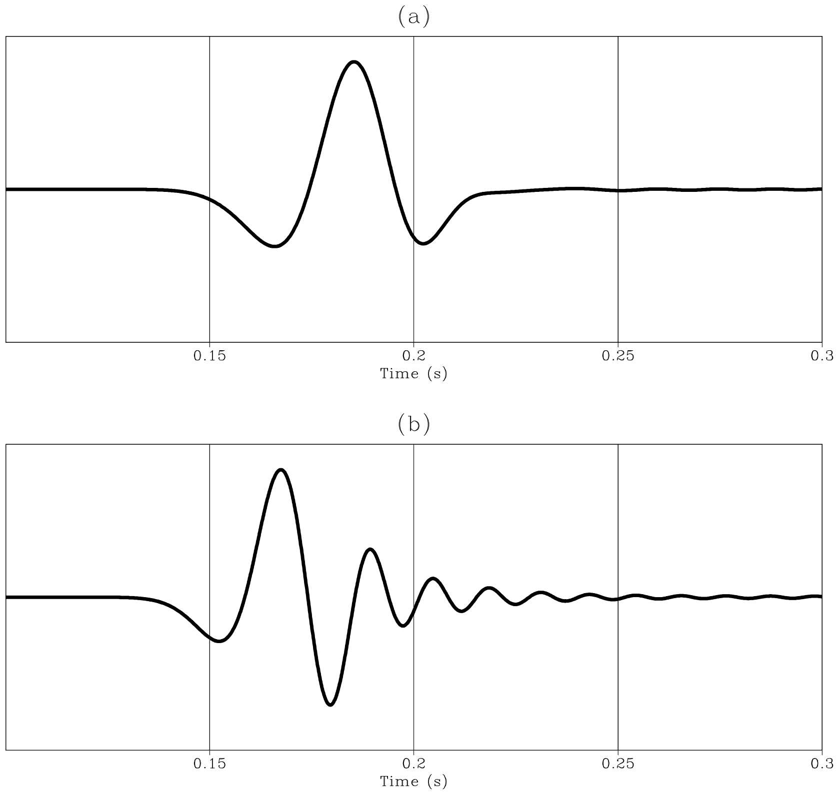


Figure 2.11: Depiction of grid dispersion effects for a vertical particle velocity recording ($\mathbf{V} = v_z$). Two zero phase Ricker wavelets were propagated within a homogenous medium [$V_p = 1500\,\frac{m}{s}$, $\rho = 1000\,\frac{kg}{m^3}$, $\Delta x = \Delta z = 5\,m$]. (a) Maximum propagated wave number $0.02m^{-1} \approx \frac{1}{5}\,k_{nyq}$ (b) Maximum propagated wave number $0.05m^{-1} \approx \frac{1}{2}\,k_{nyq}$. Notice the grid dispersion results in a phase distortion of the source function, and the generation of oscillatory artifacts.

2.4 Seismic data forward modelling

In heterogeneous elastic media, contrasts in the elastic parameters create interfaces. In the subsurface of the earth, complex interfaces exist as fluids and different layers of faulted and deformed rocks. As an incident wave-field propagates into an interface, wave-field scattering occurs. That is part of the incident wave is reflected off the interface and part of the wave is transmitted through the interface. The Zoeppritz equations (1919), calculate the reflection and transmission coefficients for an interface or reflector as a function of the elastic parameters on either side of the interface and the incident angle of the wave-field (Shuey, 1985). These reflection and transmission coefficients can be used to generate reflectivity Green's functions for source receiver pairs (Sherwood et al., 1983). Reflectivity Green's functions are impulse responses scaled by the reflection coefficients of interfaces that an incident plane wave would encounter propagating from its source position to receiver. Temporal locations of impulses must be at the time reflected energy from the corresponding interface would arrive at the receiver. Convolution of a finite-difference source function with a reflectivity Green's function will produce an analytic seismic trace that can be directly compared to a numerically modelled finite-difference seismic trace. Levander (1988) shows that the $O(\Delta t^2, h^4)$ Madariaga-Virieux finite-difference scheme solutions are in good agreement with analytic reflectivity solutions. We will not directly compare analytic and numerical solutions of wave-field reflection and transmission at a reflecting interface in this section. However, we will analyze the numerically modelled seismic phases generated from plane wave interaction with reflecting interfaces and compare with predicted phases from the analytic solutions.

Aki and Richards (2002) outlines that five types of reflecting interfaces can occur; solid/solid, fluid/fluid, solid/fluid, solid/free-surface (vacuum), and fluid/free-surface. In addition, using scalar potentials obtains equations for reflection and transmission coefficients describing the seismic phases generated from each type of interface. In this section we will investigate the reflection and transmission behaviour of solid/solid and solid/vacuum interfaces, in particular a single solid/solid planar interface underneath a solid/free-surface boundary. Table 2.1 lists all the predicted seismic phases that should be generated for a pressure source underneath a free-surface interface with nomenclature to describe the phases. The Rayleigh phase, is a surface phase with unique properties and will be discussed further below.

All of the down going phases generated from the free-surface interface and the original down going pressure wave from the source will reflect off the the planar solid/solid interface. Table 2.2 lists all predicted seismic phases that should be reflected from a solid/solid interface beneath a solid/free-surface interface for a pressure source implemented underneath the free-surface.

Incident Wave	Scattered Wave	Nomenclature
Up going P	Down going P	GP
Up going P	Down going SV	GS
Up going P	Rayleigh	R

Table 2.1: Predicted seismic phases generated from a pressure source below a solid/free-surface (vacuum) interface.

Incident Wave	Scattered Wave	Nomenclature
Down going P	Up going P	PP
Down going P	Up going SV	PS
Down going GP	Up going P	GPP
Down going GP	Up going SV	GPS
Down going GS	Up going P	GSP
Down going GS	Up going SV	GSS

Table 2.2: Predicted seismic phases reflected from solid/solid interface underneath a free-surface interface.

Figure 2.12 illustrates the predicted seismic phases in Table 2.2. The purple arrows denote converted phases P (compressional modes) to S (shear modes) or S to P, the red arrows shear phases S to S, and the blue arrows compressional phases P to P generated off the solid/solid interface (reflector). G nomenclature denotes the seismic phases initially scattered off of the free-surface.

Figure 2.13 shows snapshots of the numerically modelled wave-field, for a pressure source implemented 12 meters beneath the solid/free-surface interface. The source is placed shallow enough to generate Rayleigh wave excitation. Model parameters are shown in Figure 2.12, with the reflector located at a depth of 500 meters, the top boundary is a free-surface and the side/bottom boundaries perfectly matched layers $[N = 20, R = 0.0001]$. The horizontal component exhibits the proper polarity reversal on either sides of a vertical plane about the shot position. Due to the source implementation being close to the free-surface the phases generated off the free-surface are superimposed with the direct phases. Shearing and converted phases have near zero energy reflected at normal incidence, making them easily identifiable. It can be observed that the numerically modelled phases are in agreement with the analytically predicted seismic phases.

Figure 2.14 shows the wave-field recorded for the numerical propagation in Figure 2.13, with receivers placed at each node along $z = 0$. The Rayleigh and direct phase dominate the recordings. The direct wave on the horizontal component is more defined as the vertical components exhibits a polarity reversal on either side of a horizontal plane about the source position with the zero near the depth of the receiver array. The Rayleigh phase

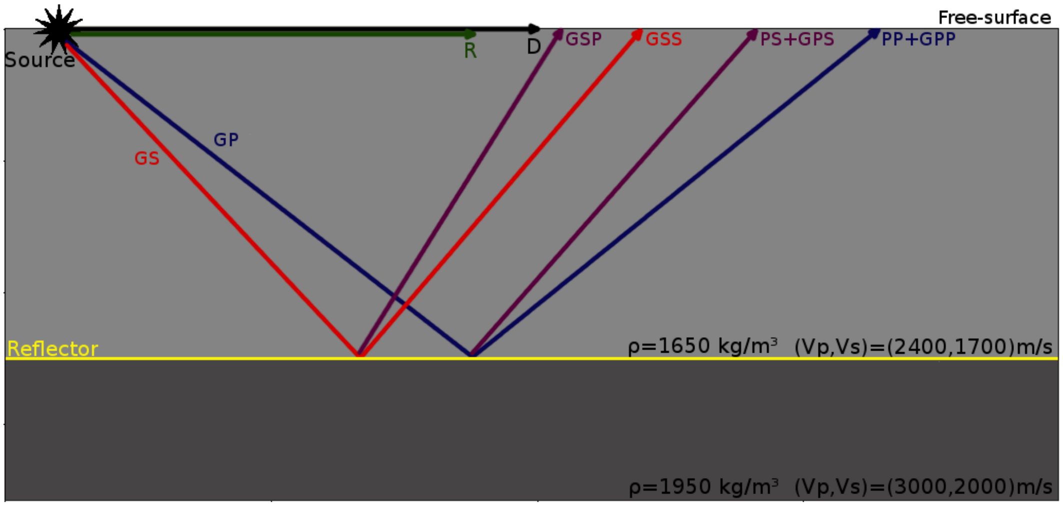


Figure 2.12: Illustration of predicted seismic phases generated from solid/solid interface beneath a solid/free-surface interface for a pressure source implemented underneath the free-surface (Table 2.1 and 2.2). Model parameters for numerical propagation are shown for above and below the reflecting interface.

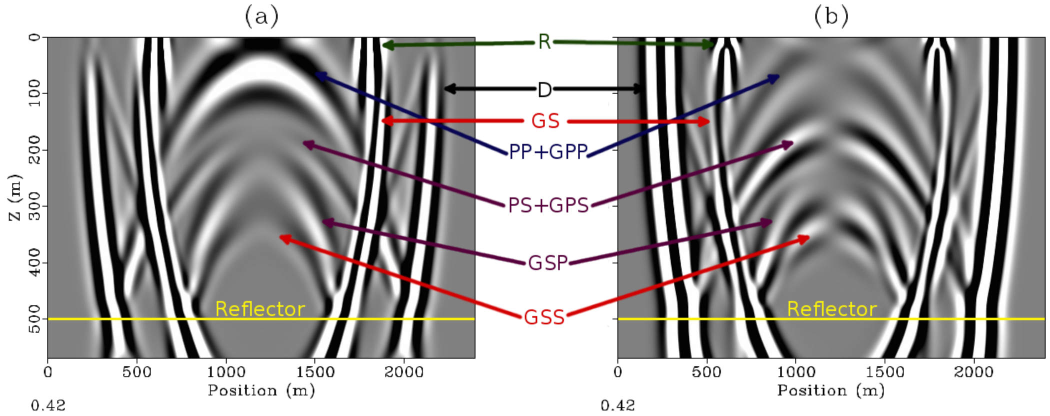


Figure 2.13: Snapshots of numerical wave-field at time 0.42 seconds for a pressure source implemented 12 meters beneath the solid/free-surface interface. (a) Vertical particle velocity snapshot ($\mathbf{V} = v_z$). (b) Horizontal particle velocity snapshot ($\mathbf{U} = v_x$).

is an evanescent surface wave that propagates along a homogenous half-space beneath a free-surface, and decays exponentially with depth. The Rayleigh phase has several unique properties. Firstly, it's speed of propagation is a few percent less than shear wave propagation. Jiang (2012) approximates the Rayleigh wave propagation speed as equation 2.54. With $\nu = \frac{\lambda}{2(\lambda+\mu)}$ being the Poisson's ratio of the propagation medium, V_R the speed of the Rayleigh wave and V_s the shear wave speed.

$$V_R = \frac{0.862 + 1.14\nu}{1+\nu} V_s \tag{2.54}$$

Poisson's ratio for the half-space beneath the free-surface is -0.0035 and thus the Rayleigh propagation speed should be numerically modelled at a speed of $V_R \approx 0.86V_s$. This property of the Rayleigh phase can be seen in Figure 2.13. Propagation of the Rayleigh phase is observed to be slower than the down going SV phase (GS) from the free-surface, lagging behind it and separating at a depth of $\approx$ 100m. In addition, Kristek et al. (2002) note that when modelling Rayleigh wave propagation the Levander dispersion relation from section 2.3.5 will no longer correctly mitigate grid dispersion of the Rayleigh phase.

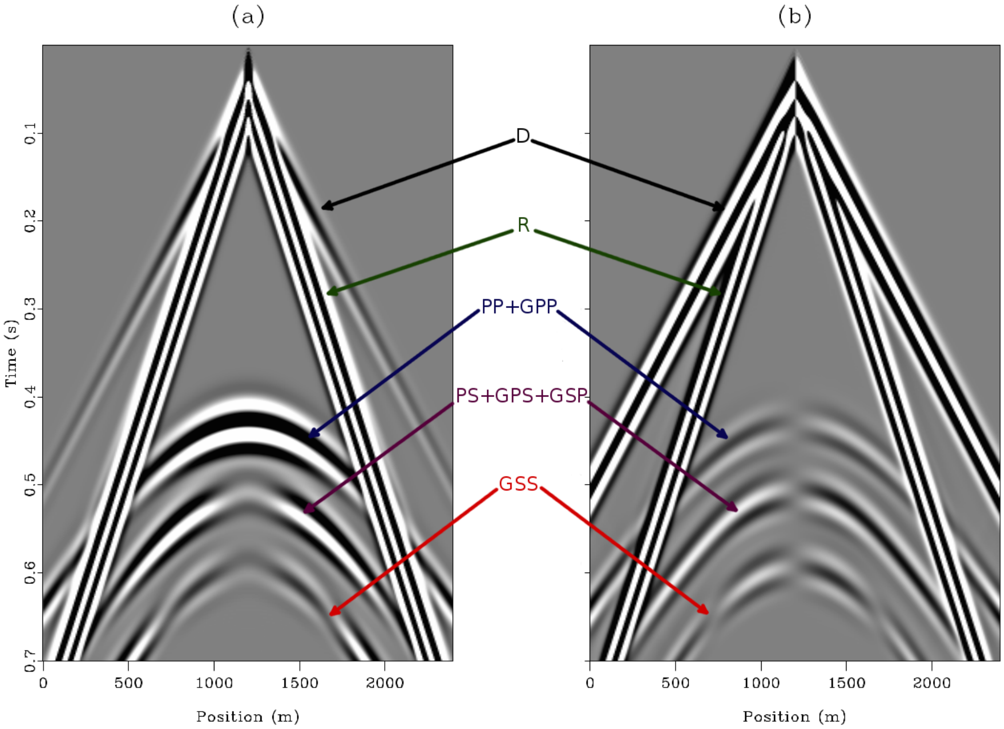


Figure 2.14: Numerical modelling of seismic phases. (a) Vertical particle velocity recording ($\mathbf{V} = v_z$). (b) Horizontal particle velocity recording ($\mathbf{U} = v_x$).

The second property of the Rayleigh phase is a retrograde elliptical particle motion at the free surface. Hodograms plot the particle motion at a fixed position for a given period of time. Figure 2.15 is the hodogram for the 750m trace in Figure 2.14 from 0.29 – 0.35s. It can be observed that the predicted elliptical particle motion is modelled correctly.

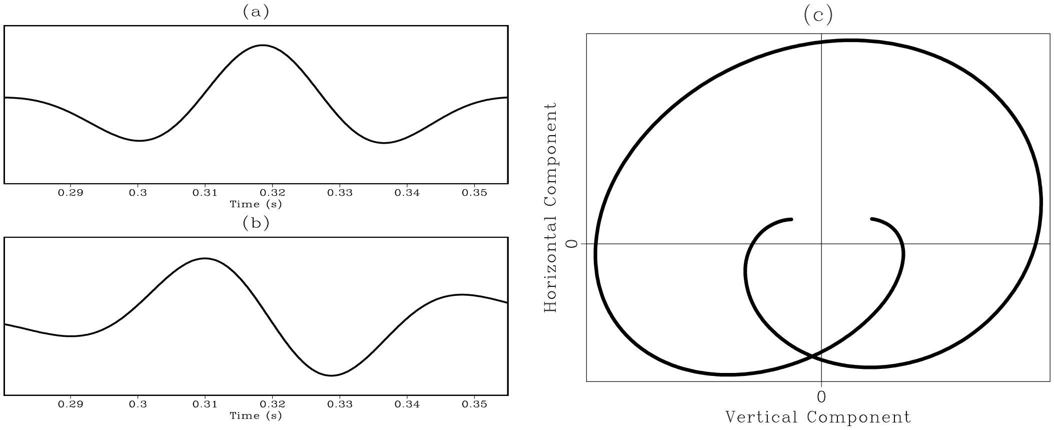


Figure 2.15: Hodogram of Rayleigh phase from 750m trace from 0.29 – 0.35s. (a) Vertical component. (b) Horizontal component. (c) Plot of horizontal component vs. vertical component displaying particle motion at free-surface. Note the clear elliptical trajectory.

2.5 Summary

In this chapter, starting with Hooke's law we derive the first-order elastodynamic system of equations. Under the assumption of continuum, this system of partial differential equations accurately describes isotropic elastic wave propagation. It is shown how explicit numerical solutions to the elastic wave equation can be generated by expressing continuous derivatives as discrete finite-difference approximations. A second-order time fourth-order space $O(\Delta t^2, h^4)$ staggered-grid finite-difference algorithm is outlined in detail with proper implementation of source functions, boundary conditions, and model parametrization. The behaviour of solutions generated via explicit finite-differences is shown to be in agreement with analytically determined solutions.

CHAPTER 3

Reverse-time migration

3.1 Introduction

Seismic data is composed of near surface recordings of a seismic wave-field at a discrete set of receiver locations. All energy within seismic recordings, apart from the direct arrival and ground roll, is source energy that has been scattered back to the receivers off of subsurface interfaces. Seismic imaging is the process of generating non-destructive structural images of the subsurface using seismic data. This is done by repositioning recorded energy to its location in the subsurface where incident source energy scattered to produce the observed seismic events. Seismic imaging or migration algorithms can be simplified into two steps, a wave-field extrapolation, and an imaging condition (Berkhout, 1981). The wave-field extrapolation requires a method to model seismic wave propagation forwards and backwards through time. The imaging condition then exploits our physical understanding of wave-field scattering to locate subsurface impedance contrasts. Thus, the accuracy of an imaging algorithm depends largely on the ability to realistically model seismic wave propagation.

This chapter will focus on the reverse-time migration method. The elastic staggered-grid finite-difference algorithm from the previous chapter will perform the wave-field extrapolation stage of the imaging algorithm. The wave-field extrapolation maps positions of source and receiver energy within a smooth background velocity model. The cross-correlation imaging condition of Claerbout (1971), then generates images of subsurface impedance contrasts. This is done under the assumption that extrapolated source and receiver energy exists at the same position spatially and temporally where incident source energy scattering occurred. For acoustic wave-fields implementation of a scalar imaging conditions is straight forward as the zero-lag cross correlation of two numerically modelled wave-fields. However, when dealing with elastic wave-fields a vector imaging condition must be implemented as compressional and shear wave modes cannot be simultaneously imaged from a single shot gather. Yan and Sava (2008) formulate a vector cross-correlation imaging condition which can be applied to isotropic elastic data which generates separate images using the elastic wave-fields scalar and vector potentials.

3.2 Migration theory

Under a weak scattering assumption, seismic data $D(\mathbf{x_r}, \mathbf{x_s}, \omega)$ can be written in frequency domain via Green's functions. The Green's functions $G(\mathbf{x}, \mathbf{x}', \omega)$, describe the wave equation for a smooth background velocity model. Green's functions can be interpreted as the impulse response measured at $\mathbf{x}$ from a point source located at $\mathbf{x}'$.

$$D(\mathbf{x_r}, \mathbf{x_s}, \omega) = \int_{\mathbf{x}'} \underbrace{P(\mathbf{x}', \omega)}_{down-going} \overbrace{G(\mathbf{x}_r, \mathbf{x}', \omega)}^{up-going} \underbrace{m(\mathbf{x}')}_{reflectivity} \, d\mathbf{x}' + \underbrace{G(\mathbf{x}_r, \mathbf{x}_s, \omega)}_{Direct\,wave} \tag{3.1}$$

The equation describes the recorded data as the volume integral of up-going Green's functions over all reflector positions, due to an incident down going wave $P(\mathbf{x}', \omega)$. The strength of the up going field is proportional to the reflectivity $m(\mathbf{x}')$, a small perturbation from the background model. An impulsive source is assumed where $\mathbf{x}_s$ describes the source position, $\mathbf{x}_r$ the receiver positions and $\mathbf{x}'$ the reflector positions.

The Born approximation estimates the scattered wave-field using only first order scattering events. This means that no free-surface or inter-layer multiples are considered in the formulation of the data. This allows the down-going component of the equation to be written as the impulse response at a reflector from incident source energy. Under the Born approximation the data can be expressed in frequency domain as the recorded first order scattered wave-field (Beylkin and Burridge, 1990).

$$D(\mathbf{x_r}, \mathbf{x_s}, \omega) = \int_{\mathbf{x}'} \underbrace{G(\mathbf{x}', \mathbf{x}_s, \omega)}_{down-going} \overbrace{G(\mathbf{x}_r, \mathbf{x}', \omega)}^{up-going} \underbrace{m(\mathbf{x}')}_{reflectivity} \, d\mathbf{x}' \tag{3.2}$$

This allows for the reflectivity term $m(\mathbf{x}')$ to be solved for as the data space integral of the product of the adjoint Green's functions and the seismic data. $G^*(\mathbf{x}, \mathbf{x}', \omega)$ denotes the adjoint or conjugate Green's functions. The adjoint Green's functions can be understood as a "flipped" Green's function which back projects energy positioned from a forward Green's function at $\mathbf{x}$ to it's point of origin $\mathbf{x}'$.

$$\hat{m}(\mathbf{x}') = \int_{\omega}\int_{\mathbf{x_s}}\int_{\mathbf{x_r}} \overbrace{G^*(\mathbf{x}', \mathbf{x}_s, \omega)}^{adjoint\, down-going} \; \underbrace{G^*(\mathbf{x}_r, \mathbf{x}', \omega)}_{adjoint\, up-going} \overbrace{D(\mathbf{x_r}, \mathbf{x_s}, \omega)}^{data} \, d\mathbf{x_r}\, d\mathbf{x_s}\, d\omega \tag{3.3}$$

Taking the complex conjugate of equation 3.3 results in equation 3.4. From the cross-correlation theorem, integration over frequency in equation 3.4 results in its time domain representation as the zero lag time cross-correlation between the down-going Green's function and the adjoint up-going Green's function convolved with the seismic data (Schuster et al., 2002). Below $\otimes$ denotes the zero lag time cross-correlation and $\star$ convolution.

$$\hat{m}(\mathbf{x}')^* = \int_{\omega}\int_{\mathbf{x_s}}\int_{\mathbf{x_r}} \overbrace{G(\mathbf{x}', \mathbf{x}_s, \omega)}^{down-going}[\; \underbrace{G^*(\mathbf{x}_r, \mathbf{x}', \omega)}_{adjoint\, up-going} \; \overbrace{D(\mathbf{x_r}, \mathbf{x_s}, \omega)}^{data}]^* \, d\mathbf{x_r}\, d\mathbf{x_s}\, d\omega \tag{3.4}$$

$$\hat{m}(\mathbf{x}') = \int_{\mathbf{x_s}}\int_{\mathbf{x_r}} \underbrace{g(\mathbf{x}', t'|\mathbf{x}_s, 0)}_{down-going} \otimes \overbrace{[g^{adj}(\mathbf{x}_r, t_r|\mathbf{x}', 0)}^{adjoint\, up-going} \star \underbrace{d(\mathbf{x_r}, t_r)}_{data}] d\mathbf{x_r}\, d\mathbf{x_s} \tag{3.5}$$

However, contemporary wave-field extrapolation migration methods do not analytically solve for the forward and adjoint Green's functions, but rather use numerical schemes to simulate wave propagation. In reverse-time migration finite-difference solutions to the two-way wave equation are traditionally used (Baysal et al., 1983; Whitmore, 1983; McMechan, 1983). Therefore, Green's functions in the above equations represent forward/adjoint modelling algorithms within a smooth background velocity model. Thus the bracketed term $[g^{adj}(\mathbf{x}_r, t_r|\mathbf{x}', 0) \star d(\mathbf{x_r}, t_r)]$ represents the back propagation of recorded seismic data from the receiver positions $\mathbf{x}_r$ to the locations in the subsurface $\mathbf{x}'$ (receiver side). Similarly the down-going wave-field term models source energy propagation within a discrete background model (source side). This allows the reflectivity for a single shot gather to be formulated as the zero lag cross-correlation of two numerically computed discrete wave-fields (Claerbout, 1971).

$$\hat{m}(\mathbf{x}) = \sum_{t}^{Nt} \underbrace{S(\mathbf{x}, t)}_{source\ side} \; \overbrace{R(\mathbf{x}, t)}^{receiver\ side} \tag{3.6}$$

The theory assumes that only first order scattering events exist within the data. Higher order scattering events and direct wave energy will therefore produce artifacts within the migrated image. In addition, it is assumed that an accurate background model is known.

3.3 Data modelling and pre-processing

3.3.1 Synthetic data modelling

Synthetic data modelling is performed by recording numerical wave propagation at a discrete set of locations within an input velocity model. The velocity model is shown below, with a complex faulted region at $1.2\,km$ in depth. The top boundary of the velocity model is set to be a perfectly matched layer such that no free-surface multiples will propagate. A $400\,m$ water layer ($V_s = 0$) exists at the top of the velocity model. A synthetic survey was conducted as an ocean bottom node experiment with receivers placed at a $10\,m$ interval along the first node in the layer beneath the water. A 75° phase rotated $17.5\,Hz$ central frequency Ricker wavelet as shown in Figure 2.6 acts as the source function. A total of 96 shots gathers were modelled with a shot spacing of $37.5\,m$ each at a depth of $40\,m$ below the surface of the water layer. Figure 3.2 shows a multi-component shot gather for a source function implemented at $x = 1.3\,km$. An acoustic data set was also modelled by setting the shearing velocity to zero everywhere within the model ($\mu = 0$), resulting in a separate data set with no shear waves.

3.3.2 Data pre-processing

Chang and McMechan (1994) outline three processing steps to be performed prior to the the wave-field extrapolation stage of the migration algorithm. Firstly, the direct wave and ground roll are removed from each shot gather. This is due to the fact that these phases are not associated with scattering of incident source energy, and therefore do not contribute to the migrated image. To remove the direct wave and ground-roll, numerical replicas of these phases are generated and subtracted from the data recordings. The replicas are produced using a modified velocity model, with velocities and density made to be a homogenous half space below the receiver positions. Secondly the spacial and temporal edges of the shot gathers must be tapered prior to back propagation. This is to eliminate the high frequencies of a hard edge. Applying a taper to the edges of the data allows the numerical dispersion relation of the finite-difference algorithm to be satisfied during data extrapolation. A simple cosine taper is an effective method to smooth the edges of the shot recordings. The final processing step is to reverse the time order of the shot recordings for input into the finite-difference algorithm. Figure 3.3 shows the data from Figure 3.2 with the direct wave and ground roll removed, because we are able to numerically replicate the direct wave and ground-roll present in the synthetic data they are perfectly removed.

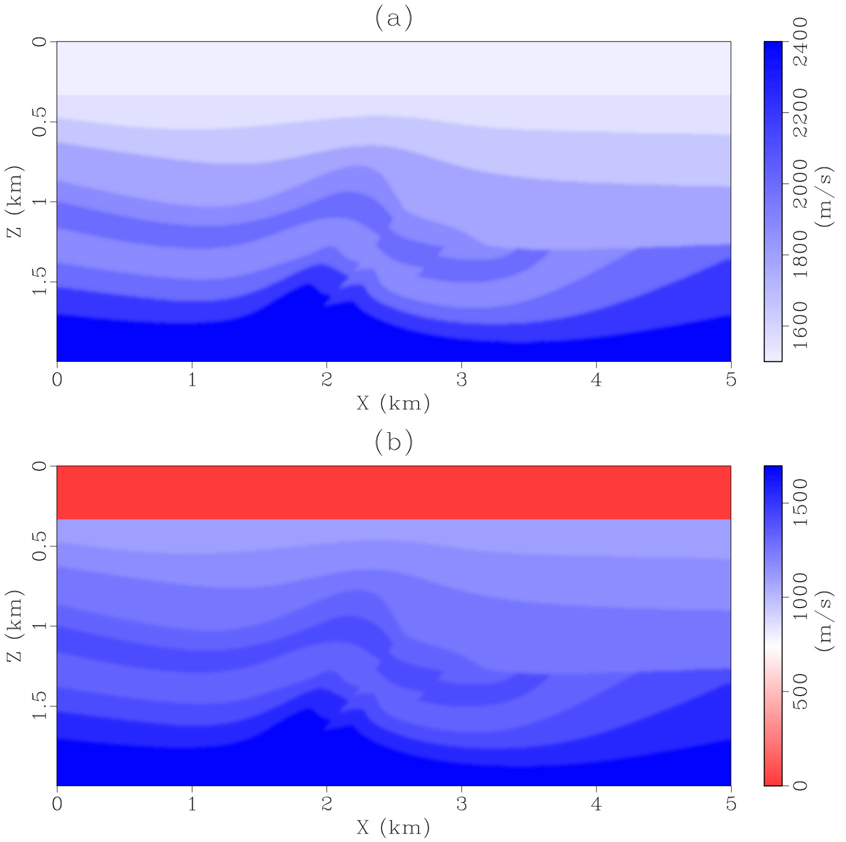


Figure 3.1: Faulted velocity model [$nx = 1000 \quad nz = 400 \quad \Delta x = 5m \quad \Delta z = 5m$], $\frac{V_p}{V_s} = \sqrt{2}$ and density $1000 - 1750\,kg/m^3$. (a) Compressional velocities V_p. (b) Shearing velocities V_s.

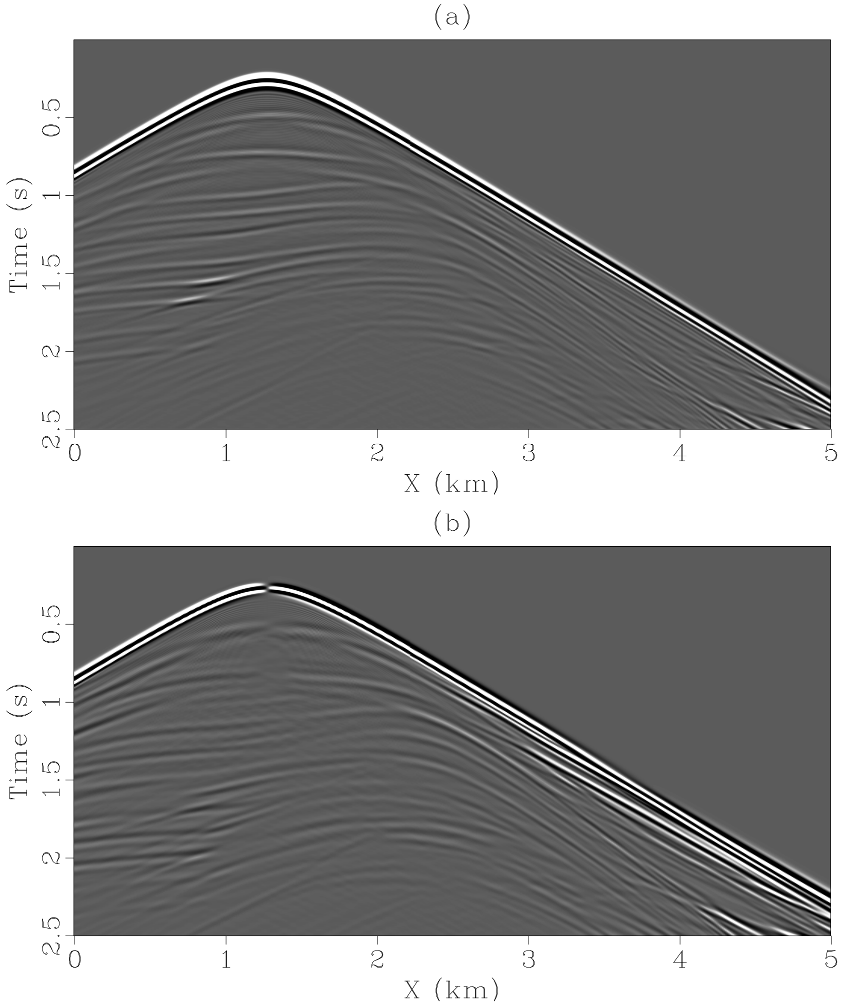


Figure 3.2: Recorded elastic data for source located at $1.3\,km$. (a) Vertical particle velocity component ($\mathbf{V} = v_z$). (b) Horizontal particle velocity component ($\mathbf{U} = v_x$).

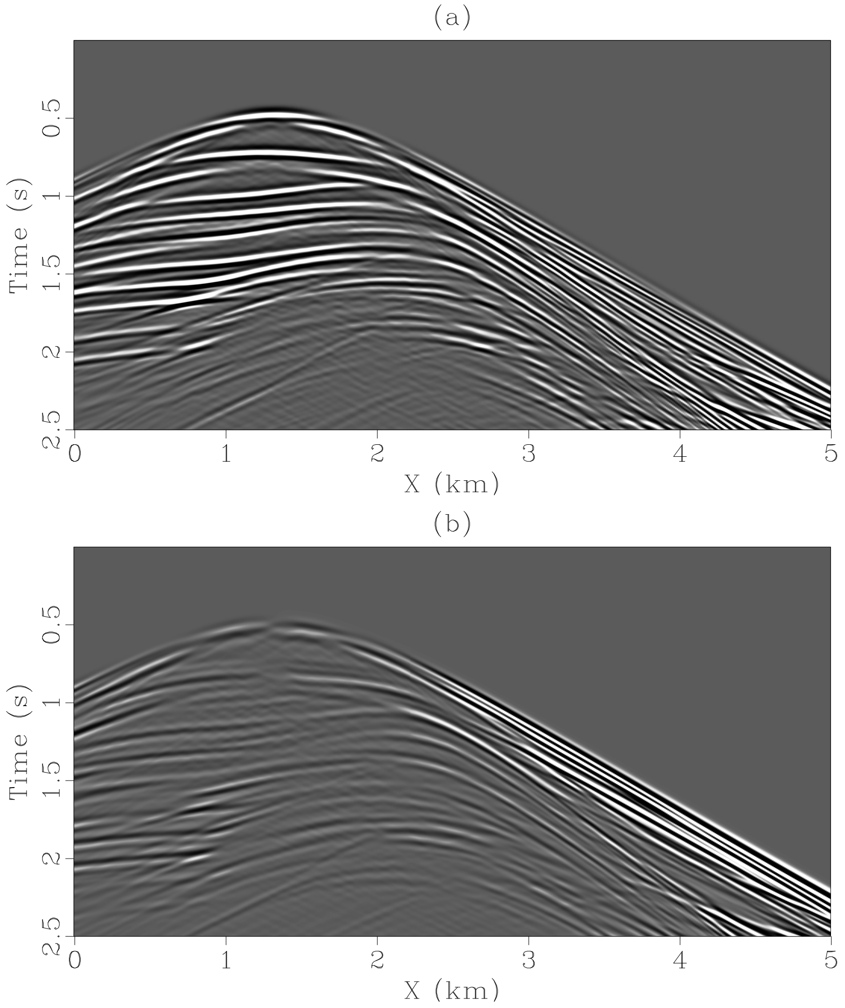


Figure 3.3: Recorded elastic data for source located at $1.3\,km$ with the direct wave removed. (a) Vertical particle velocity component ($\mathbf{V} = v_z$). (b) Horizontal particle velocity component ($\mathbf{U} = v_x$).

3.4 Wave-field extrapolation

For a given shot profile k we can visualize seismic recordings as a series of snapshots of the seismic wave-field restricted to only the receiver locations $d_k(\mathbf{x_r}, t)$. We can extrapolate our restricted view of the wave-field back through time via finite-difference modelling. This can be achieved by setting the sampling function of the finite-difference source to be the receiver locations $\mathbf{x_r} = (x_r, z_r)$, and the source function $s(t)$ to be the pre-processed data at each receiver location. Care must be taken to ensure that each sample is injected at the correct receiver position and as the component of the wave-field it was recorded as.

If we have an accurate background velocity model, the back propagated data will be properly positioned in the subsurface model $\mathbf{x} = (x, z)$ at each time step. This is referred to as the receiver side wave-field $R_k(\mathbf{x}, t)$, a partial reconstruction of the wave-field that was propagating during the generation of the synthetic seismogram. Similarly, if we have prior knowledge of the position $\mathbf{x_s} = (x_s, z_s)$ and temporal shape of the source function (seismic wavelet) $s_k(\mathbf{x_s}, t)$, we can extrapolate the source wave-field forward through time. This is referred to as the source side wave-field $S_k(\mathbf{x}, t)$, a reconstruction of incident source energy that was scattered to produce the recorded seismogram.

The source side wave-field is extrapolated forward in time from $t = 0$ to the maximum recording time $t = t_{max}$ and the receiver side wave-field is extrapolated backwards in time from $t = t_{max}$ to $t = 0$. This means that the zero lag time cross-correlation of the two wave-fields cannot be performed in place as the first computed time slice from source side wave-field must be multiplied with the last computed time slice from the receiver side wave-field. The source wave-field must therefore be stored in memory/disk to implement the imaging condition. For two-dimensional problems the data transfer of writing the wave-field to disk is not overly problematic. However, for the three-dimensional case the wave-fields can become very large and having to write them to disk results in bottlenecks in system input/output. Check-pointing schemes can be implemented for large three-dimensional problems to remedy this issue by replacing writing the source wave-field to disk with an additional finite-difference propagation stage (Symes, 2007).

Figure 3.4 illustrates the wave-field extrapolation process for a single shot profile k. The acoustic data set with no converted waves was used to produce this illustration. Only the vertical particle velocity component of the data $d_k(\mathbf{x_r}, t)$ is shown, however both horizontal and vertical particle velocity components were injected as the source function to reconstruct the receiver side wave-field. The background velocity model is a smooth version of Figure 3.1 (a). The wave-field extrapolation produces three dimensional snapshot cubes of the reconstructed wave-fields. Only the pressure component snapshot cubes are shown (average of normal stresses), but all components of the wave-field are propagated.

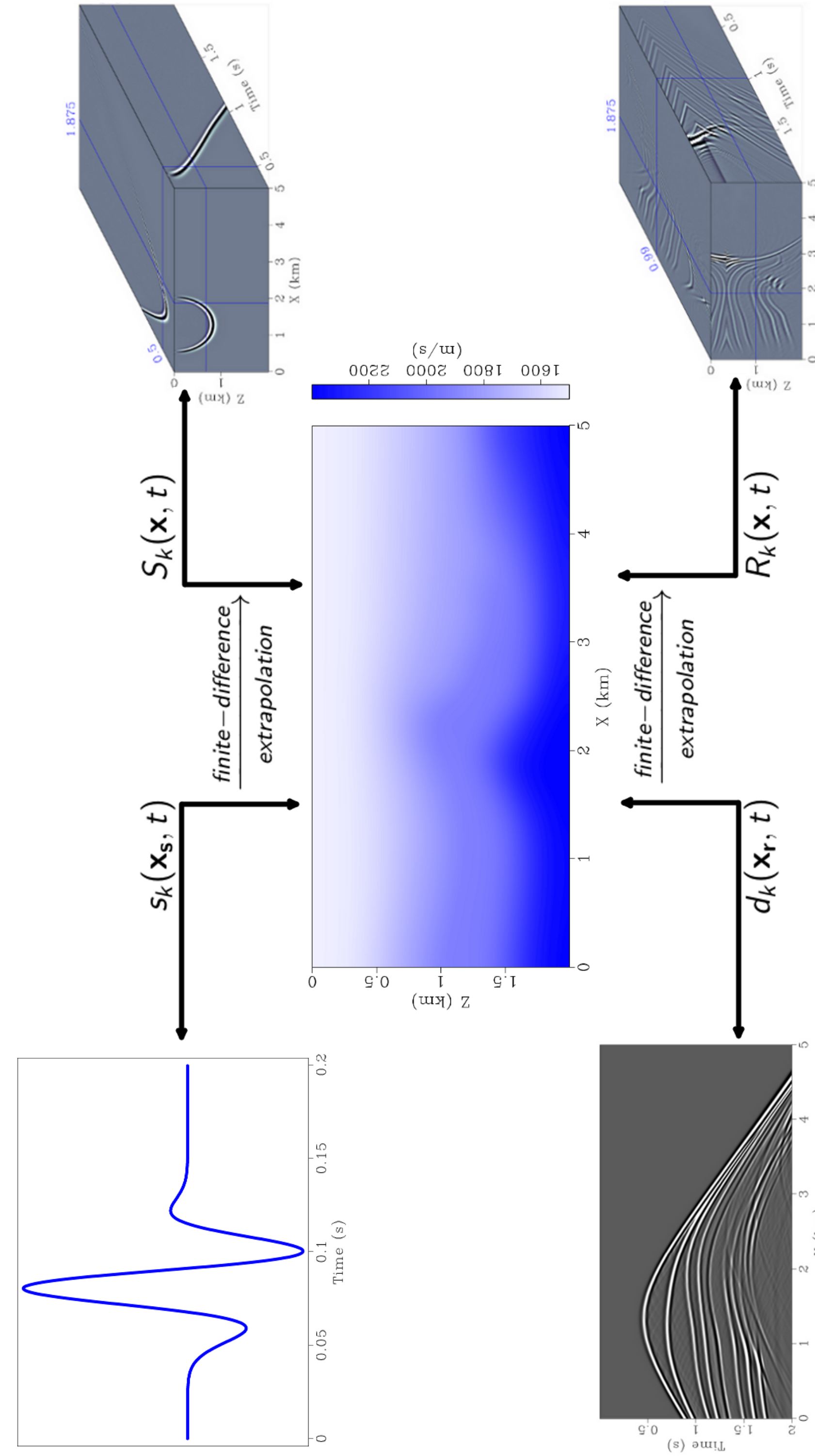


Figure 3.4: Depiction of wave-field extrapolation for a single acoustic shot gather via two-dimensional finite-difference.

3.5 Scalar reverse-time migration

For acoustic data ($\mu = 0$) only compressional wave modes exist within the data. This allows the wave-field to be described as a scalar field with values corresponding to volumetric compression and dilatation. The pressure field (average of normal stresses) describes this type of deformation. Therefore, the source and receiver wave-fields should be represented as the pressure field when imaging with acoustic data.

If source side and receiver side energy exist at the same position in the subsurface at time step t, this reveals a subsurface reflector where source energy scattering has occurred. This is the concept of the cross-correlation imaging condition. The mathematical representation is delineated by $\hat{m}_k(\mathbf{x})$, in equation 3.7, as the zero lag cross-correlation of the source and receiver side snapshot cubes $S_k(\mathbf{x}, t)$ and $R_k(\mathbf{x}, t)$. This produces an image of subsurface impedance contrasts $\hat{m}_k(\mathbf{x})$ by migrating the recorded wave-field to positions in the subsurface where scattering of the incident source wave-field occurred.

In conventional shot-profile migration extrapolation of individual shot recordings are cross-correlated with there corresponding source side wave-fields. Migrations of a single shot are represented by the indice k. These separate migrations are only capable of imaging the subsurface region illuminated by a lone source, where the region of illumination is dependent on the sources spatial position. Therefore, in order to image an extensive area multiple migrations must be performed for shots spatially distributed over the region of interest. Calculating the sample mean of all migrations (summation over k) produces an image of the subsurface region illuminated by all shots $\mathrm{I}(\mathbf{x})$.

$$\hat{m}_k(\mathbf{x}) = \sum_{t}^{Nt} S_k(\mathbf{x}, t) R_k(\mathbf{x}, t) \tag{3.7}$$

$$\mathrm{I}(\mathbf{x}) = \frac{1}{Nk} \sum_{k}^{Nk} \hat{m}_k \tag{3.8}$$

The two-way wave equation extrapolation in reverse-time migration allows for upward and downward wave propagation within the source side and receiver side wave-fields. This results in advantages over other imaging methods, such as the being able image steeply dipping structures, and models all types of waveforms including turning waves. However, it also results in unwanted low wave number artifacts in the image from the cross-correlation of down going and upcoming wave paths not occurring at scattering locations. The artifacts typically only occur to a depth in the model where a the first large velocity contrast exists such that the critical angle may be reached. After this the wave-field fails to penetrate below this depth for certain angles and the artifacts are greatly reduced (Jones, 2014). Guitton et al. (2006) outlines three categories for attenuating these artifacts, wave-field extrapolation methods, imaging condition methods, and post imaging condition methods. Post imaging condition methods have the most straight forward implementation. A simple and effective way to remove these artifacts is a to apply a high pass wave number filter ($k_x\, k_z$) to the contaminated image. Derivative and Lapacian filters are other alternative post imaging condition methods to remove the low wave number artifacts.

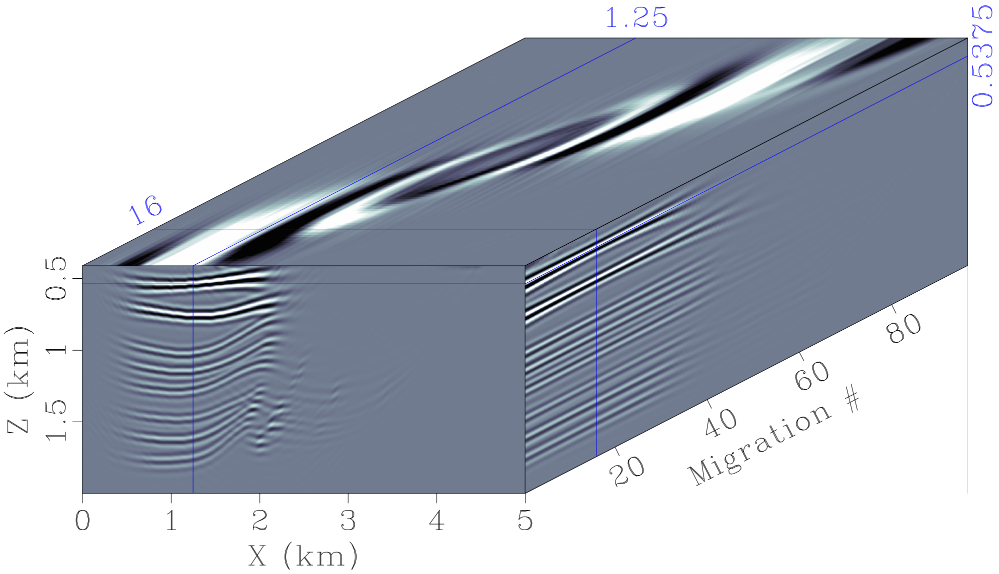


Figure 3.5: Partial image cube containing all 96 migrated synthetic shot profiles $\hat{m}_k(\mathbf{x})$. The images have been k_z filtered to remove low wavenumber artifacts. The $x = 1.3km$ shot profile is shown on the face of the cube. It can be observed how the migration of each shot profile only resolves a region of the model.

Due to geometric spreading of the elastic wave-field and energy loss during transmission through interfaces, the amplitudes of reflectors within the migrated images are not well balanced. The shallow reflectors corresponding to early arrivals have larger amplitude than deeper reflectors corresponding to later arrivals. Kaelin et al. (2006) show that normalizing the image by the source side wave-field can help suppress the low wave number artifacts and better balance the amplitudes of the deeper reflectors. The source normalized imaging condition is shown below as equation 3.9. A small regularization parameter ε must be added to the denominator to avoid division by zero.

$$\mathrm{I}(\mathbf{x}) = \frac{1}{Nk} \sum_{k}^{Nk} \sum_{t}^{Nt} \frac{S_k(\mathbf{x},t) R_k(\mathbf{x},t)}{S_k(\mathbf{x},t)^2 + \varepsilon} \tag{3.9}$$

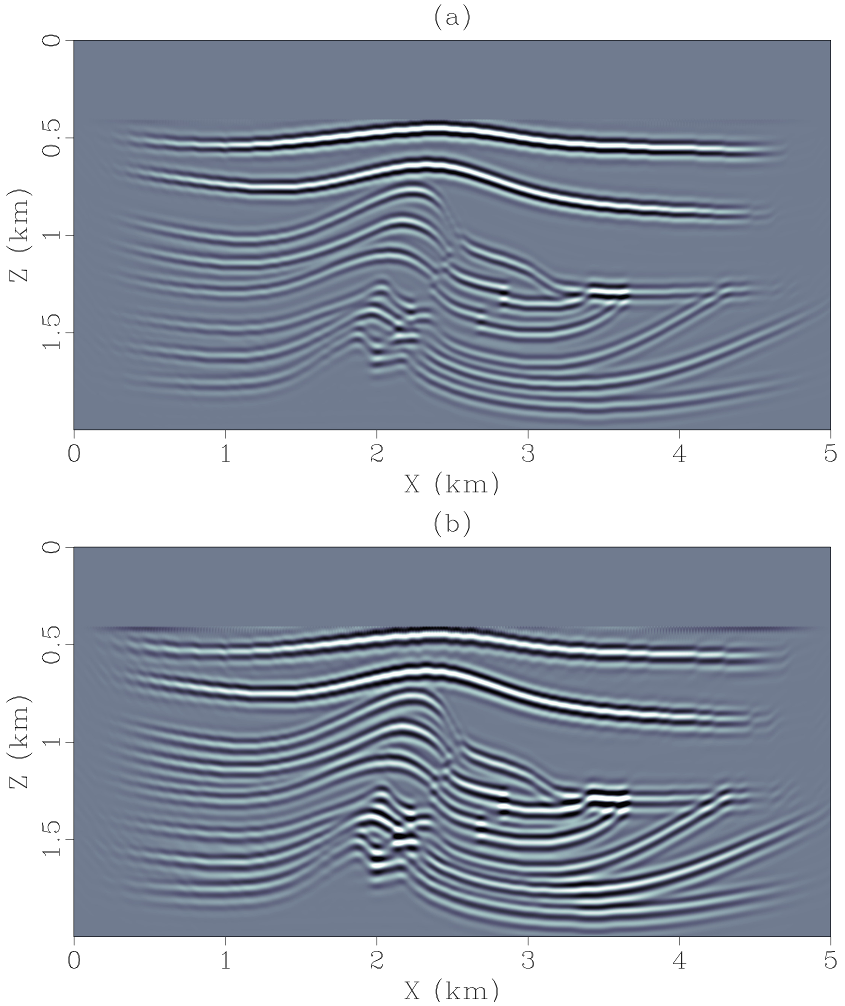


Figure 3.6: Acoustic image results for the 96 shot-profile data set taking the average of all partial images. The images have been k_z filtered to remove low wave number artifacts, and muted above the receivers. (a) No normalization. (b) Image normalized by the source side wave-field, notice the amplitudes of reflectors are better balanced.

3.6 Vector reverse-time migration

The seismic wave-field consists of both compressional and shear wave modes. This makes a migration algorithm which can image using both wave modes desirable. Chang and McMechan (1987) formulate a two way elastic wave-equation reverse time migration for pre-stack data, where the imaging condition incorporates ray tracing to calculate the image time. The migration algorithm of Botelho and Stoffa (1991) also uses an elastic-wave equation for computing image times and receiver extrapolation; however, does not consider generating separate images for primary and converted seismic phases. More contemporary techniques propose generating independent images for primary and converted phases via decomposing the elastic wave-field into scalar and vector potentials (Sun and McMechan, 2001; Yan and Sava, 2008)

A corollary of Helmholtz Theorem states that for any divergence free vector field Ψ and curl free scalar field Φ there exists a unique vector field $\mathbf{v}$ such that:

$$\nabla \cdot \mathbf{v} = \Phi \quad \text{and,} \quad \nabla \times \mathbf{v} = \Psi \tag{3.10}$$

In two-dimensional isotropic media the compressional wave mode $P_w(\mathbf{x}, t)$ corresponds purely to volumetric perturbations (curl free). Similarly, the shear wave mode $S_w(\mathbf{x}, t)$ corresponds purely to shearing/shape perturbations (divergence free). Therefore, the isotropic elastic wave-field $\mathbf{v}(\mathbf{x}, t)$ composed of compressional and shear wave modes can be described as a vector field with a prescribed divergence and curl where:

$$P_w(\mathbf{x}, t) = \nabla \cdot \mathbf{v}(\mathbf{x}, t) = \frac{\partial v_x(\mathbf{x}, t)}{\partial x} + \frac{\partial v_z(\mathbf{x}, t)}{\partial z} \tag{3.11}$$

$$S_w(\mathbf{x}, t) = \nabla \times \mathbf{v}(\mathbf{x}, t) = \frac{\partial v_x(\mathbf{x}, t)}{\partial z} - \frac{\partial v_z(\mathbf{x}, t)}{\partial x} \tag{3.12}$$

Dellinger and Etgen (1990) exploit this property of the elastodynamic equations in isotropic media, separating finite-difference modelled elastic wave-fields into compressional and shear modes via discrete divergence and curl operators. Equations 3.11 and 3.12 illustrate these operators which can be implemented via finite-differences. Figure 3.7 illustrates the application of the operators to a simple elastic wave-field.

Applying wave mode separation to the entire receiver side wave-field cube, Yan and Sava (2008) formulate an expanded cross-correlation imaging condition for vector wave-fields. This is shown in equation 3.13 where the superscripts on the source side and receiver side snapshot cubes denotes the corresponding wave mode. This imaging condition has a clear physical interpretation, as cross-correlation of a source side and receiver side compressional

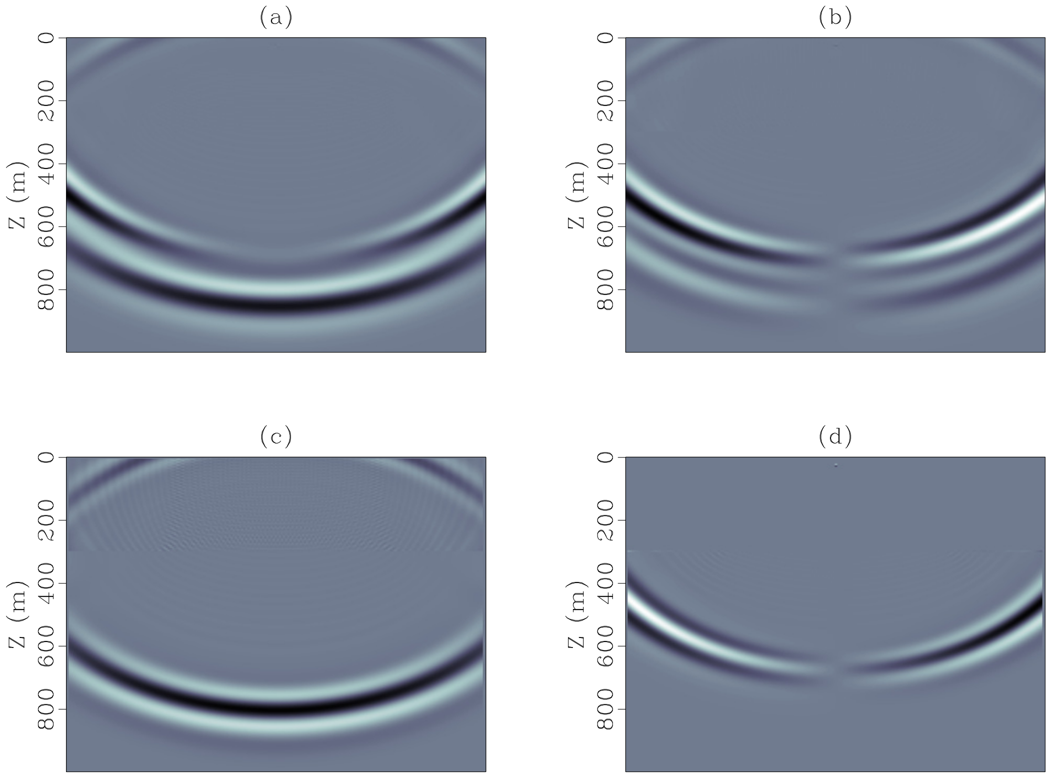


Figure 3.7: Separation of wave-field at a single time step, for a 300m water layer $[V_p = 1500\,\frac{m}{s} \;\; V_s = 0\,\frac{m}{s} \;\; \rho = 1000\,\frac{kg}{m^3}]$, over a 700m half space $[V_p = 2000\,\frac{m}{s} \;\; V_s = 1500\,\frac{m}{s} \;\; \rho = 1250\,\frac{kg}{m^3}]$. a.) Vertical particle velocity v_z. b.) Horizontal particle velocity v_x. c.) Compressional modes P_w. d.) Shear modes S_w. Note no shear modes within the water layer (0-300m).

modes gives an image of the migrated primary seismic phases $\hat{m}_K^{PP}(\mathbf{x})$ and, cross-correlation of a source side compressional mode with a receiver side shear mode an image of the PS converted seismic phases $\hat{m}_K^{PS}(\mathbf{x})$. This results in two separate images, one where compressional phases (PP) in the data have been repositioned to scattering locations, and another where the converted phases (PS) have been repositioned to scattering locations. Source and receiver side normalization can be applied to the vector imaging formulation, as in the previous section.

$$\begin{gathered} \hat{m}_k^{PP}(\mathbf{x}) = \sum_t S_k^{P_w}(\mathbf{x},t) R_k^{P_w}(\mathbf{x},t) \quad \text{and,} \quad \hat{m}_k^{PS}(\mathbf{x}) = \sum_t S_k^{P_w}(\mathbf{x},t) R_k^{S_w}(\mathbf{x},t) \\ \mathrm{I}^{PP}(\mathbf{x}) = \sum_k \hat{m}_k^{PP}(\mathbf{x}) \quad \text{and,} \quad \mathrm{I}^{PS}(\mathbf{x}) = \sum_k \hat{m}_k^{PS}(\mathbf{x}) \end{gathered} \tag{3.13}$$

In addition to the low wave number artifacts in the migrated images, the converted phase images $\hat{m}_K^{PS}(\mathbf{x})$ also exhibit a polarity reversals along each imaged reflector. This is problematic as destructive interference occurs when taking the sample mean of many partial images to produce the final image. This polarity reversal is the result of the polarity reversal present in the shear modes as observed in Figure 3.7 (d). Balch and Erdemir (1994) illustrate that for isotropic elastic media the location of the polarity reversal within the shear wave can be determined if the angle of incident energy with respect to the reflecting boundary is known. Duan and Sava (2014) outline an imaging condition applied to separated wave-fields to correct for this polarity reversal observed in the PS images. For the PS image shown below in Figure 3.8 (b), the polarity of partial PS images were reversed about a vertical plane beneath the source function prior to summation. Although this method partially corrects for the polarity reversal is it not exact as the model exhibits interfaces with variable dips. The destructive interference produced from not perfectly correcting the aforementioned polarity reversals is why the converted wave image is slightly lower quality than the compressional wave image.

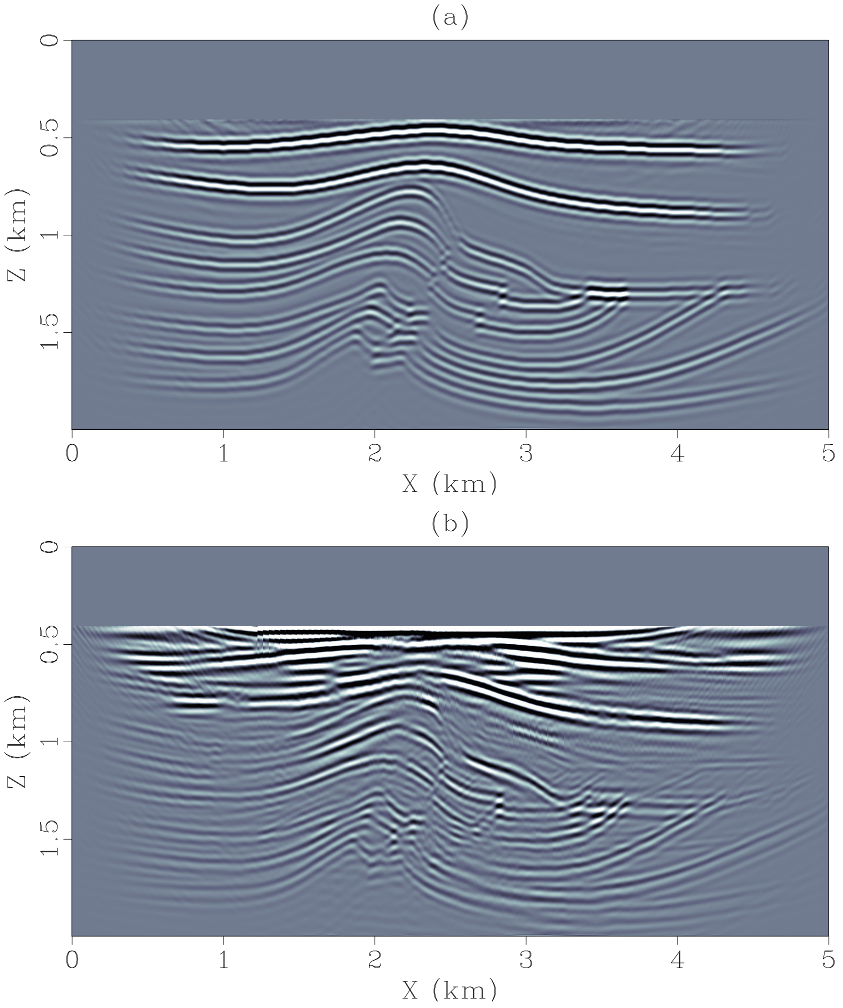


Figure 3.8: Vector image results for 96 shot-profile elastic data set. The images have been k_z filtered to remove low wavenumber artifacts, and muted above the receivers. (a) Image from compressional wave modes $\mathrm{I}^{PP}(\mathbf{x})$. (b) Image from converted wave modes $\mathrm{I}^{PS}(\mathbf{x})$.

3.7 Summary

In this chapter, we investigate a reverse time migration algorithm for two-dimensional prestack data. An acoustic data set was used to outline the two stage procedure. The staggered-grid finite-difference algorithm from the previous chapter was used to extrapolate the source side and receiver side wave-fields. The zero-lag cross-correlation of the two corresponding numerical wave-fields was then implemented producing an image of subsurface scattering locations. To account for data with both primary and converted wave modes the source side and receiver side wave-fields were decomposed into their scalar and vector potentials. The vector imaging condition of Yan and Sava (2008) was then used generated images for both primary (PP) and converted (PS) wave modes.

CHAPTER 4

Simultaneous source reverse-time migration

4.1 Introduction

The desire to image complex subsurface structure has propelled the advancement of acquisition and imaging technologies. This has resulted in larger volumes of data and the implementation of more expensive imaging algorithms. In efforts to increase economic efficiency multiple shot recordings can be blended together prior to migration, reducing the computational cost of the migration by the number of combined shots. However, the final image is degraded by crosstalk interferences generated from the cross-correlation of unrelated source side and receiver side wave-fields (Romero et al., 2000). A number of techniques referred to as shot encodings attempt to remedy the degradation of the final image, by blending shots such that the crosstalk is dispersed during the imaging process. Taking the mean of many blended images attenuates the crosstalk noise as the incoherent crosstalk interferences stack out, while the coherent reflector positions are reinforced. Godwin and Sava (2013) offer a detailed comparison and review of many shot-encoding methods. It is shown blending shots encoded with random time delays, and decimating the shot record prior to migration produce results with the lowest misfit to the conventional image.

The Maximum likelihood estimator of location (M-estimator) technique is a method to estimate a constant value from within a set of repeatable observations contaminated by noise. The lowest risk M-estimator, the mean, attenuates the amplitude of Gaussian noise by the square root of the number of observations (Ulrych et al., 1999). However, if the noise is not Gaussian the mean will not be as efficient at attenuating the noise. Trickett et al. (2007) Outlines the use of robust M-estimators to attenuate erratic noise during common midpoint stacking. This chapter will investigate the use of robust M-estimators to attenuate the crosstalk artifacts generated from blended migration.

4.2 Shot encoding

Shot encoding simply refers to a scheme which blends shot profiles together to form a super shot. Many different methods of shot encoding have been proposed including plane-wave (Tieman, 1997), random , chirp, linear (Romero et al., 2000), and opposite sign (Sun and Zhang, 1999). The goal of all encoding schemes is to reduce the crosstalk artifacts within the final migrated images by dispersing the locations where they occur within the migration of each super shot. Thus, stacking many images (taking the sample mean) will attenuate the artifacts as the true image will be coherent and stack together while the crosstalk artifacts will be incoherent and stack out (Romero et al., 2000). Romero et al. (2000) and Godwin and Sava (2013) show that blending shots encoded with random time delays produces the lowest misfit between the conventional and blended image for significant amounts of blending.

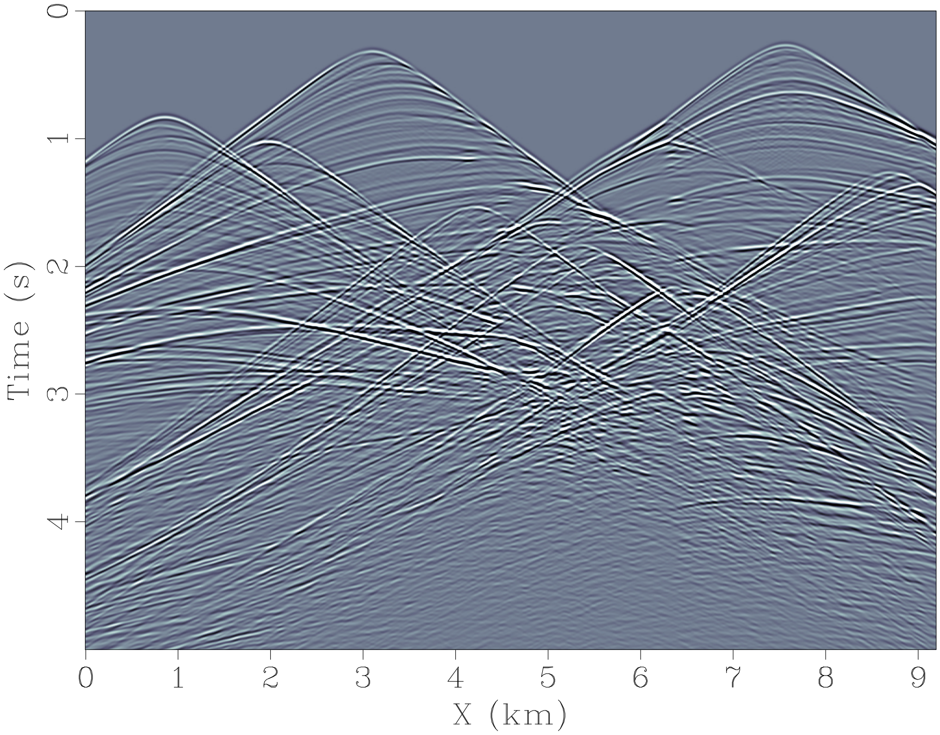


Figure 4.1: Example of random shot encoding. 8 equally spaced synthetic acoustic Marmousi shot profiles blended together with random time delays to form a super-shot.

4.3 Simultaneous source reverse-time migration

Blended migration differs from conventional migration in that shot recordings are combined into super-shots prior to wave-field extrapolation. Consequentially, corresponding source side wave-fields must contain the extrapolation of all source functions associated with the blended super-shot. This reduces the computation required by the number of shots combined. Figure 4.2 illustrates the process of extrapolating the blended wave-fields through a smooth background Marmousi model.

The cross-correlation imaging condition for a blended migration can be written as a double sum over the number of shots that compose a single super-shot.

$$\mathrm{I}^B(\mathbf{x}) = \frac{1}{Nb}\sum_{i}^{Nb}\sum_{j}^{Nb}\sum_{t}^{Nt} S_i(\mathbf{x},t)R_j(\mathbf{x},t) \tag{4.1}$$

Here i and j are the indices of the source functions and shot records that are extrapolated to produce the blended source and receiver side wave-fields $\sum_i^{Nb} S_i(\mathbf{x},t)$ and $\sum_j^{Nb} R_j(\mathbf{x},t)$. The blended migration $\mathrm{I}^B(\mathbf{x})$, can now be expressed in terms of the diagonal and off-diagonal indicies of the double sum.

$$\mathrm{I}^B(\mathbf{x}) = \frac{1}{Nb}\sum_{i}^{Nb}\sum_{j}^{Nb}\left(\delta_{ij}\hat{m}_{ij}(\mathbf{x}) + (1-\delta_{ij})\hat{m}_{ij}(\mathbf{x})\right) \tag{4.2}$$

Above, $\delta_{ij}\hat{m}_{ij}(\mathbf{x})$ denotes cross-correlation of the diagonal indices of the blended source side and receiver side wave-fields and $(1-\delta_{ij})\hat{m}_{ij}(\mathbf{x})$ the off-diagonal indicies. The blended migration in equation (4.2) can be re-written as the diagonal image component $\mathrm{I}^D(\mathbf{x})$, plus the cross-correlation of the off-diagonal wave-field components $\mathrm{I}^X(\mathbf{x})$. These off-diagonal components, are referred to as crosstalk. Cross terms between unrelated source side and receiver side wave-field terms that produce strong interferences in the migrated image.

$$\mathrm{I}^B(\mathbf{x}) = \mathrm{I}^D(\mathbf{x}) + \mathrm{I}^X(\mathbf{x}) \tag{4.3}$$

Figure 4.3 depicts the components of the blended migration from equation 4.3. The subscript m is used as an indice for multiple blended migrations.

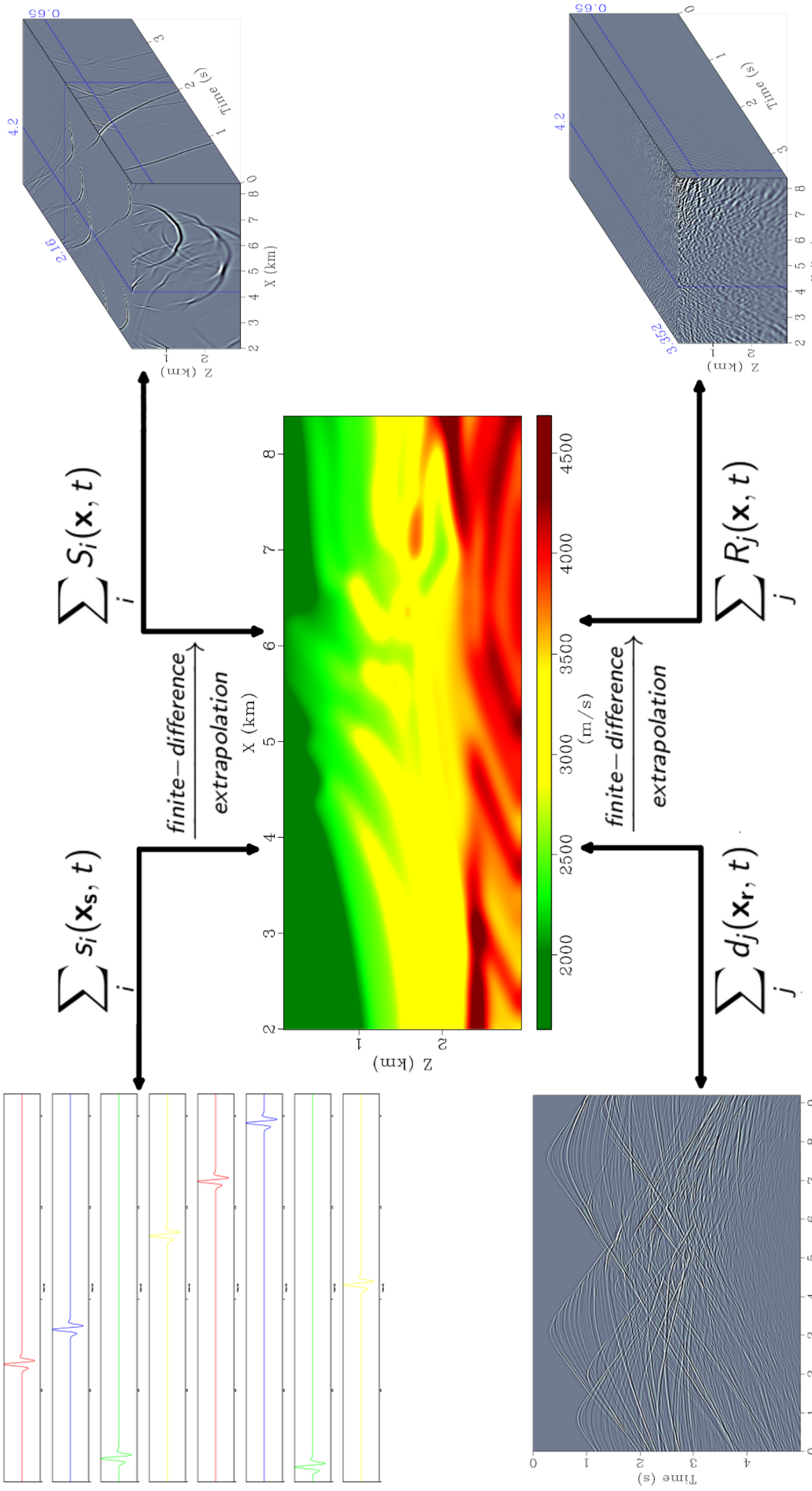
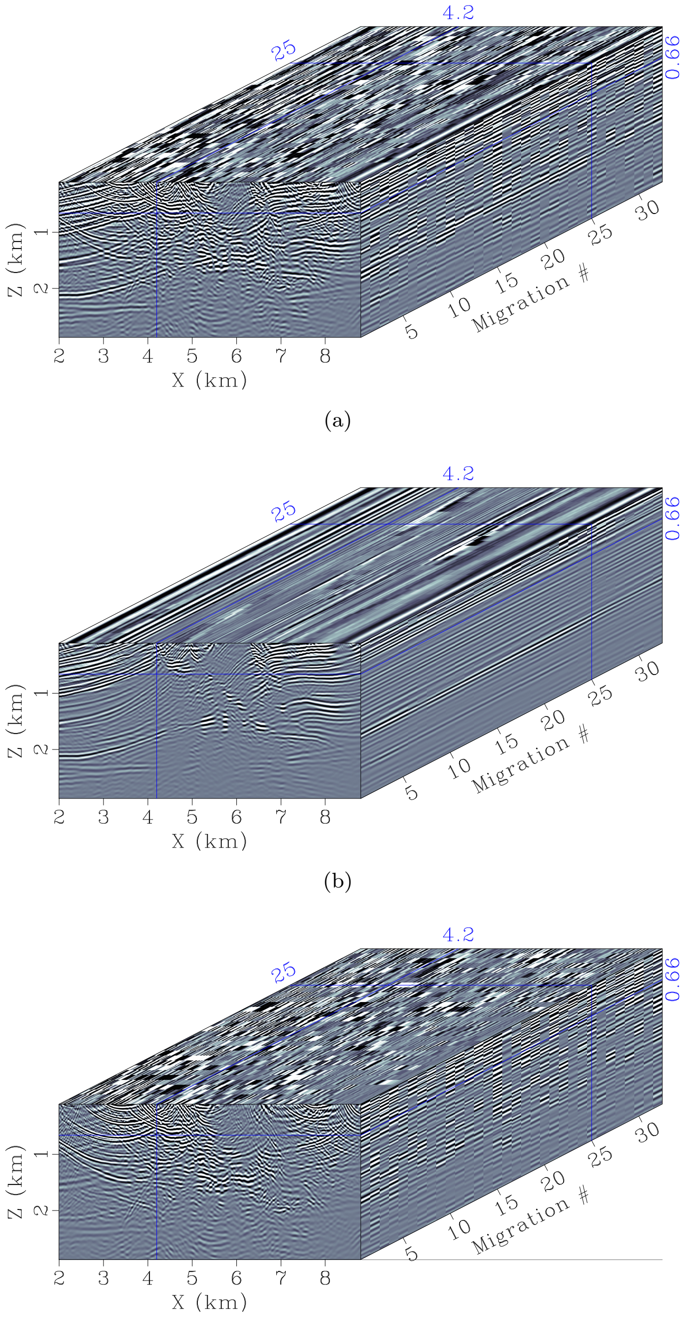


Figure 4.2: Depiction of wave-field extrapolation for a super-shot composed of 8 random phase encoded synthetic Marmousi shot-profiles via two-dimensional finite-difference. The smooth Marmousi model shown was used for wave-field extrapolation



(a)

(b)

(c)

Figure 4.3: (a) Cube plot depicting 32 migrations of 8 blended shots I_m^B. (b) The diagonal component of each blended migration $\mathrm{I}_m^D(\mathbf{x})$. (c) The crosstalk contamination present in each blended migration $\mathrm{I}_m^X(\mathbf{x})$. Intersecting blue lines project cross-sections onto faces of the cube parallel with the cross-section.

4.4 Shot decimated migration

A more trivial method to reduce the computational cost is to simply migrate fewer shots. Shot decimated migration is achieved doing a conventional migration using a subset of shots from a data set $s \subseteq k$. The subset selection is controlled by a decimation factor α, such that every α^{th} shot-record is included in the subset ($s \ni k = n\alpha$).

$$\mathrm{I}^{S}(\mathbf{x}) = \frac{1}{Ns}\sum_{s}^{Ns} P_s(\mathbf{x}) = \frac{1}{Ns}\sum_{s}^{Ns}\sum_{t}^{Nt} S_s(\mathbf{x},t)R_s(\mathbf{x},t) \tag{4.4}$$

Godwin and Sava (2013) outline that shot decimated migration has several advantages over blended migration. Firstly, it is more simplistic as blending shot-records is not required. Secondly, it is cheaper than blended imaging for the same number of migrations as a smaller migration aperture can be used. However, decimating the shot record creates the risk of inadequate illumination of the subsurface region of interest.

4.5 M-estimators

The Maximum likelihood estimator of location (M-estimator) technique is a method in robust statistic commonly used to estimate an unknown parameter within a set of observations. Consider a data set of N repeatable observations $[x_1, x_2, ...x_N]$, where x_i is contaminated by noise such that $x_i = \mu + \epsilon_i$. The problem is to estimate μ the true observation in the presence of noise ϵ_i. The maximum likelihood estimation of location of μ within the set of observation can be determined if a priori about the distribution of ϵ_i is known.

$$\hat{\mu} = \arg\min_{\mu} \sum_{i}^{N} \rho(x_i - \mu) \tag{4.5}$$

The problem can be formulated as a cost function where the objective is to find the value of μ such that the sum of its residuals with the observations best fits the assumed distribution of the noise. This can be executed by introducing a loss function $\rho(x)$, that assigns weights to the residuals according to the probability of there occurrence within the distribution $f(x)$. The loss function is computed as the natural logarithm of the probability density function $\rho(x) = -\log f(x)$. The solution to equation (4.5) is what is referred to as the maximum likelihood type estimator of location or M-estimator (Huber, 1964).

The first two entries in Table 1 shows the basic M-estimators, the sample mean and sample median. The sample mean is the M-estimator generated from a Gaussian density function and the sample median from a Laplace density function. The third entry in Table 1 is a more dynamic M-estimator known as the sample myriad. The sample myriad is the M-estimator generated by a Cauchy density function with a tunable linearity parameter γ. Geometrically γ is equivalent to half the inner quartile range of the Cauchy density function. This allows the sample myriad to be tuned to have increased resistance to outliers (small γ) or increased efficiency in Gaussian environments (large γ) (Gonzalez and Arce, 2002).

Cost Function	Estimator
$\arg\min\limits_{\mu} \sum_i^N (x_i - \mu)^2$	$mean\left\{x_i\|_{i=1}^N\right\}$
$\arg\min\limits_{\mu} \sum_i^N \|x_i - \mu\|$	$median\left\{x_i\|_{i=1}^N\right\}$
$\arg\min\limits_{\mu} \sum_i^N \log\left[\gamma^2 + (x_i - \mu)^2\right]$	$myriad\left\{x_i\|_{i=1}^N; \gamma\right\}$

Table 4.1: Cost-functions and M-estimators.

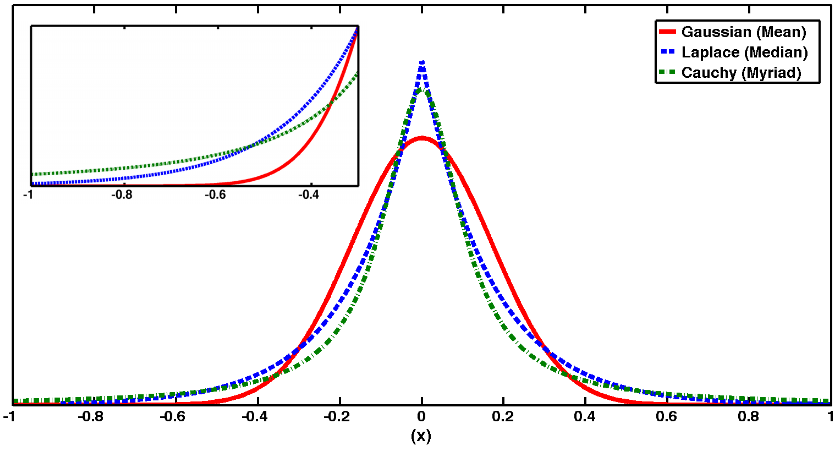
Figure 4.4 illustrates density and loss functions for the three estimators shown in table 4.1. The inner quartile ranges of all density functions are equal. It can be observed the Gaussian estimator is not robust as large residual values will be penalized harshly by its loss function. Thus, outliers or values located within the tails of the Gaussian density function will severely influence the cost function minima. The Laplace and Cauchy functions are more robust as the tails of their density functions are heavier or higher valued. This results in loss function penalties on large residuals to have a balanced influence on the minimum of their cost functions.

It can be seen from Figure 4.3 that attenuating crosstalk in blended migration can be formulated as a problem of parameter estimation, where the true subsurface image becomes an unknown parameter within a set of crosstalk contaminated observations. Equation 4.6 depicts the maximum likelihood image estimates from a set of blended images $\mathrm{I}^{MLE}(\mathbf{x})$ for the difference M-estimators (MLE) in table 4.1.

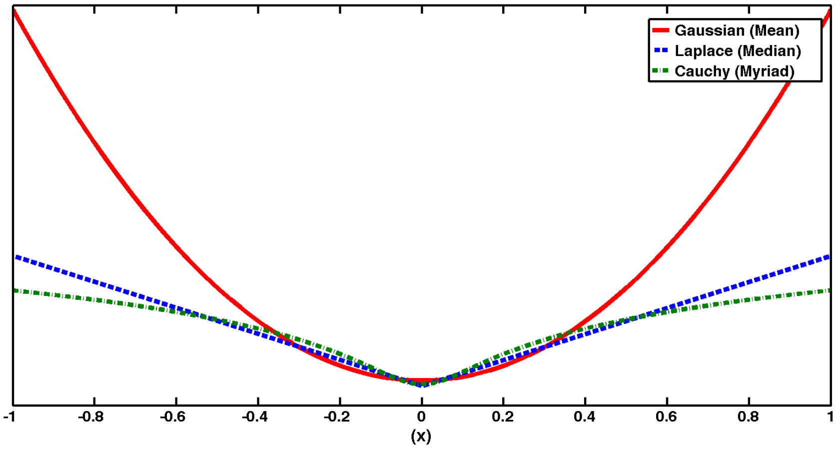
$$\mathrm{I}^{MLE}(\mathbf{x}) = MLE\left\{\mathrm{I}_m^B(\mathbf{x})|_{m=1}^{Nm}\right\} \tag{4.6}$$

4.6 Examples

The method was applied to a synthetic Marmousi data set containing 256 acoustic shot profiles. Figure 4.5 shows the Marmousi model, and Figure 4.6 the conventional unnormal-



(a)



(b)

Figure 4.4: (a) Probability density functions $f(x)$ with equal inner-quartile ranges. A zoomed in view of the tail information is shown. It can be observed the tails of the Laplace and Cauchy density functions are heavier than that of the Gaussian function. (b) Corresponding loss functions $\rho(x) = -\log f(x)$, the heavy tails of the Laplace and Cauchy functions result in smaller loss function penalties to large residual values.

ized scalar migration result. A blended data set was produced from the 256 shot profiles. Maximally separated shots were blended and given a random time delay from the uniform distribution $U(0\,s, 2\,s)$. Each shot record was used exactly once in producing a blending image set. The estimators abilities to resolve the true image were compared, calculating the estimated images $\mathrm{I}^{MLE}(\mathbf{x})$ relative L_2 norm difference with the crosstalk free conventional image $I^C(\mathbf{x})$ (4.7). Shot decimated migration results $\mathrm{I}^S(\mathbf{x})$ were compared to the blended migration results in the same way, setting the decimation factor α equal to the number of shots per super-shot such that the same number of migrations were performed.

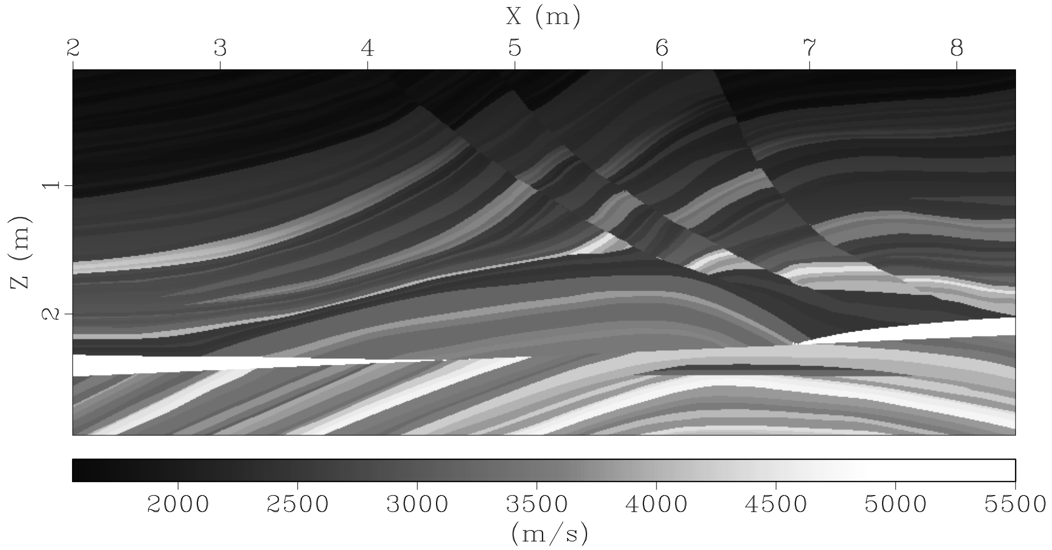


Figure 4.5: Marmousi velocity model [$nx = 4602 \quad nz = 1922 \quad \Delta x = 2\,m \quad \Delta z = 2\,m$], and density $1000 - 2150\,kg/m^3$. Compressional velocities V_p shown. Shearing velocities are set to zero to make modelling acoustic.

$$\frac{\|\mathrm{I}^C(\mathbf{x}) - \mathrm{I}^{MLE}(\mathbf{x})\|}{\|\mathrm{I}^C(\mathbf{x})\|} \quad and, \quad \frac{\|\mathrm{I}^C(\mathbf{x}) - \mathrm{I}^S(\mathbf{x})\|}{\|\mathrm{I}^C(\mathbf{x})\|} \tag{4.7}$$

Figure 4.7 shows the results for the different M-estimators and shot decimated migration. Shot decimated migration has the smallest relative L_2 norm difference for 4 blended shots per migration/64 migrations and 8 blended shots per migration/32 migrations. However, the myriad estimator performs the best for 16 blended shots per migration/16 migrations and is about a ten percent improvement over the mean and median estimators. The results for 16 blended shots per migration/16 migrations is shown by Figure 4.8. It can be observed that although the relative L_2 norm difference of the myriad estimate is very similar to that of the shot decimated migration the image residuals are dissimilar. The artifacts present in the myriad estimate are dispersed throughout the image, where as the residuals in the

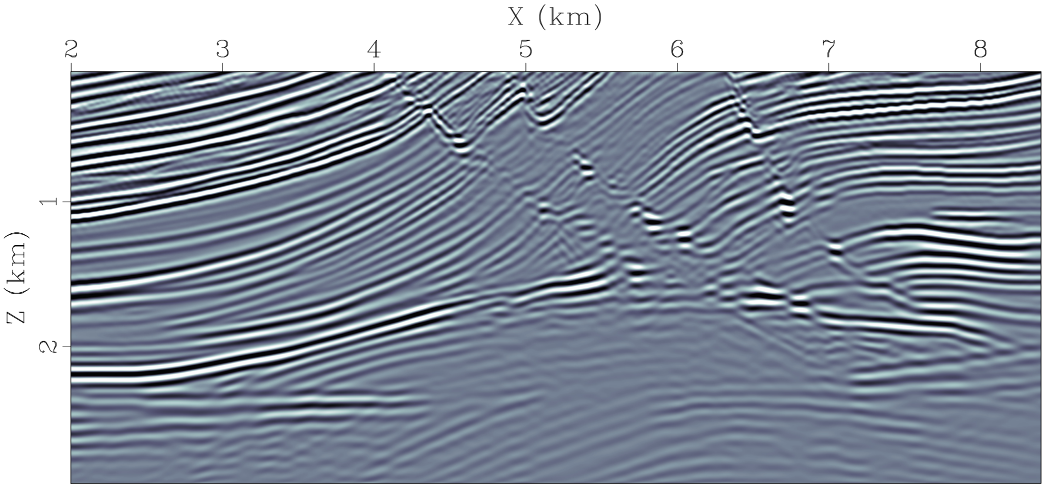


Figure 4.6: Conventional scalar migration result of Marmousi model using 256 shot profiles.

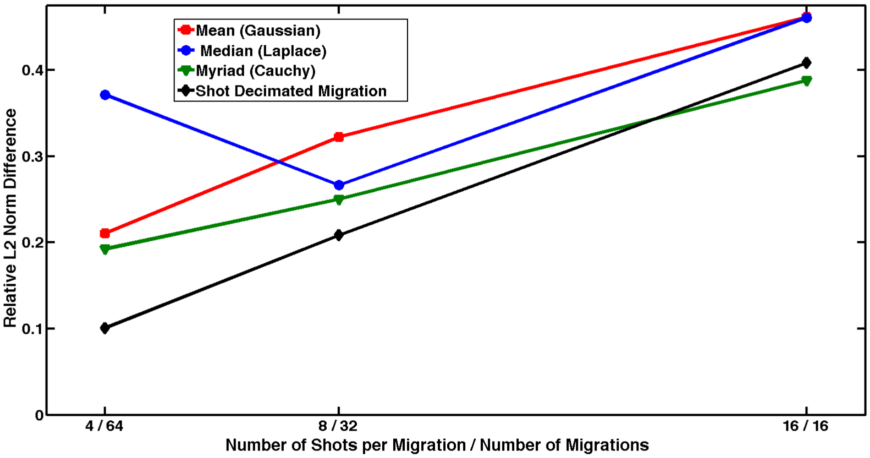
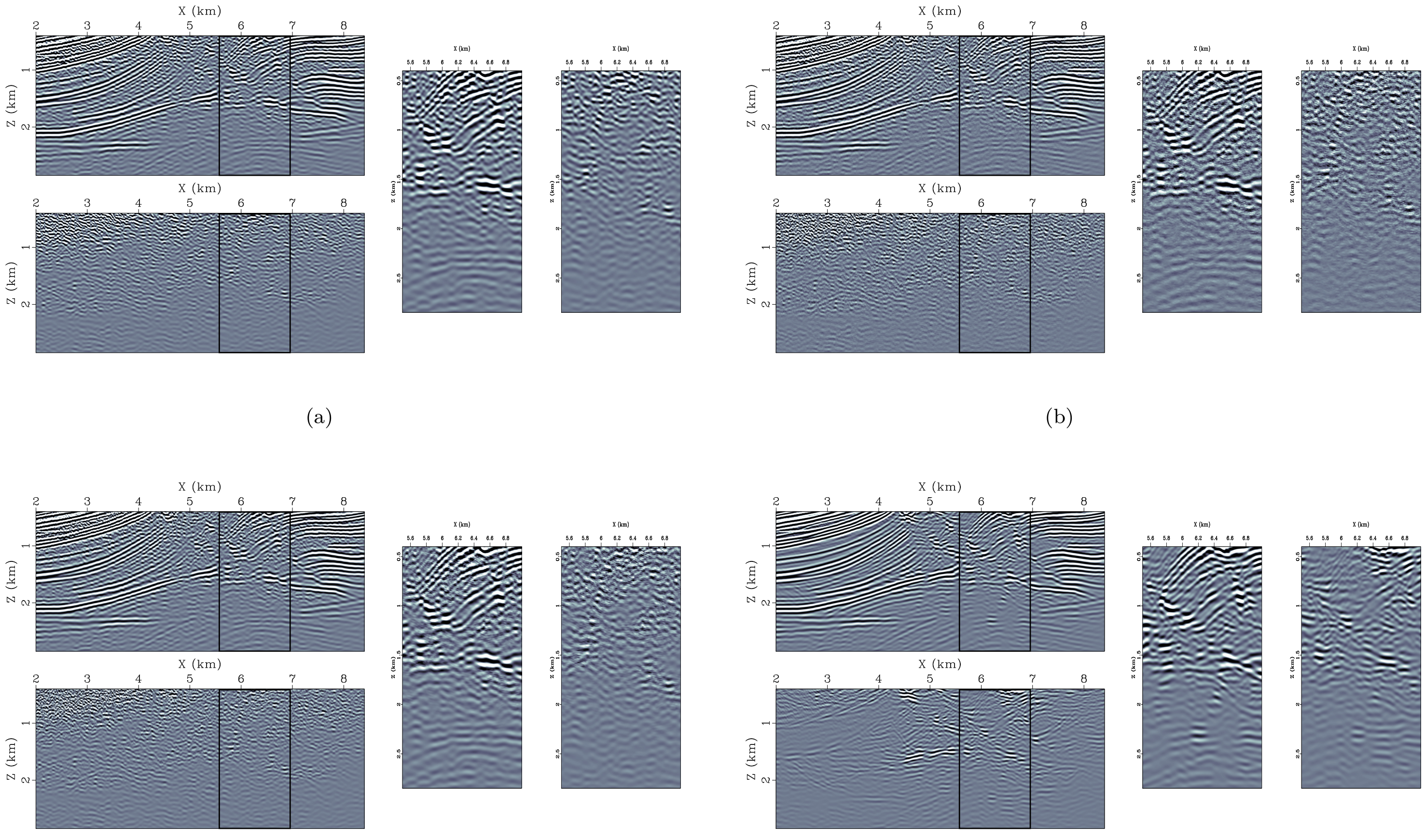


Figure 4.7: Relative L_2 norm difference for different M-estimators and shot decimated migration.

shot decimated image are localized to poorly illuminated regions in the image. For a full migration aperture synthetic the myriad estimator with γ correctly selected is a viable alternative to shot-decimated migration for both image quality and computational expense. However, in real world scenarios the shot-decimated migration could be performed at a significantly reduced cost in comparison to the blended migration. This is because for large models the migration aperture of individual shot gathers can be limited reducing the model size needed for each wave-field extrapolation.

4.7 Summary

Blending shot records prior to wave-field extrapolation reduces the amount of time required to migrate a data set. However, this comes at the expense of a degraded final image due to crosstalk generated from the cross-correlation of unrelated source side and receiver side wave-fields. It has been shown that attenuating crosstalk in blended migration can be formulated as a parameter estimation problem. A comparison of image estimates from the sample mean, median, and myriad demonstrates that robust M-estimators can be more efficient at attenuating crosstalk noise from a set of blended images than Gaussian estimators. For larger data sets the advantage of being able to limit the migration aperture of individual shot profiles, ultimately make shot-decimated migration more efficient then blended migration. Future research could be conducted to determine the effect of pairing different shot encoding schemes and M-estimators.



(a)

(b)

(c)

(d)

Figure 4.8: Image estimate (top left), residual with conventional image (bottom left), zoomed view from within boxed regions (right) for 16 shots per migration/number of migrations. For (a) mean, (b) median (c) myriad ($\gamma = 0.055$), and (d) shot decimated migration. Notice part of the reflector at $2.4\,km$ in the zoomed region is not resolved by shot decimated migration

CHAPTER 5

Conclusions

5.1 Thesis summary

Seismic migration is the process of generating non-destructive structural images of the subsurface from seismic data. This is done by repositioning events in the seismic data to there scattering locations in the subsurface. Seismic migration algorithms consist of two stages, a wave-field extrapolation, and an imaging condition (Berkhout, 1981). The wave-field extrapolation requires a method to describe seismic wave propagation, the imaging condition then exploits our physical understanding of wave-field scattering to locate subsurface interfaces. Thus, the accuracy of an imaging algorithm depends largely on the ability to realistically model seismic wave propagation. Chapter 2 of this thesis outlines a elastic staggered-grid finite-difference algorithm. The method is capable of accurately modelling both compressional and shear wave propagation within an isotropic elastic medium.

The focus of this thesis is the shot profile reverse-time migration technique. Shot profile migrations individually migrate the recordings from each source function to produce a partial image of the subsurface. The partial images only resolve the region of the subsurface illuminated by a single source function. Stacking partial images produces an image resolving the region illuminated by all of the individually migrated shots. Chapter 3 outlines a shot profile reverse-time migration algorithm. The method uses the staggered-grid finite-difference algorithm from chapter 2 to act as the wave-field extrapolation engine. The zero lag cross-correlation imaging condition of Claerbout (1971) extracts images of subsurface interfaces from reconstructed source and receiver side wave-fields. For acoustic data, the imaging condition can be implemented as the zero lag time cross-correlation of two numerically modelled scalar wave-fields. However, for data consisting of both compressional and shear wave modes the reconstructed wave-fields are separated into scalar and vector potentials prior to cross-correlation (Yan and Sava, 2008). This results in two images, one from repositioning compressional wave modes, and another from repositioning converted wave modes.

The bulk of the computational cost in the reverse time migration algorithm is due to the two wave-field extrapolations that must be performed for each shot profile. Shot encoding, blending shot records together prior to the wave-field extrapolation stage of the algorithm can reduces the computational cost by the number of combined shows. However, the nature of migrating blended data is double-edged whereby reduction in imaging cost compromises the quality of the final migrated images. This compromise comes in the form of non-physical crosstalk artifacts; interferences in the migrated images from the algorithms inability to distinguish between individual shot records that compose blended data. Maximum likelihood type estimators of location or M-estimators, use robust statistic to estimate constant amplitude signals from within noise contaminated data sets. Chapter 4 formulates the attenuation of the aforementioned crosstalk from a set of blended seismic images as an M-estimation problem. The mean, median, and myriad estimators ability to resolve the true image from a set of crosstalk contaminated images are compared using a synthetic Marmousi data set. In addition, we compare errors introduced from decimating the shot record prior to migration to those introduced from blended migration.

Future research into the reverse-time migration method could involve incorporating the algorithm into an iterative scheme. Chapter 3 outlines the adjoint reverse-time migration operator, which uses data to produce structural images. A forward reverse-time migration operator could be formulated which would produce data from structural images. Evaluating data space residuals from the forward operator would quantitatively determine how well an adjoint image fits a given data set. A forward/adjoint migration pair could therefore be used to perform a least squares reverse-time migration by iteratively updating an image to minimize the data residual.

Bibliography

Aki, K. and P. G. Richards. *Quantitative Seismology: Theory and Methods.* Geology Series. W. H. Freeman, 1980.

Aki, K. and P. G. Richards. *Quantitative Seismology.* Geology (University Science Books).: Seismology. University Science Books, 2002.

Alford, R. M., K. R. Kelly, and D. M. Boore. "Accuracy of finite-difference modeling of the acoustic wave equation." *Geophysics* 39 (1974): 834–842.

Balch, A. H. and C. Erdemir. "Sign-change correction for prestack migration of P-S converted wave reflections." *Geophysical Prospecting* 42 (1994): 637–663.

Baysal, E., D. Kosloff, and J. Sherwood. "Reverse time migration." *Geophysics* 48 (1983): 1514–1524.

Berkhout, A. J. "Wave field extrapolation techniques in seismic migration, a tutorial." *Geophysics* 46 (1981): 1638–1656.

Beylkin, G. and R. Burridge. "Linearized inverse scattering problems in acoustics and elasticity." *Wave motion* 12 (1990): 15–52.

Boore, D. M. "Finite difference methods for seismic wave propagation in heterogeneous materials." *Methods in computational physics* 11 (1972): 1–37.

Botelho, M. A. and P. L. Stoffa. "Finite-Difference Prestack Reverse time Migration Using the P-SV Wave Equation." *1991 SEG Annual Meeting.* 1991.

Chang, W. and G. A. McMechan. "Elastic reverse-time migration." *Geophysics* 52 (1987): 1365–1375.

Chang, W. and G. A. McMechan. "3-D elastic prestack, reversetime depth migration." *Geophysics* 59 (1994): 597–609.

Chew, W. C. and Q. H. Liu. “Perfectly matched layers for elastodynamics: A new absorbing boundary condition.” *Journal of Computational Acoustics* 4 (1996): 341–359.

Claerbout, J. F. “Toward a unified theory of reflector mapping.” *Geophysics* 36 (1971): 467–481.

Collino, F. and C. Tsogka. “Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media.” *Geophysics* 66 (2001): 294–307.

Coutant, O., J. Virieux, and A. Zollo. “Numerical source implementation in a 2D finite difference scheme for wave propagation.” *Bulletin of the Seismological Society of America* 85 (1995): 1507–1512.

Dablain, M. “The application of highorder differencing to the scalar wave equation.” *Geophysics* 51 (1986): 54–66.

Dellinger, J. and J. Etgen. “Wave-field separation in two-dimensional anisotropic media.” *Geophysics* 55 (1990): 914–919.

Duan, Y. and P. Sava. “Converted-waves imaging condition for elastic reverse-time migration.” *2014 SEG Annual Meeting*. 2014.

Gazdag, J. “Wave equation migration with the phase-shift method.” *Geophysics* 43 (1978): 1342–1351.

Gazdag, J. and P. Sguazzero. “Migration of seismic data by phase shift plus interpolation.” *Geophysics* 49 (1984): 124–131.

Godwin, J. and P. Sava. “A comparison of shot-encoding schemes for wave-equation migration.” *Geophysical Prospecting* 61 (2013): 391–408.

Godwin, J., P. Sava, et al. “A comparison of shot-encoding schemes for wave-equation migration.” *2011 SEG Annual Meeting*. 2011.

Gonzalez, J. G. and G. R. Arce. “Statistically-efficient Filtering in Impulsive Environments: Weighted Myriad Filters.” *EURASIP J. Appl. Signal Process.* 2002 (2002): 4–20.

Goodway, W. “AVO and Lamé constants for rock parameterization and fluid detection.” *CSEG Recorder* 26 (2001): 39–60.

Graves, R. W. “Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences.” *Bulletin of the Seismological Society of America* 86 (1996): 1091–1106.

Guitton, A., B. Kaelin, and B. Biondi. “Least-squares attenuation of reverse-time-migration artifacts.” *Geophysics* 72 (2006): S19–S23.

Hill, N. R. “Gaussian beam migration.” *Geophysics* 55 (1990): 1416–1428.

Huber, P. J. “Robust Estimation of a Location Parameter.” *The Annals of Mathematical Statistics* 35 (1964): 73–101.

Hustedt, B., S. Operto, and J. Virieux. “Mixed-grid and staggered-grid finite-difference methods for frequency-domain acoustic wave modelling.” *Geophysical Journal International* 157 (2004): 1269–1296.

Jiang, Z. *Elastic wave modelling and reverse-time migration by a staggered-grid finite-difference method*. PhD thesis, University of Calgary, 2012.

Jones, I. “Tutorial: migration imaging conditions.” *first break* 32 (2014): 45–55.

Kaelin, B., A. Guitton, et al. “Imaging condition for reverse time migration.” *2006 SEG Annual Meeting*. 2006.

Kelly, K., R. Ward, S. Treitel, and R. Alford. “Synthetic Seismograms: a finite-difference approach.” *Geophysics* 41 (1976): 2–27.

Kristek, J., P. Moczo, and R.J. Archuleta. “Efficient Methods to Simulate Planar Free Surface in the 3D 4th-Order Staggered-Grid Finite-Difference Schemes.” *Studia Geophysica et Geodaetica* 46 (2002): 355–381.

Levander, A. R. “Fourth-order finite-difference P-SV seismograms.” *Geophysics* 53 (1988): 1425–1436.

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

Marcinkovich, C. and K. Olsen. “On the implementation of perfectly matched layers in a three-dimensional fourth-order velocity-stress finite difference scheme.” *Journal of Geophysical Research: Solid Earth (1978–2012)* 108 (2003).

Mathews, J. H. and K. D. Fink. *Numerical Methods Using MATLAB*. Pearson Education, Limited, 2006.

McMechan, G. A. “Migration by extrapolation of time-dependent boundary values.” *Geophysical Prospecting* 31 (1983): 413–420.

Mitchell, A. R. *Computational methods in partial differential equations*. Introductory mathematics for scientists and engineers. J. Wiley, 1969.

Mittet, R. “Free-surface boundary conditions for elastic staggered-grid modeling schemes.” *Geophysics* 67 (2002): 1616–1623.

Nougier, J. *Méthodes de calcul numérique*. Paris: Masson, 1987.

Pujol, J. *Elastic Wave Propagation and Generation in Seismology*. Cambridge University Press, 2003.

Romero, L., D. Ghiglia, C. Ober, and S. Morton. “Phase encoding of shot records in prestack migration.” *Geophysics* 65 (2000): 426–436.

Schneider, J. B., C. L. Wagner, and S. L. Broschat. “Implementation of transparent sources embedded in acoustic finite-difference time-domain grids.” *The Journal of the Acoustical Society of America* 103 (1998): 136–142.

Schuster, G. T. et al. “Reverse-time migration= generalized diffraction stack migration.” *SEG Technical Program Expanded Abstracts*. 2002, 1280–1283.

Sherwood, J. W. C. “Elastic wave propagation in a semi-infinite solid medium.” *Proceedings of the Physical Society* 71 (1958): 207.

Sherwood, J. W. C., F. J. Hilterman, R. N. Neale, K. C. Chen, et al. “Synthetic seismograms with offset for a layered elastic medium.” *Offshore Technology Conference*. 1983.

Shuey, R. T. “A simplification of the Zoeppritz equations.” *Geophysics* 50 (1985): 609–614.

Stoffa, P. L., J. T. Fokkema, R. M. de Luna Freire, and W. P. Kessinger. “Split-step Fourier migration.” *Geophysics* 55 (1990): 410–421.

Sun, P. and S. Zhang. “Prestack Migration of Areal Shot Records with Mix Phase Encoding.” *1999 SEG Annual Meeting*. 1999.

Sun, R. and G. A. McMechan. “Scalar reverse-time depth migration of prestack elastic seismic data.” *Geophysics* 66 (2001): 1519–1527.

Symes, W. “Reverse time migration with optimal checkpointing.” *Geophysics* 72 (2007): SM213–SM221.

Taylor, J. R. *Classical Mechanics*. University Science Books, 2005.

Tieman, H. J. “Improving plane-wave decomposition and migration.” *Geophysics* 62 (1997): 195–205.

Trefethen, L. N. “Group velocity in finite difference schemes.” *SIAM review* 24 (1982): 113–136.

Trefethen, L. N. *Finite Difference and Spectral Methods for Ordinary and Partial Differential Equations.* unpublished text, 1996.

Trickett, S. et al. "Maximum-likelihood-estimation stacking." *77th Annual International Meeting.* 2007, 2640–2643.

Ulrych, T. J., M. D. Sacchi, and J. M. Graul. "Signal and noise separation: Art and science." *Geophysics* 64 (1999): 1648–1656.

Virieux, J. "SH-wave propagation in heterogeneous media: Velocity-stress finite-difference method." *Geophysics* 49 (1984): 1933–1942.

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

Whitmore, N. D. "Iterative Depth Migration By Backward Time Propagation." *1983 SEG Annual Meeting, September 11 - 15, 1983 , Las Vegas, Nevada.* 1983.

Wiggins, J. W. "Kirchhoff integral extrapolation and migration of nonplanar data." *Geophysics* 49 (1984): 1239–1248.

Yan, J. and P. Sava. "Isotropic angle-domain elastic reverse-time migration." *Geophysics* 73 (2008): S229–S239.

Yao, G. *Least-Squares Reverse-Time Migration.* PhD thesis, Imperial College London, 2013.

Yee, K. S. et al. "Numerical solution of initial boundary value problems involving Maxwells equations in isotropic media." *IEEE Trans. Antennas Propag* 14 (1966): 302–307.

APPENDIX A

Finite-difference

This appendix will outline the proofs for the second-order $O(h^2)$ and fourth-order $O(h^4)$ central finite-difference approximations to the first derivative. The proofs are similar to that of Mathews and Fink (2006) but formulated for the staggered-grid scheme. They are deduced from Taylor expansion analysis about x, for a given continuous function $f(x)$ (Nougier, 1987).

A.1 Second-order finite-difference approximation to the first derivative $O(h^2)$

The proof starts with the second-degree Taylor expansion about x, for $f(x + \frac{1}{2}h)$ and $f(x - \frac{1}{2}h)$. The last term in the expansion is the Lagrange remainder.

$$f(x + \frac{1}{2}h) = f(x) + f'(x)\frac{h}{2} + \frac{f''(x)h^2}{4(2!)} + \frac{f'''(x^*)h^3}{8(3!)} \tag{A.1}$$

$$f(x - \frac{1}{2}h) = f(x) - f'(x)\frac{h}{2} + \frac{f''(x)h^2}{4(2!)} - \frac{f'''(x^*)h^3}{8(3!)} \tag{A.2}$$

Subtraction of equation A.1 with A.2 results in A.3.

$$f(x + \frac{1}{2}h) - f(x - \frac{1}{2}h) = f'(x)h + \frac{f'''(x^*)h^3}{4(3!)} \tag{A.3}$$

Thus the first order derivative can be approximated discretely about x as A.4. With the first term on the right hand side being the second-order central finite-difference operator and the second term the truncation error.

$$f'(x) = \frac{f(x + \frac{1}{2}h) - f(x - \frac{1}{2}h)}{h} - \frac{f'''(x^*)h^2}{24} \tag{A.4}$$

The truncation error will go to zero at the same rate as h^2 expressed as $O(h^2)$, with h being the full node spacing of the discrete mesh.

A.2 Fourth-order finite-difference approximation to the first derivative $O(h^4)$

This proof starts with the fourth-degree Taylor expansion about x, for $f(x + \frac{1}{2}h)$ and $f(x - \frac{1}{2}h)$.

$$f(x + \frac{1}{2}h) = f(x) + f'(x)\frac{h}{2} + \frac{f''(x)h^2}{4(2!)} + \frac{f'''(x)h^3}{8(3!)} + \frac{f^{(4)}(x)h^4}{16(4!)} + \frac{f^{(5)}(x^*)h^5}{32(5!)} \tag{A.5}$$

$$f(x - \frac{1}{2}h) = f(x) - f'(x)\frac{h}{2} + \frac{f''(x)h^2}{4(2!)} - \frac{f'''(x)h^3}{8(3!)} + \frac{f^{(4)}(x)h^4}{16(4!)} - \frac{f^{(5)}(x^*)h^5}{32(5!)} \tag{A.6}$$

Subtraction of equation A.5 with A.6 results in A.7.

$$f(x + \frac{1}{2}h) - f(x - \frac{1}{2}h) = f'(x)h + \frac{f'''(x)h^3}{4(3!)} + \frac{f^{(5)}(x^*)h^5}{16(5!)} \tag{A.7}$$

Next increase the step size to $\frac{3}{2}h$ and do the same as above.

$$f(x + \frac{3}{2}h) - f(x - \frac{3}{2}h) = 3f'(x)h + \frac{27f'''(x)h^3}{4(3!)} + \frac{243f^{(5)}(x^*)h^5}{16(5!)} \tag{A.8}$$

Now multiple equation A.7 by 27 and subtract equation A.8 from it, eliminating the $f'''(x)$ terms.

$$-f(x + \frac{3}{2}h) + 27f(x + \frac{1}{2}h) - 27f(x - \frac{1}{2}h) + f(x - \frac{3}{2}h) = 24f'(x)h - \frac{216f^{(5)}(x^*)h^5}{16(5!)} \tag{A.9}$$

Thus the first order derivative can be approximated discretely about x as A.10. With the first term on the right hand side being the fourth-order central finite-difference operator with $C_1 = \frac{9}{8}$ and $C_2 = \frac{1}{24}$ and the second term the truncation error.

$$f'(x) = \frac{-C_2\left(f(x+\frac{3}{2}h) - f(x-\frac{3}{2}h)\right) + C_1\left(f(x+\frac{1}{2}h) - f(x-\frac{1}{2}h)\right)}{h} + \frac{9f^{(5)}(x^*)h^4}{80} \tag{A.10}$$

The truncation error in equation A.10 will go to zero at the same rate as h^4, expressed at $O(h^4)$. Comparing equations A.4 and A.10, it can be observed that the $O(h^4)$ truncation error will go to zero more quickly than the $O(h^2)$ truncation error. This means the fourth-order approximation is able to use a larger node spacing h than the second-order approximation, while introducing equal amounts truncation error.

APPENDIX B

Parallel computing via Open-MP

B.1 Parallel computing via Open-MP

Parallel programming splits the computational workload of an algorithm over many cores or processors. Open-MP (Open Multi-Processing) is a shared memory API for parallel programming in C, C++ and Fortran. The processors of a shared memory system share a common memory cache, such that all cores are able to access the same memory locations. Open-MP divides the computation of a problem into a number of separate threads or pieces. The sudo code below parallelizes a for loop in C using Open-MP. The *num_threads* variable dictates the number of threads that the for loop iterations with be divided into. Imbedded for loops will split into threads over the outer most loop. Private variables are the local variables of the algorithm. A separate private variable is created for each thread. The shared variables are interdependent between threads, and therefore all threads must be able to access the memory location of the up-to-date shared variables.

```
#ifdef_OPENMP
pragma omp parallel for\
num_threads(numthreads)\
schedule(static)\
private(declare private variables)\
shared(declare shared variables)\
#endif
```

An optimal implementation of Open-MP will minimize the computational time to run an algorithm. This is dependent on the parametrization of the algorithm and the system that it runs on. Usually only one thread should be produced per available core, unless the processors are capable of multi-threading. However, due to the Open-MP API producing some overhead along with system specific memory access penalties, generating an excess number of threads can result in a loss of efficiency. The best way to determine the proper Open-MP parametrization is to benchmark your algorithm for a given system. Since all cores or processors must share memory access, depending on the memory configuration of the system (memory access penalties) there will exist a thread number that results in maximum efficiency. Generally on machines configured with non-uniform memory access (NUMA), assuming enough memory is available setting your number of threads to the number of cores per NUMA node/socket will produce an efficient speedup. This is due to the fact that the cores that make up a NUMA node/socket share a memory cache which is accessed faster than other memory locations on the system. The figure B.1 is the benchmark results for the finite-difference algorithm outlined in chapter 2 (sfea2dfd.c), on the Compute Canada symmetric multiprocessor system Hungabee. It can be seen that the speedup of the algorithm plateaus at 8 threads for the given model size. It should be noted that Hungabee has a NUMA node size of 8 cores.

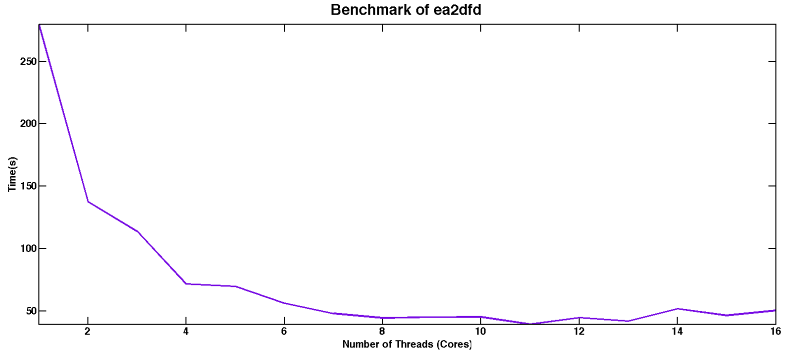


Figure B.1: Benchmark of sfea2dfd on Hungabee using Open-MP for Marmousi size model $[nx = 1502 \;\; nz = 4602 \;\; nt = 1250]$