

Important Notice

This copy may be used only for the purposes of research and private study, and any use of the copy for a purpose other than research or private study may require the authorization of the copyright owner of the work in question. Responsibility regarding questions of copyright that may arise in the use of this copy is assumed by the recipient.

UNIVERSITY OF CALGARY

Extending high-frequency asymptotic solutions to wave equations to
lower-frequency regimes

by

Chad M. Hogan

A DISSERTATION

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF GEOSCIENCE

CALGARY, ALBERTA

December, 2010

© Chad M. Hogan 2010

Abstract

All practical wave equations are derived with common assumptions and simplifications in order to make their solution tractable. In this dissertation I will explore the concept of the “high-frequency” approximation, and describe several ways in which I have attempted to extend the range of validity down into the lower frequencies which are most commonly found in seismic exploration geophysics.

I have done this in several ways. First, I have shown that the eikonal equation may be extended to give useful results in lower frequencies by simply smoothing the underlying wavespeed velocity model of the medium in a frequency-dependent fashion. Second, I have shown that a similar kind of frequency-dependent smoothing may also be applied to the design of the Generalized Phase-Shift plus Interpolation (GPSPI) algorithm wavefield extrapolation operator, and this too yields higher fidelity in the extrapolation, especially at lower frequencies. Additionally, I have taken theoretical mathematical extensions to this same operator, and developed them into a practical and useful operator. In another study, I have shown that the Early Arrival Waveform Tomography method, a low-frequency extension of more common travelttime tomography, may be feasible to use for the time-lapse monitoring of changing petroleum reservoirs. Following this, I show that planewave imaging can dramatically reduce imaging computation time by introducing a new method for measuring its convergence; this allows for more widespread usage of a method that is inherently more valid at low frequencies than many other common algorithms. Finally, I explain a new method for the stabilization and practical implementation of a faster version of the GPSPI method, which makes its overall lower-frequency validity compared to other methods more practical and economical.

Acknowledgements

I gratefully acknowledge the CREWES and POTSI research groups for their continued support of my research. I must specifically thank Kevin Hall and Rolf Maier for their tireless support and endless optimism.

I would also like to thank my supervisors, Gary Margrave and Michael Lamoureux, for their wisdom and guidance throughout my degree.

Ken Hedlin and Husky Energy Co. assisted with the research on time-lapse monitoring with EAWT, and their contributions – both scientific and financial – are gratefully acknowledged.

I received additional financial support from NSERC, the SEG, the University of Calgary's Department of Science, MITACS, and PIMS.

Finally, I would like to thank all of my friends and family, and especially Bernadette for helping me to maintain my sanity throughout the past five years.

Table of Contents

Abstract	ii
Acknowledgements	iii
Table of Contents	iv
List of Figures	vi
Glossary	xiv
1 Introduction	1
1.1 Original contributions in this dissertation	1
1.2 Petroleum exploration	2
1.3 The scalar wave equation in one spatial dimension	6
1.3.1 Beyond a one-dimensional scalar wave equation	9
1.4 Solutions to the wave equation	10
1.4.1 Direct numerical solution via finite differences	11
1.4.2 High-frequency asymptotic solutions	13
1.4.3 The Helmholtz equation	16
1.5 Seismic imaging	17
1.5.1 The profile imaging scheme	18
1.5.2 Depth extrapolation	20
1.5.3 The FOCI algorithm	31
1.5.4 Plane wave migration	31
1.5.5 Kirchhoff Migration	34
1.6 Inversion and early-arrival waveform tomography	42
1.6.1 Inverse problems	42
1.6.2 Early arrival waveform tomography	50
1.7 Chapter summary	56
2 Applying the eikonal equation to lower frequencies	58
2.1 Introduction	58
2.1.1 The hypereikonal equation	59
2.1.2 The physics of the eikonal equation	61
2.2 Frequency-dependent smoothing of a banded velocity model	64
2.3 Effects of smoothing on ray-tracing	74
2.4 Chapter summary	75
3 Frequency-dependent velocity model smoothing in GPSPI	78
3.1 Introduction	78
3.2 The square-root Helmholtz operator symbol	79
3.3 Approximating symbols	81
3.4 Testing	84
3.5 Chapter summary	87
4 The Stolk operator in GPSPI migration	89
4.1 Introduction	89

4.1.1	Recursive phase-shift wavefield propagation	89
4.1.2	Generalized PSPI	90
4.1.3	The Stolk operator	92
4.2	Testing	102
4.2.1	Impulse response	102
4.2.2	Focussing power	103
4.2.3	Marmousi migration	106
4.3	Discussion of imaging performance	114
4.4	Discussion of computational performance	115
4.5	Chapter summary	116
5	Feasibility testing of time-lapse seismic monitoring with early-arrival waveform tomography.	117
5.1	Introduction	117
5.2	Feasibility testing on synthetic data	118
5.2.1	Surface seismic modelling and inversion procedure	120
5.2.2	Surface seismic inversion results	123
5.2.3	VSP modelling and inversion procedure	123
5.3	Discussion of the results	125
5.4	Chapter summary and potential future work	127
6	Measurement of planewave migration convergence	129
6.1	Introduction	129
6.2	Theory	130
6.2.1	2D source plane-wave migration	132
6.2.2	Residuals	132
6.3	Testing	133
6.3.1	Simple synthetic	133
6.3.2	Marmousi testing	135
6.4	Discussion	140
6.5	Chapter summary	142
7	Locally WKBJ Operators	144
7.1	Introduction	144
7.1.1	The design of the IWKBJ operator	144
7.1.2	The $\omega - x$ operator	149
7.2	Marmousi migration	152
7.2.1	Parameter choice and performance	154
7.3	Conclusions and chapter summary	165
8	Conclusions	167

List of Figures

1.1	An idealized 2D seismic survey layout. In this survey, source points are placed at double the geophone spacing, with extra sources outside the survey to allow for broader subsurface coverage. Although only a few sources and receivers are shown, modern surveys typically place hundreds or more receivers and sources. Spacings are typically on the order of meters to tens of meters.	4
1.2	An idealized 3D seismic survey layout in plan view (“map view”). It is common to see one orientation with more densely spaced geophones (the “in-line” orientation), and the orthogonal orientation (the “cross-line” orientation) with more densely spaced shot locations. As with 2D surveys, spacings are typically on the order of meters or tens of meters, with a full survey frequently covering up to several kilometers on a side.	5
1.3	In (a), a single 2D shot gather from geophones placed with a centrally-located shot, over geology consisting of flat plane homogenous layers as illustrated in (b). Each vertical trace is the recording from a single geophone. Reflection events appear as approximate hyperbolæ. This shot gather has had a time-dependent amplification applied to account for the geometric spreading energy and other losses. In (a) the linear events beginning near time 0s and geophone 25, and forming a line to the end geophones at time 0.4s show the direct arrival of the wave travelling just beneath the subsurface directly towards geophones without reflection.	6
1.4	The string under tension T , with deflection from equilibrium position U , and deflection angle from horizontal θ	7
1.5	Phase-shift plus split-step corrections using a reference velocity of $3000m/s$. One large step at the reference velocity is taken for the entire width of the survey, and then smaller corrections are made for the regions that vary around the reference velocity.	25
1.6	Plane wave migration velocity model. The white region represents a relative velocity of 4, grey represents 6, and black represents 3. . . .	33
1.7	Full shot profile migration of the simulated survey. The survey consisted of 51 shots equally spaced across the entire surface shown, with 200 receivers spread equally across the entire span of the velocity model.	35
1.8	Horizontal plane wave migration	36
1.9	Three plane wave migration, orientations at 0° and $\pm 31^\circ$	36
1.10	Five plane wave migration, orientations at 0° , $\pm 11^\circ$, $\pm 31^\circ$	37
1.11	Seven plane wave migration orientations at 0° , $\pm 22^\circ$, $\pm 11^\circ$, and $\pm 31^\circ$.	37

1.12	Fifty-one plane wave migration, orientations equally spaced between -31° and $+31^\circ$	38
1.13	Schematic illustration of the principle behind Kirchhoff migration. (a) shows the calculation of total travelttime from source to scattering point and back up to receiver, through the initial hypothesized velocity model. (b) shows the identification of the reflection event at the calculated travelttime and on the appropriate receiver trace. . . .	40
1.14	A scaled version of the located energy (see Figure 1.13) is placed at its appropriate subsurface location in (a), and this process is repeated for every subsurface location in (b).	41
1.15	(Page 1 of 2) Deconvolution is an unstable process. The input signal in (a) is convolved with the filter in (b). This results in the signal shown in (c). A very small amount of random noise is added to this result, giving the signal shown in (d). This noisy signal is then deconvolved with signal (a), resulting in the (unstable) result in (e). Ideally it would result in a signal approximately like that in (f), which is the convolved signal shown in (c) deconvolved with (a).	48
1.15	(Page 2 of 2) Deconvolution is an unstable process. The input signal in (a) is convolved with the filter in (b). This results in the signal shown in (c). A very small amount of random noise is added to this result, giving the signal shown in (d). This noisy signal is then deconvolved with signal (a), resulting in the (unstable) result in (e). Ideally it would result in a signal approximately like that in (f), which is the convolved signal shown in (c) deconvolved with (a).	49
1.16	Seismic acquisition diagram showing three refracted arrivals traveling from source to receiver. Wave energy travels with wavefronts perpendicular to the black arrows, from the source down through the earth refracting at each layer interface, eventually being turned back up toward the receiver.	51
2.1	Ray tracing (in red) through a $2000m/s, \sigma = 500m/s$ medium. The eikonal solution envelope for the same propagation time is overlaid in blue.	62
2.2	Ray tracing (in red) through a $2000m/s, \sigma = 500m/s$ medium with light smoothing. The eikonal solution envelope for the same propagation time is overlaid in blue.	62
2.3	Ray tracing (in red) through a $2000m/s, \sigma = 500m/s$ medium with heavy smoothing. The eikonal solution envelope for the same propagation time is overlaid in blue.	63
2.4	Finite differences solution of a $2000m/s, \sigma = 200m/s$ medium. A distinct wavefront is visible.	63
2.5	Striped velocity model, velocity in m/s	65

2.6	Vertical sections of the velocity models used for eikonal solutions. On the left is the unsmoothed velocity model, on the right is the most heavily smoothed model.	66
2.7	60 Hz dominant source in unsmoothed velocity model, with eikonal solution to the unsmoothed model overlaid in blue.	67
2.8	2 Hz dominant source, with eikonal solution to the 25 m smoothing length model overlaid in blue.	68
2.9	2 Hz dominant source, 25 m smoothing length eikonal solution in white, unsmoothed velocity model eikonal solution in blue.	69
2.10	5 Hz dominant source, eikonal solution to 20 m smoothing length model overlaid in blue.	70
2.11	10 Hz dominant source, eikonal solution to 15 m smoothing length model overlaid in blue.	71
2.12	20 Hz dominant source, eikonal solution to 10 m smoothing length model overlaid in blue.	72
2.13	40 Hz dominant source, eikonal solution to 5 m smoothing length model overlaid in blue.	73
2.14	Smoothing length as a function of dominant frequency.	74
2.15	Ray-tracing (red) through the unsmoothed velocity model), with eikonal contour (black).	75
2.16	Ray-tracing (red) through the 25m smoothed velocity model, with eikonal contour (black).	76
2.17	Ray-tracing (red) through the unsmoothed velocity model, with 60 Hz dominant finite difference wavefield (grayscale).	76
2.18	Ray-tracing (red) through the 25m smoothed velocity model, with 2 Hz dominant finite difference wavefield (grayscale).	77
3.1	Real part of the infinite-frequency square-root Helmholtz operator symbol for a velocity model with blocks of constant velocity. Adapted from Fishman (2002). This is the symbol as used in GPSPI.	80
3.2	Ensemble of real parts of the square-root Helmholtz operator symbols for a velocity model with blocks of constant velocity. Adapted from Fishman (2002).	81
3.3	Real part of the infinite-frequency symbol for the three-block velocity model.	82
3.4	Real part of the low-frequency symbol approximation (via smoothing of the velocity model) for the three-block velocity model.	83
3.5	The Marmousi velocity model. Velocities range between approximately 1500m/s and 5500m/s.	84

3.6	Images of the Marmousi model generated a) without frequency-dependent smoothing of the symbol and b) with frequency-dependent smoothing of the symbol. The image generated with the smoothed symbol is clearly far superior to the image generated with the infinite-frequency symbol.	86
3.7	Determination of effectiveness of symbol smoothing as a function of frequency. Lower residuals indicate that the smoothing of the symbol resulted in a change in quality of the overall image, while a residual of 1 indicates that the smoothing had little to no effect in this region.	87
4.1	The damping function $h(\zeta)$	94
4.2	The damping function $h(\kappa (\omega^2/v^2 - \xi^2))$ for $\kappa = 400$ and $v = 1000m/s$	95
4.3	$ \alpha = e^{ik_z(x)} $ for $v = 1000m/s$	95
4.4	The Stolk operator symbol in $\omega - x$ and $\omega - \xi$ domain for $70Hz$ at $2000ms^{-1}$ with no horizontal velocity (slowness) gradient $-ie$ this operator is exactly equivalent to FOCL. The top panel shows the $\omega - x$ expression of the symbol, the middle panel shows the operator symbol amplitude in $\omega - \xi$, and the bottom panel shows the operator symbol phase in $\omega - \xi$	97
4.5	The Stolk operator symbol in $\omega - x$ and $\omega - \xi$ domain for $70Hz$ at $2000ms^{-1}$ for $35s^{-1}$ velocity gradient. The top panel shows the $\omega - x$ expression of the symbol, the middle panel shows the operator symbol amplitude in $\omega - \xi$, and the bottom panel shows the operator symbol phase in $\omega - \xi$	98
4.6	The Stolk operator in $\omega - x$ and $\omega - \xi$ domain for $15Hz$ at $2000ms^{-1}$ for $35s^{-1}$ velocity gradient. The top panel shows the $\omega - x$ operator, the middle panel shows the operator amplitude in $\omega - \xi$, and the bottom panel shows the operator phase in $\omega - \xi$	99
4.7	The Stolk operator in $\omega - x$ and $\omega - \xi$ domain for $30Hz$ at $2000ms^{-1}$ for $35s^{-1}$ velocity gradient. The top panel shows the $\omega - x$ operator, the middle panel shows the operator amplitude in $\omega - \xi$, and the bottom panel shows the operator phase in $\omega - \xi$	100
4.8	The Stolk operator in $\omega - x$ and $\omega - \xi$ domain for $30Hz$ at $2000ms^{-1}$ for $-4s^{-1}$ velocity gradient. The top panel shows the $\omega - x$ operator, the middle panel shows the operator amplitude in $\omega - \xi$, and the bottom panel shows the operator phase in $\omega - \xi$	101
4.9	Impulse response of standard symbol on the left compared to Stolk on the right, both symbols propagated an impulse 200m through a homogeneous medium.	102
4.10	Impulse response of standard symbol on the left compared to Stolk on the right, both symbols propagated an impulse 200m through a medium with a strong velocity gradient.	103

4.11	Difference of standard symbol and Stolk symbol, magnified 10 times.	104
4.12	High-velocity lens model. The centre is 4500 m/s , with sides at 1500 m/s . Total width of the central portion is approximately 300 m .	105
4.13	Focussing power of GPSPI and Stolk operators for the high-velocity lens model on the left. Radial energy density for the high-velocity lens model on the right. The GPSPI and Stolk lines are nearly indistinguishable.	105
4.14	Marmousi chunk model, with velocity in m/s	106
4.15	Focussing power of GPSPI and Stolk operators for the Marmousi chunk model on the left. Radial energy density for the Marmousi chunk model on the right.	107
4.16	Marmousi migration in standard FOCI with 41/51/41 windowing.	108
4.17	Marmousi migration using Stolk correction, run with 41/51/41 windowing.	109
4.18	Marmousi migration, Stolk vs. FOCI migration with 41/51/41 operators. Region 1.	110
4.19	Marmousi migration, Stolk vs. FOCI migration with 41/51/41 operators. Region 2.	111
4.20	Marmousi migration, Stolk vs. FOCI migration with 41/51/41 operators. Region 3.	112
4.21	Marmousi migration, Stolk vs. FOCI migration with 41/51/41 operators. Region 4.	113
5.1	Original velocity model with no time-lapse change.	119
5.2	Original velocity model plus steam-injection effect. The location of the VSP receivers is marked as a vertical green line at the injection site. The approximate wavelength at 5 Hz in the injected region is displayed for scale.	119
5.3	Original velocity model plus steam-injection effect. The location of the VSP receivers is marked as a vertical green line at the injection site.	120
5.4	Raytracing through the original model convolved with a 20m Gaussian smoother	121
5.5	Sample time slices of the forward-modelled data on the perturbed model showing the wavefront periodically as it progresses through the medium. The white rectangle delineates the extent of the perturbation. Very little energy penetrates beneath the horizontal reflector at depth $\sim 700m$	122

5.6	The resulting surface-seismic velocity estimate from inversion at 5 Hz, as a difference from the starting (background) velocity model. Ideally, the amplitude of the perturbation would be exactly +500 m/s. One wavelength at 5 Hz would approximately span this entire displayed region at the velocity of the target region. The black box marks the spatial extent of the actual perturbation.	123
5.7	The resulting surface-seismic velocity estimate from inversion at 5 Hz, as a difference from the starting (background) velocity model. Ideally, the amplitude of the perturbation would be exactly +500 m/s. One wavelength at 5 Hz would approximately span this entire displayed region at the velocity of the target region. The black box marks the spatial extent of the actual perturbation.	124
5.8	The resulting VSP velocity estimate from inversion at 5 Hz, as a difference from the starting (background) velocity model. Ideally, the amplitude of the perturbation would be exactly +500 m/s. One wavelength at 5 Hz would approximately span this entire displayed region at the velocity of the target region. The black box marks the spatial extent of the actual perturbation.	125
5.9	The resulting VSP velocity estimate from inversion at 5 and 6 Hz, as a difference from the starting (background) velocity model. Ideally, the amplitude of the perturbation would be exactly +500 m/s. One wavelength at 5 Hz would approximately span this entire displayed region at the velocity of the target region. The black box marks the spatial extent of the actual perturbation.	126
6.1	Velocity model. The white region represents a relative velocity of 4, grey represents 6, and black represents 3.	134
6.2	An image of the velocity model calculated with shot-profile migration.	135
6.3	The horizontal plane wave image.	136
6.4	Plane-wave images generated with a) plane waves oriented at $0^\circ, \pm 31^\circ$, b) $0^\circ, \pm 11^\circ, \pm 31^\circ$	137
6.5	Plane-wave images generated with a) plane waves at $0^\circ, \pm 11^\circ, \pm 22^\circ, \pm 31^\circ$, b) 51 plane waves between -31° and $+31^\circ$	137
6.6	The Marmousi velocity model. Velocity ranges from 1500 <i>m/s</i> to 6000 <i>m/s</i>	138
6.7	Residuals for shot-profile and plane-wave migration as a function of number of individual migrations (i.e. number of plane waves or number of shot records migrated). The plane-wave residuals show dramatically faster reduction in residual, compared to the slower decline in shot-profile migration.	138
6.8	Marmousi shot-profile migration with a) 110 shot records and b) 240 shot records.	139

6.9	Marmousi migration with a) 41 plane waves and b) 41 shot records.	139
6.10	Marmousi migration with a) 81 plane waves and b) 81 shot records.	140
6.11	Residual calculated between an intermediate image and the final “best” image calculated with 240 shots.	141
7.1	In the top panel, raytracing through a homogeneous medium is shown. Note that rays leaving horizontally will travel to infinity. In the bottom panel, rays are traced through a LWKBJ medium. Aperture radius was limited to 20 meters, and takeoff angles from -90° to 90° are displayed in ten meter increments	146
7.2	Symbol amplitudes for a locally homogeneous medium and for a LWKBJ medium are compared. $x_r = 30m, v_{loc} = 2000m/s, v_0 = 1785m/s, m = 45s^{-1}, dz = 10m$	147
7.3	Symbol phases for a locally homogeneous medium and for a LWKBJ medium are compared. $x_r = 30m, v_{loc} = 2000m/s, v_0 = 1785m/s, m = 45s^{-1}, dz = 10m$	148
7.4	$\omega - x$ operator kernel for FOCI(a), LWKBJ with normal stabilization (b), and LWKBJ with over stabilization (c). The kernel is not compactly supported, only a 400m extent is shown here.	150
7.5	Equivalent truncated $\omega - \xi$ symbol amplitude for FOCI(a), LWKBJ with normal stabilization (b), and LWKBJ with over stabilization (c).	151
7.6	Equivalent truncated $\omega - \xi$ symbol phase for FOCI(a), LWKBJ with normal stabilization (b), and LWKBJ with over stabilization (c).	153
7.7	Marmousi image, FOCI migration with full stabilization. (31 point forward operator, 41 point inverse operator, 31 point final window, 187.5m effective operator length.)	155
7.8	Marmousi image, LWKBJ migration. (31 point (187.5m) operator, 40m aperture radius.)	156
7.9	Marmousi region 1. a) Velocity model of the region b) FOCI image c) LWKBJ image. Note the improved continuity in reflectors near (3600, 1300) and (4400, 1000)	157
7.10	Marmousi region 2. a) Velocity model of the region b) FOCI image c) LWKBJ image. The LWKBJ may have better reflector continuity, but seems to lose some focus compared to the FOCI image.	158
7.11	Marmousi region 3. a) Velocity model of the region b) FOCI image c) LWKBJ image. The LWKBJ image appears to have better continuity, with a more prominent fault line.	159
7.12	Marmousi region 4. a) Velocity model of the region b) FOCI image c) LWKBJ image. The reservoir target is well imaged by both algorithms.	160
7.13	Marmousi image, LWKBJ migration. (101 point (625m) operator, 80m aperture radius.)	163

7.14 Marmousi image, lwkBJ migration. (15 point (87.5*m*) operator, 23*m* aperture radius.) Instability artifacts contaminate the image near the reservoir target, (5000 – 8000, 2500+). 164

Glossary

Definitions

air blast: A class of coherent noise associated with the sound/pressure wave of the source explosion travelling through air and affecting the geophones.

asymptotic solution: An approximate solution to an equation that becomes exact in some limit of a parameter. A high-frequency asymptotic solution, for example, becomes an exact solution in the limit where frequency approaches infinity.

evanescent waves: Waves which are generated at a boundary, and typically exhibit exponential decay with distance from the boundary. A common example of evanescent waves in geophysics is ground roll.

ground roll: A class of coherent noise typically associated with surface wave phenomena. See Aki and Richards (2002) for more details on surface waves.

high-frequency approximation: An approximation, often to a solution to a wave equation, that assumes that the scale of variability of wavespeed in a medium is much smaller than the dominant wavelength under consideration.

imaging condition: In shot profile migration, an algorithm which compares the down-going wavefield to the upgoing wavefield to derive an approximate reflectivity at a location.

linear operator: An operator \mathcal{G} is a linear operator if it is defined so that $\mathcal{G}(m_1) + \mathcal{G}(m_2) = \mathcal{G}(m_1 + m_2)$ and $\mathcal{G}(\alpha m) = \alpha \mathcal{G}(m)$.

poststack migration: Seismic imaging from partly reduced data, in which the data from many individual shot records are summed prior to the imaging step. Poststack imaging typically is seen as a correction to an existing image which was constructed by assuming *e.g.* that the geologic features to be imaged are composed of entirely flat plane homogenous layers. Under these assumptions, any non-flat features will have imaging artifacts associated with them. Post-stack imaging attempts to correct imaging errors introduced by overbroad assumptions.

prestack depth migration: Seismic imaging using minimally-reduced data. Input data typically includes a (smooth) velocity model and shot gathers. Other gathers may be used, but in principle they are generally derivable from shot-gather data. Prestack migration attempts to make as few assumptions about the geologic features as possible, and makes use of as much of the data as

possible. It is a direct imaging process – the image is directly calculated from input data and an existing velocity model. *cf.* Poststack migration.

profile imaging A method of seismic data migration (imaging) in which a surface-recorded seismic wavefield is considered to be the upgoing wavefield, and a modelled source approximating the original survey source is considered to be a downgoing wavefield. These wavefields are compared at depth.

seismic imaging: The process of transforming seismic survey data into an image of the geologic subsurface below the survey. “Migration” will often be used interchangeably with “imaging”.

shot gather: All data from a single source recorded by all active geophones or hydrophones.

specular raypaths: Specular raypaths are those raypaths which obey Snell’s law and the law of reflection.

Symbols

(x, y, z) Spatial dimensions.

\mathcal{G} An operator representing the physics of a system.

∇ The Laplacian operator.

\mathbb{R} The field of real numbers.

\mathbb{R}^n The real coordinate space in n dimensions, with elements $\vec{x} = (x_1, x_2, \dots, x_n)$.

μ Linear mass density.

ω Temporal frequency in radians per second.

\vec{x} A spatial vector, equivalent to (x, y, z) in three dimensional space.

t Time dimension.

Chapter 1

Introduction

1.1 Original contributions in this dissertation

In this dissertation, I will detail my original contributions to the field of exploration seismic imaging and wave-equation studies in general.

In this introduction, I will introduce the non-specialist reader to all of the fundamental concepts that are required to appreciate the work found in the rest of the dissertation. Most of the ideas are developed at least heuristically, and the most common and fundamental references are given for those looking for a deeper understanding of the underlying issues. In Chapter 2, I will describe my work regarding the extension of high frequency asymptotic solutions to the wave equation to lower frequency regimes, specifically in the context of eikonal equations and ray-tracing solutions. In Chapter 3, I will detail my contribution to the study of frequency-dependent smoothing in the wavefield extrapolation method FOCI (Forward Operator, Conjugate Inverse; Margrave et al., 2005), and show how this smoothing greatly improves the fidelity of the extrapolated wavefield. I will demonstrate that this may be due to a better match of the operator symbol with the full-frequency “exact” operator symbol. In Chapter 4, I will describe the adaptation of a theoretical improvement upon the operator symbol used in GPSP (Generalized Phase Shift Plus Interpolation; Margrave and Ferguson, 1999) wavefield extrapolation into a practical algorithm, useful for migration. This effectively extends the locally-homogeneous high-frequency operator symbol, hopefully yielding higher fidelity for wavefield ex-

trapolation. In Chapter 5, I will discuss our application of early-arrival waveform tomography to time-lapse monitoring of steam injection for heavy-oil production. This method uses primarily low-frequency information to identify and characterize anomalies in a velocity model, as opposed to many other time-lapse methods which typically rely on much higher-frequency data. In Chapter 6, I will discuss my investigation into the measurement of the convergence of plane wave imaging. The primary significance of this contribution is that it may allow for more efficient application of a method which has higher fidelity than many other methods, especially for relatively low frequencies. In Chapter 7, I will discuss my extension of the GPSPI operator to stabilize the FOCI operator in highly-variable media. This allows for easier practical use of a GPSPI method in cases where previously a less accurate algorithm may have been considered.

1.2 Petroleum exploration

Petroleum exploration is often conducted with the assistance of *seismic imaging*. This process uses a source of energy, frequently either a blast of dynamite or a large vibrating mass, to shake the ground at a particular surface location. This shaking travels through the ground as a wave, and reflects and refracts throughout the subsurface. These reflections and refractions are recorded at the surface with an array of *geophones*, instruments which can accurately measure vibrations via direct measurements of either ground acceleration or velocity.

A current seismic survey typically consists of recording this information for hundreds to thousands of individual “shots”; one shot is typically a single dynamite blast or mass-vibration at one location. Shot locations are spaced out over the area

to be investigated, with typical spacing between shot locations on the order of perhaps five or ten meters. Shots at any given location may be repeated, and the results summed to reduce random noise. This is particularly common for mass-vibration source surveys. Sources are often spaced linearly for a “2D” survey, or on a grid for a “3D” survey. See Figures 1.1 and 1.2 for diagrams of typical “2D” and “3D” survey geometry respectively.

Each shot is recorded into an array of geophones, with hundreds to thousands of individual geophones in a given survey. Geophones are often placed with a spacing of the order of five to ten meters. As with sources, the geophones may be placed linearly or on a grid, depending on the survey type (“2D” or “3D”).

Although most discussion in this dissertation is in the context of geophones and land-based acquisition, pressure-sensitive *hydrophones* may also be used to record data in water, for marine surveys. In most respects, these may be taken as giving information equivalent to that given by geophones, especially when an acoustic wave equation is employed for imaging.

Data recorded from a single shot into all active geophones forms a single *shot gather* – that is, all the data from a single *shot* is *gathered* together. A single shot gather can be seen in Figure 1.3a, as recorded from the idealized geometry shown in Figure 1.3b.

The measured disturbances at the surface are used to develop an image of the rock formations in the subsurface. This is accomplished via *seismic imaging*, which is also commonly known as *migration*. While there are many kinds of seismic imaging, this dissertation will only focus on a certain type known as *prestack depth migration*. Many versions of this method use shot gathers directly with minimal processing to develop an image of the subsurface. While the data within these gathers may be

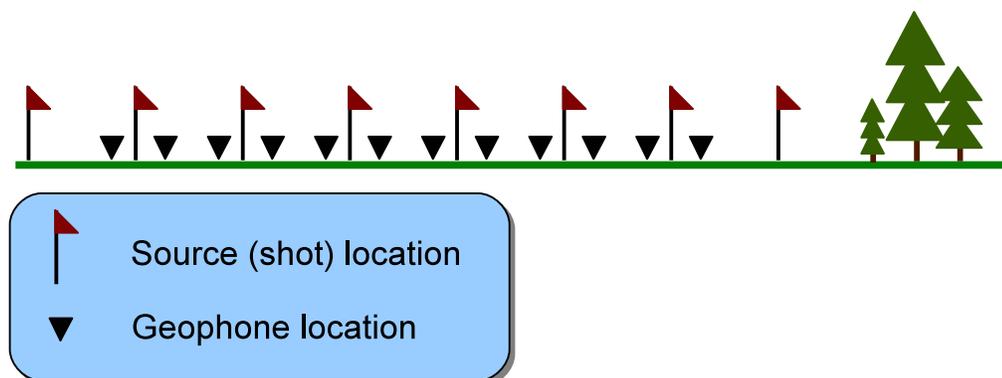


Figure 1.1: An idealized 2D seismic survey layout. In this survey, source points are placed at double the geophone spacing, with extra sources outside the survey to allow for broader subsurface coverage. Although only a few sources and receivers are shown, modern surveys typically place hundreds or more receivers and sources. Spacings are typically on the order of meters to tens of meters.

resorted into other types of gathers for the imaging and/or filtered for noise, the term *prestack depth migration* implies that information is only minimally reduced before imaging. In prestack depth migration, each individual shot gather is typically imaged individually and then all of these resulting individual images are combined to give a final image. Other processes, for example *poststack migration*, have had significant data reduction processes applied before imaging begins. In poststack migration, typically the shot records are combined together and this combination is then imaged in one operation. The prestack migration usually allows for much more detailed imaging, while the poststack method is far less expensive computationally. See *e.g.* Robein (2004) for more information on other migration methods.

Many other details about imaging and seismic data processing in general may be found in Yilmaz (2001).

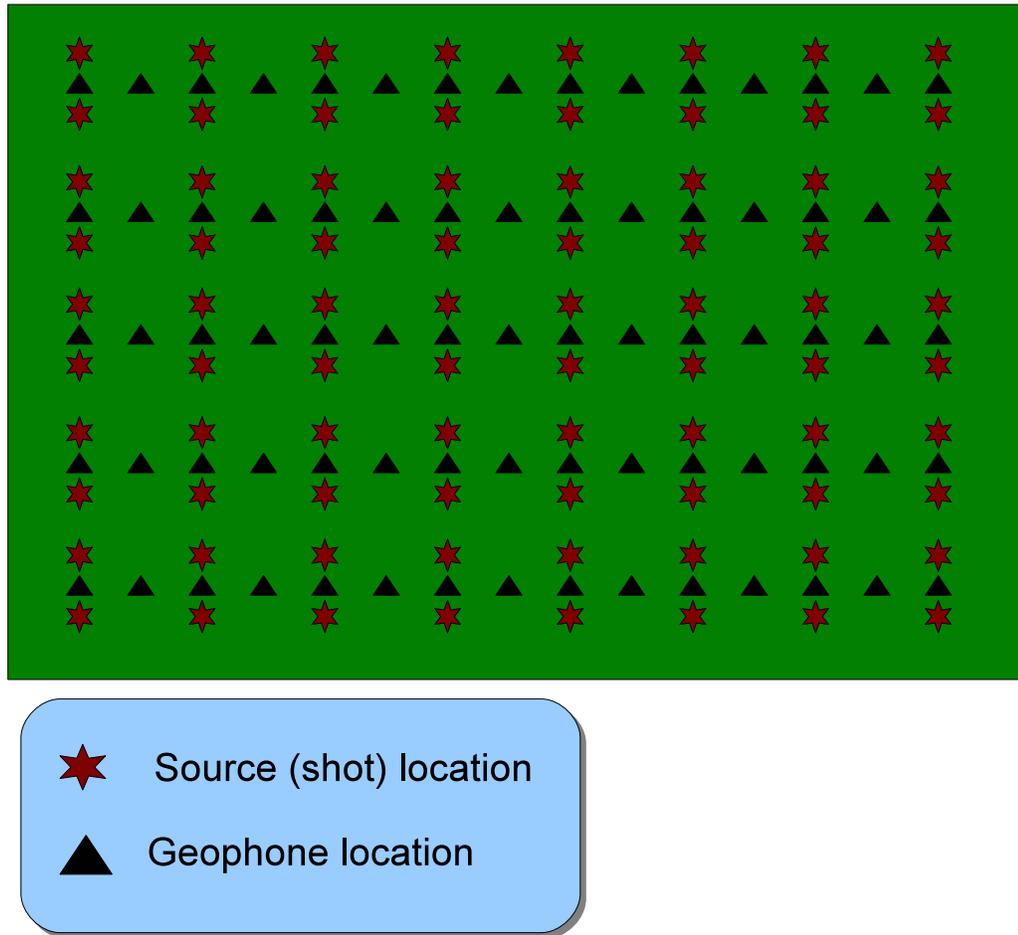


Figure 1.2: An idealized 3D seismic survey layout in plan view (“map view”). It is common to see one orientation with more densely spaced geophones (the “in-line” orientation), and the orthogonal orientation (the “cross-line” orientation) with more densely spaced shot locations. As with 2D surveys, spacings are typically on the order of meters or tens of meters, with a full survey frequently covering up to several kilometers on a side.

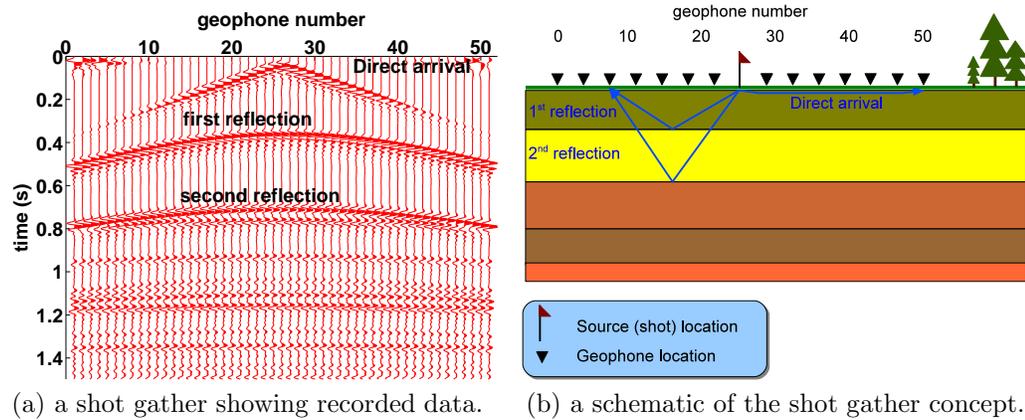


Figure 1.3: In (a), a single 2D shot gather from geophones placed with a centrally-located shot, over geology consisting of flat plane homogenous layers as illustrated in (b). Each vertical trace is the recording from a single geophone. Reflection events appear as approximate hyperbolæ. This shot gather has had a time-dependent amplification applied to account for the geometric spreading energy and other losses. In (a) the linear events beginning near time 0s and geophone 25, and forming a line to the end geophones at time 0.4s show the direct arrival of the wave travelling just beneath the subsurface directly towards geophones without reflection.

1.3 The scalar wave equation in one spatial dimension

Wave equations form a class of partial differential equations that describe the propagation of disturbances through a medium. There are many different specific wave equations that have been developed to handle many specific cases of type of disturbance, type of medium, and details of propagation. Although wave equations are relatively straight-forward to derive and impose upon a medium, their solution is virtually intractable in all but the most trivial of situations.

Consider a string under tension of magnitude $T(x, t)$ and fixed at both ends, with linear mass density μ . Assume the absence of external forces (e.g. gravity). Trivially, this string will align linearly, motionless. However, with the introduction of a small transient deflection from equilibrium at some point between the fixed

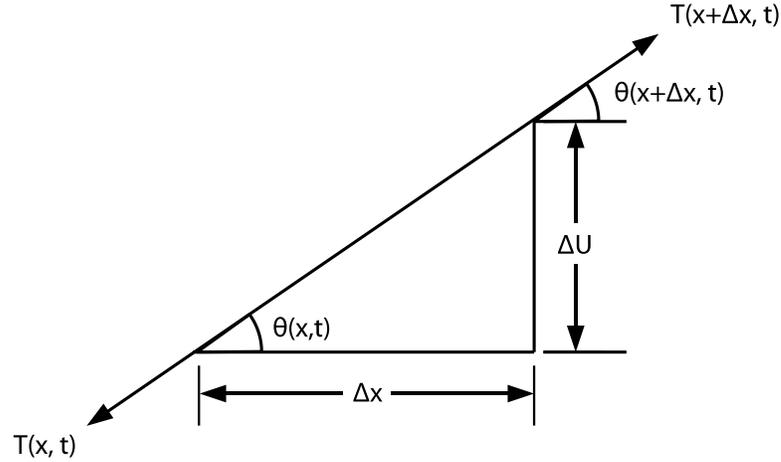


Figure 1.4: The string under tension T , with deflection from equilibrium position U , and deflection angle from horizontal θ .

ends, the system becomes dynamic. Figure 1.4 shows the situation for this transient deflection on a small portion of the string. x denotes position along the string, t is time, $u(x, t)$ is the displacement of the string perpendicular from the linear equilibrium position, $\theta(x, t)$ is the angle between the string and the linear equilibrium position.

We will consider a small section of the deflected string, small enough that the string is approximately linear. Since the deflection is small, $\theta(x, t)$ is small as well. The net force on the segment of string is $T(x, t) - T(x + \Delta x, t)$. Application of Newton's second law (Newton, 1687) in the vertical direction yields:

$$\mu\sqrt{\Delta x^2 + \Delta u^2}\partial_t^2 u = T(x + \Delta x, t) \sin \theta(x + \Delta x, t) - T(x, t) \sin \theta(x, t). \quad (1.1)$$

Dividing equation 1.1 by Δx and taking the limit as $\Delta x \rightarrow 0$,

$$\mu\sqrt{1 + (\partial_x u)^2}\partial_t^2 u = \partial_x (T(x, t) \sin \theta(x, t)). \quad (1.2)$$

By assuming that the deflection of the string is small, we effectively assume that

$\sin \theta \approx \theta$. Since $\partial u / \partial x \ll 1$, $(\partial u / \partial x)^2 \approx 0$, so equation 1.2 simplifies to

$$\mu \frac{\partial^2 u}{\partial t^2}(x, t) = \left(\frac{\partial T}{\partial x}(x, t) \right) \left(\frac{\partial u}{\partial x}(x, t) \right) + T(x, t) \frac{\partial^2 u}{\partial x^2}(x, t). \quad (1.3)$$

Since the string is fixed on both ends, the string does not move horizontally. Therefore the net local horizontal force is zero such that,

$$T(x + \Delta x, t) \cos \theta(x_0 + \Delta x, t) - T(x, t) \cos \theta(x, t) = 0. \quad (1.4)$$

In the limit as $\Delta x \rightarrow 0$,

$$\partial_x (T(x, t) \cos \theta(x, t)) = 0. \quad (1.5)$$

For small θ , $\cos \theta(x, t) \approx 1$,

$$\partial_x T(x, t) \approx 0 \quad (1.6)$$

Inserting equation 1.6 into equation 1.3 and rearranging,

$$\frac{\partial^2 u}{\partial t^2}(x, t) = \frac{T(x, t)}{\mu} \frac{\partial^2 u}{\partial x^2}(x, t). \quad (1.7)$$

In equation 1.7, $T(x, t) / \mu$ is dimensionally equivalent to [velocity]², and is in fact interpreted as the wavespeed velocity, or the propagation velocity of a wave on the string.

Equation 1.7 is a scalar one-dimensional wave equation. Note that even in this simple case, we have had to restrict ourselves to an approximate description of the phenomenon: we have assumed a “small” deflection from equilibrium. In reality, this means that any finite deflection strictly invalidates this wave equation. Equivalently, any real disturbance cannot possibly obey the wave equation exactly. Nevertheless, it is a useful approximate solution.

Wave equations may be thought of as the combination of three concepts: Newton’s second law, in this case applied giving the result in equation 1.1; a constitutive

relation, one example of which is Hooke’s Law; and a liberal application of linearization. The application of linearization was clearly present in the treatment of “small” deflections which allowed several simplifying approximations. In the absence of these linearizations, we would have a non-linear wave equation, which is significantly more challenging to solve.

Where, then, was the application of a constitutive relation like Hooke’s Law? Generally speaking, a constitutive relation is an expression that describes how a particular medium responds to a disturbance that distorts the medium away from its equilibrium resting configuration. Hooke’s Law states that the force exerted by a spring is linearly proportional to the distance through which the spring is compressed or extended: $F = kx$, where F is the magnitude of the force, x is the distance of extension or compression, and k is the so-called spring constant.

We may recognize the constitutive relation implicit in this wave equation as being the source of the natural restoring force. As the string is deflected from horizontal, the difference in tensions across a small section of the string results in a net local vertical force directed toward restoring the string to its equilibrium zero-deflection state. This restoring force is a function of the deflection, and so is analogous to Hooke’s Law.

1.3.1 Beyond a one-dimensional scalar wave equation

Although a one-dimensional scalar wave equation is interesting mathematically, and while it may be useful to approximately describe the behaviour of the string on a guitar or the vibration of a stretched spring, it does not adequately describe the reality of a seismic disturbance travelling through the subsurface.

It is possible to extend a comparable derivation to include three spatial dimen-

sions in an acoustic medium to derive a 3D scalar wave equation. Careful treatment of a three-dimensional perfectly elastic solid material leads to a vector wave equation, allowing both dilatation and rotational motion within the medium (e.g. Aki and Richards, 2002). Allowing a viscosity or damping within the material will lead to a viscoelastic wave equation (e.g. Lakes, 1998). Although these wave equations increase the complexity of the physics and subsequently the nominal fidelity with which the imaging may be conducted, designing seismic imaging algorithms based on an acoustic wave equation is often a reasonable compromise between complexity and simplicity, both in terms of computational cost and intensity of physics.

It is worth explicitly noting that the derivation of the three-dimensional elastic wave equation (and every other commonly used seismic wave equation) also requires the assumption of a “small” perturbation of the medium, such that any finite disturbance strictly invalidates these equations as well.

1.4 Solutions to the wave equation

Though a wave equation is relatively simple to pose, as with many partial differential equations that model real physical phenomena, solutions to all but the most trivial of cases can be extremely difficult to calculate. Three common approaches to approximate solutions are direct numerical solution via finite differences, high-frequency asymptotic solutions to the wavefield, and pseudodifferential solutions. Each method of solution has its own unique strengths and weaknesses. Although I do not directly address the pseudospectral method for solving wave equations in this dissertation, this method is common in exploration geophysics. Many interesting topics, including pseudospectral methods and more, are presented by Boyd (2001).

It is important to emphasize that these are approximate solutions to a wave equation which was derived using an approximation to reality – even in the most optimistic case we can calculate *almost* the correct answer to *not quite* the correct question. It is always important to question the validity of these calculated results in the larger context of the approximations.

1.4.1 Direct numerical solution via finite differences

Direct numerical solution typically involves the approximation of the derivatives in the equation with “finite differences”. For example, the derivative $\partial_x f(x, \cdot)$ may be approximated by softening the implied limit that defines it: $\lim_{\Delta x \rightarrow 0} \frac{f(x+\Delta x, \cdot) - f(x, \cdot)}{\Delta x}$. A finite (but small) Δx can be used directly to approximately calculate the derivative of a function. Typically, the choice of this Δx size means that the derivative of a function is approximated on a discrete grid of spacing Δx .

A second derivative may be calculated with a similar process, leading to an example approximate second derivative,

$$\partial_x^2 f(x, \cdot) \approx \frac{f(x + \Delta x, \cdot) - 2f(x, \cdot) + f(x - \Delta x, \cdot)}{\Delta x^2}. \quad (1.8)$$

There are many introductory texts that cover finite differences and other direct numerical solutions techniques, including Råde and Westergren (2004) and Burden and Faires (2004).

The main benefit of direct numerical solutions, especially with finite differences, is that solutions can be arbitrarily accurate given sufficient computing resources. This can be shown by considering a finite difference approximation as an implementation of the Taylor series expansion of the solution.

Given a function f which is n times continuously differentiable on the interval

$[a, x]$, and $n + 1$ times differentiable on the open interval (a, x) , this function may be expanded into a polynomial called the *Taylor series*:

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x - a)^n + R_n(x). \quad (1.9)$$

R_n is the remainder term, and quantifies the difference between the n^{th} degree Taylor polynomial and the original function f . One form of this remainder term says that there exists a c between a and x such that

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x - a)^{n+1}. \quad (1.10)$$

A frequent application of the Taylor series is to extend an existing solution to an equation (including derivatives) at the point a to a small neighbourhood around a . That is, $f(x)$ may be calculated given the values $f(a)$, $f'(a)$, $f''(a)$, \dots , $f^{(n)}(a)$. As long as $x - a$ is small, the remainder will be small. More on the Taylor series may be found in many texts, including Morse and Feshbach (2005).

A first-order Taylor series for the function f at a may be rearranged,

$$f'(a) = \frac{f(x) - f(a)}{(x - a)} + R_1(x) \quad (1.11)$$

This, however, is simply the first-order finite difference approximation for the derivative, plus the remainder term R . Explicitly,

$$R_1(x) = \frac{f''(c)}{2}(x - a)^2. \quad (1.12)$$

Since c is between a and x , and since f is sufficiently smooth, this remainder may be made arbitrarily small by keeping the interval $[a, x]$ small. Practically speaking, this means that the accuracy of the approximation depends on the size of the grid spacing Δx . In partial differential equations, the different derivative approximations

for each partial derivative can often interact in a complicated fashion. For solutions to a wave equation, the time-axis grid spacing and spatial-axis grid spacing are related, and must be chosen in a fashion that satisfies a *grid dispersion* restriction. See Cho et al. (2007) for an introduction to this concept, and Holberg (1987) for a thorough discussion. Additionally, finite difference schemes must have these spacings chosen to satisfy a *stability* requirement. That is, small errors in the computation must not grow with time as this will cause the solution to diverge as the errors swamp the results. This topic is discussed broadly in numerical analysis, with one recent and applicable example given by Lines et al. (1998).

1.4.2 High-frequency asymptotic solutions

In addition to a frontal assault via computational methods, wave equations may also be attacked indirectly via further approximations and assumptions. A “high frequency” assumption is a popular and convenient approach in many cases. This treatment follows one given by Pujol (2003). To derive the eikonal equation, begin with a scalar wave equation with position coordinate \vec{x} and time t ,

$$\left(\nabla^2 - \frac{1}{v^2(\vec{x})} \partial_t^2 \right) \Psi(\vec{x}, t) = 0. \quad (1.13)$$

Fourier transform $t \rightarrow \omega$:

$$\nabla^2 \psi(\vec{x}, \omega) + \frac{\omega^2}{v^2(\vec{x})} \psi(\vec{x}, \omega) = 0, \quad (1.14)$$

where the Fourier transform is defined by,

$$\psi(\vec{x}, \omega) = \frac{1}{2\pi} \int_{\mathbb{R}} \Psi(\vec{x}, t) e^{-i\omega t} dt. \quad (1.15)$$

Now assume trial solutions of (harmonic) form:

$$\psi(\vec{x}, \omega) = A(\vec{x}, \omega) e^{i\omega\phi(\vec{x})}. \quad (1.16)$$

Now we calculate the component of the ∇^2 operator for each spatial axis j ,

$$\partial_j^2 \psi = (\partial_j^2 A + 2i\partial_j A \omega \partial_j \phi + i\omega A \partial_j^2 \phi - A\omega^2(\partial_j \phi)^2) e^{i\omega\phi} \quad (1.17)$$

By substituting this expression and the harmonic solutions from equation 1.16 into equation 1.14, we have

$$(\partial_j^2 A + 2i\partial_j A \omega \partial_j \phi + iA\omega \partial_j^2 \phi - A\omega^2(\partial_j \phi)^2) e^{i\omega\phi} + \frac{\omega^2}{v^2} A e^{i\omega\phi} = 0 \quad (1.18)$$

for one spatial axis. Cancelling the exponential, dividing by $A\omega^2$, and rearranging yields

$$\left((\partial_j \phi)^2 - \frac{1}{v^2} \right) - \frac{i}{\omega} \left(\frac{2}{A} \partial_j A \omega \partial_j \phi + \partial_j^2 \phi \right) - \frac{1}{\omega^2 A} \partial_j^2 A = 0. \quad (1.19)$$

This solution can be simplified by considering real and imaginary parts separately.

The real part must be zero,

$$(\partial_j \phi)^2 - \frac{1}{v^2} - \frac{1}{\omega^2 A} \partial_j^2 A = 0. \quad (1.20)$$

If we consider the limiting case where $\omega \rightarrow \infty$, the last term in equation 1.20 becomes negligible. By restoring to full dimension, we reveal the *eikonal equation*,

$$|\nabla \phi(\vec{x})|^2 = \frac{1}{v^2(\vec{x})}. \quad (1.21)$$

This equation is fundamentally valid only in this high-frequency limit. This implies that the eikonal equation may only be used when variations in velocity are negligible on spatial scales that are comparable to the wavelengths of the propagating waves. Since seismic data typically contains useful frequencies in the range 5 – 100 *Hz* over media with acoustic wavespeeds varying between 1500 – 6000 *m/s*, this implies a wavelength range of at least 15 – 1200 *m*. This in turn implies that this

“high frequency” approximation will be valid when the medium variability is only on scales of hundreds of meters or larger.

The eikonal equation allows for solution of $\phi(\vec{x})$ throughout the medium. Reference to equation 1.16 reveals that this $\phi(\vec{x})$ is in fact the *traveltime* of the wavefield. By solving for the evolution of the phase of the wavefield, we are effectively solving for the first-arrival traveltime of the high-frequency wavefront. That is, the eikonal equation allows us to determine exactly the minimum time that it will take for theoretical high-frequency wave energy to arrive at any given point in the medium from some chosen starting location.

As a simple example, consider $v(\vec{x}) = v_0$ as a constant for all waves within the medium, then the magnitude of the gradient of the traveltime is simply a constant,

$$|\nabla\phi(r)| = |\partial_r\phi(r)| = v_0^{-1}, \quad (1.22)$$

where we have used the polar form of the gradient, noting that the angular derivatives are zero for this isotropic medium. This gives a traveltime from the origin to the distance r ,

$$\phi(r) = rv_0^{-1} + \phi_0. \quad (1.23)$$

The *transport equation* is derived by considering only the second term in equation 1.19, noting that the imaginary part must be equal to zero, and multiplying by $A\omega/i$. Upon restoring to full dimension,

$$2\nabla A \cdot \nabla\phi + A\nabla^2\phi = 0. \quad (1.24)$$

Again, with reference to equation 1.16, it is clear that this transport equation is solving for the amplitudes, $A(\vec{x}, \omega)$, of the assumed harmonic solutions throughout the domain. Note that solution of this transport equation requires prior solution of the eikonal equation (or some other determination of ϕ).

By solving for both $A(\vec{x}, \omega)$ and $\phi(\vec{x})$, we can approximately solve for the full wavefield, $\psi(\vec{x}, \omega)$, with validity in the high-frequency limit. The eikonal equation will be discussed, along with extensions to improve its applicability to the exploration seismic context, in Chapter 2.

1.4.3 The Helmholtz equation

An alternate form of approximate solution to the wave equation comes via the pseudodifferential operator calculus, applied to the Helmholtz equation.

Throughout this dissertation we will consider a simple (but powerful) transformation of the scalar wave equation: the Helmholtz equation. The Helmholtz equation is a frequency-domain representation of the scalar wave equation:

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \Psi(\vec{x}, t) = 0. \quad (1.25)$$

In a homogeneous medium, solutions to the scalar wave equation are quite simple. Solutions will be a simple function of the form $f(|\vec{x}| \pm ct)/|\vec{x}|$. These solutions are frequently taken to be exponential functions, in part to take advantage of the great power and flexibility of Fourier analysis. In fact, if we assume that the wave solutions are harmonic in time, this equation can be transformed into the Fourier domain. Once transformed via $t \rightarrow \omega$, the scalar wave equation becomes the Helmholtz equation,

$$\left(\frac{\omega^2}{c^2} + \nabla^2 \right) \psi(\vec{x}, \omega) = 0, \quad (1.26)$$

This simple transform dramatically changes the nature of the equation. The scalar wave equation propagates disturbances or discontinuities, such as those created by an instantaneous impulsive source modelled by a delta function, through

the spatial domain in time. The Helmholtz equation smoothes and spreads the frequency spectrum across the spatial domain and converges on a stable solution. That is, the wave equation marches a wavefield forward in time across space, while the Helmholtz equation spreads frequency content across space in a sort of static solution.

Once transformed into the Helmholtz equation, it is possible to conceptually resolve the solutions into two distinct components: the up- and down-going wavefields. In a homogenous medium, this is a simple operation. Since there are no changes in the medium whatsoever, waves travelling in one direction will continue travelling in that direction forever. When there are changes in the medium, however, some components of the wave will reflect or refract back in the opposite direction. If there is more than one spatial dimension, changes in the medium will also spawn evanescent waves. This complicates matters significantly. Pseudodifferential solutions that address the Helmholtz equation will be explained in §1.5.2.

1.5 Seismic imaging

Seismic imaging may be defined as any process that transforms seismic reflection data into a useful image of the subsurface over which the survey was performed. There are many different imaging algorithms, but virtually all of them rely upon the approximate solution of the wave equation within an estimated version of the subsurface below the location of the survey.

1.5.1 The profile imaging scheme

“Profile imaging” in the context of seismic exploration is fully described by Claerbout (1985). Briefly, profile imaging is a method of seismic data migration (imaging) in which a surface-recorded seismic wavefield is considered to be the upgoing (“upcoming” in Claerbout’s words) wavefield, and a modelled source approximating the original survey source is considered to be a downgoing wavefield. These wavefields are compared at depth in the subsurface in an attempt to recover the reflection coefficient at this depth. This method requires an approximate model of the subsurface wavespeeds at all locations.

The principle relies on the concept that, at each depth in the subsurface, a downgoing wavefield is transformed into an upgoing wavefield via reflection. A disturbance that is coincident at the same place and time in both wavefields is such an event which means that there is a finite reflection coefficient at that location.

The reflection coefficient ratio commonly referred to as “Claerbout’s Imaging Principle” is $\frac{U(\omega, x, z)}{D(\omega, x, z)}$, where U and D represent upgoing and downgoing wavefields respectively, Fourier transformed $t \rightarrow \omega$ (Claerbout, 1971). This ratio is not practical, however, as (for example) the denominator may be zero, leading to instabilities. These instabilities may be reduced by adding a small amount of white noise to the signal – a “prewhitening” or “water level”. Another approach is to use the zero lag of the cross-correlation, $U(\omega, x, z)D^*(\omega, x, z)$, where D^* represents the complex conjugate of D . This preserves the phase of the original ratio, and improves the stability of the operation. As long as the phase is preserved, the time/space location of the coefficient will be preserved, though the exact amplitude may be somewhat altered. In the case of seismic imaging, the accurate location of anomalies is typi-

cally considered a higher priority than the absolute preservation of the amplitude of the reflection coefficient. Amplitudes are certainly important, however, and there are many other *imaging conditions* which attempt to protect amplitude information while providing a stable result. See, for example, Rietveld (1995).

In exploration reflection seismology, geophones record the reflected wavefield. Neglecting surface waves (*e.g.* ground roll), air blast, and other sources of noise; and ignoring surface topography, recording problems, and other real-world issues, the recorded wavefield in an ideal survey is effectively the upgoing wavefield at depth $z = 0$. That is, we record directly $U(x, z = 0, \omega)$ – or at least, we record noisy time-domain data that may be Fourier-transformed directly to approximately U with appropriate filtering and noise-removal. The downgoing wavefield at $z = 0$ may be estimated by mathematically modelling the source. Given both wavefields at the surface, all that remains is to extrapolate these wavefields to depth, and compare them at each step along the way. It is this depth extrapolation step that encompasses the solution of the wave equation throughout the estimated subsurface. Also, this extrapolation requires an approximate model of the subsurface wavespeed at all subsurface imaging locations (the “velocity model”).

Because of the requirement for a velocity model, the profile imaging scheme essentially is a method of locating the boundaries between rock layers. Equivalently, the profile imaging scheme locates the reflectors within a smoothed version of the original velocity model – an accurate starting velocity model is required for profile imaging and many other migration schemes. For details on velocity modelling procedures used in modern seismic exploration, see Yilmaz (2001) or Robein (2004) and references within.

It may seem trivial to migrate data when an existing velocity model already

appears to tell exactly where the layer boundaries occur – why go through the imaging process when we already have a good enough idea of the geology to build a velocity model? In practice, however, the migration effectively reveals where the seismic data *disagrees* with the velocity model in some fashion. For example, if we assume flat plane homogeneous layers in the model, but in reality one layer has a slight departure from this assumption – perhaps a small localized channel in one layer, formed by erosion from the flow of an ancient river – then the migration will locate not only the layer that we know about, but it will also locate with reasonable accuracy the size and shape of this channel. Some slight error will occur due to the initial modelling error, but this can be either ignored or accommodated by repeated iterations of the imaging/velocity-modelling stage. Furthermore, velocity models used for migration are often smooth, without hard boundaries between layers. The migration is often used to precisely locate the boundaries between the layers.

1.5.2 Depth extrapolation

Many current wavefield extrapolation migration algorithms (see e.g. Gazdag, 1980; Gazdag and Sguazzero, 1984; Berkhout, 1984; Holberg, 1988; Stoffa et al., 1990; Hale, 1991; Wu, 1994) are space-frequency methods related to or derived from the phase-shift method introduced by Gazdag (1978). In this method, the wavefield is Fourier-transformed over time ($t \rightarrow \omega$) and the lateral spatial coordinates $(x, y) \rightarrow (\xi, \nu)$, resulting in a plane-wave decomposition. This plane-wave version of the equation can greatly simplify solution.

Phase-shift

Under the assumption of a constant velocity (i.e., acoustic wavespeed) at a given depth, an appropriate phase-shift is applied to each Fourier plane wave to extrapolate the wavefield a single step in the vertical (z) direction, that is across a homogeneous layer. In geologic regions where this approximation is valid, for example in a region with flat plane homogeneous layers commonly found in sedimentary basins such as the Western Canadian Sedimentary Basin, this is a useful approximation.

The scalar Helmholtz equation, 1.26, may be easily solved when the velocity c is constant. A two-dimensional wavefield at depth z may be extrapolated to a wavefield at depth $z + \Delta z$ by simply phase-shifting each frequency through the appropriate Δz step,

$$\hat{\psi}(\xi, z + \Delta z, \omega) = \hat{\psi}(\xi, z, \omega) \exp(i\Delta z \sqrt{\omega^2/c^2 - \xi^2}). \quad (1.27)$$

This solution is nothing more than the convolution (*i.e.* Fourier-multiplication) of the input wavefield with the delta-function solution of extrapolation from z to $z + \Delta z$, and it assumes that waves are propagating in the positive z direction. This concept will be more thoroughly and generally discussed later in this section during the discussion of GPSPI.

This method is inexpensive computationally to calculate: simply Fourier transform the original recorded data from $t \rightarrow \omega$ and $x \rightarrow \xi$, and multiply by an exponentiated value (which may be precalculated and tabulated). This yields the new wavefield $\hat{\psi}(\xi, z + \Delta z, \omega)$. This new data may be again multiplied by this same exponential, successively advancing the wavefield deeper and deeper into the section.

So far we have made no restriction on the relationship between ω , c and ξ . This suggests that it is possible that $\sqrt{\omega^2/c^2 - \xi^2}$ may take imaginary values. In this case, the argument of the exponential becomes real and negative. Then propagation of

the wavefield changes from an effective phase rotation into an exponentially damped propagation – *i.e.* evanescent propagation. At the point where $\sqrt{\omega^2/c^2 - \xi^2} = 0$, we define the evanescent boundary.

This method may be expanded beyond a homogeneous medium into one that varies in the z direction simply by allowing c to vary in each step of Δz . In this case, however, only *transmission* is modelled. This method ignores all reflections, as it can only propagate a wavefield in a single direction – *i.e.* in the positive z direction, which is traditionally taken to be the downward vertical direction. This method also ignores transmission coefficients, as no effort is taken to diminish the amplitude of the transmitted wavefield between each Δz step.

Split-step Fourier

The constant-velocity-in- x phase shift method may be extended simply but powerfully by the split step method of Hardin and Tappert (1973). If we are to extend equation 1.27 to handle a variable velocity in x , we must first allow for c to be a variable in the result. Therefore, we will inverse-transform the result, and allow for variation in the velocity so that $c = c(x)$,

$$\psi(x, z + \Delta z, \omega) = \int_{\mathbb{R}} e^{i\Delta z \sqrt{\omega^2/c(x)^2 - \xi^2}} \hat{\psi}(\xi, z, \omega) e^{i\xi x} d\xi. \quad (1.28)$$

This integral is extremely costly to compute, though it will be later discussed in detail as Generalized Phase-Shift Plus Interpolation (GPSPI).

This integral may be simplified considerably if the x dependence can be separated from the ξ dependence, thus removing it from the integral. To do this, we begin with the square root term in equation 1.28 as,

$$\sqrt{\frac{\omega^2}{c(x)^2} - \xi^2}. \quad (1.29)$$

Identifying slowness $S(x) = 1/c(x)$, and moving the factor $\omega S(x)$ out of the square root,

$$\sqrt{\frac{\omega^2}{c(x)^2} - \xi^2} = \omega S(x) \sqrt{1 - \frac{\xi^2}{\omega^2 S(x)^2}}. \quad (1.30)$$

Now we treat slowness as a perturbation around a reference slowness, such that $S(x) = S_0 + \Delta S(x)$,

$$\sqrt{\frac{\omega^2}{c(x)^2} - \xi^2} = \omega S_0 \sqrt{1 - \frac{\xi^2}{\omega^2 S_0^2}} + \omega \Delta S(x) \sqrt{1 - \frac{\xi^2}{\omega^2 S(x)^2}}. \quad (1.31)$$

Now we take the first term in expression 1.31 and multiply ωS_0 back into the square root. Also, we expand the second term in a binomial approximation, so that

$$\sqrt{\frac{\omega^2}{c(x)^2} - \xi^2} \approx \left(\omega^2 S_0^2 - \frac{\xi^2 S_0^2}{S(x)^2} \right)^{1/2} + \omega \Delta S(x) \left(1 - \frac{\xi^2}{2\omega^2 S(x)^2} \right). \quad (1.32)$$

If we assume that the perturbation $\Delta S(x) \ll S_0$, then $S_0^2 \approx S(x)^2$, simplifying 1.32 to

$$\sqrt{\frac{\omega^2}{c(x)^2} - \xi^2} \approx (\omega^2 S_0^2 - \xi^2)^{1/2} + \omega \Delta S(x) - \frac{\xi^2 \Delta S(x)}{2\omega^2 S(x)^2}. \quad (1.33)$$

The final term in expression 1.33 may be ignored in the case where horizontal wavenumber ξ is small, where the perturbation $\Delta S(x) \ll S(x)$, or when $\omega \rightarrow \infty$. Equivalently, we may expect this approximation to be poor when we are interested in low-frequency wavefields, when wavefields are propagating at high horizontal wavenumbers, or when the perturbations to the reference slowness S_0 are very large.

This approximation leaves the square root term in equation 1.28 greatly simplified:

$$\sqrt{\frac{\omega^2}{c(x)^2} - \xi^2} \approx \omega \Delta S(x) + \sqrt{\frac{\omega^2}{c_0^2} - \xi^2}, \quad (1.34)$$

where $c_0 = 1/S_0$. This equation is preferable to the original because it explicitly separates x and ξ dependence into two terms which are summed together, thereby allowing a simplification of the original integral.

Inserting this approximation back into equation 1.28,

$$\psi(x, z + \Delta z, \omega) \approx e^{i\Delta z \omega \Delta S(x)} \int_{\mathbb{R}} e^{i\Delta z \sqrt{\omega^2/c_0^2 - \xi^2}} \hat{\psi}(\xi, z, \omega) e^{i\xi x} d\xi. \quad (1.35)$$

This approximation results in a dramatic reduction in the cost of computation of the integral, because all x dependence is moved outside of the integral. If x dependence is left inside the integral, then we must perform this integral for every point x at which we would like to evaluate this equation. In practice, this amounts to performing an inverse Fourier transform for every x . However, with this x dependence outside, only one inverse Fourier transform must be calculated, followed by simple multiplication with an exponential for each output point x .

In this formulation, we refer to c_0 as the reference velocity. The algorithm extrapolates the entire wavefield with this one reference velocity, and then applies an inexpensive spatially-dependent correction at each output x location that simply requires the calculation of an exponential. This correction is closely related to the thin lens correction in optics (Hecht, 1997). A diagram showing the idea of the split-step method applied to phase-shift extrapolation can be seen in Figure 1.5. When the lateral velocity variation in the medium is small, the split-step correction offers an inexpensive method to achieve greater fidelity with the base phase-shift method.

There are many other published methods (*e.g.* pseudo-screens, phase-screens) that can be seen as generalizations and improvements upon the split-step method (see *e.g.*, and references within: de Hoop et al., 2000).

PSPI

Gazdag and Sguazzero (1984) extended the phase-shift method to accommodate lateral inhomogeneity in the layer by interpolating a final result from a suite of

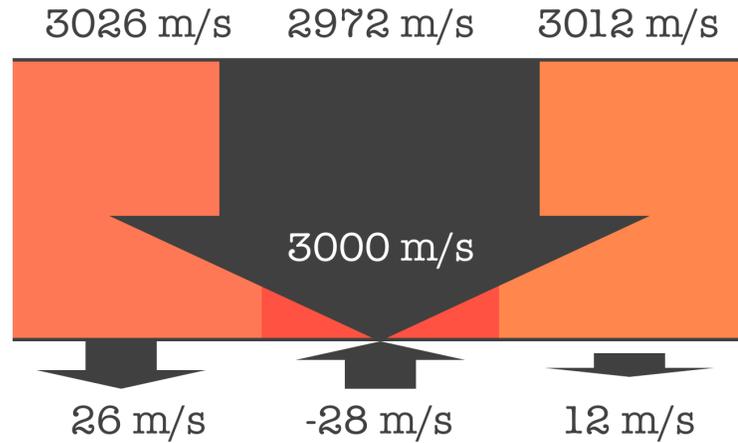


Figure 1.5: Phase-shift plus split-step corrections using a reference velocity of 3000m/s . One large step at the reference velocity is taken for the entire width of the survey, and then smaller corrections are made for the regions that vary around the reference velocity.

reference wavefields each extrapolated with a distinct well-chosen constant *reference velocity*. Typically, the number of reference wavefields calculated is much smaller than the number of individual unique velocities found in the layer. This extended method was called PSPI, an acronym for *Phase-Shift-Plus-Interpolation*. A wavefield can be marched through a highly heterogeneous medium by quantizing the velocity model in the vertical direction into a suite of laterally heterogeneous layers (i.e., no vertical variation) and recursively applying the PSPI method to step across each one. PSPI extends the usual phase-shift method by breaking a single step at a reference velocity c_0 into several steps, with the resulting wavefields interpolated together.

Naturally, the split-step correction may also be applied within the PSPI framework. PSPI plus split-step correction can yield impressive results even in extremely complicated media.

Generalized PSPI

Later Margrave and Ferguson (1999) developed GPSPI (generalized phase-shift plus interpolation), which can be considered the limiting form of PSPI when a unique reference velocity is used at each output point. Although the GPSPI algorithm is inherently unstable (Etgen, 1994; Wapenaar and Grimbergen, 1998; Le Rousseau and de Hoop, 2001; Ferguson and Margrave, 2006), the instabilities do not become significant when the extrapolation extends less than several hundred wavelengths, which is the case in most practical seismic imaging applications. Although GPSPI may be considered a limit of PSPI, and so heuristically “derived”, we will develop a more careful and mathematically enlightening derivation of it here.

Mathematically, 2D wavefield extrapolation from $z = 0$ to $z = \Delta z$ may be represented in abstract operator notation as

$$\Psi(x, z = \Delta z, \omega) = (\mathbf{T}_{\alpha(0:\Delta z)}\Psi)(x, z = 0, \omega) \quad (1.36)$$

where ω is temporal frequency, x is the lateral spatial coordinate, z is depth, and $\mathbf{T}_{\alpha(0:\Delta z)}$ is the wavefield extrapolation operator for a single step through a laterally variable medium from depth $z = 0$ to $z = \Delta z$. $\mathbf{T}_{\alpha(0:\Delta z)}$ operates at constant ω along the transverse (x) coordinates of the wavefield. In our approach we construct \mathbf{T}_{α} as a pseudodifferential operator (Martinez, 2002; Saint Raymond, 1991; Grigis and Sjöstrand, 1994) where the subscript α is called the operator symbol and is a function of position, wavenumber, and frequency that describes the physics of the propagating waves. In principle, it is possible to find exact symbols for highly complex lateral velocity variations (Fishman et al., 1997) which describe all internal scattering as well as primary transmitted waves. Here we are concerned with more approximate expressions.

Pseudodifferential operators

There are several texts which give an introduction to the formal mathematics of pseudodifferential operators, including Martinez (2002), Saint Raymond (1991), and Grigis and Sjöstrand (1994).

Generally, a pseudodifferential (PSDO) operator may be represented as

$$P(x, D)u(x) = \frac{1}{2\pi^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i(x-y)\xi} P(x, \xi) u(y) dy d\xi \quad (1.37)$$

where $P(x, D)$ denotes a pseudodifferential operator on \mathbb{R}^n . x is a coordinate, and D signifies that the operator includes derivatives. y represents the spatial coordinate variable before the operator's action. Thus, we will Fourier-transform from $y \rightarrow \xi$, and then inverse Fourier-transform $\xi \rightarrow x$. The distinction between x and y is made to explicitly set the functional dependence of $P(x, D)$ onto the output coordinate x rather than input coordinate y . Alternatives exist for P such that it may depend upon y or both x and y , and details of this may be seen in the standard PSDO references.

The square-root Helmholtz symbol

From the Helmholtz equation we may represent the second partial derivative of the wavefield in depth in PSDO form,

$$\frac{\partial^2}{\partial z^2} \Psi(x, z, \omega) = \int_{\mathbb{R}} \phi(\xi, z, \omega) \eta_2(x, \xi) e^{i\xi x} d\xi, \quad (1.38)$$

where $\phi(\xi, z, \omega)$ is the Fourier transformation of the wavefield $\Psi(x, z, \omega)$ over lateral spatial coordinate $x \rightarrow \xi$ and $\eta_2 = \left(\xi^2 - \frac{\omega^2}{v(x)^2} \right)$ is called the *symbol* of $\partial^2/\partial z^2$. Note that for this operator, we have effectively defined $P(x, D)$ to be the second partial derivative in z , and we have already Fourier-transformed Ψ into ϕ , thus implicitly one of the integrals in equation 1.37 has already been performed.

The notion of an operator symbol is familiar to most geophysicists in the simpler context of stationary filters. There the operator is convolution of a signal with a filter impulse response. The corresponding operator symbol is the multiplier which is applied in the Fourier domain to achieve the same result as convolution. Elliptic PSDOs are the most well studied. Informally, elliptic operators effectively are those which have symbols that grow quadratically with their arguments. A part of this includes that they are bounded away from zero except possibly at the origin. We assume ellipticity to allow the use of the asymptotic operator composition theorem which is a generalization of the convolution theorem for stationary filters (Saint Raymond, 1991, pg. 37). We note that η_2 is strictly *not* elliptic as it vanishes when $\left| \frac{\omega}{v(x)} \right| = |\xi|$, which corresponds to the evanescent boundary. Thus it is expected that the use of elliptic PSDO theory will lead to formulae that are not generally valid near the evanescent boundary.

The operator composition theorem gives an asymptotic expansion formula for the symbol of the PSDO resulting from the composition of two other PSDOs. The first term of this asymptotic expansion is simply the product of the two original symbols. Further terms may be considered, though careful treatment of the evanescent boundary must be undertaken (e.g. Stolk, 2004). We may treat the $\partial^2/\partial z^2$ operator as the result of the composition of $\partial/\partial z$ with itself (that is, the derivative operator is applied to the result of the derivative operator applied to a function),

$$\left(\frac{\partial}{\partial z} \frac{\partial}{\partial z} \right) \Psi(x, z, \omega) = \int_{\mathbb{R}} \phi(\xi, z, \omega) (\eta_1 \# \eta_1)(x, \xi) e^{i\xi x} d\xi, \quad (1.39)$$

where η_1 is the symbol of $\partial/\partial z$ and $\eta_1 \# \eta_1$ abstractly denotes the composition of operator symbols underlying the composition of their respective operators. Away from the evanescent boundary these operators are effectively elliptic. The leading

order term of the asymptotic expansion of $\eta_1 \# \eta_1$ is the product of the two symbols.

That is, $\eta_2 = \eta_1 \# \eta_1 \sim (\eta_1)^2$. So we conclude

$$\eta_1 \sim \sqrt{\eta_2} = \pm i \sqrt{\frac{\omega^2}{v(x)^2} - \xi^2}, \quad |\xi| < \left| \frac{\omega}{v(x)} \right|, \quad (1.40)$$

and we have an expression for the PSDO,

$$\frac{\partial}{\partial z} \Psi(x, z, \omega) \approx \int_{\mathbb{R}} \phi(\xi, z, \omega) \sqrt{\eta_2} e^{i\xi x} d\xi. \quad (1.41)$$

Choice of the square root branch is motivated by physical circumstances such as the direction of wavefield propagation or extrapolation. For the remainder, we choose the positive sign for convenience.

The GPSPI integral

To derive the GPSPI integral, we write an expansion for $\Psi(z + \Delta z)$ using a Taylor series,

$$\Psi(x, z + \Delta z, \omega) = \sum_{n=0}^{\infty} \Delta z^n \left(\frac{\partial}{\partial z} \right)^n \Psi(x, z, \omega), \quad (1.42)$$

where $\left(\frac{\partial}{\partial z} \right)^n$ denotes the n -fold self-composition of $\partial/\partial z$. Encouraged and emboldened by equations 1.40 and 1.41, we further assume that

$$\left(\frac{\partial}{\partial z} \right)^n \Psi(x, z, \omega) \approx \int_{\mathbb{R}} \phi(\xi, z, \omega) (\sqrt{\eta_2})^n e^{i\xi x} d\xi. \quad (1.43)$$

With equation 1.43 we can treat each term of equation 1.42 as if it were a PSDO.

If the order of integration and summation are exchanged we obtain an integrand which can be summed in closed form to an exponential,

$$\Psi(x, z + \Delta z, \omega) \approx \int_{\mathbb{R}} \phi(\xi, z, \omega) \exp(\Delta z \sqrt{\eta_2}) e^{i\xi x} d\xi. \quad (1.44)$$

Equation 1.44 represents an explicit form of equation 1.36, is known as the *locally homogeneous approximation* (Fishman et al., 1997), and is written by Margrave and

Ferguson (1999) as,

$$\Psi(x, z = \Delta z, \omega) \approx \Psi_{LH}(x, z = \Delta z, \omega) = \int_{\mathbb{R}} \phi(\xi, z = 0, \omega) \alpha(k(x), \xi, \omega, 0 : \Delta z) e^{i\xi x} d\xi, \quad (1.45)$$

where ξ is the wavenumber dual to x , and the symbol $\alpha(k(x), \xi, \omega, 0 : \Delta z)$ is

$$\alpha(k(x), \xi, \omega, 0 : \Delta z) = \begin{cases} e^{\Delta z \eta(x)}, & |\xi| \leq k(x) \\ e^{-|\Delta z \eta(x)|}, & |\xi| > k(x) \end{cases} \quad (1.46)$$

where $\eta(x) = i\sqrt{k(x)^2 - \xi^2}$ and $k(x) = \frac{\omega}{v(x)}$.

This representation is called ‘‘locally homogeneous’’ because the form of the symbol is mathematically the same as for the exact homogeneous case (Gazdag, 1978) except that the actual velocity function $v(x)$ is substituted for the homogeneous velocity. Margrave and Ferguson (1999) call the algorithm derived using the locally homogeneous operator *generalized phase-shift plus interpolation* (GPSPI) because the formula can be derived by eliminating the explicit interpolation in PSPI by taking the limiting form when a unique reference velocity is used for each output point.

There is one minor mathematical detail that must be addressed. Equation 1.40 explicitly excludes the evanescent boundary, *i.e.* when $\omega^2/v(x)^2 = \xi^2$. This suggests that our approach may be strictly invalid for propagation at the evanescent boundary. This is not surprising from a physical point of view. Physically, this one-way extrapolation of wavefields is designed to propagate vertically. The evanescent boundary, however, is reached in the limit where the propagation travels perpendicular to our design direction, *i.e.* horizontally. Intuitively we can understand why this causes a problem: it is nonsensical to calculate the relative vertical phase shift that is incurred by a monochromatic plane wave component travelling horizontally. As the planewave orientation approaches horizontal, the vertical phase

shift approaches infinity, and is strictly undefined. Beyond horizontal, however, the physics of extrapolation transforms from wavelike behaviour into evanescent, exponentially decaying behaviour as we expect. In practice, it is usually enough to suppress extrapolation in a small neighbourhood around the evanescent boundary with an exponential-decay (damping) term.

1.5.3 The FOCI algorithm

The GPSP algorithm describes a full Fourier-domain multiplication. That is, in order to find the wavefield $\Psi(x, z + \Delta z, \omega)$, we take the original wavefield at depth z , Fourier-transform it over time and space to frequency and wavenumber, multiply by the symbol α , and then inverse-transform back to time and space. Every Fourier domain multiplication can also be implemented as a convolution in its conjugate domain.

The FOCI algorithm described in Margrave et al. (2005) and Margrave et al. (2006) does exactly this for the space-frequency domain. That is, it uses an equivalent operation that takes place in the $\omega - x$ domain as

$$\Psi(x, z + \Delta z, \omega) = \int_{\mathbb{R}} \Psi(x', z, \omega) W(k(x'), x - x', z : z + \Delta z) dx' \quad (1.47)$$

where

$$W(k(x'), x - x', z : z + \Delta z) = \int_{\mathbb{R}} \alpha(k(x'), \xi, \omega, z : z + \Delta z) e^{i\xi(x-x')} d\xi \quad (1.48)$$

is the nonstationary convolution kernel of the lwKBJ operator in the $\omega - x$ domain.

1.5.4 Plane wave migration

The previous sections dealt with the extrapolation of wavefield via plane wave decomposition. That is, we break down the wavefield into individual monochromatic

plane waves – points in $\omega - \xi$ space. Data in this domain is transformed via extrapolation and filtering operators. However, in many cases, we may suspect that very few plane waves would be required to actually image the subsurface.

For example, consider geology consisting mainly of horizontal planar homogenous layers, as might be found in a sedimentary basin. In this circumstance, if a source wavefield approximating a single horizontally oriented (vertically travelling) plane wave was generated for seismic imaging, the recorded wavefield would also be a horizontally-oriented plane wave. In this case, all relevant geologic information would be imaged, and the wavefield extrapolation algorithm would be reduced to a one-dimensional solution of the wave equation, thus drastically reducing the required computation time.

Unfortunately, practical concerns make this a challenging imaging procedure to implement. The actual physical generation of a coherent plane wave hundreds to thousands of meters across is difficult to imagine, though it could perhaps be accomplished approximately by the simultaneous detonation of hundreds or thousands of point sources¹. In this insight, however, there is a clue.

We can simulate this synthesis of a plane wave by the combination of individual shot gathers. That is, the recorded seismic data from all individual point-source generated shots may be directly summed to approximate what *would* have been recorded had all shots been fired simultaneously. Thus, if all seismic data from all shots is summed, and if a source model is generated based on all of the locations of all source points used, then profile imaging may proceed as usual. The computation of the image from this single simulated record is approximately the same cost as the cost of the computation of the image from a single shot. In a survey of hundreds to

¹or the simultaneous operation of hundreds or thousands of controlled vibration sources.

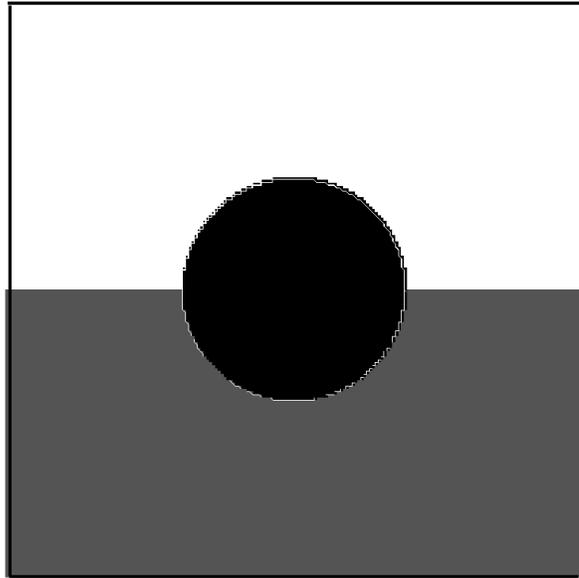


Figure 1.6: Plane wave migration velocity model. The white region represents a relative velocity of 4, grey represents 6, and black represents 3.

thousands of shots, we have therefore reduced our computation costs by a factor of hundreds to thousands. This approach is only valid, obviously, if the data resulting from the superposition of two individual shot gathers is equivalent to the data resulting from shooting both individual sources simultaneously. The validity of this linearity assumption in the wavefields is not proven, but is widely believed to be true in exploration geophysics.

Of course, an economy is only realized if the single plane wave migration yields a useful image. Consider a seismic survey over a geologic structure as shown in Figure 1.6. Obviously this is not a realistic velocity model, but the shapes of the structures will highlight various strengths and weaknesses of the plane wave imaging approach.

This velocity model will not be adequately illuminated entirely by a single hori-

zontal plane wave travelling vertically. However, if we were to additionally consider plane waves oriented at angles, we could add more information to the imaging. In the exhaustive limit in which all orientations were added, we would recover essentially the full quality of the traditional shot-profile method. These variously oriented plane waves are constructed by summing the shot gathers with a spatially-varying time delay, guided by the same time-delay method that would be used to construct an approximate oriented plane wave source from the point sources.

A standard shot-profile image generated by the migration of all 51 shots into 200 receivers is shown in Figure 1.7, which may be compared to Figure 1.8, an image constructed of only the horizontal plane wave. Shots and receivers were both equally spaced across the entire width of the velocity model. The horizontal reflector is imaged in the migration, as are the top and bottom of the circular anomaly. As there is very little direct energy travelling directly from source to near-vertical features on the circular and back to receivers, these features are not well imaged in this migration.

The single horizontal plane wave image shows significant structure but does not have the definition throughout, nor does it show much detail of the circular feature in the centre. However, as we add more plane waves, the image develops. Images derived from three, five, seven, and fifty-one plane waves may be seen in Figures 1.9, 1.10, 1.11, and 1.12 respectively. Although numerical artifacts are present in the image, the features within the model are clear.

1.5.5 Kirchhoff Migration

An alternative method of migration is known as *Kirchhoff migration*. Bleistein et al. (2001) provides a thorough introduction to the theory of this method, and

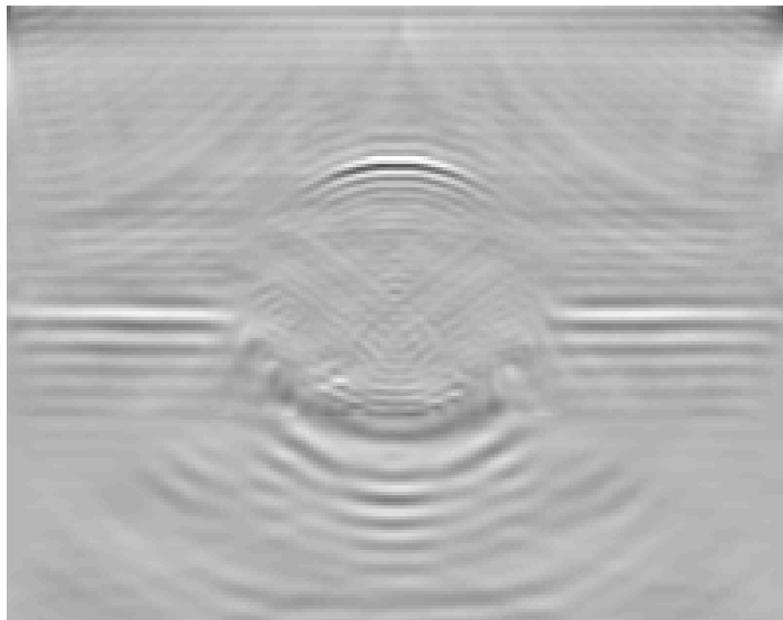


Figure 1.7: Full shot profile migration of the simulated survey. The survey consisted of 51 shots equally spaced across the entire surface shown, with 200 receivers spread equally across the entire span of the velocity model.

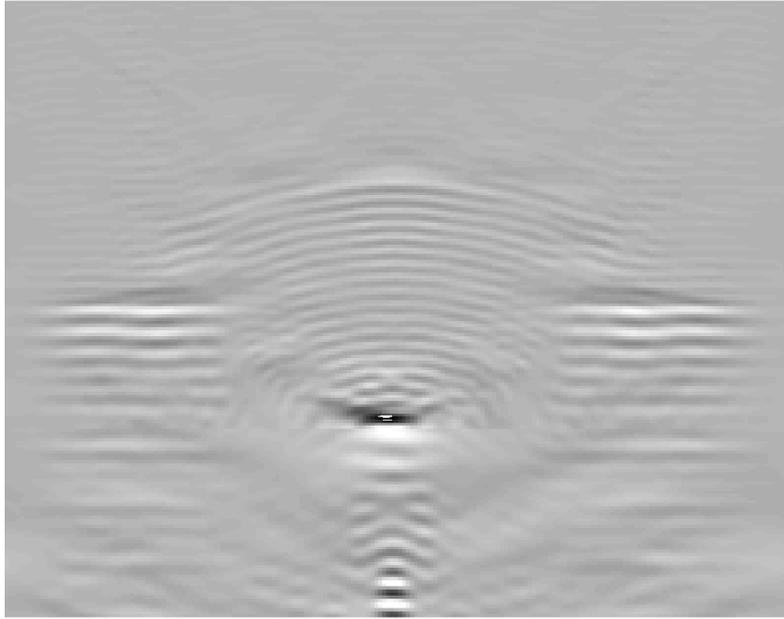


Figure 1.8: Horizontal plane wave migration

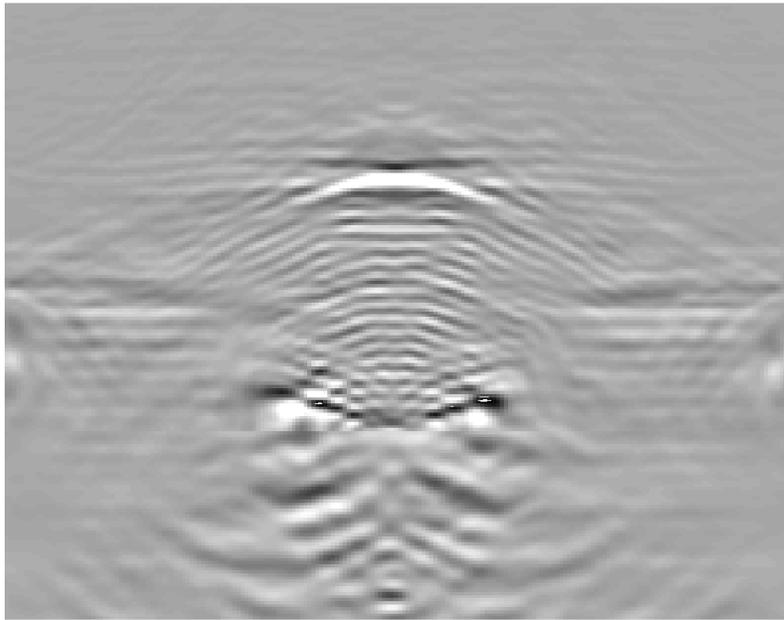


Figure 1.9: Three plane wave migration, orientations at 0° and $\pm 31^\circ$.

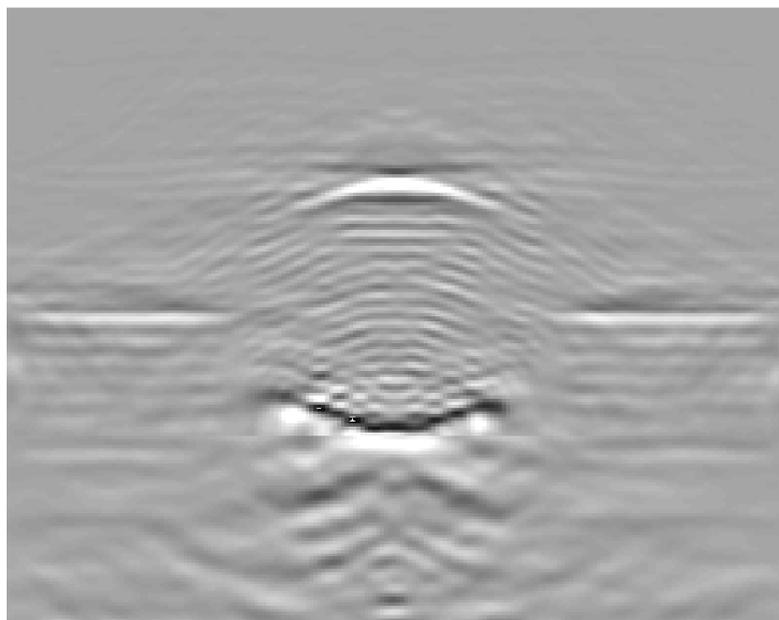


Figure 1.10: Five plane wave migration, orientations at 0° , $\pm 11^\circ$, $\pm 31^\circ$.

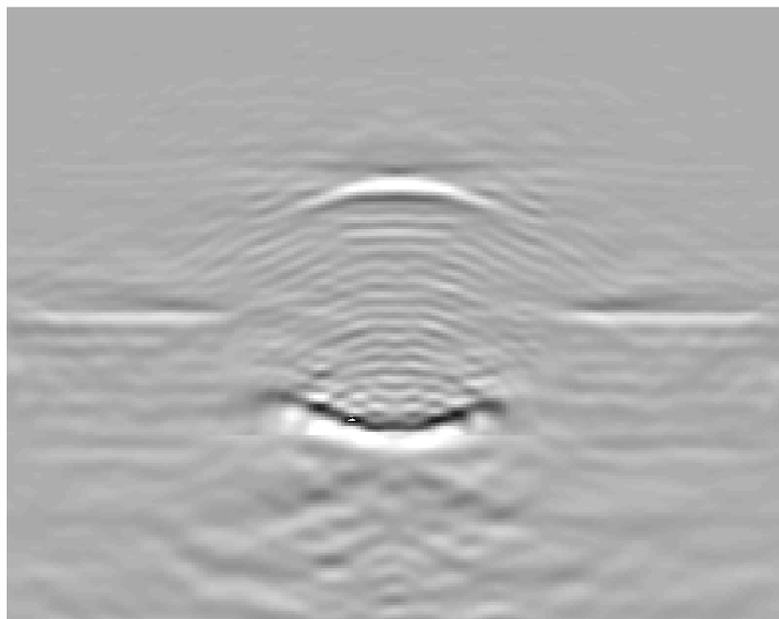


Figure 1.11: Seven plane wave migration orientations at 0° , $\pm 22^\circ$, $\pm 11^\circ$, and $\pm 31^\circ$.

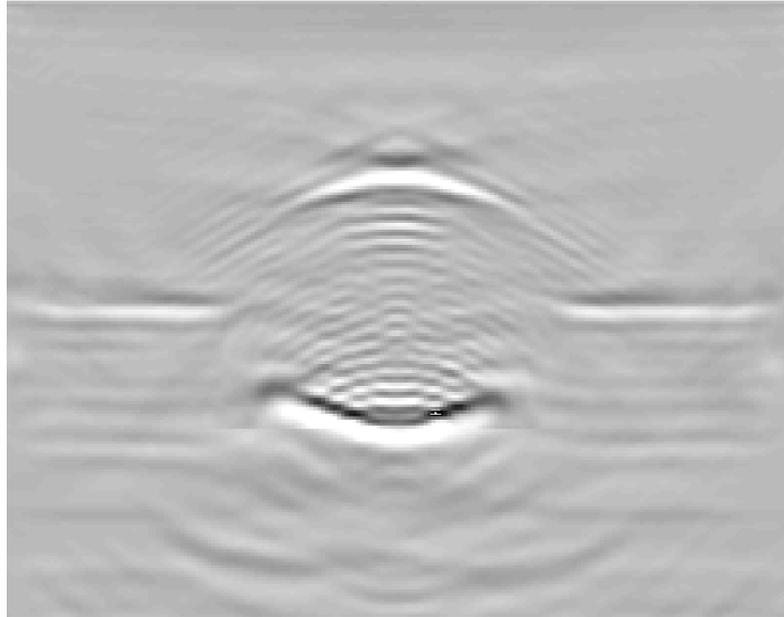


Figure 1.12: Fifty-one plane wave migration, orientations equally spaced between -31° and $+31^\circ$.

this will not be reproduced here. However, I will provide a heuristic introduction to the concepts behind this method by describing a particular implementation of Kirchhoff migration.

Essentially, the Kirchhoff algorithm requires the transformation of time-recorded reflection data into an accurate image of the subsurface reflectors. However, this is trivially true of all imaging algorithms. The Kirchhoff method attempts to accomplish this goal directly, rather than by recursive marching.

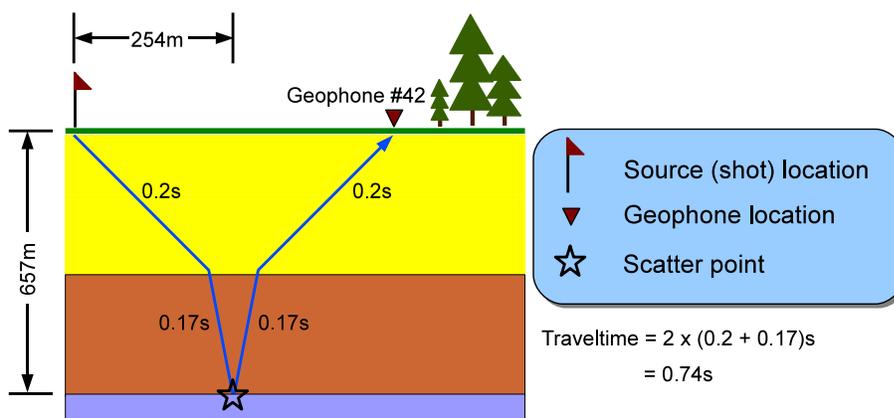
Given the locations of sources, the locations of receivers, and a model of the underlying geologic features, it is possible to predict where and when on a *e.g.* shot gather the energy from a given reflection event will be located – *i.e.* at the scatter point. Figure 1.13 illustrates this idea. Once the appropriate energy is located, this energy may be scaled according to the appropriate physics model used, and

then placed at the scatter point location in the image matrix (Figure 1.14). Note that although Figures 1.13 and 1.14 show the results from the specular raypaths, all possible raypaths may be used to allow for the accurate imaging of diffracted events.

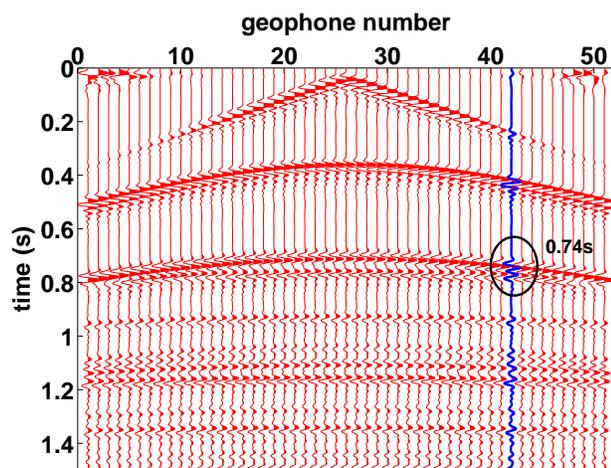
With a mental model of this concept in mind, we may be more specific about the actual calculation details. Given the source location at $(x_s, 0)$, a receiver location $(x_r, 0)$, a subsurface location (x_0, z_0) , and a model of the underlying geology, we can estimate the time taken for the source energy to travel from source location, down through the earth to a scatterpoint at (x_0, z_0) , and back up to the receiver. The reflection energy appearing on the receiver's recording at the appropriate time can be placed into an image, with this energy at the subsurface "scattering" location. As with shot-profile imaging schemes, we assume that there is a reflection (or scattering) at every point in the subsurface. Then we calculate the approximate reflectivity at each point. This calculated reflectivity may be zero if there is no scattered data found at the recorded data. Figure 1.14 (b) reveals the top two reflecting layers placed at their correct location in the subsurface.

Therefore, the first part of Kirchhoff migration requires the computation of the estimated traveltimes from source to scatterpoint, and from scatterpoint to receiver. This traveltimes may be calculated in different ways, including solution of the eikonal equation and ray tracing. I will not cover ray tracing in this dissertation, but a thorough treatment may be found in Červený (2001).

Simply taking the data straight from the geophone recording is not sufficient to estimate reflectivity, however. Shot profile imaging based on downward continuation has an imaging condition to compare the downgoing wavefield to the upgoing wavefield and generate an estimate of reflectivity, and Kirchhoff migration has a

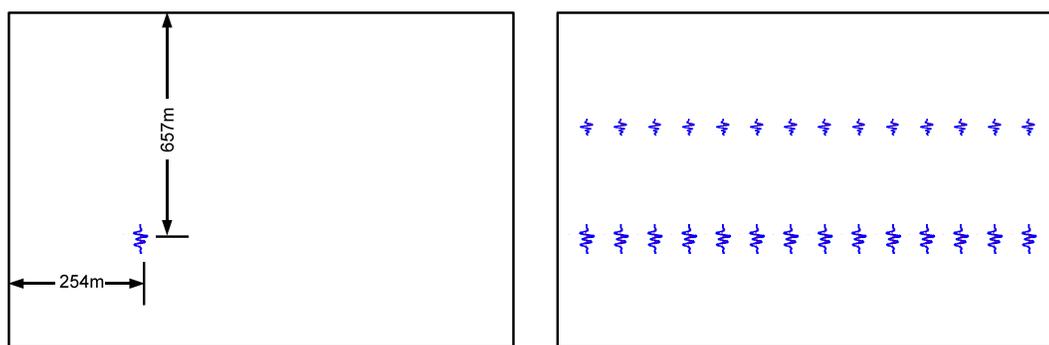


(a) Calculating traveltime from source to a subsurface scattering point at a known location, and back up to receiver.



(b) Identify the data at the calculated traveltime for the appropriate recorded trace.

Figure 1.13: Schematic illustration of the principle behind Kirchhoff migration. (a) shows the calculation of total traveltime from source to scattering point and back up to receiver, through the initial hypothesized velocity model. (b) shows the identification of the reflection event at the calculated traveltime and on the appropriate receiver trace.



(a) Place a scaled version of the data identified into the image matrix.

(b) Repeat for every subsurface location within the model.

Figure 1.14: A scaled version of the located energy (see Figure 1.13) is placed at its appropriate subsurface location in (a), and this process is repeated for every subsurface location in (b).

similar concept. In the Kirchhoff case, the outgoing wave is assumed to scatter at the scatterpoint, and the travel back up towards the geophone. In this situation, extensive physical arguments may be used to relate the amplitude of the recorded event to the reflection coefficient at the scatterpoint based on the orientation of the incoming and outgoing waves (rays), the assumed orientation of the reflector, and other assumed physical characteristics of the subsurface.

As with shot-profile migration, the Kirchhoff method treats the *location* of a reflection event separately from the *strength* of the reflection. The location of the event loosely corresponds to an accurate measurement of the phase of the recorded wavefield. The phase of the wavefield follows directly from the accurate estimation of the traveltime between source and receiver. As long as a reasonable background model is provided, the location of the reflection events (typically geological layer boundaries) can be determined. The amplitude of the reflection event holds information about the reflectivity – the stronger the event, the higher the reflectivity.

The information contained within the amplitude of the recorded wavefield is

historically more difficult to extract reliably than the information within the phase. For example, it is more robust to identify the arrival time of a particular reflection event within noisy data than it is to recover the “true” amplitude of this event in the presence of noise. Additionally, some of the physics involved with the determination of the reflectivity based on the recorded wavefields are sensitive to small errors – for example, a small error in travelttime or distance travelled due to an error in the velocity model will result in an amplitude that is anomalously high or low due to too much or too little correction for the usual geometric spreading found in wavefield propagation.

1.6 Inversion and early-arrival waveform tomography

1.6.1 Inverse problems

Inverse problems comprise an important part of applied mathematics. This brief introduction to inverse problems may be supplemented by other standard texts such as Aster et al. (2005) and Tarantola (2005).

Inverse problems can be defined in terms of *forward problems*. Typically, in physical sciences and deterministic systems in general, a certain measurement based on the configuration of a system can lead to a set of unique and predictable data. For example, dropping an unknown massive object through an unknown height under the influence of a gravitational field in the absence of all other forces will result in that object experiencing a predictable acceleration. This constant of proportionality is defined as g , the acceleration of gravity. A mass m experiencing a force will accelerate according to Newton’s second law, $f = ma$, where in this case g takes the role of a .

The potential energy E_g of a mass suspended at a height h is defined as $E_g = hmg$. The kinetic energy K gained by the mass falling through this height is equivalent to the potential energy lost, and can be directly measured using the velocity v relative to the velocity of the mass before falling. This kinetic energy is defined as $K = 1/2mv^2$. Since the potential energy is transformed to kinetic energy, the values are equal, that is E_g (before falling) equals K (after falling). Equivalently, $mgh = 1/2mv^2$, or $v = \sqrt{2gh}$.

In this case, the *forward problem* requires taking the physical parameters of the system, m , g , and h , and determining the resulting falling velocity v . We note that actual knowledge of m is not necessary, but nonetheless it is a physical parameter of the initial system. The *inverse problem* for this system is the inversion of this process: given the resulting falling velocity v , can we reconstruct some or all of the initial physical parameters of the system?

Given that $v = \sqrt{2gh}$, we can simply invert this equation to solve for the normally-independent variables,

$$gh = \frac{v^2}{2}. \quad (1.49)$$

Equation 1.49 can be easily solved if we add some additional constraints – namely, that this observation was taken in a place of known gravitational constant g , for example, the surface of the earth. If we add this constraint, this equation becomes

$$h = \frac{v^2}{2g}. \quad (1.50)$$

Now we have a simple equation that can directly tell us an unknown physical parameter about the original system, h , simply by measuring the falling velocity v of the mass after it has accelerated within the assumed gravitational field g . However, there is one physical parameter in the initial system that is beyond our reach: the

mass of the object, m . No measurement of the falling velocity can ever reveal this mass, no matter how careful. Conversely, this may also be taken from the point of view that, no matter what the original mass of the object, we can uniquely determine the height through which it fell simply by measuring its falling velocity.

Modification of the original scenario can reveal more interesting information, however. Consider that, instead of measuring *velocity*, we somehow directly measure the *kinetic energy* after falling. Perhaps the falling mass slows by compressing an ideal spring, compressing the spring by a distance x . Imagine that we arrange the situation so that the height of falling is independent from the distance of compression of the spring: *e.g.* the mass is a sphere, and it falls into an ideal frictionless tube with a 90° bend that redirects the ball's velocity from a vertical drop into a horizontal direction without losses.

If the maximal compression of the spring is measured, then the kinetic energy is transformed into potential energy of the spring, E_s . Given the spring constant k , the potential energy of a spring is $E_s = 1/2kx^2$. This must equal the initial potential gravitational energy, E_g . Therefore,

$$1/2kx^2 = mgh. \quad (1.51)$$

If we rearrange this equation to put measurements on the left hand side, and system configuration and known variables on the right hand side,

$$x = \sqrt{\frac{2mgh}{k}}. \quad (1.52)$$

Again, the forward problem is entirely deterministic and easily solved. Solution of the inverse problem, however, is significantly more challenging.

The inverse problem for this system would be to determine m and h from measurement of the spring compression x , including knowledge of the spring constant

k and an assumption about the gravitational constant g . This yields an inverse problem of the form,

$$mh = \frac{kx^2}{2g}. \quad (1.53)$$

Now we have an interesting problem: mass and height are both unknown. The evaluation of the right hand side of Equation 1.53 tells us the product of m and h , but it cannot uniquely determine either m or h . No measurement of the spring compression (or indeed, any other direct measurement of the falling kinetic energy) can yield a unique m or h for the original system.

Now let us consider the combination of both approaches: measurement of falling distance, and direct measurement of kinetic energy. In this case, first we can directly obtain the falling distance via Equation 1.50. Now equation 1.53 becomes $h = \frac{kx^2}{2gm}$, and the height through which the mass has fallen is easily and uniquely determined.

General inverse problems

The unique determination of mass and falling distance is one simple example of a general class of mathematical inverse problems. Generally, an inverse problem is to find m given the equation,

$$d = \mathcal{G}(m), \quad (1.54)$$

where d is “data” (or measurements as in the previous example), m is a “model”, (or system configuration and known parameters as in the previous example), and \mathcal{G} is an operator that describes some general physical process or explicit relationship that transforms the model into the data under some observation process. In the falling mass scenario, data includes the falling velocity and kinetic energy of the mass. The model includes the gravitational constant. The operator would be the physics describing gravitational force, acceleration of a mass due to a force, and the

compression of a spring.

For another example, we may consider wave propagation. In this new scenario, the model m may consist of the underlying wavespeed of the medium through which the wave is propagating and an original source of energy, the data d would be the resulting propagated wavefield after a certain time, and \mathcal{G} would be the usual wave equation operator acting upon the original source for the duration of the specified time throughout the medium specified by the wavespeed.

Inverse problems are difficult for several reasons. Aster et al. (2005) cites several essential issues that must be carefully considered when solving any inverse problem. These include **existence**, **uniqueness**, and **stability**. There is also a challenge of **practical computability**.

For a solution to **exist**, this requires a model that *exactly* reproduces the observed data when transformed by the physics described by \mathcal{G} . Strictly speaking, existence is almost never entirely satisfied in real inverse problems. There are several reasons for this. For example, the data almost always includes some kind of noise that is not accounted for, or the physics is almost always an approximation of the real world situation under which the data was collected (*cf.* §1.3, in which a scalar wave equation is derived under the assumption that there is actually no finite displacement of the material. This is clearly a non-physical assumption, and therefore any operator \mathcal{G} based on this physics is an approximation to reality).

For a solution to be **unique**, the model must be the *only* model that could possibly reproduce the observed data when the operator \mathcal{G} is applied. In the case of the earlier example involving the falling mass, we saw a clear example of non-unique solutions when only the falling velocity of the mass was observed. In this case, it was possible to reproduce the data (*i.e.* the falling velocity) by dropping

any mass m through a height h . In this case, the m was not unique. In the case where we directly measured the kinetic energy K of the falling mass, the product mh was the result, leaving both mass and height without uniqueness. However, the combination of *both* direct kinetic energy measurement *and* falling velocity led to a unique constraint on both h and m .

For a solution to be **stable**, a small change in the data should lead to only a “small” change in the resulting model. If we consider a version of equation 1.54 that includes a true model, m_{true} , a small amount of noise, μ , an imaginary data measurement without any noise or error, d_{ideal} , and a “real world” measurement that includes noise d_{rw} , then

$$d_{\text{rw}} = \mathcal{G}(m_{\text{true}}) + \mu \quad (1.55)$$

$$= d_{\text{ideal}} + \mu \quad (1.56)$$

$$= \mathcal{G}(m_{\text{est}}), \quad (1.57)$$

where m_{est} is the model estimated from the noisy data. Stability requires that when μ is small, $m_{\text{est}} - m_{\text{true}}$ should also be small. In the case of our falling mass example, this is clearly a stable solution. If we add a small amount of error to the falling velocity or kinetic energy measured, the resulting mass and/or heights change by a finite amount.

A naturally unstable problem, however, is the process of seismic deconvolution. This procedure is illustrated in Figures 1.15a through 1.15f. Briefly, a seemingly innocuous signal (Figure 1.15a) is convolved with the filter shown in Figure 1.15b. When this result is deconvolved with the original signal, the exact filter is recovered (Figure 1.15f). However, when small random noise is added to the filtered signal, as in Figure 1.15d, the recovered signal (Figure 1.15e) is completely unrecognizable.

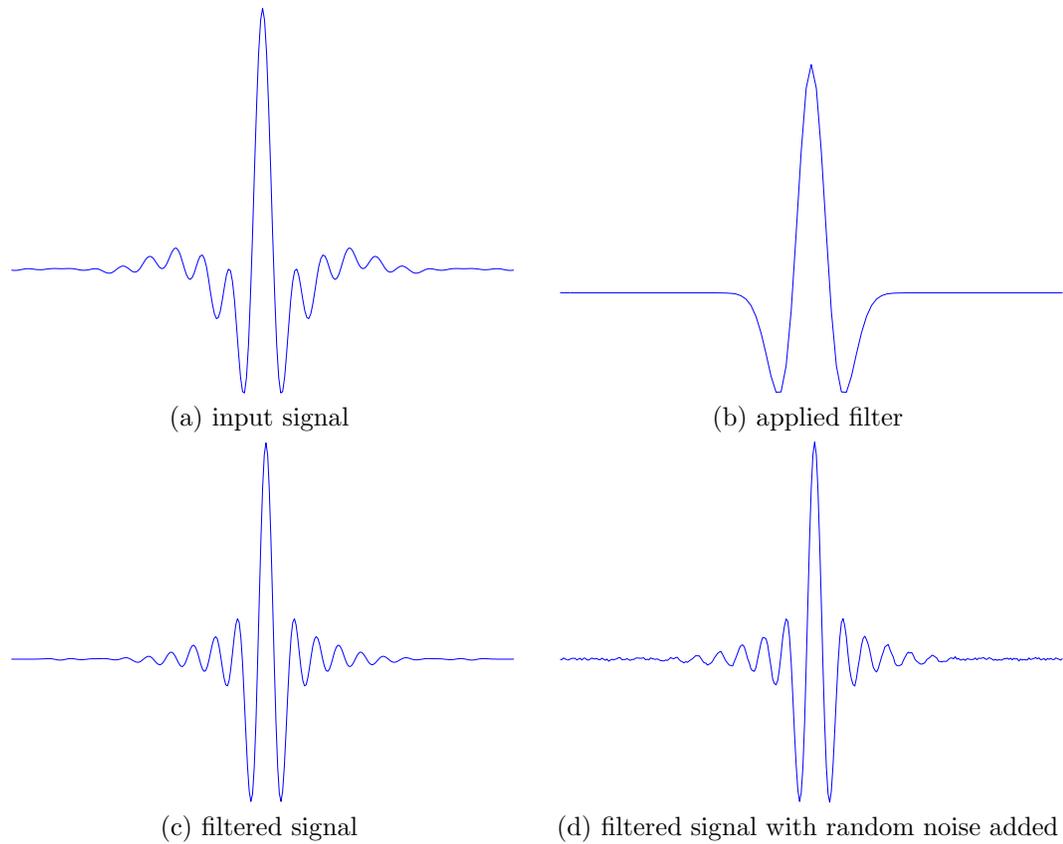


Figure 1.15: (Page 1 of 2) Deconvolution is an unstable process. The input signal in (a) is convolved with the filter in (b). This results in the signal shown in (c). A very small amount of random noise is added to this result, giving the signal shown in (d). This noisy signal is then deconvolved with signal (a), resulting in the (unstable) result in (e). Ideally it would result in a signal approximately like that in (f), which is the convolved signal shown in (c) deconvolved with (a).

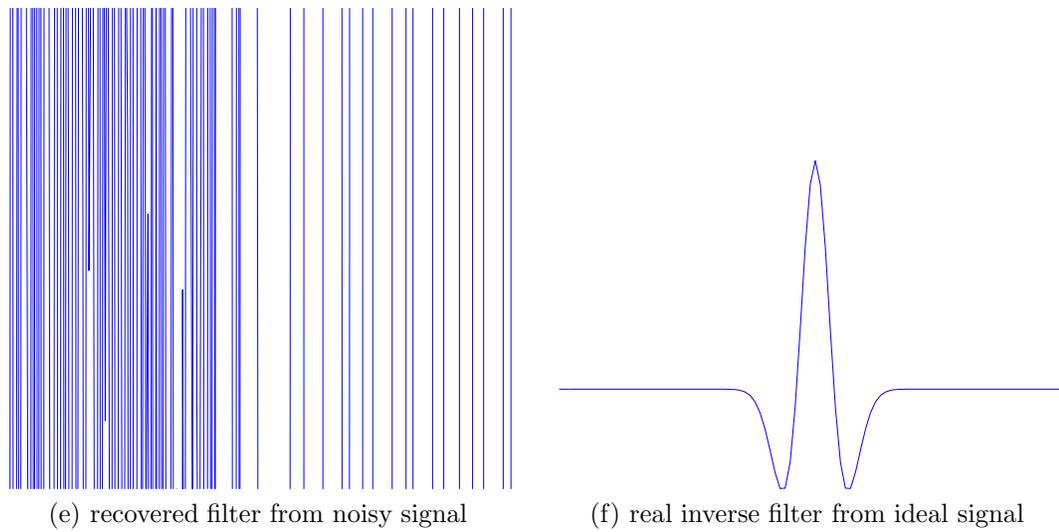


Figure 1.15: (Page 2 of 2) Deconvolution is an unstable process. The input signal in (a) is convolved with the filter in (b). This results in the signal shown in (c). A very small amount of random noise is added to this result, giving the signal shown in (d). This noisy signal is then deconvolved with signal (a), resulting in the (unstable) result in (e). Ideally it would result in a signal approximately like that in (f), which is the convolved signal shown in (c) deconvolved with (a).

For a solution to be **practically computable**, a numerical implementation of the solution must be computable within practical constraints such as time and computing resources. This is necessarily a poorly defined concept, and changes based on context and with computational technologies. However, in many cases in real world applications of inverse problem solutions, it is extremely important. Many theoretical solutions to inverse problems elegantly deal with existence, uniqueness, and stability challenges, but are infeasible to compute.

1.6.2 Early arrival waveform tomography

The Early Arrival Waveform Tomography (EAWT) method is an algorithm for treating seismic imaging as an inverse problem. The data collected d is composed of refracted (as opposed to reflected) seismic waves that have travelled from the source down into the subsurface and refracted back up towards the surface (Figure 1.16). The model is a model of the underlying subsurface wavespeed, describing the geology over which the survey is undertaken. The operator \mathcal{G} is effectively a simplified wave equation.

The initial work in this area was performed by Lailly (1983) and Tarantola (1984); a brief introduction to the history of the method is given by Pratt (1999). A full description of the core method is described in Pratt (1999), and an example of its application to crosshole data is given in the continuation of this paper, (Pratt and Shipp, 1999). In this introduction, I will briefly outline the method generally and heuristically, but for a mathematically thorough treatment of an example implementation, please consult these two papers.

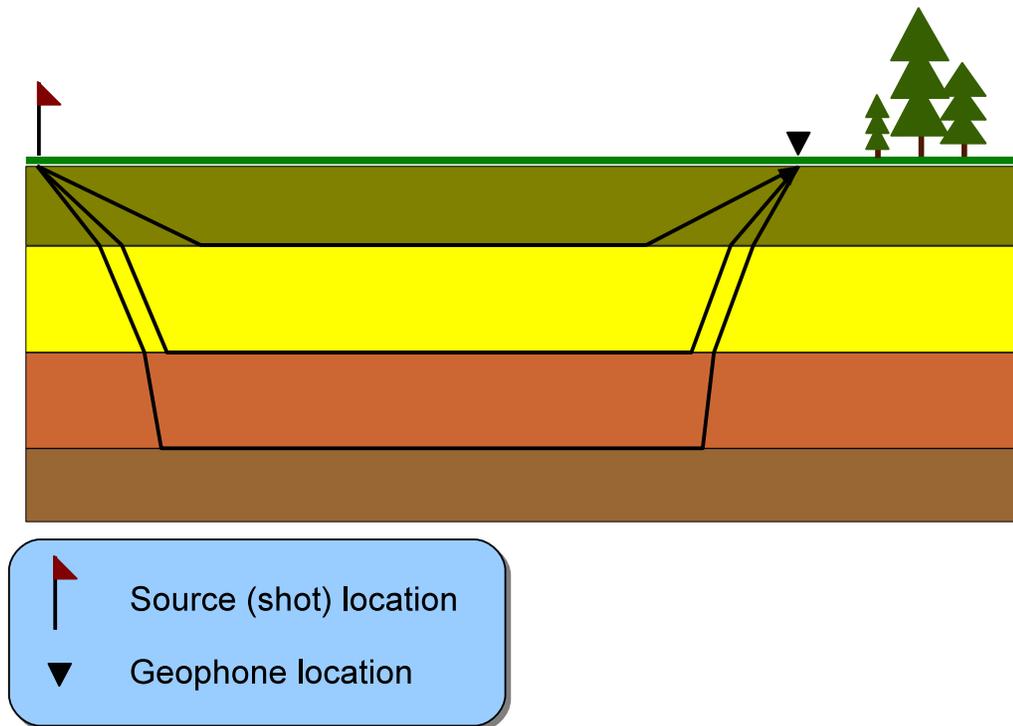


Figure 1.16: Seismic acquisition diagram showing three refracted arrivals traveling from source to receiver. Wave energy travels with wavefronts perpendicular to the black arrows, from the source down through the earth refracting at each layer interface, eventually being turned back up toward the receiver.

The inversion algorithm in EAWT

Here I will introduce the basic ideas supporting the process of inversion as applied to EAWT. For a more thorough introduction to the fundamental concepts, see Aster et al. (2005). This discussion is heavily inspired by Pratt (2008).

Consider the (ideal) recorded data $d = \mathcal{G}(m_{true})$, and synthetic data $u = \mathcal{G}(m_{est})$. Ideally, we hope to improve m_{est} so that the difference between u and d is reduced to a minimum. Inverse theory often introduces a *misfit* function, $E(\mathbf{m}) = \frac{1}{2}\delta\mathbf{d}^T\delta\mathbf{d}$, where \mathbf{d} is a vector representation of d . Minimization of E is a least-squares solution for the minimization of $\delta d_i = u_i - d_i$, which are known as the *residuals*. T signifies a conjugate transpose. In this context, then, solving an inverse problem amounts to finding a minimal $E(\mathbf{m})$.

As with all inverse problems, it is conceptually possible to test every \mathbf{m} and therefore exhaustively map the function $E(\mathbf{m})$. “Exhaustive” is the correct word for this, however, and for any but the most trivial problem this is practically impossible. Another approach would be to sample $E(\mathbf{m})$ sparsely throughout its domain – if a randomized approach is used to select points, this would correspond to a Monte Carlo method. Alternatively, we may attempt to search for the minimum “intelligently”. We may consider taking the value of $E(\mathbf{m})$ at a given location and then use other local information, such as derivatives, to find a likely direction in which to travel where we would expect to find a minimum. If we are minimizing this misfit, then perhaps we can simply drive “down hill” towards this minimum!

This approach naturally suggests an abstract Taylor series expansion of the misfit, denoting transpose as t ,

$$E(\mathbf{m} + \delta\mathbf{m}) = E(\mathbf{m}) + \delta\mathbf{m}^t\nabla_m E(\mathbf{m}) + \frac{1}{2}\delta\mathbf{m}^t\mathbf{H}\delta\mathbf{m} + \dots \quad (1.58)$$

where ∇_m is the gradient of the misfit with respect to \mathbf{m} ,

$$(\nabla_m E)_i = \partial_{m_i} E, \quad (1.59)$$

and \mathbf{H} is the Hessian – a matrix of second derivatives,

$$\mathbf{H}_{ij} = \frac{\partial^2 E}{\partial_{m_i} \partial_{m_j}}. \quad (1.60)$$

As with any minimization problem, to solve this we simply differentiate $E(\mathbf{m} + \delta\mathbf{m})$ with respect to model perturbations $\delta\mathbf{m}$, and then set the result equal to zero. This results in a standard expression,

$$\mathbf{H}\delta\mathbf{m} = -\nabla_m E(\mathbf{m}). \quad (1.61)$$

Since $\delta\mathbf{m}$ is the model perturbation that would modify the current approximate model into the model for the E approximated by the Taylor series, it is essentially the solution we are looking for. Obviously, the solution to this equation is

$$\delta\mathbf{m} = -\mathbf{H}^{-1}\nabla E. \quad (1.62)$$

By simple application of the inverted Hessian, the solution is complete. Unfortunately, actual inversion of the Hessian is completely impractical in any but the most trivial problem. The Hessian contains one column and one row for each parameter in the model. If our problem is the inversion of a recorded wavefield propagating through a velocity model with, say, 100 x 100 grid points, then this inversion would correspond to inverting the Hessian of size 10 000 x 10 000. This is numerically challenging, even if the Hessian is actually invertible mathematically. Furthermore, even this grid size is laughably inadequate to treat any realistic modern exploration seismic problem.

This Hessian inversion method is frequently referred to as Newton’s method, and in its simpler forms it sees frequent application in many solvers (e.g. Burden and Faires, 2004). In general, this method may be understood by recognizing that it is approximating the E function with a quadratic function (during the initial Taylor series expansion). Then it solves for the minimum of this approximating polynomial. In the case where the original function E is quadratic itself, then this method will converge in exactly one step.

While Newton’s method is impractical for large models, we still may gain insight from it, as we retreat into our original desire for a method that simply “drives downhill”. In fact, we can use this concept directly in an iterative method,

$$\mathbf{m}_k = \mathbf{m}_{k-1} - \alpha_{k-1} \nabla E_{k-1}. \quad (1.63)$$

That is, we update the model by driving into the direction of the gradient², and we take a step size α . By repeating this n times, we can approach the solution as long as the shape of the misfit function is convex.

The key, then, is to calculate both $\nabla_m E$ and α efficiently. Accomplishing these goals is strongly dependent on the exact nature of the inverse problem. For the EAWT problem, Pratt (2008) explains that the calculation of $\nabla_m E$ amounts to three steps:

1. Time-reverse the partial derivatives of the wavefield – which are closely related to diffracted wavefields.
2. Convolve these diffractions with residuals.
3. Sum these results over all receivers.

²or more correctly, exactly *against* the gradient, as the gradient points in the direction of greatest increase.

The α term may be calculated in several different ways. It would be possible to execute a line-search strategy to find an approximate α by simply testing several values. Pratt (2008) suggests the use of an estimate,

$$\alpha = \frac{|\nabla E_m|^2}{|\mathbf{J}\nabla_m E|^2}, \quad (1.64)$$

where $\mathbf{J}_{ij} = \frac{\partial u_i}{\partial m_j}$ is the Fréchet derivative matrix, and each column in \mathbf{J} contains a partial derivative wavefield. These partial derivative wavefields may be thought of as diffracted wavefields, representing changes in the wavefield due to an infinitesimal perturbation of the underlying velocity model.

Pratt (2008) also describes the actual implementation of the gradient calculation as consisting of two main steps

1. The data residuals are used as effective sources into the forward-modelling code, using the current best velocity model. These residuals correspond to those parts of the wavefield which exist in the actual recorded data, but do not exist in the forward-modelled data using the velocity model. In this manner, the residuals will naturally back-propagate to the location of the differences in the velocity model which cause scattering in the “true” model, but which are not present in the estimated model.
2. This back-propagated wavefield is multiplied (in the Fourier domain, so cross-correlated in the time-domain) by virtual sources at each grid point.

These virtual sources arise by local perturbations of the velocity model at a given point. If we think of the forward-modelled wavefield travelling through the current estimated velocity model as some kind of “reference” wavefield, then by introducing a small perturbation in the model at some point, this point will act like a new point

scatterer and generate a perturbation in the wavefield that was not in the original “reference” wavefield. By comparison of these forward-propagated virtual sources with the backward-propagated residual wavefields, we migrate the residuals.

This new “imaged” perturbation of the wavefield gives the location and “direction”, and even magnitude of the mismatch between the true and estimated velocity models. We may imagine that the amplitude of this velocity anomaly image, which bears some resemblance to a reflection coefficient as explained in §1.5.1, may be used to estimate an approximate α with an analogous concept of an imaging condition.

In Chapter 5, I introduce a novel application of the EAWT method to time-lapse seismic monitoring. Although this method was originally developed to invert for (static) velocity models to improve or even replace seismic migration, in my research I have shown that it may also be useful in highlighting time-evolving changes in the structure of the subsurface relative to some initial baseline survey. This is a use of low-frequency information in a field where it is common for industry to attempt to use mainly high-frequency information in an effort to delineate a small spatial anomaly.

1.7 Chapter summary

Within this chapter is an introduction to the main concepts underlying the aspects of modern seismic imaging that I have chosen to address in my research. Although many topics are introduced heuristically and approximately, an understanding at this level is all that is necessary to fully appreciate the contributions that are detailed in the following chapters.

The underlying theme of my dissertation could be summarized as, “what as-

sumptions and simplifications are we making in order to solve these problems, and are they justified? Could we perhaps do a little better in some cases, in order to save a little more computation time, or perhaps get a more accurate result with the same computation time?”

Frequently, in terms of the wave equation, the fundamental approximation that is most commonly made is the high-frequency approximation. Equivalently, we often assume that the waves travelling through a medium are of a significantly shorter wavelength than the scale of variation of the medium. In seismic imaging, this is unfortunately very far from the case. In some cases, I have chosen to address this broad gap empirically and practically, as in Chapters 2 and 3. In another case, I have approached the issue from a more analytic and theoretical direction, by adapting advanced mathematical theory into a working imaging scheme, as detailed in Chapter 4. In Chapter 5, I investigate a new and unintended realm of application of an existing method that is explicitly designed for low-frequency validity. Finally, in Chapters 6 and 7 I develop improvements to a fundamental imaging algorithm that has a stronger validity in lower-frequency solutions (GPSPI), and so it is more practical and useful for seismic imaging.

Chapter 2

Applying the eikonal equation to lower frequencies

2.1 Introduction

The eikonal equation is derived as a high-frequency asymptotic solution to a scalar wave equation in §1.4.2. This eikonal equation is frequently used in seismic imaging, with application to Kirchhoff migration as described in §1.5.5. The eikonal equation is used to calculate first-arrival traveltimes between any two points within a velocity model. A typical application of this equation to Kirchhoff migration would be to use it to calculate traveltimes from a source or geophone location to all points in the subsurface. Then for the stage in Kirchhoff migration in which the total traveltime from source to scattering point and then back up to receiver is required, it is simply a matter of checking the solutions from the source location eikonal solution to the scatterpoint, and then from the geophone location eikonal solution to the scatterpoint, and summing. This satisfies the traveltime requirement of the Kirchhoff algorithm.

The eikonal equation is a useful approximation to solving traveltimes economically, especially when implemented with a high-speed method *e.g.* fast marching (Sethian and Popovici, 1999). However, its high-frequency nature fundamentally conflicts with the reality of seismic exploration. P-wave propagation speeds within competent solid rock are typically at least 2000m/s (more often 3000-5000m/s), while exploration-target bedding thicknesses are often tens of meters thick at most. A single wavelength of 100m in 2000m/s rock oscillates at 20Hz. The “high-

frequency” assumption, however, effectively asserts that the wavelength of propagating waves must be much smaller than the characteristic size of variability in the medium. Therefore, to be valid, we must consider only frequencies “much higher” than 20 Hz. Even if we consider wavelengths that are only one fifth of the characteristic length scale of the medium, this restricts us to using only the portion of the signal above 100Hz when we are exploring in 100m-thick bedding with 2000m/s speed of propagation – all generous and almost unrealistic relaxations of our restrictions¹.

Clearly, “high-frequency” does not apply to the vast majority of seismic exploration data. Nevertheless, these approximations are extremely useful, and are much simpler to calculate than the full-frequency versions. For this reason, we are motivated to find methods that take advantage of the performance of the high-frequency approximations, while retaining more of the low-frequency validity of the full expression.

2.1.1 The hypereikonal equation

One classical attempt to extend the eikonal equation concept to lower frequencies results in the *hypereikonal* equation, also known as the *frequency-dependent eikonal equation*.

We will define reference slowness $S_0(\vec{x})$ such that

$$S_0(\vec{x})^2 = \frac{1}{v(\vec{x})^2} = |\nabla\phi(\vec{x})|^2. \quad (2.1)$$

¹Seismic data rarely contains significant information higher than 100 Hz. A typical useful band of signal will be between perhaps 4 Hz and 80 Hz, though this varies widely depending on acquisition particulars.

By inserting this into the full-dimensional version of equation 1.19,

$$\left((\partial_j \phi)^2 - \frac{1}{v^2} \right) - \frac{i}{\omega} \left(\frac{2}{A} \partial_j A \partial_j \phi + \partial_j^2 \phi \right) - \frac{1}{\omega^2 A} \partial_j^2 A = 0,$$

simplifying, and considering only the real part, we arrive at the *hypereikonal* or *frequency-dependent eikonal* equation,

$$\left| \nabla \tilde{\phi}(\vec{x}, \omega) \right|^2 = S(\vec{x}, \omega)^2 = S_0(\vec{x})^2 + \frac{1}{\omega^2} \frac{\nabla^2 A(\vec{x}, \omega)}{A(\vec{x}, \omega)}. \quad (2.2)$$

This reference slowness $S_0(\vec{x})$ is the “natural” slowness model of the medium, while this new slowness $S(\vec{x}, \omega)$ is a frequency-dependent effective slowness. We have also extended ϕ as defined in equation 1.16 to depend upon both ω and \vec{x} .

This hypereikonal equation allows for the solution of frequency-dependent traveltimes through a medium. Reference slowness S_0 is used to find traveltimes for the limiting case $\omega \rightarrow \infty$, while the frequency-dependent slownesses give traveltimes for finite frequencies. Note that the only assumption made in this hypereikonal equation is that harmonic solutions of the form given in equation 1.16 may be found – this is *not* another high frequency approximation. However, the hypereikonal equation requires the simultaneous solution for A , which contributes to the complexity of solving the entire equation.

The physical meaning of this equation is interesting to ponder. Biondi (1997) has shown that it is feasible to solve numerically for these frequency-dependent slownesses, with some moderate linearization in the equation. The solutions reveal a sort of frequency-dependent smoothing, with heavy smoothing at low frequencies, and very little smoothing at high frequencies. Unfortunately, these solutions are extremely complicated to calculate accurately and in a stable fashion. There is, however, one important conclusion: even this simple approach to a scalar wave equa-

tion in an idealized acoustic medium reveals a sort of intrinsic chromatic dispersion that is rarely considered.

2.1.2 The physics of the eikonal equation

Physically, the (high-frequency) eikonal equation may be thought of as defining an outer envelope which would approximately contain all rays traced from the start time $t = 0$ to some later travel time t_0 . Consider a medium consisting of a homogeneous background velocity of 2000 m/s but with random fluctuations of a nearly-normal distribution with a standard deviation of 500 m/s but with velocities outside of 2σ set to 2σ . In Figure 2.1, ray tracing through this medium is shown, along with the eikonal solution for the same propagation time. Ray tracing was accomplished with the CREWES raytracing tools, which are implemented in MATLAB (CREWES, 2010). Eikonal solutions were developed in MATLAB following the algorithm presented by Sethian and Popovici (1999). Note how very few of the rays actually approach the travelttime given by the eikonal solution. If this medium is lightly smoothed, the rays much more readily agree with the eikonal solution in this smoothed medium, as in Figure 2.2. With heavier smoothing, the match is even better (Figure 2.3). This, however, raises the question: is this *better physics*?

Finite difference methods were used to benchmark the solutions. The code used for these solutions is a part of the CREWES Matlab Library (CREWES, 2010). In a velocity model with extreme random variation, a finite-difference solution as shown in Figure 2.4 does reveal a wavefront in the expected location. However, a significant portion of the energy of the wavefield is represented by the multiply-scattered waves inside the wavefront. Although the *wavefront* location is distinct, it is not necessarily where the vast majority of the *energy* may always be found. Physically, then,

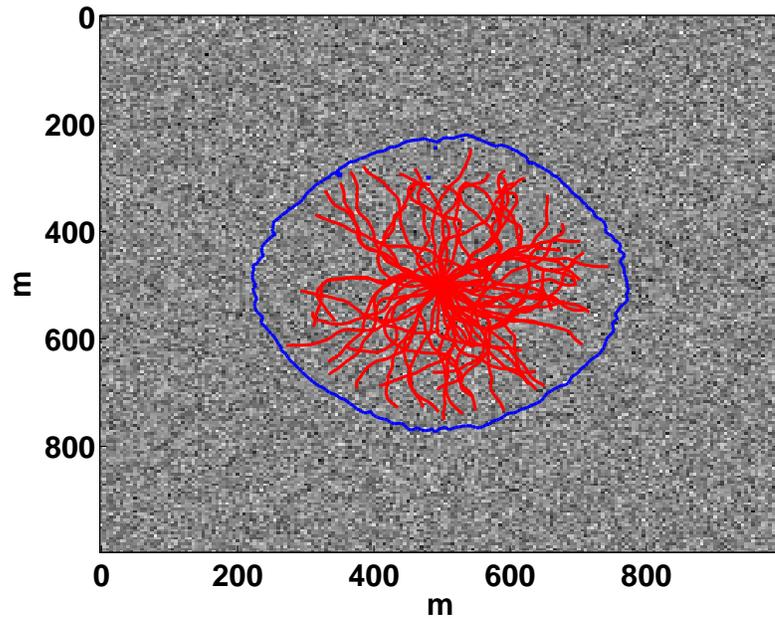


Figure 2.1: Ray tracing (in red) through a $2000m/s, \sigma = 500m/s$ medium. The eikonal solution envelope for the same propagation time is overlaid in blue.

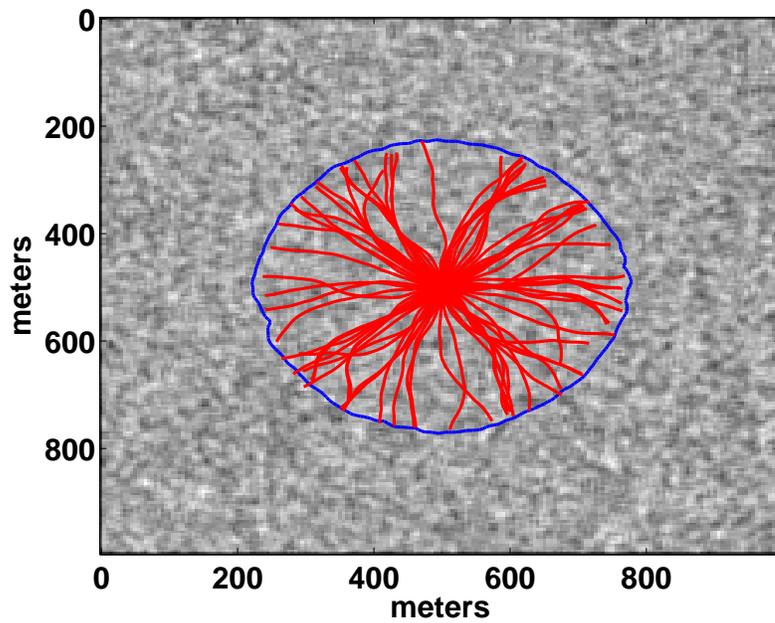


Figure 2.2: Ray tracing (in red) through a $2000m/s, \sigma = 500m/s$ medium with light smoothing. The eikonal solution envelope for the same propagation time is overlaid in blue.

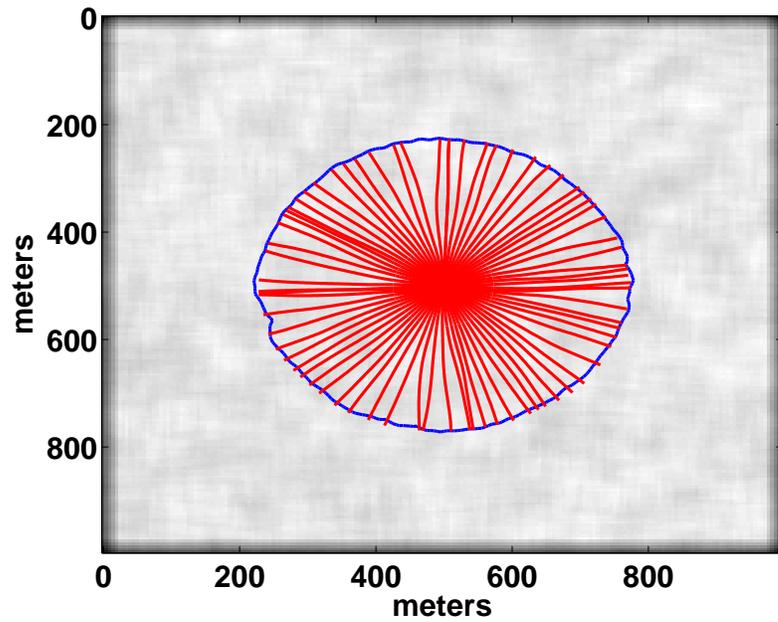


Figure 2.3: Ray tracing (in red) through a $2000m/s, \sigma = 500m/s$ medium with heavy smoothing. The eikonal solution envelope for the same propagation time is overlaid in blue.

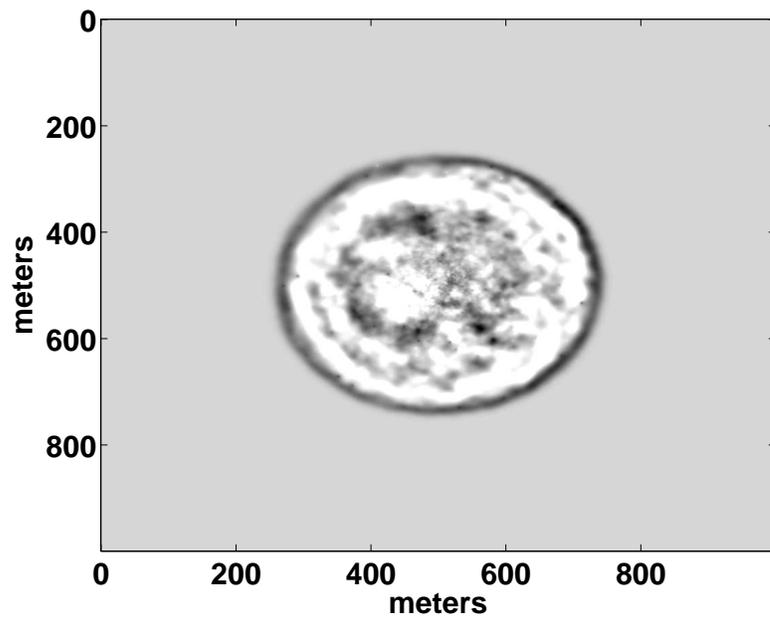


Figure 2.4: Finite differences solution of a $2000m/s, \sigma = 200m/s$ medium. A distinct wavefront is visible.

perhaps the ray-tracing in Figure 2.1 is an appropriate approximation of the sort of propagation seen in Figure 2.4 in that the ends of the ray paths proportionally represent the location of the energy of the propagation. Since the raypaths do not entirely terminate on the eikonal solution, we may infer that the bulk of the energy of the wavefield is still confined to a much smaller radius than the theoretical fastest outgoing wavefront represented by the eikonal solution.

In terms of seismic migration, however, this question is not always so important. Velocity models actually used for ray-tracing are not usually contaminated with high-amplitude random variations, so we would expect that ray-tracing-derived traveltimes and solutions to the eikonal equation would be a closer match – at least in the case of simple out-going wavefields. Therefore, we will treat the eikonal solution as a limiting case of ray-tracing, valid when ray-tracing through media free from extensive local variability, and for this reason we will now only consider velocity models without a high degree of local variability throughout.

2.2 Frequency-dependent smoothing of a banded velocity model

In chapter 3, frequency-dependent smoothing of the velocity model is shown to be effective in improving imaging via GPSP (Hogan and Margrave, 2006). Here we will investigate whether or not the same is true for eikonal-equation based approaches to solution of the wave equation.

To test this, we generated a velocity model consisting of horizontal stripes of alternating high-velocity/low-velocity bands (Figure 2.5). This original velocity model was smoothed several times with Gaussian smoothers, with σ varying between 0 and 25m. A representation of these variously-smoothed velocity models is shown in Fig-

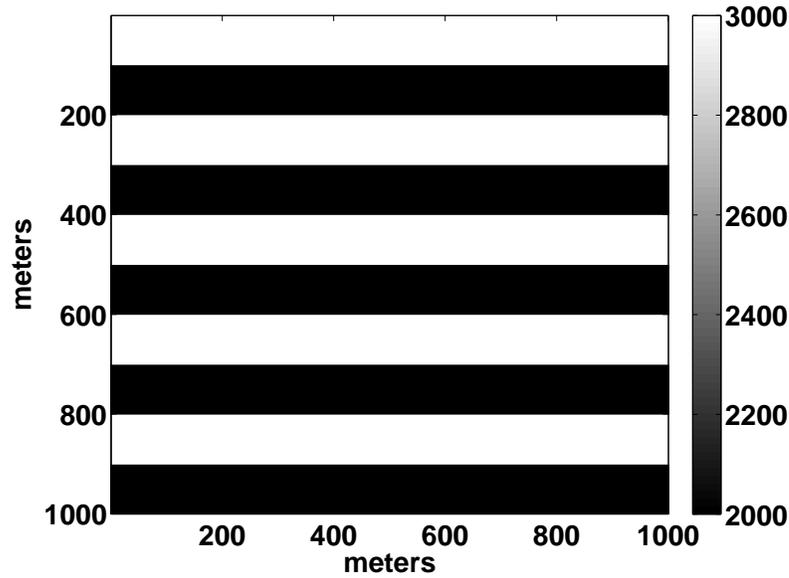


Figure 2.5: Striped velocity model, velocity in m/s .

ure 2.6.

A finite-difference solution of the wave equation over this domain with a high-frequency (60 Hz dominant minimum-phase) source at $(500,500)$ was calculated to a final time of $t = 0.136s$, and a solution of the eikonal equation over the same domain was also calculated. These solutions may be seen in Figure 2.7. The position of the outgoing wavefront may be seen to agree well with the eikonal contour.

Next, lower-frequency finite-difference wavefields were calculated over the same unsmoothed medium. At the extreme low-frequency end (2 Hz dominant minimum phase source), the outermost envelope of the wavefield coincides best with an eikonal solution over the original velocity model smoothed by convolving with a Gaussian smoother with $\sigma=25m$, *i.e.* a smoothing length of $25m$. For comparison, Figure 2.9 shows eikonal solutions for both unsmoothed and smoothed velocity models overlaid on the finite-difference solution with a 2 Hz dominant source. At approximately

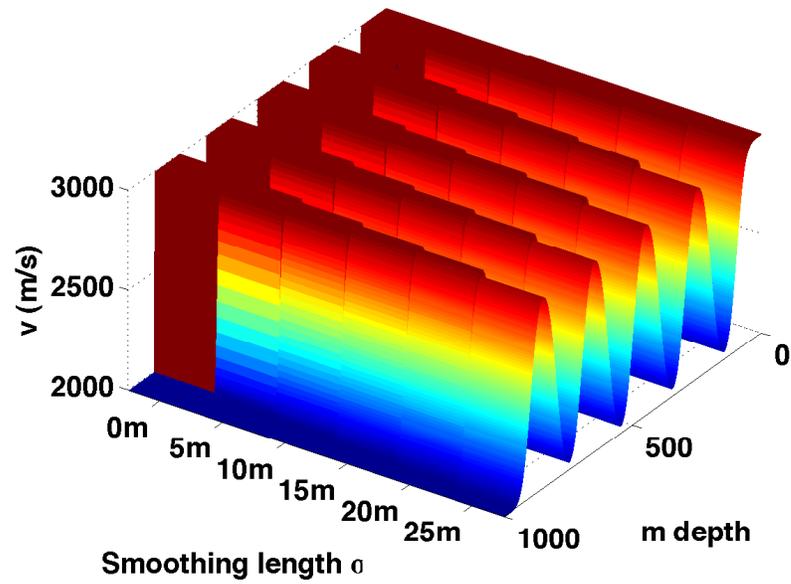


Figure 2.6: Vertical sections of the velocity models used for eikonal solutions. On the left is the unsmoothed velocity model, on the right is the most heavily smoothed model.

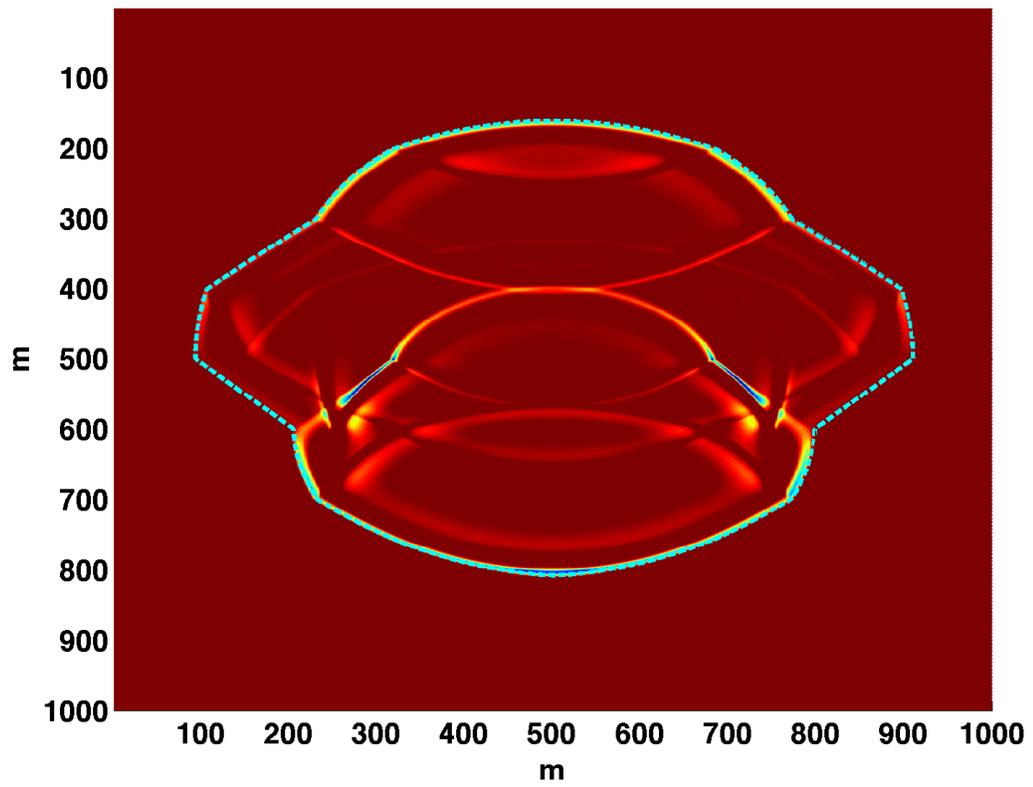


Figure 2.7: 60 Hz dominant source in unsmoothed velocity model, with eikonal solution to the unsmoothed model overlaid in blue.

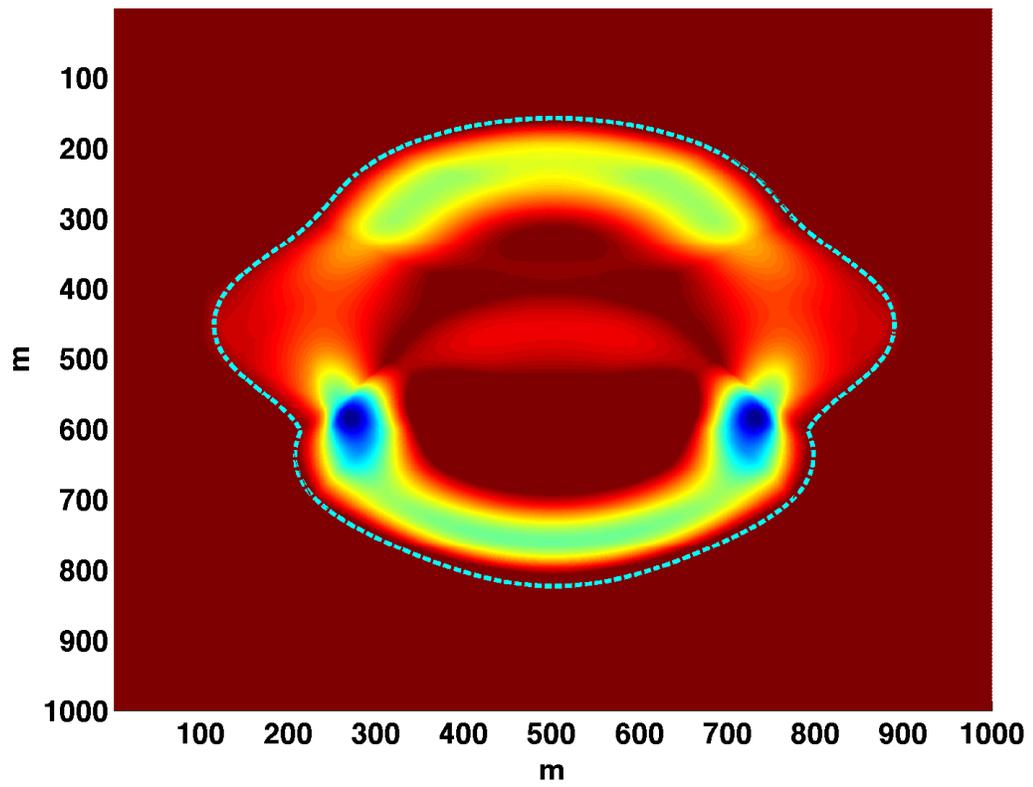


Figure 2.8: 2 Hz dominant source, with eikonal solution to the 25 m smoothing length model overlaid in blue.

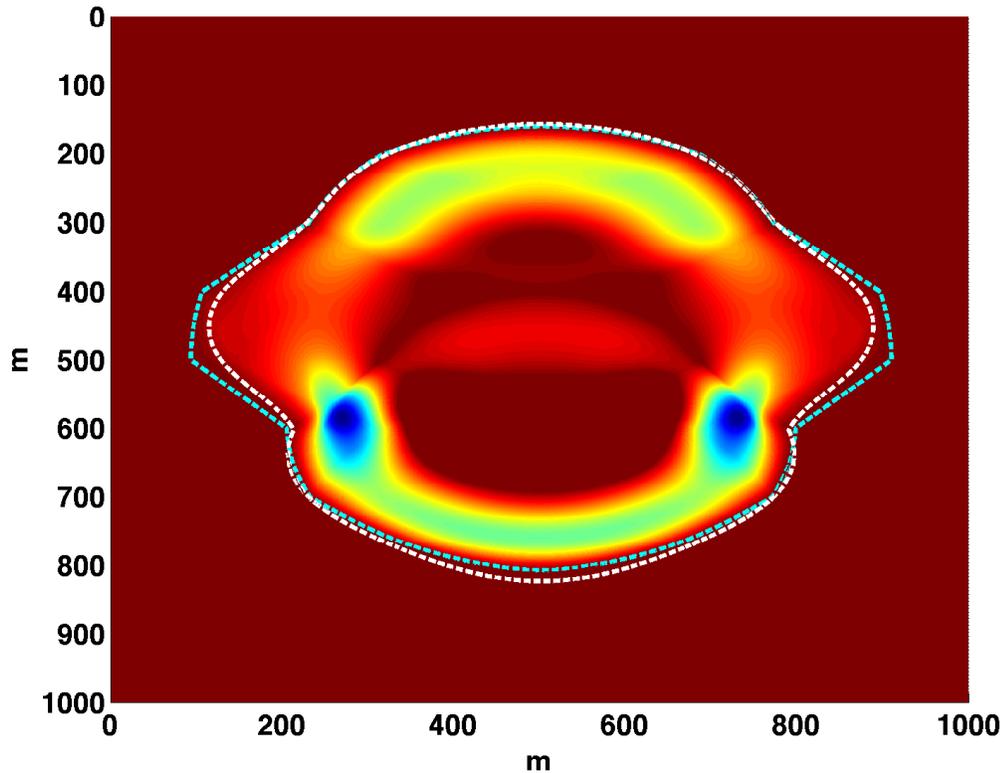


Figure 2.9: 2 Hz dominant source, 25 m smoothing length eikonal solution in white, unsmoothed velocity model eikonal solution in blue.

(500, 800) on Figure 2.9, it may be seen that the smoothed-model eikonal contour extends farther down in the section than the amplitude of the original wavefield. This may be explained by considering the velocity model in Figure 2.6. The 25m smoother smears the high-velocity band between 800m and 900m depth up into the low-velocity band between 700m and 800m. Thus the eikonal contour locally out-paces the actual velocity of propagation in this direction. This is purely a local effect, and is corrected upon farther propagation. Similar smoothed-model eikonal solutions are overlaid on finite-difference solutions for a 5 Hz dominant source (Figure 2.10), 10 Hz dominant source (Figure 2.11), 20 Hz dominant source (Figure 2.12), and a 40 Hz dominant source (Figure 2.13).

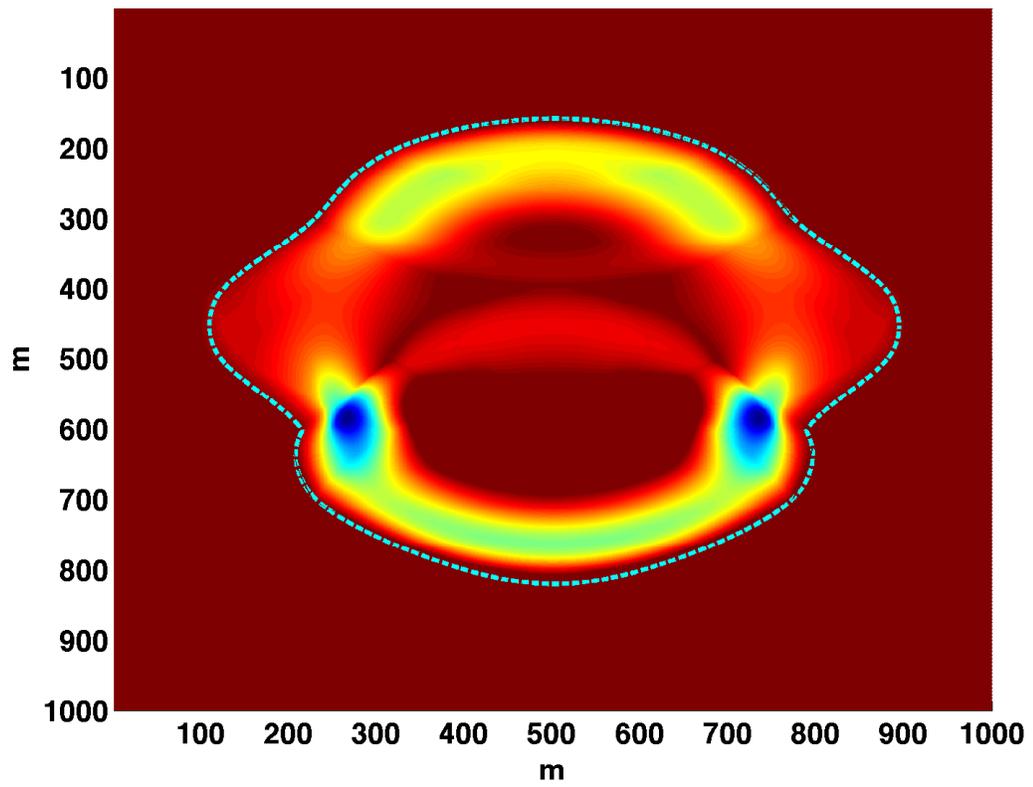


Figure 2.10: 5 Hz dominant source, eikonal solution to 20 m smoothing length model overlaid in blue.

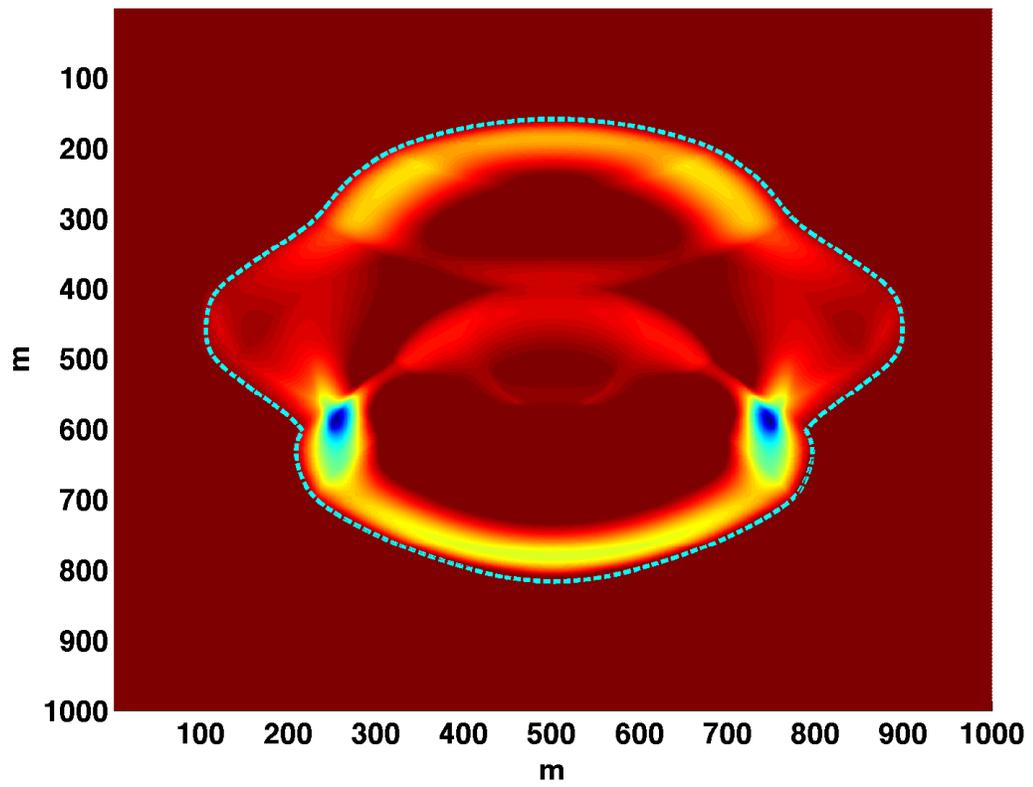


Figure 2.11: 10 Hz dominant source, eikonal solution to 15 m smoothing length model overlaid in blue.

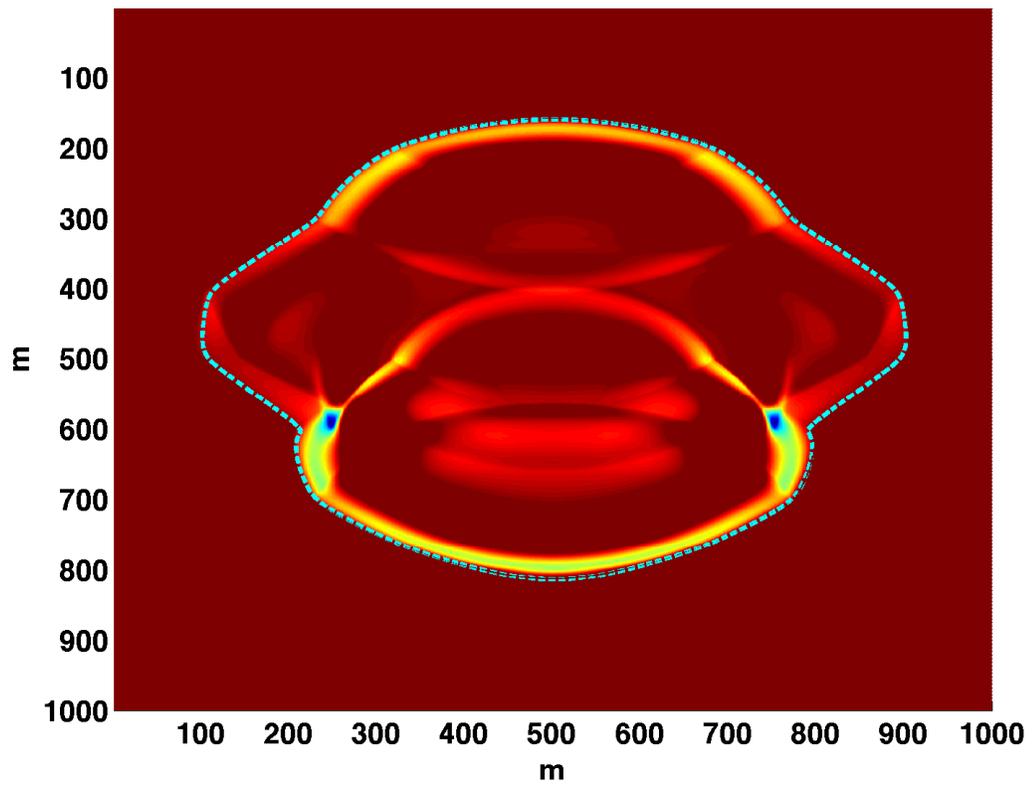


Figure 2.12: 20 Hz dominant source, eikonal solution to 10 m smoothing length model overlaid in blue.

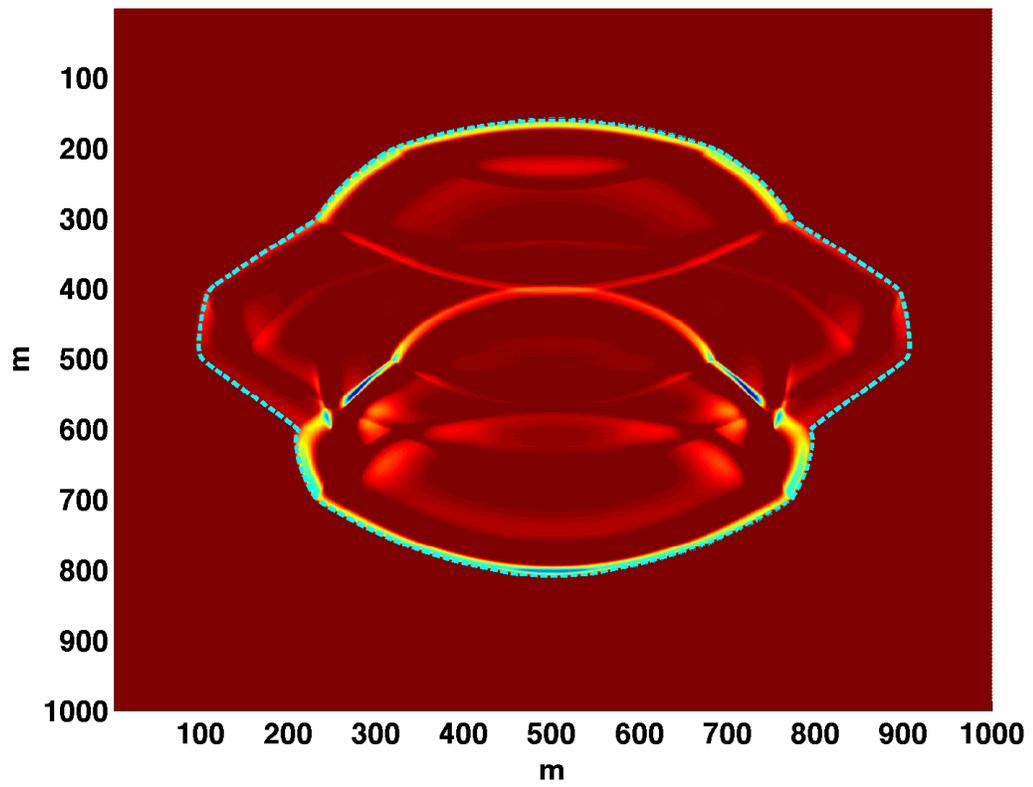


Figure 2.13: 40 Hz dominant source, eikonal solution to 5 m smoothing length model overlaid in blue.

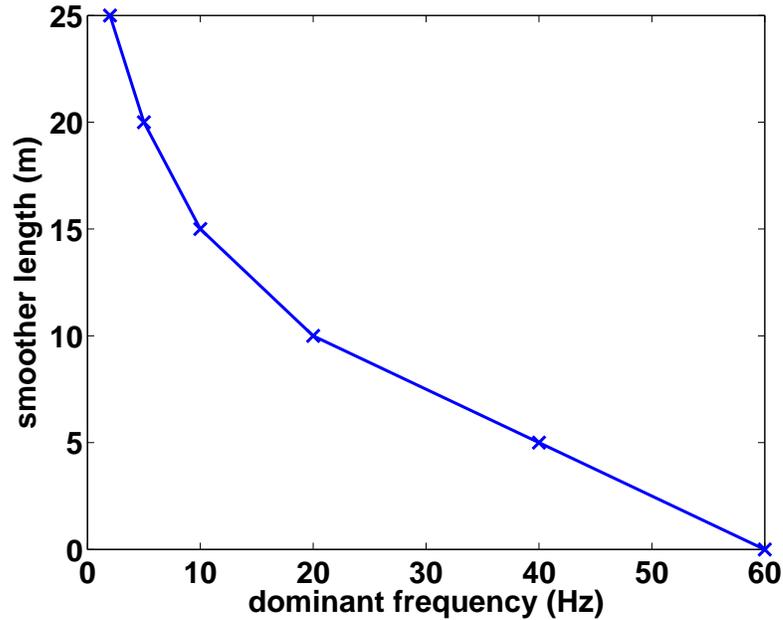


Figure 2.14: Smoothing length as a function of dominant frequency.

Overall results are summarized in Figure 2.14. At high frequency, the eikonal solutions match the finite difference wavefield well. At lower frequencies, smoothing lengths up to $\sigma = 25\text{m}$ become necessary.

2.3 Effects of smoothing on ray-tracing

Figure 2.15 shows the result of tracing rays through the unsmoothed medium. Many of the rays terminate on the eikonal contour as expected. However, several rays appear to reflect and terminate inside the contour. This highlights one major difference between eikonal-based modelling and ray-tracing modelling: ray-tracing allows for more natural inclusion of the modelling of multiple reflections, while standard eikonal solutions are insensitive to multiple reflections². Whether or not either is a

²though this can depend on the actual algorithm employed to solve the eikonal equation.

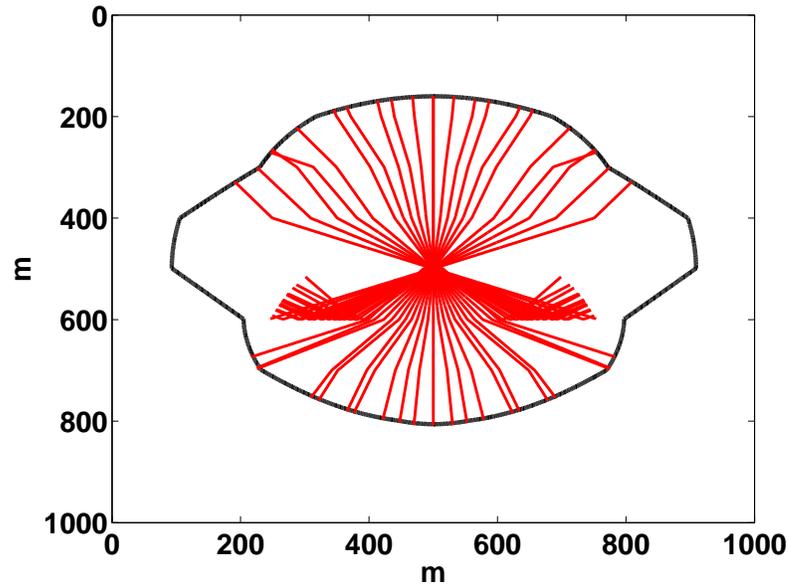


Figure 2.15: Ray-tracing (red) through the unsmoothed velocity model), with eikonal contour (black).

benefit or hindrance is dependent on the goals of the modelling. When the velocity model is smoothed, ray-tracing and eikonal solutions converge: that is, more rays terminate exactly at the eikonal contour as in Figure 2.16.

In Figure 2.17, the 60 Hz source wavefield is shown with rays traced. Here it is clear that the reflected rays are following the high-amplitude reflected event in the full wavefield. By contrast, in Figure 2.18, the smoothed velocity model allows the raypaths to accurately track the envelope of the low-frequency wavefield.

2.4 Chapter summary

Frequency-dependent smoothing of a velocity model can lead to eikonal and ray-tracing solutions which match the propagation envelope of a full wavefield better at low frequencies, where these approximations are not strictly valid. At these lower

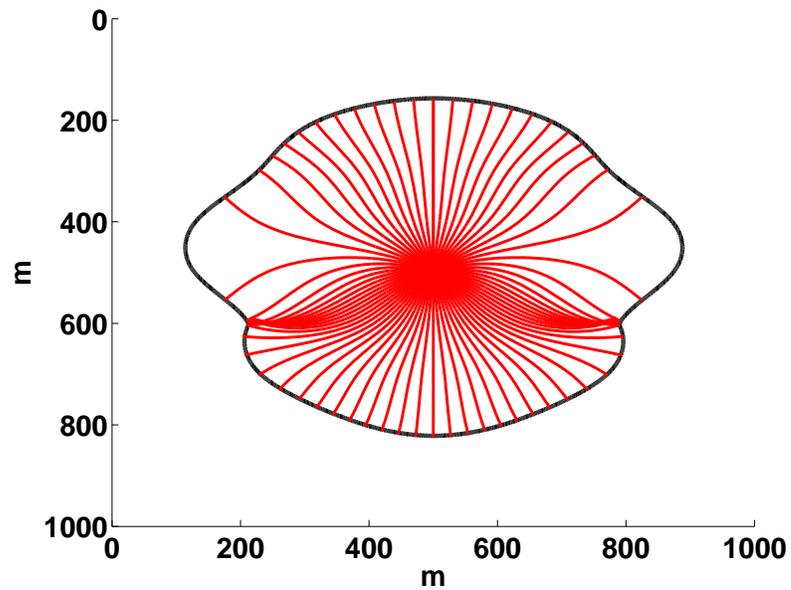


Figure 2.16: Ray-tracing (red) through the 25m smoothed velocity model, with eikonal contour (black).

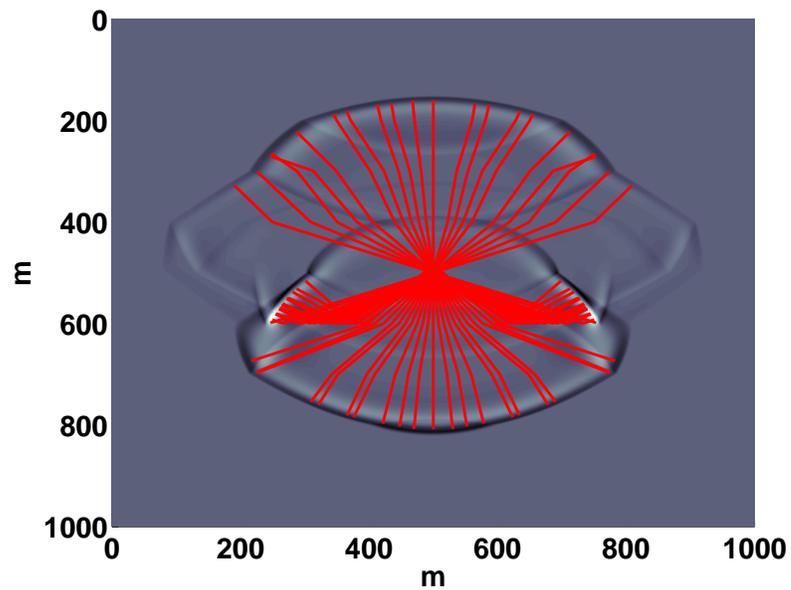


Figure 2.17: Ray-tracing (red) through the unsmoothed velocity model, with 60 Hz dominant finite difference wavefield (grayscale).

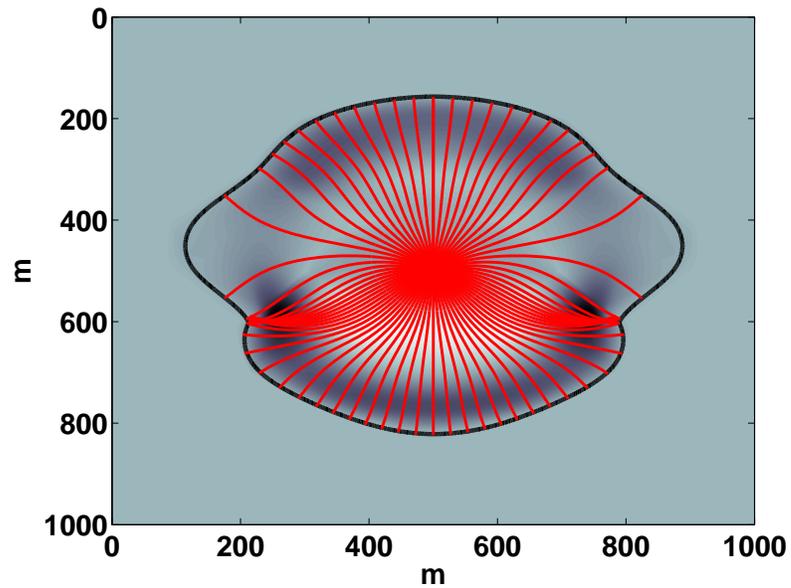


Figure 2.18: Ray-tracing (red) through the 25m smoothed velocity model, with 2 Hz dominant finite difference wavefield (grayscale).

frequencies, the smoothed velocity model also leads to ray-tracing solutions which converge to give the same kinematics as the eikonal solutions. These solutions are very simply calculated, and do not require any new code development. Simply inserting the appropriately-smoothed velocity model into existing eikonal solvers will yield immediately useful results.

Chapter 3

Frequency-dependent velocity model smoothing in GPSPI

3.1 Introduction

The GPSPI algorithm (Margrave and Ferguson, 1999) provides a highly-accurate wavefield extrapolation method for use in pre-stack migration seismic imaging. This algorithm depends upon a numerical implementation of the “infinitesimal extrapolator”,

$$\begin{aligned}\Psi(x, z + \Delta z, \omega) &= \mathbf{T}_{\alpha(z:z+\Delta z)}\Psi(x, z, \omega) \\ &\approx \int_{\mathbb{R}} \phi(\xi, z, \omega)\alpha(x, \xi, \omega, z : z + \Delta z)e^{i\xi x}d\xi\end{aligned}\quad (3.1)$$

where

$$\alpha(x, \xi, \omega, z : z + \Delta z) = \begin{cases} \exp\left(i\Delta z\sqrt{\frac{\omega^2}{v(x)^2} - \xi^2}\right), & |\xi| \leq \frac{|\omega|}{v(x)} \\ \exp\left(-\left|\Delta z\sqrt{\frac{\omega^2}{v(x)^2} - \xi^2}\right|\right), & |\xi| > \frac{|\omega|}{v(x)} \end{cases}. \quad (3.2)$$

Here Ψ represents the wavefield as a function of horizontal position x , vertical depth z , depth-step Δz , and temporal frequency ω . \mathbf{T} is the infinitesimal extrapolation operator characterized by symbol α , where α is a function of x , horizontal wavenumber ξ , ω , and z . ϕ is the wavefield Ψ after Fourier-transforming from x to ξ . In this case, \mathbf{T} extrapolates the wavefield from depth z to depth $z + \Delta z$.

Since this algorithm is intimately connected to seismic imaging, a robust, efficient, and accurate method is key to successful application. In this chapter, the high-frequency nature of the locally-homogeneous approximation to the square-root Helmholtz operator symbol is investigated. Here, we use knowledge of the overall

evolution of the true full (not approximate) operator symbol to guide an estimation process that supports the use of frequency-dependent smoothing of the underlying velocity model to design GPSPI operator symbols.

3.2 The square-root Helmholtz operator symbol

Fishman (2002) identifies the term $\sqrt{\omega^2/v(x)^2 - \xi^2}$ in equation 1.46 as the limiting form of a high-frequency approximation to the square-root Helmholtz operator symbol (i.e. an “infinite-frequency symbol”). Figure 3.1 shows the real part of the infinite-frequency symbol as used in GPSPI. This symbol is calculated for a velocity model consisting of three blocks of constant velocity. On the left, velocity v_1 is relatively high; in the middle, velocity v_2 is relatively low; and on the right, velocity v_3 is moderate. Figure 3.2 shows an ensemble of exact square-root Helmholtz operator symbols for high, moderate, low, and zero frequency, also due to Fishman (2002). These symbols have been rotated to show the detail within the symbol. At high frequency, the symbol appears to be quite similar to the infinite-frequency symbol. The added ripple features capture the physics of the trapped horizontal modes – effectively the multiple horizontal reflections. The symbol at moderate frequency retains many of these features, but appears somewhat smoothed compared to the higher frequency symbol. The low-frequency symbol appears heavily smoothed, and the zero-frequency symbol shows only a single effective velocity. In an intuitive sense, these symbols are demonstrating that the various frequencies “see” different media. The highest frequencies are affected by all scales of structure within the velocity model, while the lowest frequencies see only the largest-scale overall velocity – any finite velocity changes are lost and only the background average velocity appears.

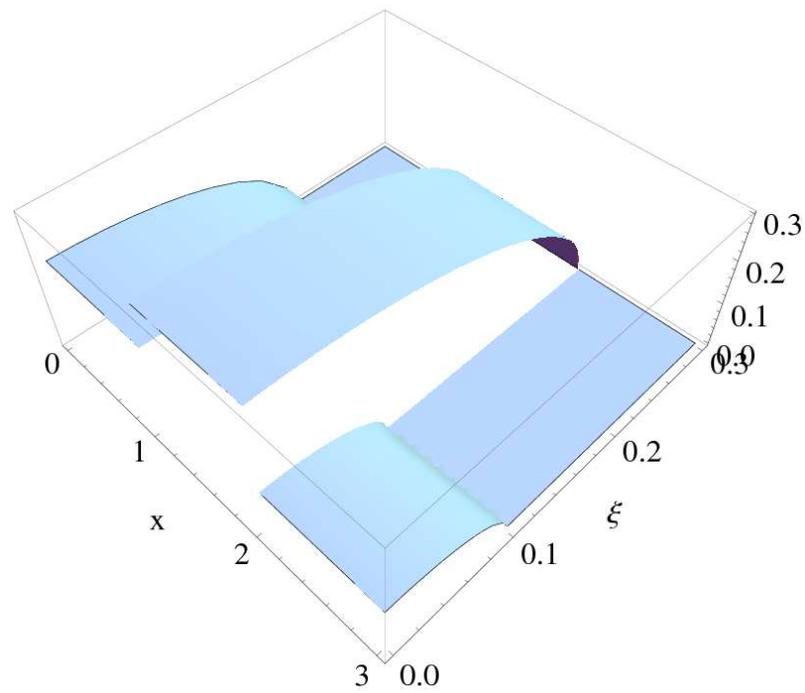


Figure 3.1: Real part of the infinite-frequency square-root Helmholtz operator symbol for a velocity model with blocks of constant velocity. Adapted from Fishman (2002). This is the symbol as used in GPSPI.

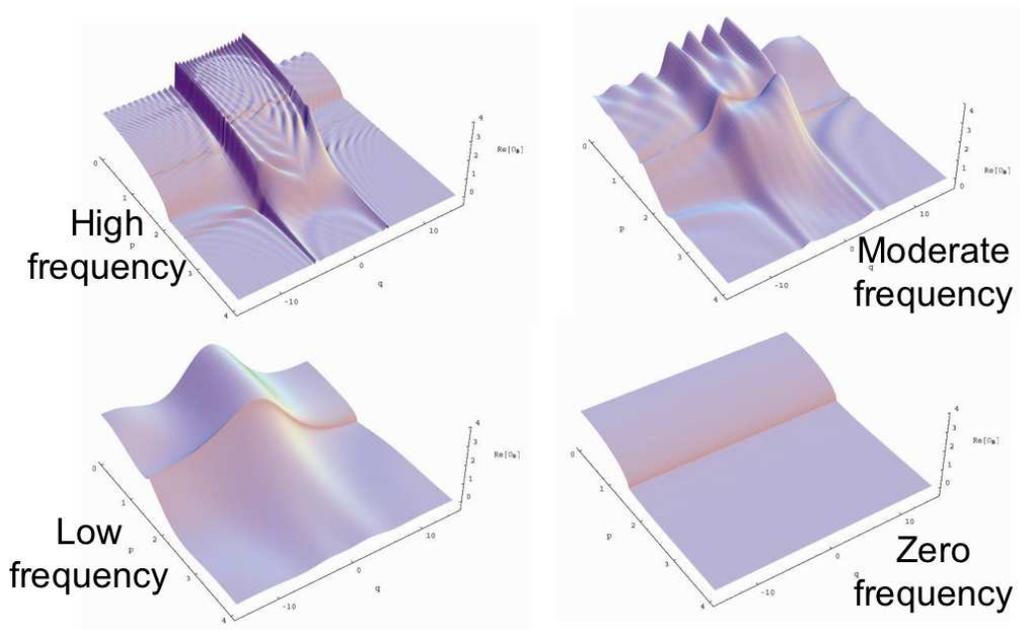


Figure 3.2: Ensemble of real parts of the square-root Helmholtz operator symbols for a velocity model with blocks of constant velocity. Adapted from Fishman (2002).

3.3 Approximating symbols

Although the symbol calculations given by Fishman (2002) are, at this point, far too computationally expensive for practical migration purposes, approximations may be considered. Specifically, we choose to preserve the apparent smoothing displayed by the symbols at various frequencies. Figure 3.3 again displays a three-block velocity model symbol at infinite frequency. Figure 3.4 shows this same symbol calculated with a velocity model that is smoothed with a smoothing length determined by the wavelength of low-frequency waves. The optimal method for smoothing remains an open problem, and at this point smoothing is done purely empirically.

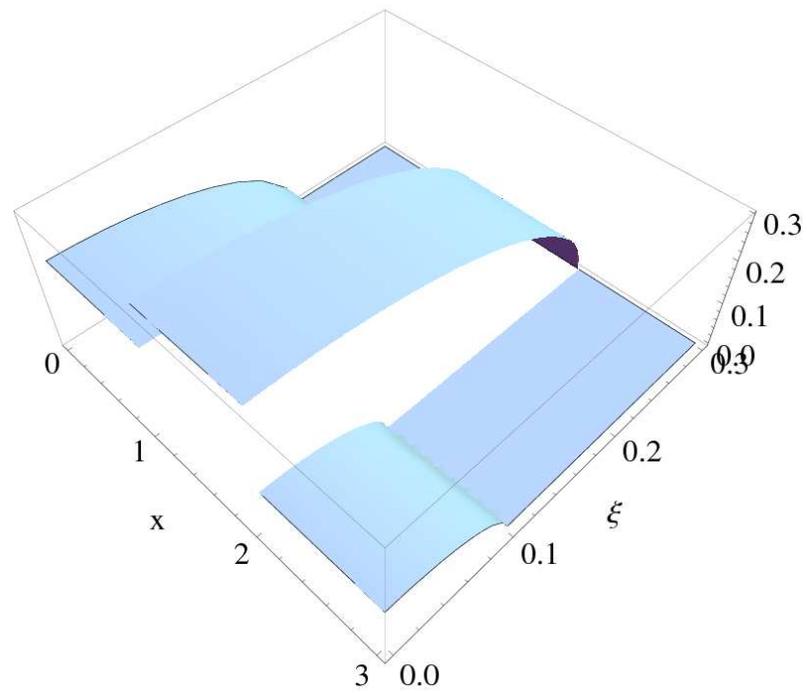


Figure 3.3: Real part of the infinite-frequency symbol for the three-block velocity model.

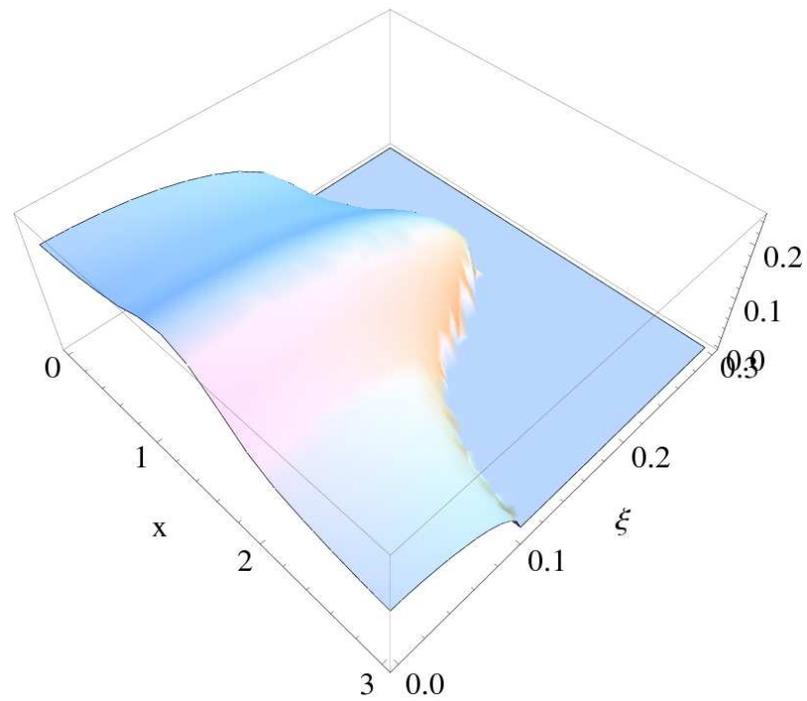


Figure 3.4: Real part of the low-frequency symbol approximation (via smoothing of the velocity model) for the three-block velocity model.

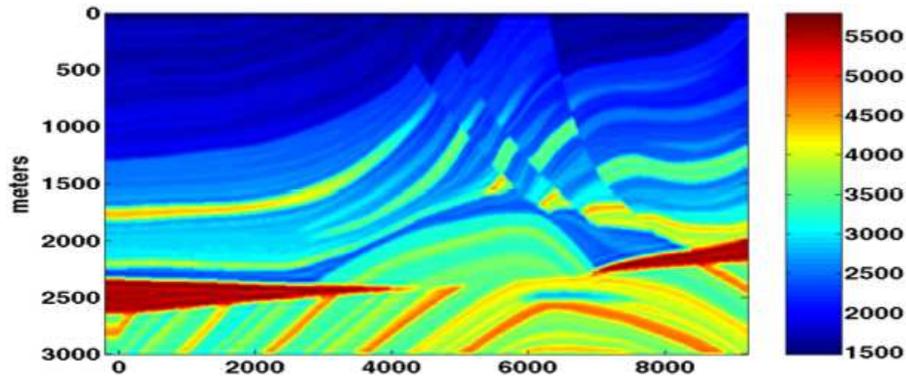


Figure 3.5: The Marmousi velocity model. Velocities range between approximately $1500m/s$ and $5500m/s$.

3.4 Testing

Testing of smoothed vs. unsmoothed symbols was performed using the FOCI (Margrave et al., 2006) method of GPSP migration. The FOCI code used was the reference implementation developed by CREWES using MATLAB. The smoothing was accomplished using the spatial-resampling portion of the FOCI method. Although this spatial resampling was originally developed to address numerical stability and efficiency concerns, it effectively does perform a frequency-dependent smoothing of the velocity model. The smoothing method within spatial resampling is a simple spatial averaging. Migration tests were performed using synthetic data from the Marmousi model (Figure 3.5).

The FOCI spatial resampling typically decomposes the data into on the order of 10 frequency bands, ranging from low frequency ($\approx 4 - 6Hz$) to high frequency ($\approx 42 - 60Hz$). We may recognize that the low-frequency components are travelling with a much longer wavelength, and so require far fewer effective geophone traces to adequately sample them than the higher frequencies. Therefore, significant computation time may be saved since only a fraction of the original number of traces

in the chunk must be considered in the extrapolation. However, by down-sampling the original traces, we also must anti-alias filter and then downsample the underlying velocity model which is used for extrapolation in these lower-frequency bands. This filtering and downsampling effects a smoothing with a spatial averaging that is characteristic for each frequency band.

Migrations were run both with and without spatial resampling. As the FOCI algorithm requires spatial resampling for stability and efficiency, special care was taken to ensure that these features were preserved in the unsmoothed case. Specifically, the image calculated using smoothed symbols had a final operator length of 31 points¹. In order to preserve stability and fidelity of the operator in the absence of the spatial resampling, the operator length had to be scaled in order to provide the same operator control within the wavelike region of propagation for each frequency band. Therefore, for the lowest frequency block, a final operator length of more than 200 points was required. Full details of the FOCI algorithm may be found in Margrave et al. (2006).

The resulting images are shown in Figure 3.6. The image calculated with the smoothed symbol is clearly significantly better than the image calculated with the infinite-frequency symbol.

In order to determine the frequency-dependence of the difference in image quality, the concept of “residual” as defined in §6.2.2 was used. The smoothed- and unsmoothed-symbol images were broken into frequency bands 10 Hz wide, centred every 5 Hz from 5 to 55 Hz. 11 new images were generated, each image consisting of

¹The number of FOCI operator “points” corresponds to the number of spatial grid points over which the convolutional operator is non-zero. A larger number of points corresponds to a more accurate and stable operator, while a smaller number of points corresponds to a faster but possibly more unstable operator due to the truncation of the operator and resulting Gibbs phenomenon.

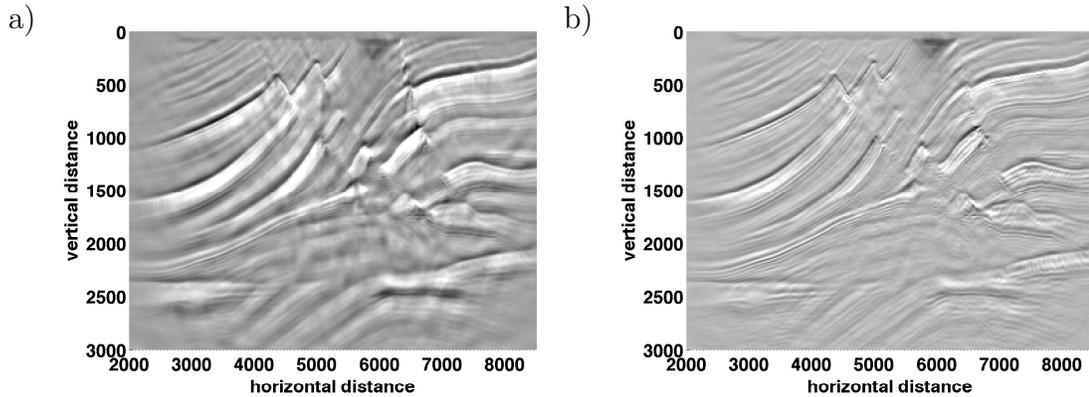


Figure 3.6: Images of the Marmousi model generated a) without frequency-dependent smoothing of the symbol and b) with frequency-dependent smoothing of the symbol. The image generated with the smoothed symbol is clearly far superior to the image generated with the infinite-frequency symbol.

the unsmoothed-symbol image with one frequency band replaced by the respective frequency band from the smoothed-symbol image. For example, the first image was generated with the 5-15 Hz band from the smoothed-symbol image, along with all other 10 bands from the unsmoothed-symbol image. Each of these images was compared with an image composed of all 11 bands of the unsmoothed-symbol image. The residuals are plotted in Figure 3.7.

These residuals indicate that the smoothing of the symbol had the most effect on the image in the lower frequencies, and had very little effect on higher frequencies. This matches our expectations that the biggest differences in the symbol will be found in the lower frequency bands, and so the biggest improvement in overall quality will be found there as well.

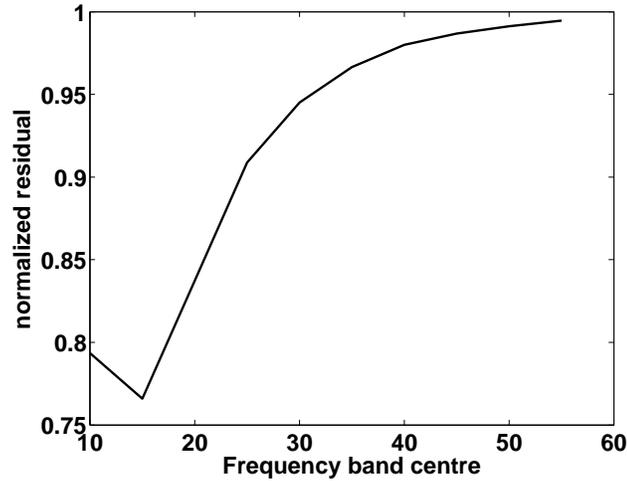


Figure 3.7: Determination of effectiveness of symbol smoothing as a function of frequency. Lower residuals indicate that the smoothing of the symbol resulted in a change in quality of the overall image, while a residual of 1 indicates that the smoothing had little to no effect in this region.

3.5 Chapter summary

The frequency dependence found in the exact square-root Helmholtz operator symbol is striking and leads to symbols that are dramatically different from the infinite-frequency symbols that are used in standard GPSP migration. Although these exact operator symbols are extremely difficult to compute, a qualitative and possibly naïve approximation to these symbols based on frequency-dependent smoothing of the underlying velocity model results in an image that is far superior to one that is calculated with just the infinite-frequency symbol.

The FOCI algorithm may already be implemented with a simple but effective form of symbol smoothing: spatial resampling following Margrave et al. (2006). Therefore, in addition to providing greater computational efficiency and numerical stability to the imaging, this spatial resampling also adds somewhat more accurate

physics to the wavefield extrapolation and therefore a better image.

Chapter 4

The Stolk operator in GPSPI migration

4.1 Introduction

Much of the recent research in “wavefield extrapolation methods” (WEM) is directed towards improving the quality of the WEM code without sacrificing numerical efficiency. Modern algorithms that make use of WEM, such as the Generalized Phase-Shift-Plus-Interpolation (GPSPI) migration algorithm due to Margrave and Ferguson (1999), have produced high-quality seismic images at affordable computational costs. Moving beyond GPSPI migration will require either significantly better images, or significantly improved numerical performance without a loss of imaging quality (or both). The current GPSPI migration operator uses an approximation in which the local wavefield propagation velocity is assumed to be a constant, which is effectively a high-frequency approximation. In this chapter I evaluate a new migration operator that mathematically extends the capabilities of the GPSPI operator beyond this assumption to account for a local first derivative in (horizontal) slowness, thus incorporating more physics and hopefully a higher-fidelity wavefield extrapolation over a broader range of frequencies.

4.1.1 Recursive phase-shift wavefield propagation

Many current WEM algorithms are space-frequency methods related to or derived from the phase-shift method introduced by Gazdag (1978). Fourier-transforming the scalar wave equation results in the Helmholtz equation, which results in an

interpretation of the wavefield as a sum of plane waves. By applying an appropriate phase-shift to each plane wave, the wavefield may be propagated through a distance accurately.

4.1.2 Generalized PSPI

The GPSPI algorithm is an accurate, efficient implementation of these phase-shift concepts. This algorithm takes an entire constant-depth-slice of the wavefield and propagates it to a deeper (or shallower) level. The z (depth) direction is typically referred to as the *range* direction, while horizontal directions x and sometimes y are referred to as *transverse* directions. In addition to Fourier-transforming the time series of the data, it is often convenient to represent the transverse directions x and y in their Fourier-transformed coordinates, ξ and ν .

Mathematically, the GPSPI algorithm may be represented in compact operator notation as:

$$\Psi(x, z = \Delta z, \omega) = \mathbf{T}_\alpha \Psi(x, z = 0, \omega) \quad (4.1)$$

Or, more explicitly:

$$\Psi(x, z = \Delta z, \omega) \approx \int_{\mathbb{R}} \phi(\xi, z = 0, \omega) \alpha(k(x), \xi, \omega) e^{i\xi x} d\xi \quad (4.2)$$

where

$$\alpha(k(x), \xi, \omega) = \begin{cases} e^{i\Delta z k_z(x)}, & |\xi| \leq \frac{\omega}{v(x)}, & k_z(x) = \sqrt{k(x)^2 - \xi^2} \\ e^{-|\Delta z k_z(x)|}, & |\xi| > \frac{\omega}{v(x)}, & k(x) = \frac{\omega}{v(x)} \end{cases}, \quad (4.3)$$

and

$$\phi(\xi, z = 0, \omega) = \frac{1}{2\pi} \int_{\mathbb{R}} \Psi(x, z = 0, \omega) e^{-i\xi x} dx. \quad (4.4)$$

This version of GPSPI is mathematically identical to that given in §3.1, but is written slightly differently to accommodate the work developed here.

The algorithm applies a nonstationary filter to the data (Margrave, 1998), is a pseudodifferential operator in the Standard Calculus (Hörmander, 1985) and is the limiting form of the PSPI algorithm described by Gazdag and Sguazzero (1984) when taken to the extreme of using a separate velocity for each and every output point. (Margrave and Ferguson, 1999). In the pseudodifferential calculus, α is known as the symbol of the operator \mathbf{T}_α , which is explicitly defined in equation (4.3).

It may be seen that the wavefield $\Psi(x, z, \omega)$ at each output point $(x_0, \Delta z)$ is the result of the wavefield extrapolation operator \mathbf{T}_α applied to the entire line $(x, 0) \ x \in \mathbb{R}$. A new operator is constructed for each required output point along the line using the local velocity at the output point $v(x_0, \Delta z)$. The core assumption at the heart of the GPSPI method is its dependence on the fact that the wavefield at each output point $(x_0, \Delta z)$ is calculated by propagating the wavefield from the line at depth $z = 0$ assuming that the entire wavefield at depth $z = 0$ propagates through a medium of constant velocity $v(x_0, \Delta z)$. That is, the local region between the line $(x, 0) \ x \in \mathbb{R}$ and the point $(x_0, \Delta z)$ is assumed to be homogeneous. For this reason, equation (4.2) with symbol as defined in equation (4.3) is often referred to as the locally-homogeneous approximation (Fishman et al., 1997). The magnitude of Δz is typically on the order of $\sim 10m$, which is assumed to be much smaller than the scale of the range velocity variation and so a range-homogeneous approximation is often reasonable. In numerical implementations the length of the line $(x, 0) \ x \in \mathbb{R}$ is frequently many hundreds of meters or more, so a transversely-homogeneous approximation is far from ideal. Dipping events with strong velocity gradients will likely not be imaged accurately.

4.1.3 The Stolk operator

It is hoped that mathematically-sound extensions to the GPSPI migration scheme that allow for locally-inhomogeneous velocity models will result in improved image quality compared to computational cost. We consider the specific case of transverse velocity variations. The hope is that this will allow for high-quality images in media that feature strong lateral velocity variations – salt-dome flanks, for example. This extension of the operator is accomplished by adapting mathematical theory introduced by Stolk (2004) into a useful GPSPI-type operator.

The formal mathematical approximations to the solutions of the Helmholtz equation developed by Stolk (2004) form the inspiration for a new operator that may be tested in comparison to the GPSPI operator. This theory is especially attractive as it presents itself as a correction to the standard operator (the “square-root”) used in GPSPI migration. Specifically, the Stolk correction adds a term involving the derivative of the slowness $s(x) = 1/v(x)$, so that the k_z term within α in equation (4.3) becomes:

$$k_z(x) = \left(\sqrt{\frac{\omega^2}{v(x)^2} - \xi^2} \right) + \frac{i\xi}{2} \frac{\omega^2}{v(x)} \left(\frac{\partial s}{\partial x} \right) \left(\frac{\omega^2}{v(x)^2} - \xi^2 \right)^{-3/2} \quad (4.5)$$

This means that it is relatively straight-forward to adapt existing GPSPI code to make use of this extended theory: we simply replace the $k_z(x) = \sqrt{\omega^2/v(x)^2 - \xi^2}$ term in the original implementation with the new expression in equation 4.5. We will refer to $k_z(x)$ as the operator symbol, and this is what is used within GPSPI, which applies this as a full Fourier-domain multiplier. We may also take the inverse Fourier transform of this symbol for a constant x , and transform $\xi \rightarrow x'$. This is referred to as the operator kernel, and is applied within the FOCI algorithm as an $\omega - x$ convolution.

This correction is interesting theoretically, as it is directly applicable to the pure rigorous PSDO interpretation of the GPSPI algorithm, rather than being an *ad hoc* approximation built upon physical intuition or signal-processing approaches. Most other improvements to the basic GPSPI-oriented algorithm address computational efficiency (*e.g.* FOCI) or computational stability (*e.g.* IWKBJ operators as in §7).

Some care must be taken to ensure a valid implementation within fast versions of the GPSPI ideas such as the FOCI algorithm (Margrave et al., 2004) so that the correction is not lost via some assumption. For example, the FOCI method normally assumes a symmetric $k_z(x)$ in x , while the Stolk correction results in a $k_z(x)$ which may be asymmetric.

Equation (4.5) has one obvious problem: there is a singularity at the evanescent boundary, that is where $\omega/v(x) = \xi$. Stolk (2004) corrects for this by using a damping term $h(\varsigma)$ that masks the singularity:

$$h(\varsigma) = \begin{cases} 0 & \varsigma \leq 0, \\ \exp(-1/\varsigma)/(\exp(-1/\varsigma) + \exp(-1/(1-\varsigma))) & 0 < \varsigma < 1, \\ 1 & \varsigma \geq 1. \end{cases} \quad (4.6)$$

This damping function is shown in Figure 4.1. We can use $\kappa |(\omega^2/v(x)^2 - \xi^2)| = \varsigma$, where κ is a scaling constant chosen to broaden or narrow the width of the damping function. This will shape an appropriate function built from equation (4.6), and the resulting damping term effectively removes the singularity but passes the bulk of the wavelike region. A sample plot of this damping function is shown in Figure 4.2 for $\kappa = 400$ and $v = 1000\text{m/s}$. Here κ is chosen to be relatively small in order to exaggerate the damping effect. Note that the function is nearly zero with a broad range around the evanescent boundary. A larger κ value will obviously narrow the damping breadth by scaling the range of the argument to $h(\varsigma)$. Practical values of

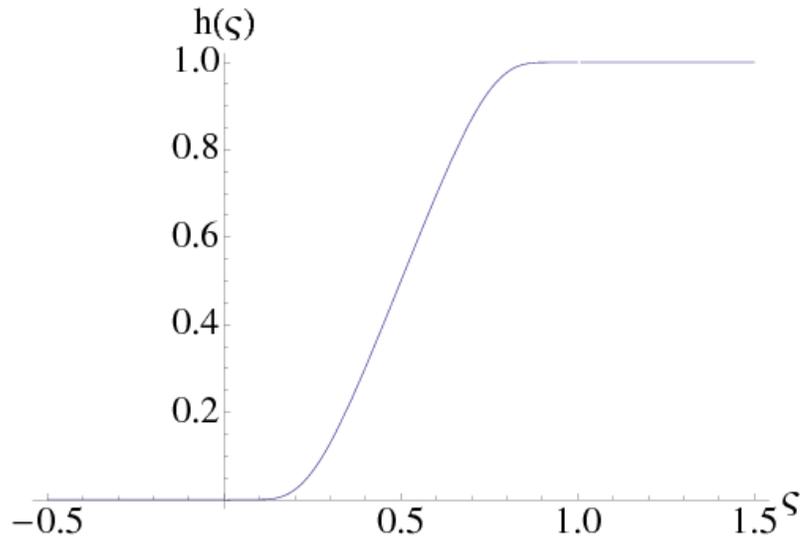


Figure 4.1: The damping function $h(\varsigma)$.

κ range between 1000 and 10 000. Within the wavelike region, the magnitude of α is 1, and within the evanescent region it decays exponentially. At the evanescent boundary there is a sharp corner in the magnitude of α . It is at this point where $|(\omega^2/v^2 - \xi^2)| = 0$ where the singularity occurs, and so it must be suppressed with the damping term. Figure 4.3 shows $|\alpha| = |e^{ik_z(x)}|$ for $v = 1000m/s$. Note that the evanescent boundary, shown as the sharp corner bounding the region where $|\alpha| = 1$, exactly corresponds to the locations where the damping function $h(\varsigma)$ takes values of 0, as shown in Figure 4.2. The damping function in equation 4.6 approaches zero when ς approaches zero. ς approaches zero as we approach the evanescent boundary, which is where the instabilities in k_z occur. Thus, the instabilities are effectively suppressed.

As previously stated, GPSPi algorithms may be implemented in the $\omega - \xi$ domain as a full Fourier-domain multiplication process. FOCI is the $\omega - x$ convolutional equivalent algorithm. Therefore, it makes sense to consider the resulting operator

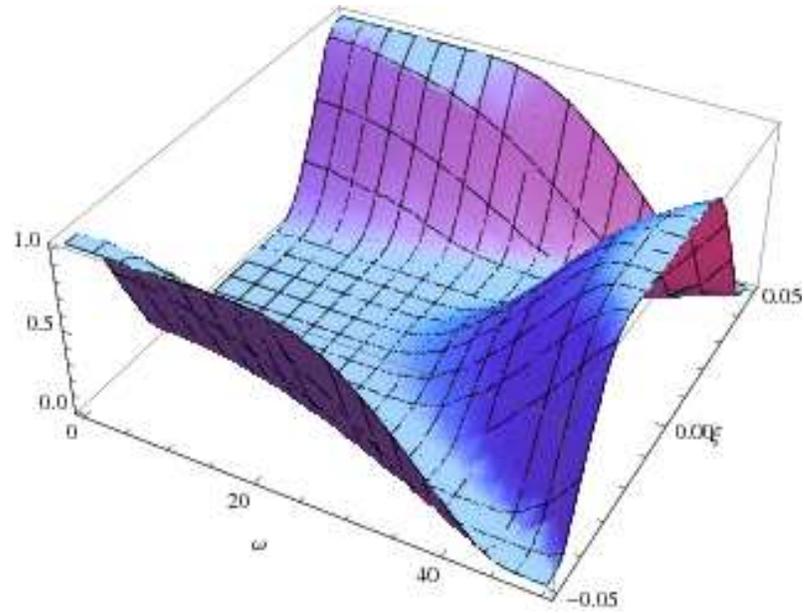


Figure 4.2: The damping function $h(\kappa |(\omega^2/v^2 - \xi^2)|)$ for $\kappa = 400$ and $v = 1000\text{m/s}$.

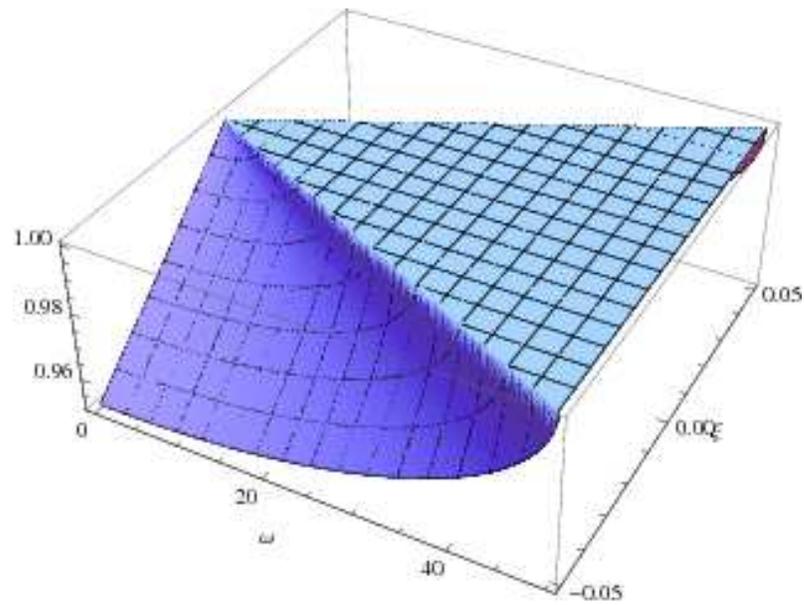


Figure 4.3: $|\alpha| = |e^{ik_z(x)}|$ for $v = 1000\text{m/s}$.

symbol/kernel in both the $\omega - x$ and $\omega - \xi$ domains. In Figure 4.4, the Stolk operator symbol and kernel for frequency $70Hz$, $v = 2000ms^{-1}$, with no horizontal velocity (slowness) gradient is displayed. This operator is exactly equivalent to the usual FOCI operator for these parameters. The top panel in this figure shows the convolutional kernel in $\omega - x$. This kernel is convolved with the Fourier-transformed traces in space. The middle panel shows the amplitude spectrum of the equivalent symbol – *ie* the equivalent $\omega - \xi$ domain multiplier. This amplitude spectrum is ideally identically one throughout the wavelike region, and decays exponentially in the evanescent region. In this figure, we see that the amplitude spectrum is one throughout most of the wavelike region, but instabilities at the evanescent boundary (around ± 0.035) are visible. The bottom panel in this figure shows the applied phase-rotation for each wavenumber ξ . The phases match well throughout most of the region, though they differ slightly near the evanescent boundary.

Referring to Figure 4.5, this operator may be directly compared to the Stolk operator for precisely the same medium but with one difference: a local slowness gradient equivalent to a velocity gradient of approximately $35s^{-1}$. Notice the dramatic change in both the $\omega - x$ and the $\omega - \xi$ representations. The operator is clearly asymmetric, reflecting the asymmetric nature of the underlying medium.

The effect is obviously frequency dependent. At low ($15Hz$) frequency, the resulting operator can be seen in Figure 4.6, and at intermediate frequency ($30Hz$) the operator can be seen in Figure 4.7. Changing the sign of the velocity gradient also reverses the asymmetry of the operator, as seen in Figure 4.8, in which the $30Hz$ case is recalculated with an effective velocity gradient of $-4s^{-1}$. Additionally, it can be seen in this figure that a weaker velocity gradient ($-4s^{-1}$ compared to $35s^{-1}$) results in a weaker distortion of the otherwise-flat operator in the $\omega - \xi$ plots.

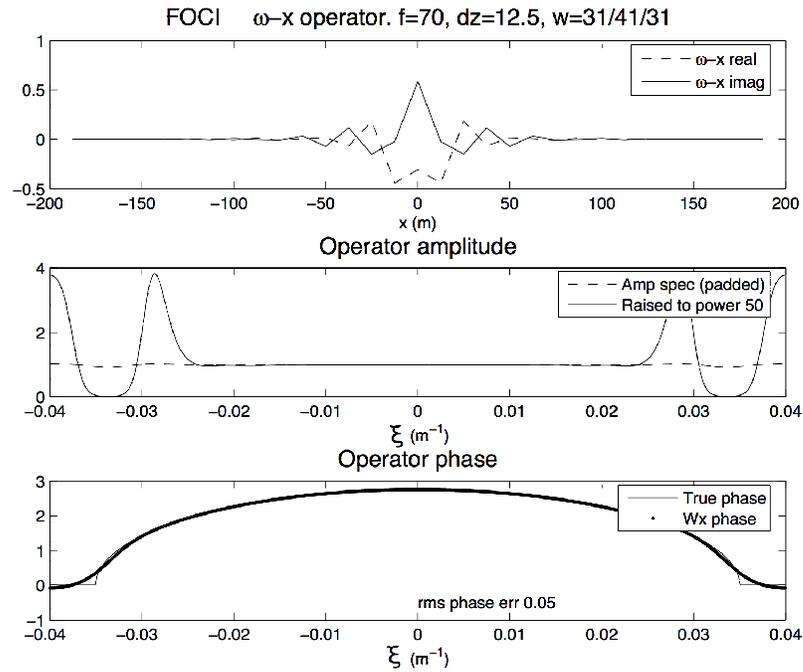


Figure 4.4: The Stolk operator symbol in $\omega - x$ and $\omega - \xi$ domain for $70Hz$ at $2000ms^{-1}$ with no horizontal velocity (slowness) gradient – *ie* this operator is exactly equivalent to FOCI. The top panel shows the $\omega - x$ expression of the symbol, the middle panel shows the operator symbol amplitude in $\omega - \xi$, and the bottom panel shows the operator symbol phase in $\omega - \xi$.

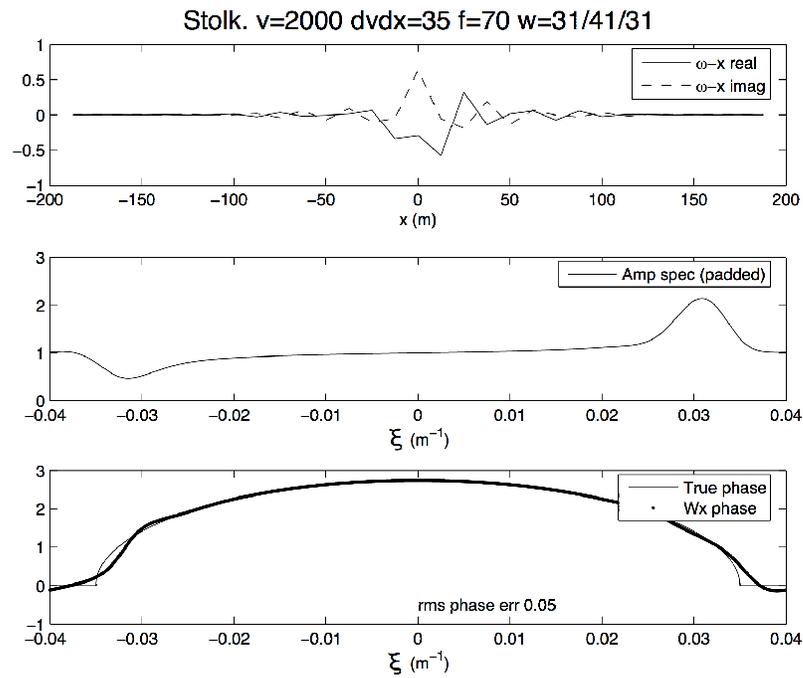


Figure 4.5: The Stolk operator symbol in $\omega - x$ and $\omega - \xi$ domain for $70Hz$ at $2000ms^{-1}$ for $35s^{-1}$ velocity gradient. The top panel shows the $\omega - x$ expression of the symbol, the middle panel shows the operator symbol amplitude in $\omega - \xi$, and the bottom panel shows the operator symbol phase in $\omega - \xi$.

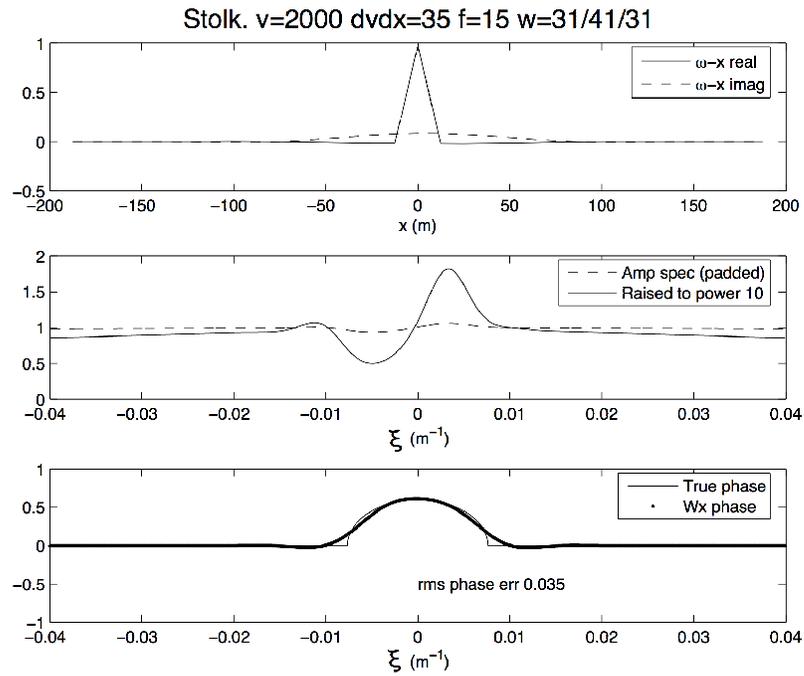


Figure 4.6: The Stolk operator in $\omega-x$ and $\omega-\xi$ domain for $15Hz$ at $2000ms^{-1}$ for $35s^{-1}$ velocity gradient. The top panel shows the $\omega-x$ operator, the middle panel shows the operator amplitude in $\omega-\xi$, and the bottom panel shows the operator phase in $\omega-\xi$.

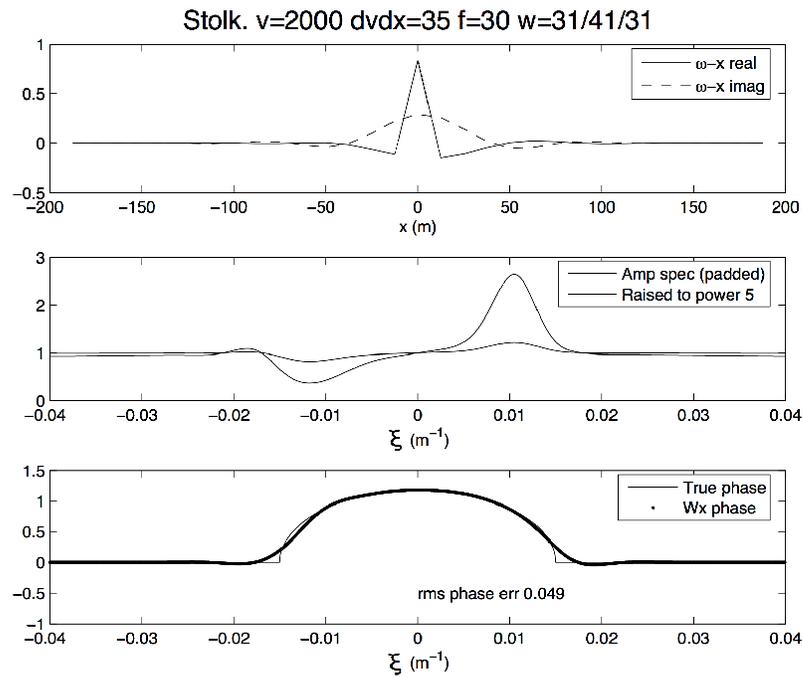


Figure 4.7: The Stolk operator in $\omega - x$ and $\omega - \xi$ domain for 30Hz at 2000ms^{-1} for 35s^{-1} velocity gradient. The top panel shows the $\omega - x$ operator, the middle panel shows the operator amplitude in $\omega - \xi$, and the bottom panel shows the operator phase in $\omega - \xi$.

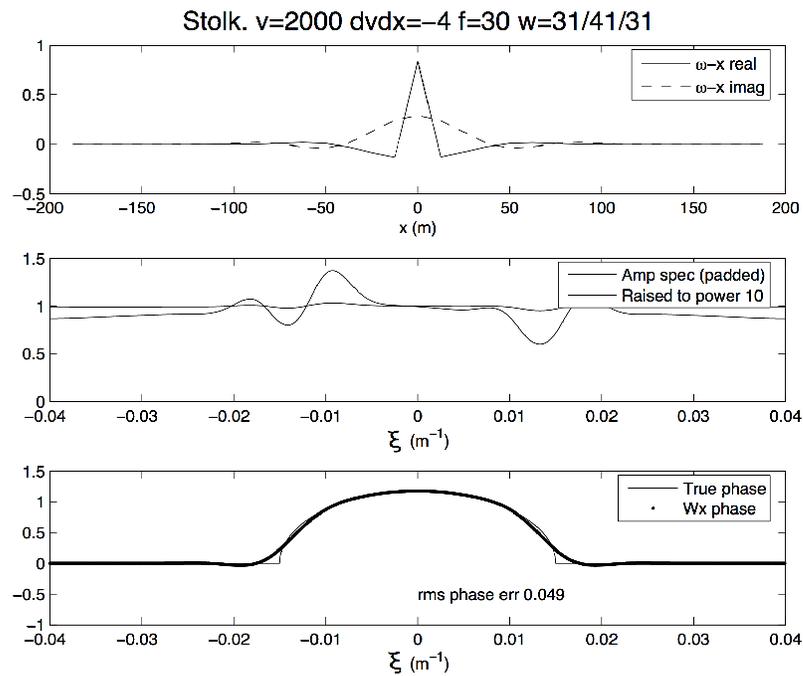


Figure 4.8: The Stolk operator in $\omega - x$ and $\omega - \xi$ domain for $30Hz$ at $2000ms^{-1}$ for $-4s^{-1}$ velocity gradient. The top panel shows the $\omega - x$ operator, the middle panel shows the operator amplitude in $\omega - \xi$, and the bottom panel shows the operator phase in $\omega - \xi$.

4.2 Testing

The Stolk symbol was implemented in both the standard GPSPI and FOCI fashion, using existing MATLAB codes developed by CREWES and POTSI. Impulse response and focussing-power tests were calculated using standard GPSPI methods, while full pre-stack depth migrations were calculated using the FOCI algorithm.

4.2.1 Impulse response

The impulse response of the Stolk operator in a homogeneous medium is compared to the response of the operator using the standard square-root in Figure 4.9. The

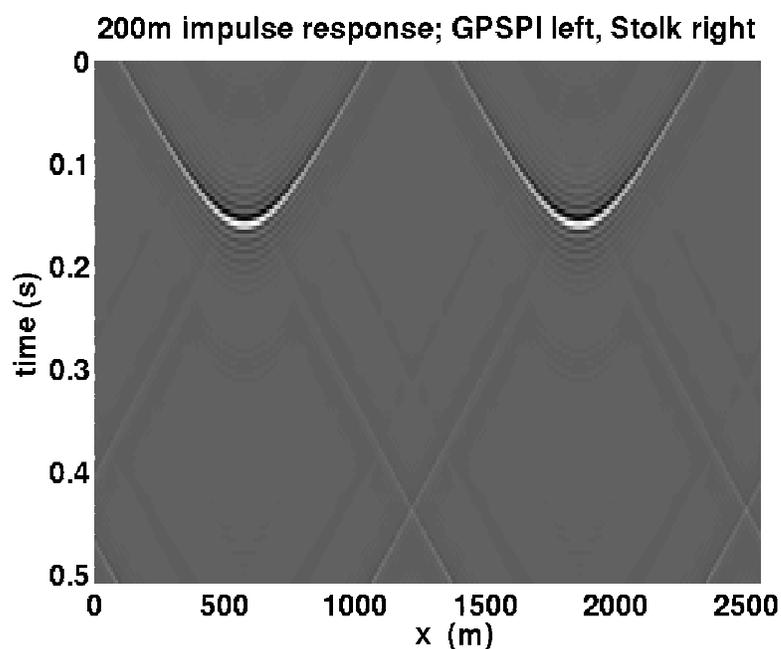


Figure 4.9: Impulse response of standard symbol on the left compared to Stolk on the right, both symbols propagated an impulse 200m through a homogeneous medium.

responses are identical in appearance as they should be for a homogeneous medium.

They are numerically identical as well. The impulse response of the symbols through

an inhomogeneous medium, specifically one with a strong horizontal velocity profile, is shown in Figure 4.10. Once again, the symbols look very similar. The difference,

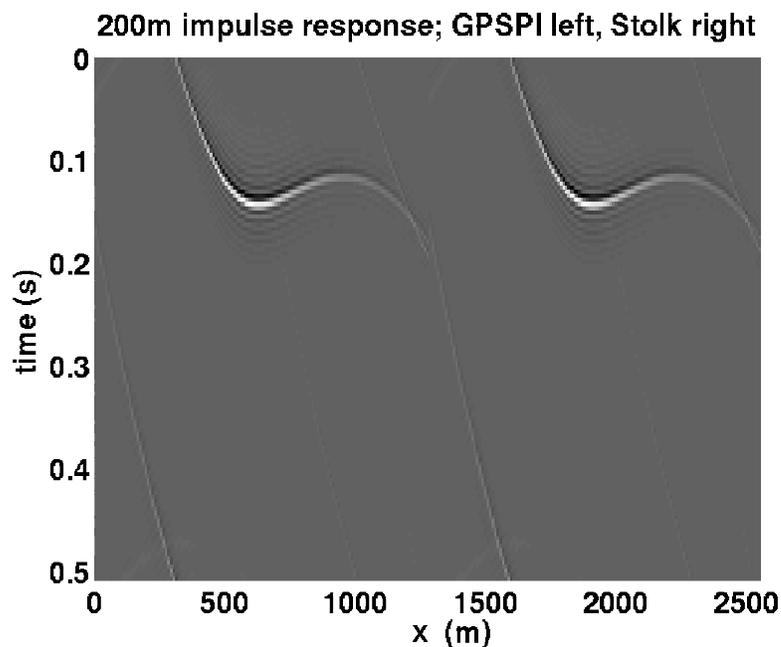


Figure 4.10: Impulse response of standard symbol on the left compared to Stolk on the right, both symbols propagated an impulse 200m through a medium with a strong velocity gradient.

however, is only apparent upon subtraction of the two impulse responses as seen in Figure 4.11. Overall, the character of both impulse responses is extremely similar. In Figure 4.10 the strength of both impulses is nearly identical. The Stolk response appears to be somewhat smoother qualitatively, and this makes up the bulk of the difference shown in Figure 4.11.

4.2.2 Focussing power

While it is interesting to observe the impulse response of the operator, the focussing power may be measured by back-propagating a wavefield that is the result of an

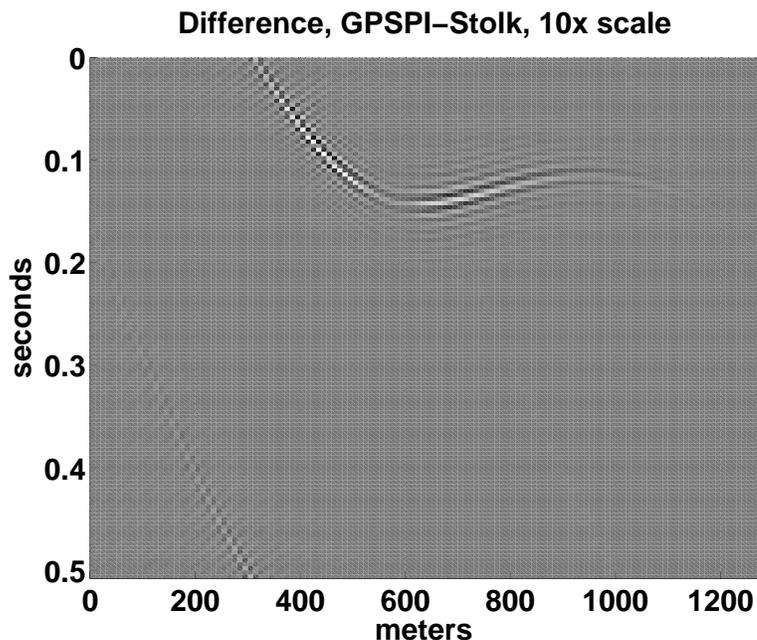


Figure 4.11: Difference of standard symbol and Stolk symbol, magnified 10 times.

impulse that has been forward-propagated by an exact¹ extrapolator. In this case, the exact forward propagation was accomplished using the method of Pai (1988). After the back-propagation is executed, in principle the impulse should be recovered.

First, the “high-velocity lens” model (Figure 4.12) was tested. After a propagation through 200m with the exact extrapolator, the result was refocussed using GPSPi and Stolk operators. The resulting point is shown in Figure 4.13.

Visually, there is little difference between the focussing power of the two operators. Numerically there is very little difference as well. The energy density as a function of radius from the central point was measured and plotted in Figure 4.13. The energy density of the Stolk-focussed point is virtually indistinguishable from the GPSPi-focussed point.

A more complicated velocity model was also tested. In this case, it is a section

¹...and very computationally expensive

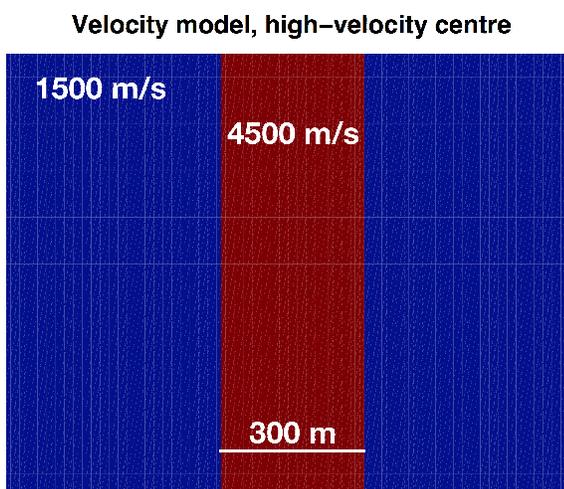


Figure 4.12: High-velocity lens model. The centre is 4500 m/s , with sides at 1500 m/s . Total width of the central portion is approximately 300 m .

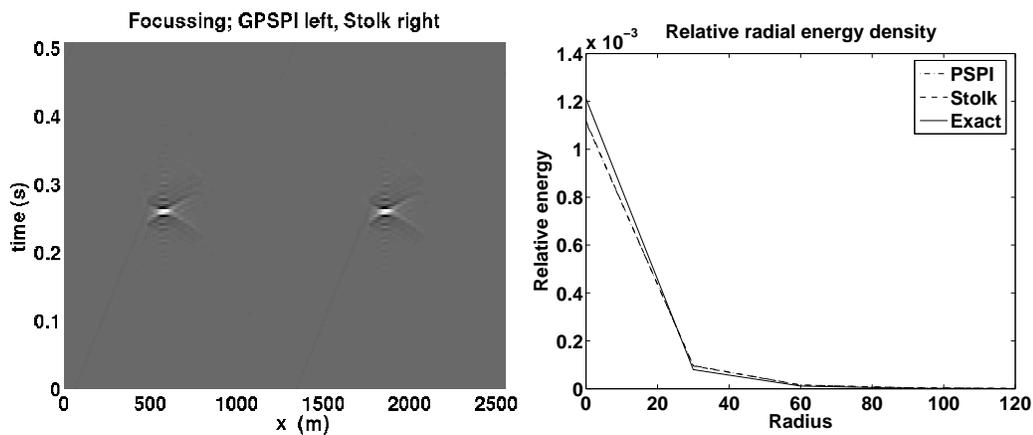


Figure 4.13: Focussing power of GPSPI and Stolk operators for the high-velocity lens model on the left. Radial energy density for the high-velocity lens model on the right. The GPSPI and Stolk lines are nearly indistinguishable.

of the Marmousi model as described by Bourgeois et al. (1991). The section used is shown in Figure 4.14. The resulting focussed points are compared in Figure 4.15,

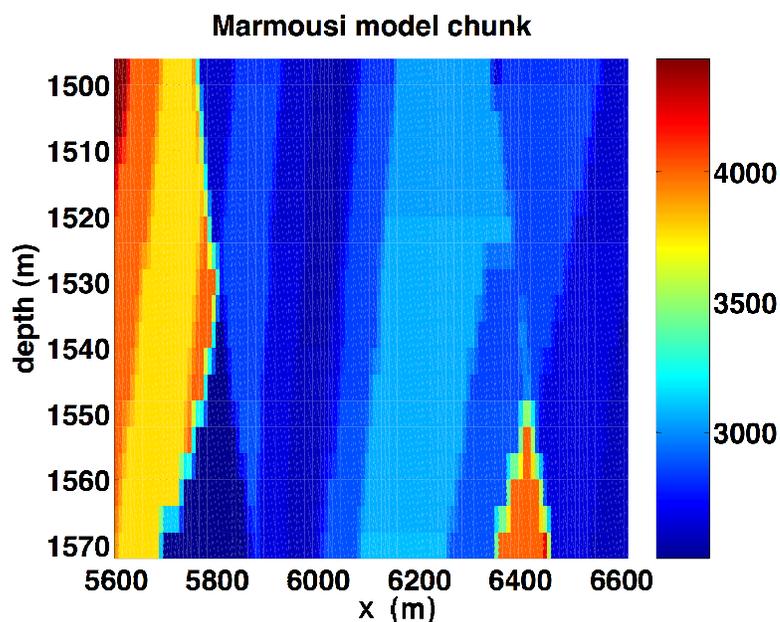


Figure 4.14: Marmousi chunk model, with velocity in m/s

with relative energy densities compared in Figure 4.15. Once again, the focussing powers of the operators are very nearly identical.

4.2.3 Marmousi migration

In addition to impulse response and focussing power tests, full prestack migrations of the Marmousi model were calculated using FOCI. The standard FOCI operator was compared to a modified version that included the Stolk correction term in the operator design phase but was otherwise identical in all operational parameters. Specifically, the operators were calculated with a forward operator size of 41 points, an inverse operator size of 51 points, and a final window size of 41 points². A

²The number of FOCI operator “points” corresponds to the number of spatial grid points over which the convolutional operator is non-zero. A larger number of points corresponds to a more

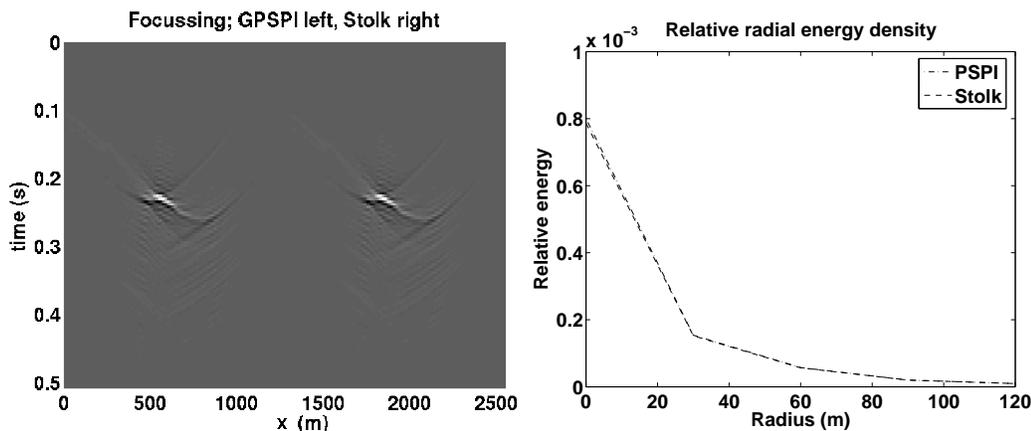


Figure 4.15: Focussing power of GPSPi and Stolk operators for the Marmousi chunk model on the left. Radial energy density for the Marmousi chunk model on the right.

standard FOCI image is shown in Figure 4.16, and the resulting Stolk image is shown in Figure 4.17. The two images are very similar.

Several subsections of the image are compared more closely in Figures 4.18–4.21. Each figure contains four panels. On the top left, the $v(z)$ image is displayed. On the bottom left, the FOCI image is displayed. On the top right, the corresponding region from the Marmousi velocity model is displayed. On the bottom right, a histogram of the colourbar usage of both images is shown, in order to allow for an accurate comparison of the relative strengths and weaknesses of amplitudes within the images. In the first region (Figure 4.18) there is almost no visual difference between the two images at all. The second region (Figure 4.19) shows slightly better definition and continuity, especially in the neighbourhoods of (5400, 1300) and (5700, 1300). These differences are extremely small, however. The third region (Figure 4.20) reveals a very slight improvement in continuity near the top left, near (6200, 500). The fourth

accurate and stable operator, while a smaller number of points corresponds to a faster but possibly more unstable operator due to the truncation of the operator and resulting Gibbs phenomenon.

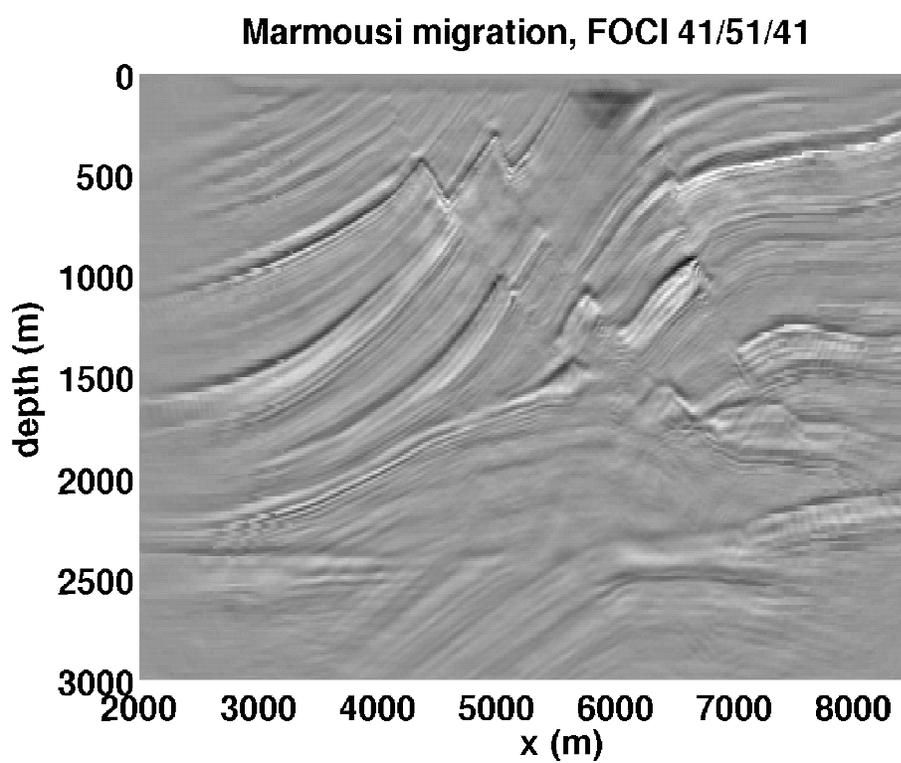


Figure 4.16: Marmousi migration in standard FOCI with 41/51/41 windowing.

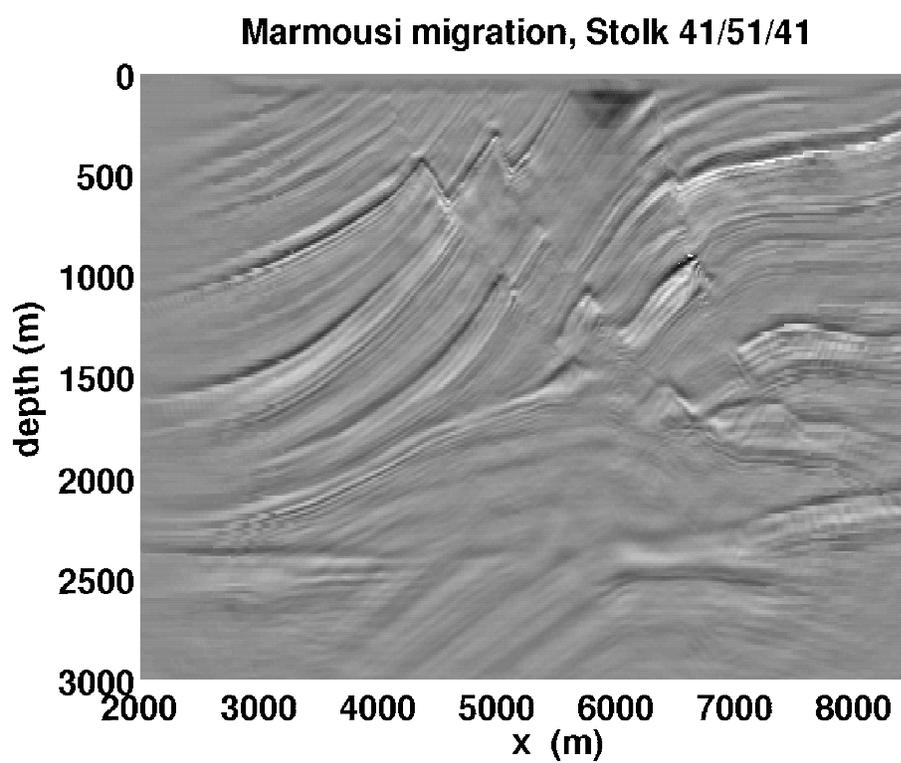


Figure 4.17: Marmousi migration using Stolk correction, run with 41/51/41 windowing.

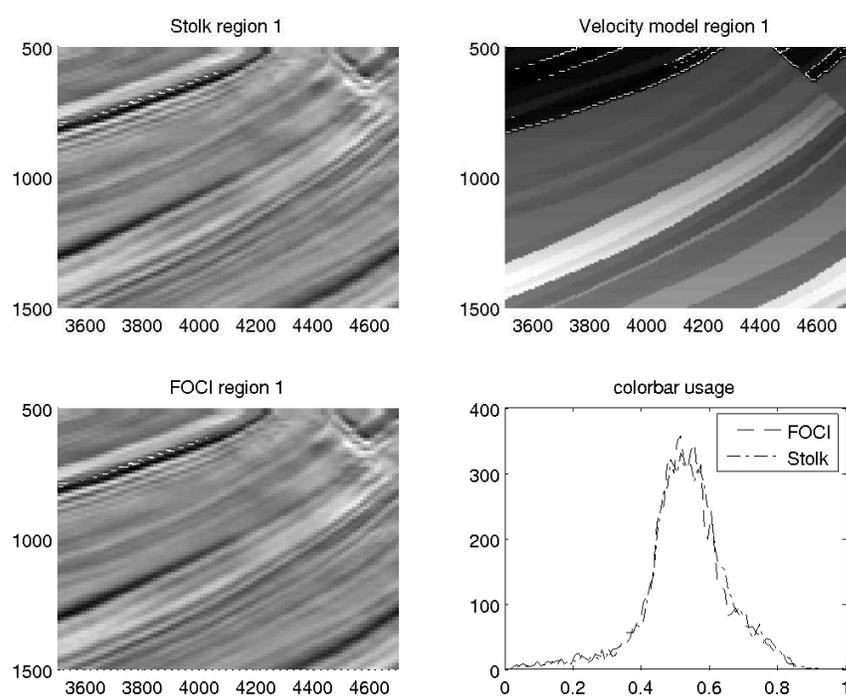


Figure 4.18: Marmousi migration, Stolk vs. FOCI migration with 41/51/41 operators. Region 1.

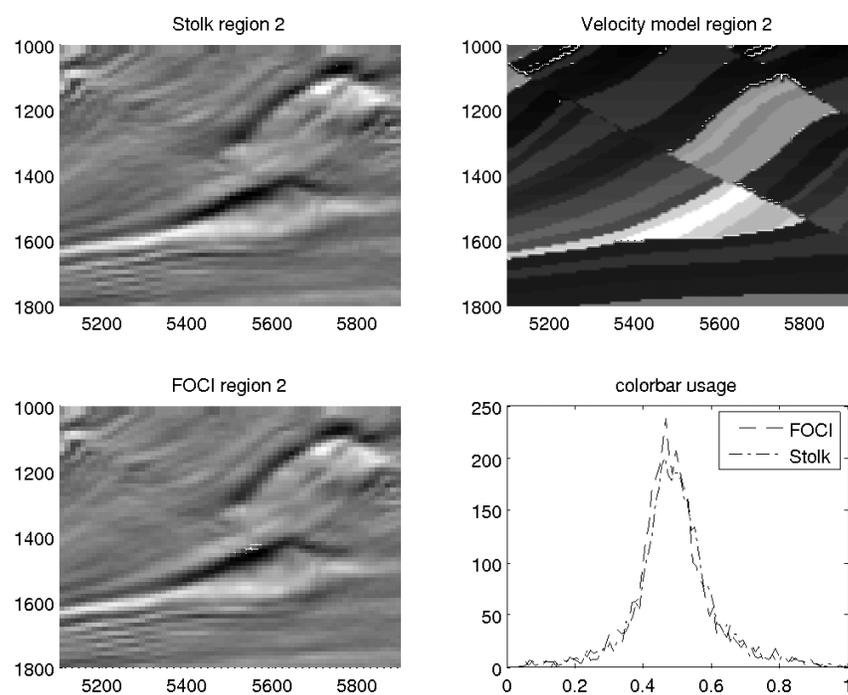


Figure 4.19: Marmousi migration, Stolk vs. FOCI migration with 41/51/41 operators. Region 2.

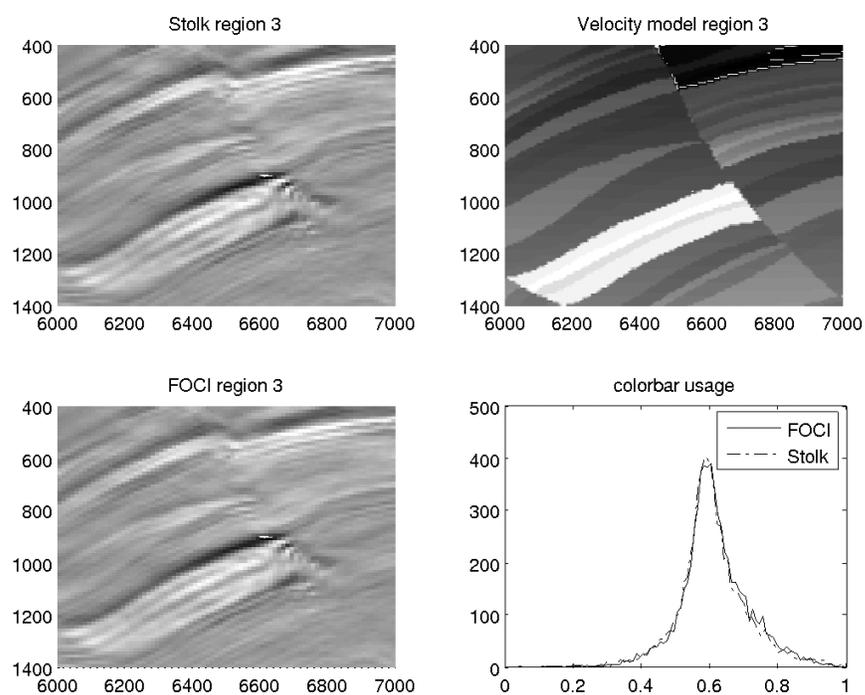


Figure 4.20: Marmousi migration, Stolk vs. FOCI migration with 41/51/41 operators. Region 3.

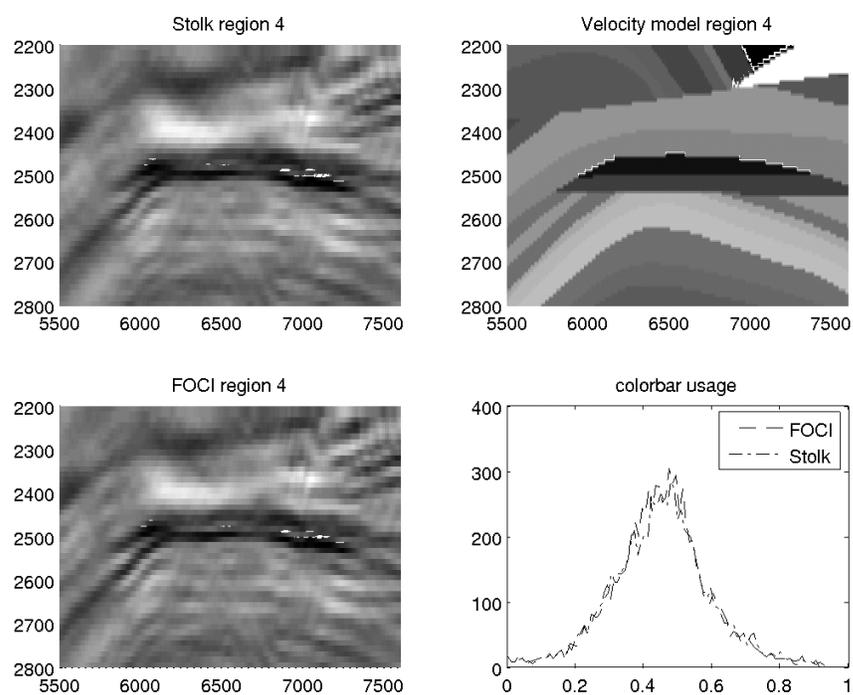


Figure 4.21: Marmousi migration, Stolk vs. FOCI migration with 41/51/41 operators. Region 4.

region (Figure 4.21) shows perhaps slightly less definition in the Stolk image. This is evident in a small loss of resolution around (6500, 2500).

Overall, the Stolk image seems to have the theoretical potential for better imaging, but in the migration of the Marmousi image it has produced very little actual improvement.

4.3 Discussion of imaging performance

The overall imaging performance of the Stolk operator in the migration of the Marmousi model was disappointing. It seemed to produce a very slight improvement. There were cases, however, in which the Stolk operator actually produced worse results. It was noted that performance of the Stolk operator seemed to be improved in the shallower regions, and degraded deeper down. It is possible that instabilities in the $\omega - x$ representation of the operator make it less suitable for use. This could be addressed by an alternate implementation of this method that does not make use of the FOCI algorithm, or by tuning of the FOCI algorithm as more is understood about the Stolk operator's behaviour within the FOCI framework.

Additionally, it should be noted that the Stolk correction is dependent on strong high-frequency assumptions in its derivation. The Marmousi model is not an especially high-frequency data set. We are trying to image complicated and finely-layered structures with frequencies of at most $\sim 60Hz$ with velocities between $1500ms^{-1}$ and $4500ms^{-1}$. This corresponds to wavelengths between $25m$ and $75m$ in the best case. For this structural model, a "high-frequency" assumption really only has validity in a regime where the frequencies are significantly higher than $60Hz$ if we are interested in imaging with $25m$ wavelengths

4.4 Discussion of computational performance

All calculations were performed using MATLAB 7.1 on Linux-based PC computers with 3.06 GHz Intel Pentium 4 CPUs. The running performance of the FOCI-based Stolk migration is comparable to the running performance of standard FOCI. Standard FOCI completes the Marmousi migration in approximately 16 hours. Experience suggests that using Stolk operators adds approximately 10% more time to the imaging. The extra time is largely consumed in the overhead of selecting the correct operator from a two-dimensional table (velocity and slowness derivative) instead of just a one-dimensional (velocity only) table as for standard FOCI.

Unfortunately, the precalculation of the Stolk operators requires more computational time than the migration itself. In fact, the operator calculations require approximately 1.5 to 2 times as much calculation as the generation of the image! Standard FOCI operator calculation times are insignificant compared to the calculation time of the total image.

We believe, however, that there is significant room for improvement here. A more thorough analysis of the table calculation requirements may reveal a far more efficient manner of determining the optimal partitioning of the slowness gradient range into discrete operators. Currently, the procedure is to divide the slowness derivative dimension into the same number of elements as are contained in the velocity dimension. This procedure, combined with the substantial increase in computation time for each individual operator, makes the table calculations more than an n^2 operation. Perhaps a more sparse slowness gradient dimension would produce an adequate image, or perhaps a non-linear coverage of the slowness gradient dimension would produce an even higher quality image with less overall table computation.

Even the problem as it stands is not unmanageable, however. Once a table is calculated, it never need be calculated again (for the same parameters). Therefore, a database may store these calculations for future reuse, allowing an asymptotic approach to standard FOCI calculation times.

4.5 Chapter summary

By incorporating the local derivative in slowness, the Stolk correction is a first step towards the inclusion of more information about the local propagation region within the GPSP migration framework. This may eventually provide a useful imaging benefit for high-quality data sets. However, at this moment it is not clear that the benefit is worth the additional computational cost, and instabilities in the propagation make it difficult to compute optimally. Further research may reveal much more efficient ways to calculate and use the required operators more effectively.

Chapter 5

Feasibility testing of time-lapse seismic monitoring with early-arrival waveform tomography.

5.1 Introduction

Early-arrival waveform tomography represents a natural low-frequency extension of the concept of standard travelttime tomography. In this approach, the whole waveform of the transmitted signal is used in the tomographic process, rather than simply the first arrival time. This method originated 25 years ago with the work of Lailly (1983), Tarantola (1984), and Mora (1987). It has been subsequently developed by many others, including Woodward (1992). Lately the primary champion of this method has been Gerhard Pratt and his research group. A subset of these publications includes Pratt (1990); Pratt and Worthington (1990); Pratt et al. (1998); Pratt (1999); Pratt and Shipp (1999); Sirgue and Pratt (2004); Brenders and Pratt (2006, 2007). An introduction to this method is given in §1.6.2.

To date, applications of this method to the analysis of time-lapse seismic data have been sparse. We believe that this method is ideally suited to this analysis, however. In time-lapse analysis, it is common to have a well-established velocity model of the relevant geology, due to past surface seismic surveys, VSP investigations, and well-log data. The waveform tomography process requires a sufficiently accurate starting model, typically constrained by requiring that diving waves travelling from source to receiver in the starting model must have first-arrival times within roughly

half the period of a given frequency multiplied by the average velocity along the travel path (Pratt, 1999).

If it is possible to construct a velocity model that accurately models the first-arrival of seismic waveforms on a baseline survey, then it may be possible to use waveform tomography to analyze time-lapse changes in the imaged region. Small, subtle local perturbations to this velocity field should in principle be recoverable given sufficient data.

The purpose of this study is to investigate two major questions. First, is waveform tomography at all feasible for time-lapse monitoring? Second, if so, what acquisition parameters would maximize the effectiveness of this method?

5.2 Feasibility testing on synthetic data

Our primary test model is a laterally-homogeneous section derived from the horizontal extension of a P-wave sonic log from the Pikes Peak field. This 2D model is displayed in Figure 5.1. A perturbation of -500 m/s over an area of approximately 100m horizontally and 30m vertically is introduced to simulate the effect of steam injection or some other physical process with a similar net result on the seismic velocity of the region. This perturbation may be seen in Figures 5.2 and 5.3. Raytracing, shown in Figure 5.4, through a smoothed version of this velocity model reveals that our survey size is sufficient (barely) to capture the diving waves that pass through the perturbed region. Without this smoothing, these raypaths are reflected before they reach the target depth. This implies that only relatively low frequency diving waves will reach the target, and that the longest offsets will be required for inversion. The model dimensions were constrained by computational

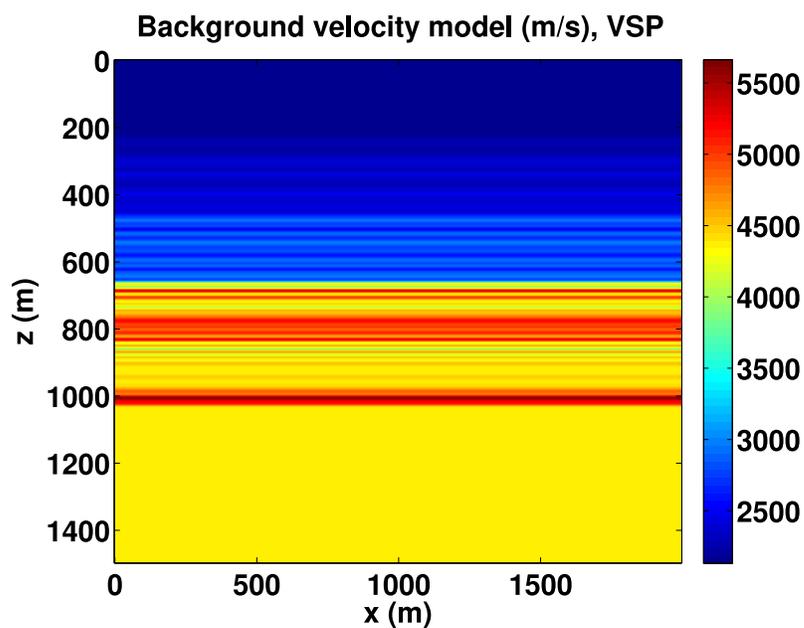


Figure 5.1: Original velocity model with no time-lapse change.

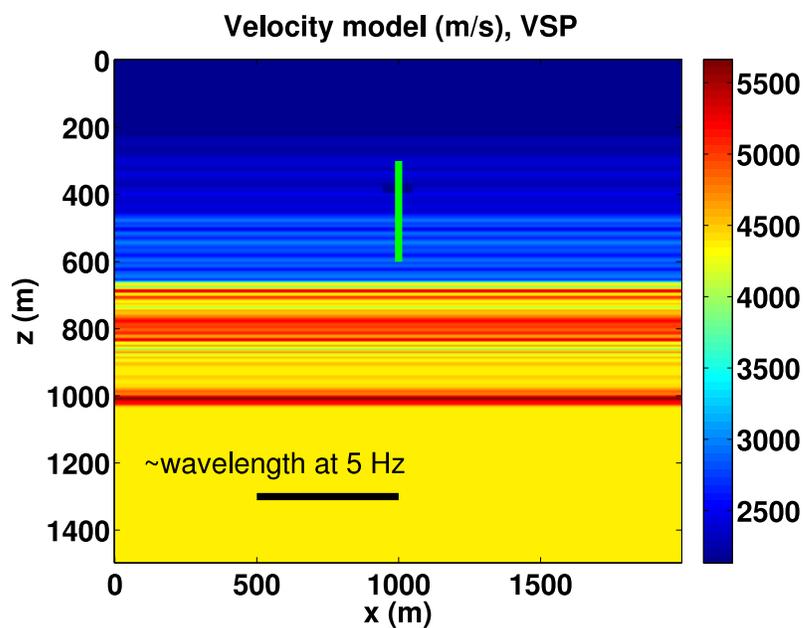


Figure 5.2: Original velocity model plus steam-injection effect. The location of the VSP receivers is marked as a vertical green line at the injection site. The approximate wavelength at 5 Hz in the injected region is displayed for scale.

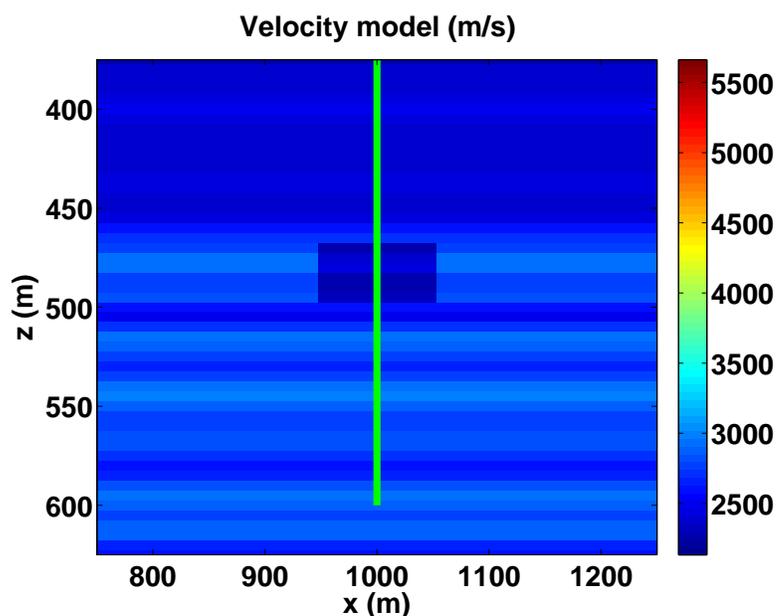


Figure 5.3: Original velocity model plus steam-injection effect. The location of the VSP receivers is marked as a vertical green line at the injection site.

limitations.

5.2.1 Surface seismic modelling and inversion procedure

The forward modelling was performed on the perturbed model using a 2D acoustic frequency-domain finite difference code as provided by Gerhard Pratt. This is the same forward-modelling code that is used in the inversion algorithm, and so is committing the so-called “inverse crime”. Future work could involve extending this to using external acoustic and elastic codes. A sample time slice of this modelling is shown in Figure 5.5.

The simulated seismic surface reflection survey was recorded with receivers placed along the surface of the model at 10 m spacing. Sources were placed with 20 m spacing. Sources and receivers were located across the entire 2000 m extent of the survey.

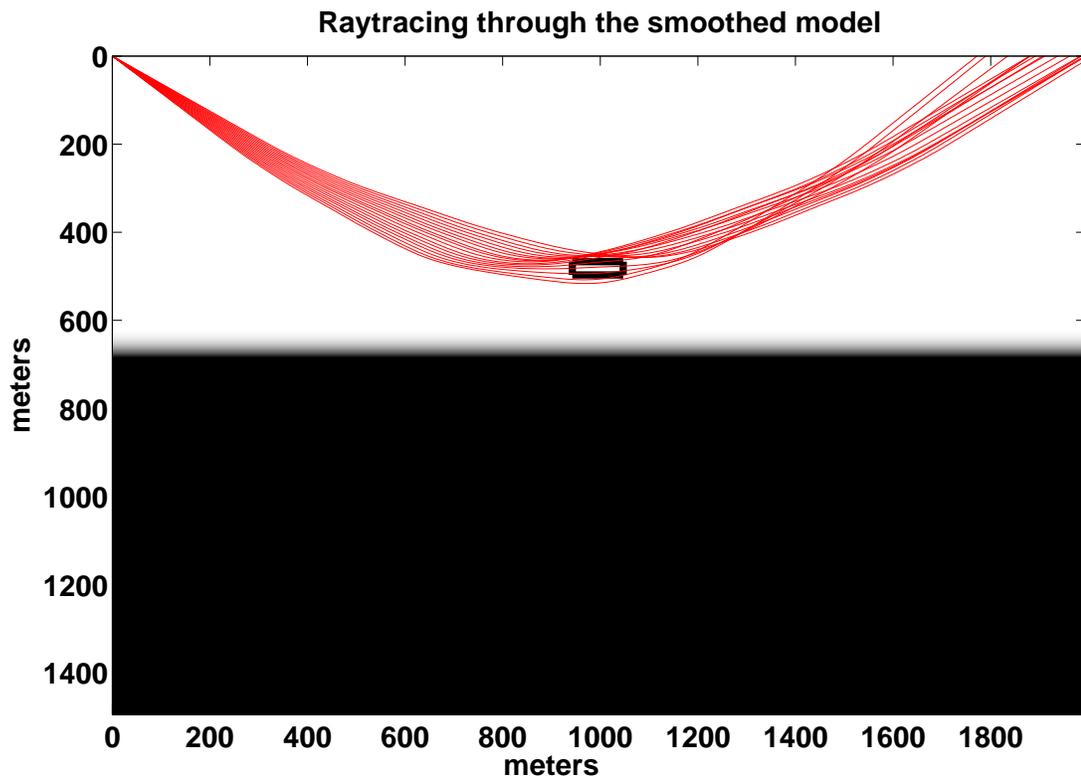


Figure 5.4: Raytracing through the original model convolved with a 20m Gaussian smoother

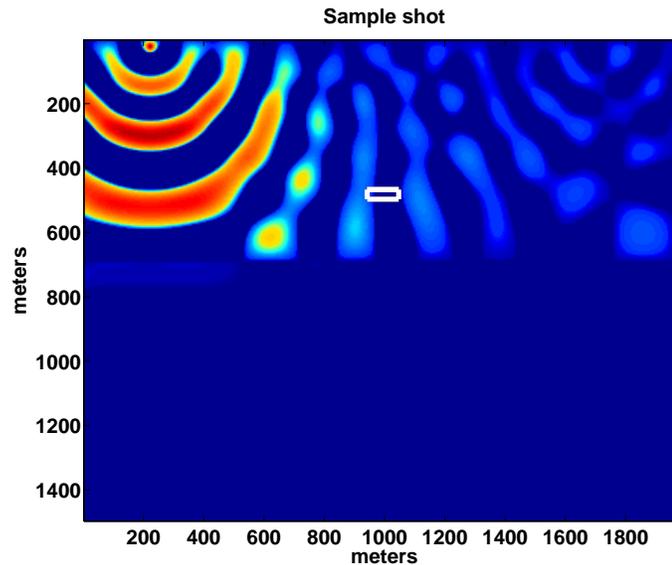


Figure 5.5: Sample time slices of the forward-modelled data on the perturbed model showing the wavefront periodically as it progresses through the medium. The white rectangle delineates the extent of the perturbation. Very little energy penetrates beneath the horizontal reflector at depth $\sim 700m$

The waveform tomography inversion was then performed using the original (unperturbed) background velocity model as its starting point, carried out with the reference implementation of the algorithm provided by Gerhard Pratt. Constant-frequency inversions were carried out beginning at 5 Hz, and then using this result as input into a 6 Hz inversion. Although in many cases it is possible to use many (or few) frequencies to optimize the convergence (Sirgue and Pratt, 2004), for this inversion we found that results were best with an inversion beginning no higher than 5 Hz, and that beyond 6 Hz no appreciable improvement was detectable.

All inversions were constrained to update the model only within a region of 500 m by 500 m, centred at the anomaly. This stabilizes the inversion, and would be a reasonable (perhaps even overly conservative) constraint for the inversion of realistic data. This constrained region is shown in all difference plots of the inversion results.

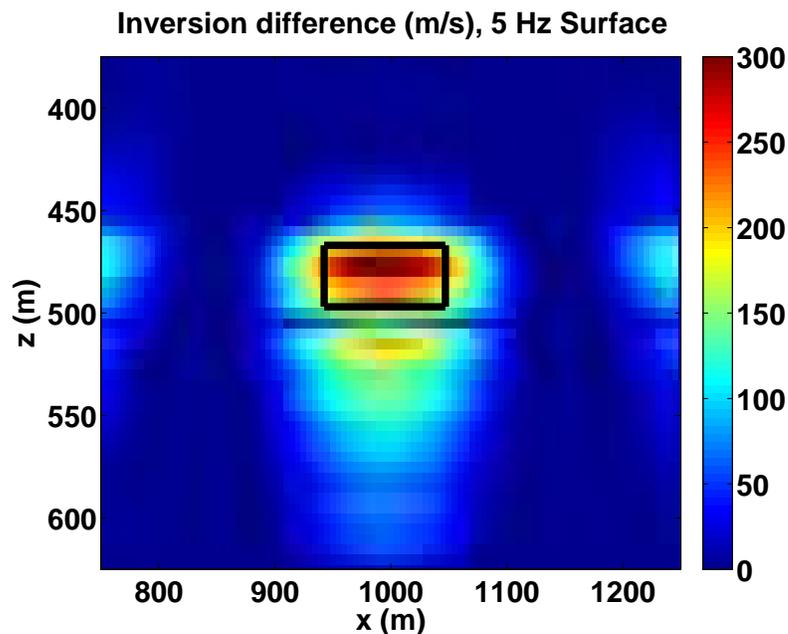


Figure 5.6: The resulting surface-seismic velocity estimate from inversion at 5 Hz, as a difference from the starting (background) velocity model. Ideally, the amplitude of the perturbation would be exactly +500 m/s. One wavelength at 5 Hz would approximately span this entire displayed region at the velocity of the target region. The black box marks the spatial extent of the actual perturbation.

5.2.2 Surface seismic inversion results

The updated velocity model with the 5 Hz (Figure 5.6) and 5, 6 Hz (Figure 5.7) inversions are shown as a difference-plot with respect to the starting (background) velocity model, zoomed into the region of interest shown in Figure 5.2.

5.2.3 VSP modelling and inversion procedure

The same starting velocity models were also used in a simulated VSP survey. In this survey, source locations across the 2000m extent of the model were used at 20m spacing. Receivers were placed in a borehole from 300m to 600m deep, at 10m spacing. This well bore bisected the perturbed (steam-injection) site. This borehole

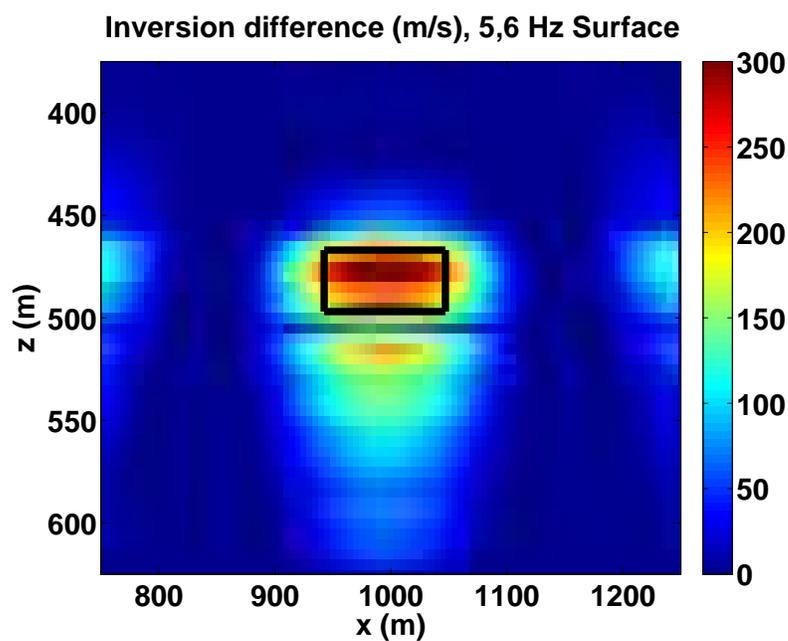


Figure 5.7: The resulting surface-seismic velocity estimate from inversion at 5 Hz, as a difference from the starting (background) velocity model. Ideally, the amplitude of the perturbation would be exactly +500 m/s. One wavelength at 5 Hz would approximately span this entire displayed region at the velocity of the target region. The black box marks the spatial extent of the actual perturbation.

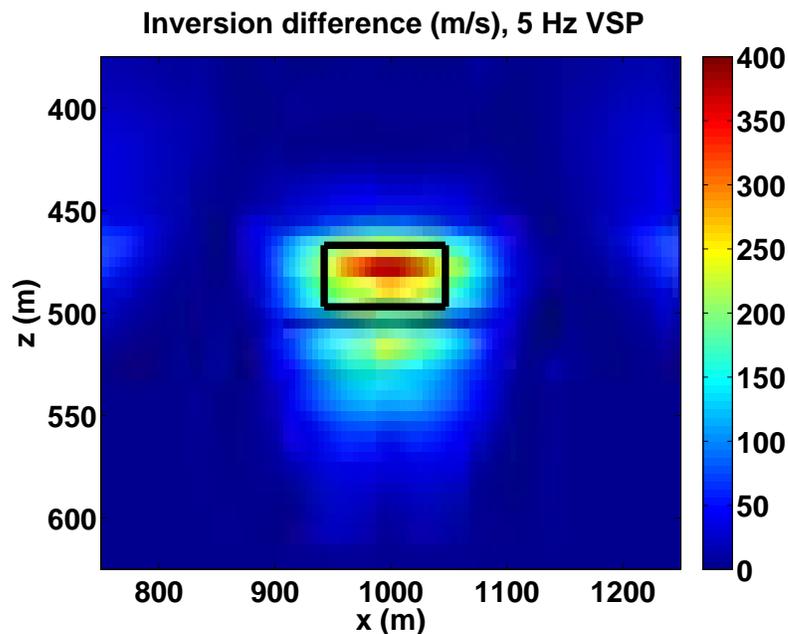


Figure 5.8: The resulting VSP velocity estimate from inversion at 5 Hz, as a difference from the starting (background) velocity model. Ideally, the amplitude of the perturbation would be exactly +500 m/s. One wavelength at 5 Hz would approximately span this entire displayed region at the velocity of the target region. The black box marks the spatial extent of the actual perturbation.

is marked in green in Figure 5.2.

The updated velocity model with the 5 Hz and 5, 6 Hz inversions are shown in Figures 5.8 and 5.9 respectively, again as a difference-plot with respect to the starting (background) velocity model zoomed into the region of interest shown in Figure 5.2.

5.3 Discussion of the results

The waveform inversion procedure is providing significant updates to the background velocity model at well below the wavelength scale. Both surface seismic and VSP approaches yielded useful updates to the model that were consistent with the true

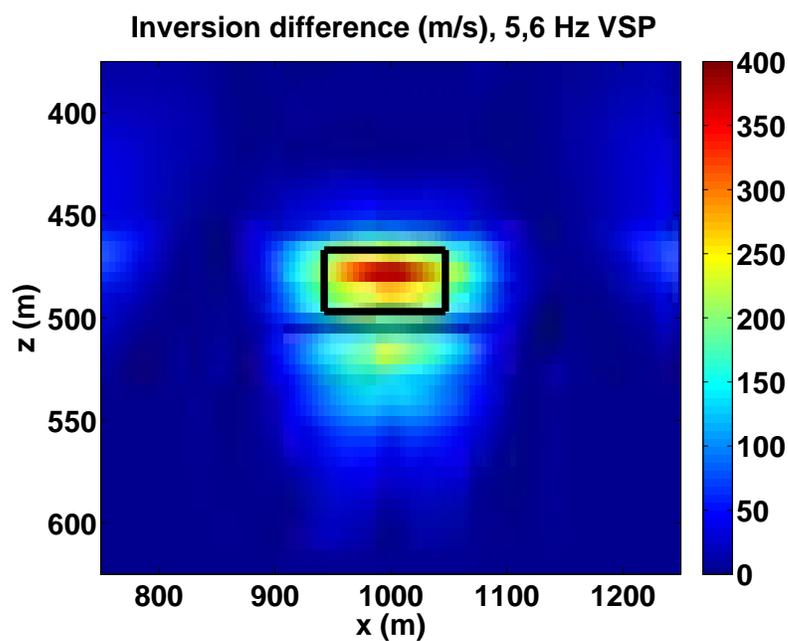


Figure 5.9: The resulting VSP velocity estimate from inversion at 5 and 6 Hz, as a difference from the starting (background) velocity model. Ideally, the amplitude of the perturbation would be exactly +500 m/s. One wavelength at 5 Hz would approximately span this entire displayed region at the velocity of the target region. The black box marks the spatial extent of the actual perturbation.

anomaly both in terms of spatial extent and in magnitude. Although the maximum amplitude of the anomaly (~ 300 m/s for the surface seismic, and ~ 400 m/s for the VSP survey) is somewhat less than the true amplitude (500 m/s), this is not surprising as the discovered spatial extent of the anomaly is somewhat larger than the true anomaly.

Both images gained the most benefit from 5 Hz data, with some minor improvement with the addition of 6 Hz. Although it was hoped that higher frequencies would focus the image better, in practice, higher frequencies did not converge effectively within the region of constraint. We speculate that this is due to the higher frequency components of the wavefields reflecting off strong contrasts in shallower regions.

5.4 Chapter summary and potential future work

The results we have seen strongly indicate that waveform tomography merits further investigation as a viable method for analyzing a time-lapse signal in seismic data. Both surface and VSP methods provided comparable images, with slightly better results from the VSP method.

Seismic source considerations are significant. First, this “ideal conditions” inversion required 5 Hz data. Although explosive-source surveys easily contain this frequency and lower, vibration-source surveys often begin their sweep at frequencies higher than 5 Hz.

Acquisition geometry is also significant. In this test case, raytracing revealed that only the longest offsets (nearly 2 km for a 500 m deep target) contributed significantly to the inversion. Also, although VSP surveys with geophones directly

in the zone of interest are useful for reflection surveys, and can in some cases provide more information than surface seismic, in this case there was very little difference in results. We speculate that a crosswell survey or a VSP in a nearby observation well, either providing many raypaths travelling through the zone of interest, will yield improved results.

There are many open issues that could be addressed in the near future. These are broadly grouped into at least two sections: modelling and inversion methods, and acquisition and practical concerns.

For potential future work in modelling and inversion, other migration/imaging approaches besides Pratt's frequency-domain inversion may be attempted. It may prove worthwhile to directly migrate residuals within a conventional depth imaging algorithm, for example.

In terms of acquisition and practical concerns, there are at least two main potential areas of focus for future work. First, one could investigate the effect of shallower layers in shadowing the perturbations at higher frequencies. This may allow for the optimization of acquisition geometries, including testing cross-well surveys, to allow broad-band signals to propagate effectively through the target region more easily. Second, there is potential for an investigation into more practical scenarios for the inversion, such as beginning the inversion with a more realistic seismic-derived background velocity model rather than the "perfect" background velocity model used in this investigation, including random and coherent noise, unknown source waveforms, and unknown near-surface layers.

Chapter 6

Measurement of planewave migration convergence

6.1 Introduction

Prestack depth migration is costly for complex regions with strong lateral velocity variations. In these regions, it is desirable to use a wave-equation migration algorithm such as FOCI (Margrave et al., 2006). Plane-wave migration was in part developed to preserve the fidelity benefits of prestack wave equation techniques while adding the benefits of poststack processing economy (see e.g. Rietveld et al., 1992; Whitmore, 1995; Mosher and Foster, 1998; Duquet et al., 2001; Liu et al., 2006). Physically, the method may be seen as an application of Huygens' principle. A plane wave is synthesized by the superposition of numerous point sources. In terms of seismic imaging, this may be accomplished by stacking common shot gathers that are time-delayed by a linear function of the shot location. This stack is imaged using a similarly constructed plane wave source model. A zero time-delay corresponds to a horizontal plane wave (i.e. with 0° orientation). Positive and negative time-delays correspond to plane waves with positive and negative orientation.

In contrast to usual shot-profile migration, plane-wave migration has the benefit that, in many cases, a useful image can be developed from relatively few individual plane waves. In the case of flat homogeneous layers and a seismic survey with numerous shots and receivers, it is conceivable that only the 0° plane wave could be required to generate a usable image. This is roughly equivalent in cost to a poststack migration. However, plane-wave migration has the added benefit that more plane

waves with varying orientation may be added at any time to selectively improve the image. This allows fine control of the overall cost of imaging, and allows individuals to choose precisely where they would like to spend their time in imaging.

The algorithm we have implemented is an extension of the CREWES FOCI code. It retains all features of FOCI including operator stabilization and spatial resampling, but adds the ability to stack shot records into plane-wave gathers and use the requisite plane-wave source model.

6.2 Theory

The theory of plane-wave migration is explained by several authors. Here we select several important concepts as described by Liu et al. (2006). A similar treatment may also be found in Romero et al. (2000).

Consider a source wavefield of a shot $S_j(\omega, x, z)$, where ω is temporal frequency, x is the lateral spatial coordinate, z is spatial coordinate below the surface, and index $j = 1, 2, \dots, N$ where N is the total number of shots. A composite wavefield $\bar{S}(\omega, x, z)$ is expressed as

$$\bar{S}(\omega, x, z) = \sum_{j=1}^N a_j(\omega) S_j(\omega, x, z) \quad (6.1)$$

where the $a_j(\omega)$ are N functions that serve to time-delay shots as required via time-delay/phase-shift equivalency. Similarly, we may consider a composite receiver wavefield $\bar{R}(\omega, x, z)$,

$$\bar{R}(\omega, x, z) = \sum_{j=1}^N a_j(\omega) R_j(\omega, x, z). \quad (6.2)$$

$R_j(\omega, x, z)$ is the backward-extrapolated receiver wavefield that corresponds to $S_j(\omega, x, z)$.

Compose a 2D plane-wave section simulating a line-source wavefield with ray

parameter p ,

$$a_j(\omega) = f(\omega)e^{i\omega p(x_j-x_0)} \quad (6.3)$$

where $f(\omega)$ is a real-valued function describing the amplitude of the plane-wave components, and x_0 is the plane wave origin at the surface.

For wavefield extrapolation operators, Liu et al. (2006) define \mathcal{L} and its conjugate operator \mathcal{L}^* such that

$$S(\omega, x, z) = \mathcal{L}^*[S(\omega, x, z - \Delta z)] \quad (6.4)$$

$$R(\omega, x, z) = \mathcal{L}[R(\omega, x, z - \Delta z)] \quad (6.5)$$

Application of \mathcal{L} to \bar{S} and \bar{R} gives

$$\bar{S}(\omega, x, z) = \mathcal{L}^*[\bar{S}(\omega, x, z - \Delta z)], \quad (6.6)$$

$$\bar{R}(\omega, x, z) = \mathcal{L}[\bar{R}(\omega, x, z - \Delta z)]. \quad (6.7)$$

Use of a cross-correlation imaging condition (Claerbout, 1985) yields an image $I(x, z)$,

$$I(x, z) = \sum_{\omega} \bar{S}^*(\omega, x, z) \bar{R}(\omega, x, z) \quad (6.8)$$

$$\begin{aligned} &= \sum_{j=1}^N \sum_{\omega} |a_j(\omega)|^2 S_j^*(\omega, x, z) R_j(\omega, x, z) \\ &\quad + \sum_{j \neq k}^N \sum_{\omega} a_j^*(\omega) a_k(\omega) S_j^*(\omega, x, z) R_k(\omega, x, z) \end{aligned} \quad (6.9)$$

Liu et al. (2006) describe each term in equation 6.9. The first is the stack of images for each individual shot, which is the expected output from a shot-profile migration. The second term is described as the results of the crosscorrelation of source wavefields with the receiver wavefields from different shots – “cross terms”. This results in an imaging artifact which is traditionally reduced with phase encoding techniques (e.g. Romero et al., 2000).

6.2.1 2D source plane-wave migration

If equation 6.3 is substituted into equation 6.9, Liu et al. (2006) show that the image generated by a single plane-wave section $I_p(x, z)$ is

$$I_p(x, z) = \sum_{\omega} f^2(\omega) \sum_{j=1}^N \sum_{k=1}^N e^{i\omega p(x_j - x_k)} S_k^*(\omega, x, z) R_j(\omega, x, z), \quad (6.10)$$

that, upon stacking all source plane waves, the final image $I(x, z)$ is

$$\begin{aligned} I(x, z) &= \sum_{p=-N_p}^{N_p} I_p(x, z) \\ &= \sum_{j=1}^N \sum_{k=1}^N \sum_{\omega} f^2(\omega) S_k^*(\omega, x, z) R_j(\omega, x, z) \sum_{l=-N_p}^{N_p} e^{i\omega l \Delta p(x_j - x_k)}, \end{aligned} \quad (6.11)$$

and that the final sum in equation 6.11 approximates a delta function,

$$\lim_{N_p \rightarrow \infty} \sum_{l=-N_p}^{N_p} e^{i\omega l \Delta p(x_j - x_k)} = |\omega|^{-1} \delta(x_j - x_k). \quad (6.12)$$

This important result demonstrates that, given enough plane waves, the cross terms are suppressed. Also, plane-wave migration is valid even in cases of irregular and sparsely-sampled data sets. Liu et al. (2006) also make the point that the computational savings in plane-wave migration are likely to be found in large data sets. The number of plane waves required to suppress these artifacts is roughly a constant for a given physical volume to image, independent of the actual number of shot records in that volume.

6.2.2 Residuals

The question remains, how many plane waves are enough? We propose a simple measure of convergence. This residual is defined in terms of two successive plane-wave images, $I_N(x, z)$ and $I_{N+1}(x, z)$. N may refer to any ordering of plane-wave

images. For example, $I_N(x, z)$ refers to the image generated by the stacking of 11 distinct plane wave images, while $I_{N+1}(x, z)$ refers to the image generated by the stacking of 13 distinct plane wave images. In this implementation for the sequence, plane waves are added symmetrically, two at a time, of mirror-symmetric orientation. Therefore, the 13 plane-wave image is the successor to the 11 plane-wave image. This avoids any biasing of the image quality improvement due to a true image that contains significant energy of a dipping-left or dipping-right orientation. As a more extended example, the first image, I_1 , will contain the plane-wave image generated with only a single horizontal plane-wave $- 0^\circ$. I_2 may correspond to the image generated with I_1 plus the addition of $\pm 11^\circ$ plane-wave images. I_3 may correspond to the image generated with I_2 plus the addition of $\pm 33^\circ$ plane-wave images, and so on.

The two successive images are spatially localized by a window $\Omega(x, z)$, with a value of one inside the region of interest, zero when well outside, with a smooth transition between inside and outside. This windowing allows the algorithm to focus on a specific portion of the image, perhaps shallow or deep, perhaps in complex or simple structure.

Specifically, the residual $\mathcal{R}(x, z)$ is calculated as

$$\mathcal{R}(x, z) = \frac{\sqrt{\sum_{x,z} \Omega(x, z) (I_{N+1}(x, z) - I_N(x, z))^2}}{\sqrt{\sum_{x,z} \Omega(x, z) (I_N(x, z))^2}} \quad (6.13)$$

6.3 Testing

6.3.1 Simple synthetic

We tested the algorithm on the velocity model shown in Figure 6.1. This model was chosen to provide a continuous range of dip in order to highlight the various

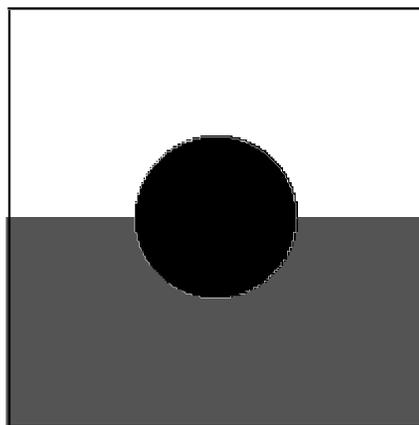


Figure 6.1: Velocity model. The white region represents a relative velocity of 4, grey represents 6, and black represents 3.

plane-wave incident angles, and the effectiveness of varying numbers of plane-wave images.

The simulated seismic model data consisted of 51 equally spaced shots spanning the surface of the model recorded into 200 equally spaced receivers which also spanned the surface. The data was generated with the CREWES `afd_shotrec` finite difference modelling facility.

Figure 6.2 shows an image of this velocity model calculated with a standard FOCI shot-profile migration. The horizontal contact is easily visible. It appears that dips up to approximately 30° of the circle are visible as well. The shot-profile image here does not show the steep sides of the circle due to a lack of adequate direct scattering of energy back to the surface.

Figure 6.3 shows the plane wave image generated using only the horizontal plane wave. This is roughly equivalent in computation time to a poststack migration. A significant portion of the image is recognizable, though it is hardly equivalent to the full shot-profile image. Figure 6.4 shows the addition of plane waves at $\pm 31^\circ$. With

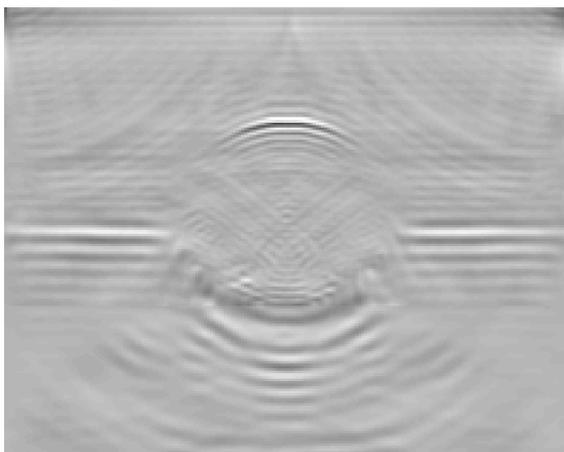


Figure 6.2: An image of the velocity model calculated with shot-profile migration.

just three plane waves, we have clearly revealed the gross structure of the model including the same $\pm 30^\circ$ dip limits on the top of the circle. As seen in Figure 6.4, the addition of plane waves oriented at $\pm 11^\circ$ clarifies the image and removes noise, especially for the horizontal contact. With two more plane waves at $\pm 22^\circ$, the image in Figure 6.5 clarifies even more, though no significant structure is revealed.

Finally, in Figure 6.5, 51 plane waves ranging between $+31^\circ$ and -31° are used to generate an image of comparable computational cost to the shot-profile migration, which required the migration of 51 shot records.

6.3.2 Marmousi testing

Testing of this method on the Marmousi velocity model (Figure 6.6) showed an interpretable image that emerged from very few plane waves. Residuals were calculated following equation 6.13 for the region at approximately (6500, 2400), and are shown in Figure 6.7. First, we notice that the plane-wave migration seems to converge to its final form much faster than the shot-profile migration. Second, we notice that

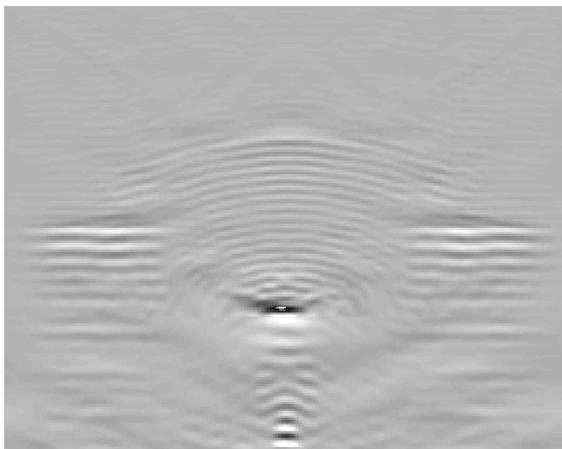


Figure 6.3: The horizontal plane wave image.

the plane-wave image stops improving dramatically with the use of 41 plane waves, and essentially stops improving at 81 plane waves. At this point, the plane-wave migration was halted as no improvement was evident. The shot-profile migration continues to benefit from additional shots throughout the entire 240 shots, though the major improvement slows at around 110 shots. Figure 6.8 shows shot-profile migration with 110 shots compared to its final state at 240 shots. Figure 6.9 compares the result from 41 plane waves to the result from 41 shots, and Figure 6.10 compares the result from 81 plane waves to the result from 81 shots.

However, the plane-wave migration images seem to suffer from a lack of high-frequency resolution in the fine details of the images, and very little improvement in the image quality is observed with more than 81 plane waves.

Converging, but to what?

Figure 6.7 clearly shows the plane-wave imaging converging to its final result faster than shot-profile migration. The obvious question is then, is plane-wave migration converging to the “correct” final image? To test this, we first accept that the shot-

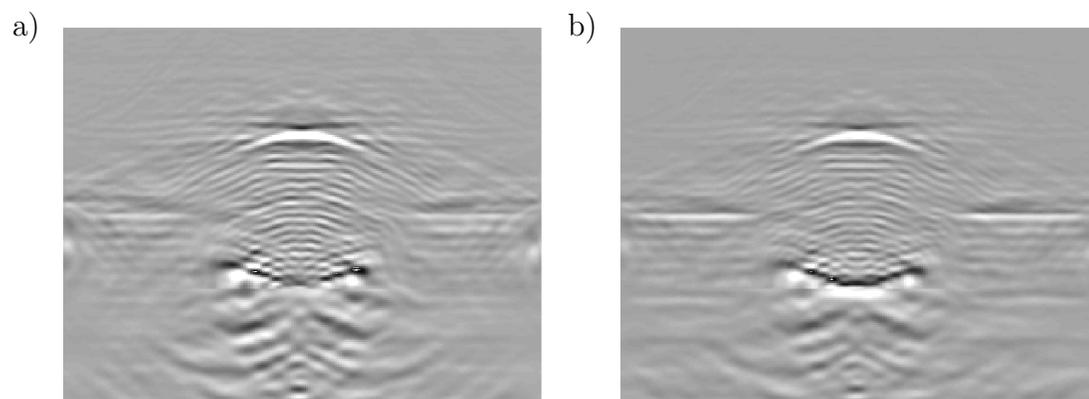


Figure 6.4: Plane-wave images generated with a) plane waves oriented at $0^\circ, \pm 31^\circ$, b) $0^\circ, \pm 11^\circ, \pm 31^\circ$.

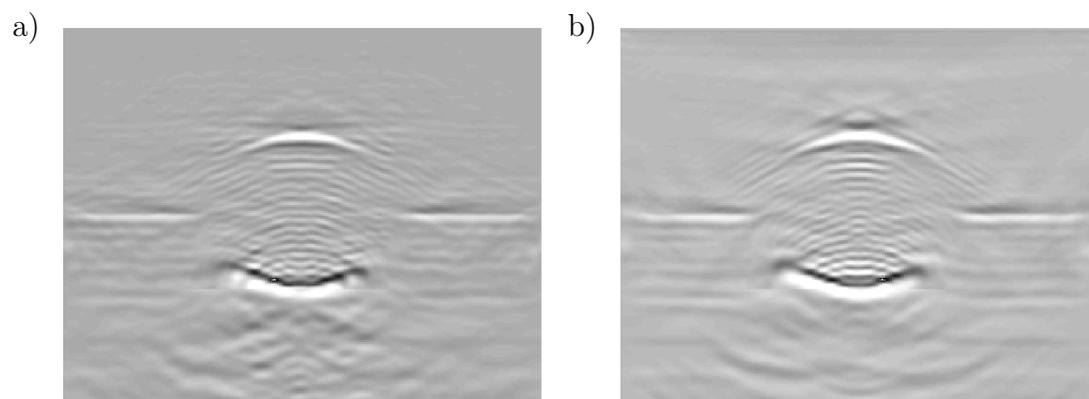


Figure 6.5: Plane-wave images generated with a) plane waves at $0^\circ, \pm 11^\circ, \pm 22^\circ, \pm 31^\circ$, b) 51 plane waves between -31° and $+31^\circ$.

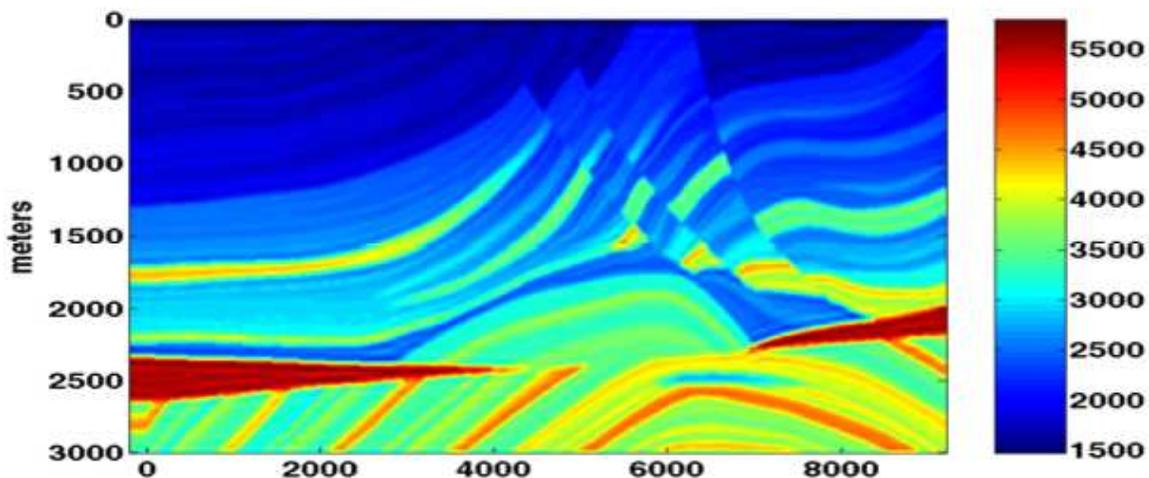


Figure 6.6: The Marmousi velocity model. Velocity ranges from 1500 m/s to 6000 m/s .

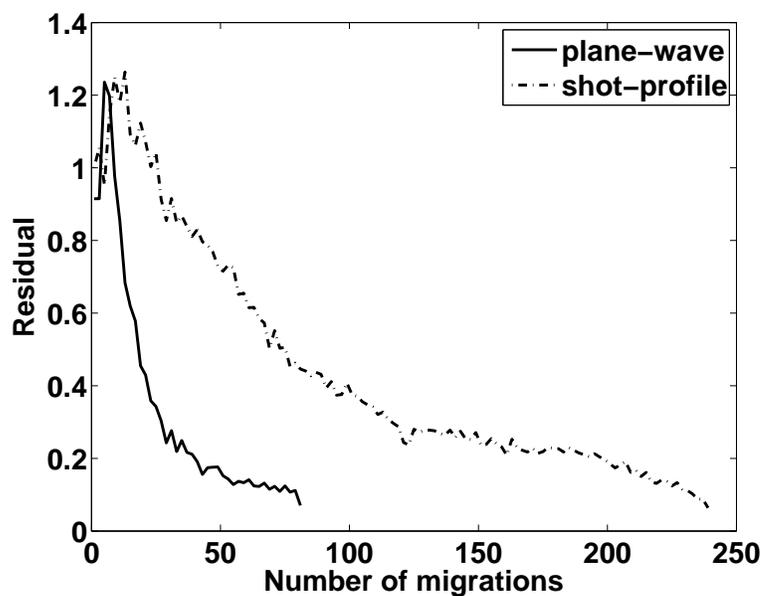


Figure 6.7: Residuals for shot-profile and plane-wave migration as a function of number of individual migrations (i.e. number of plane waves or number of shot records migrated). The plane-wave residuals show dramatically faster reduction in residual, compared to the slower decline in shot-profile migration.

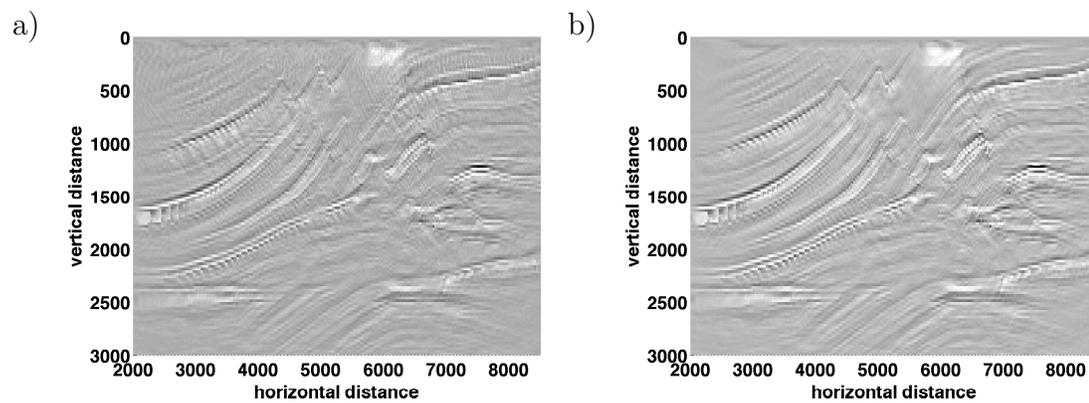


Figure 6.8: Marmousi shot-profile migration with a) 110 shot records and b) 240 shot records.

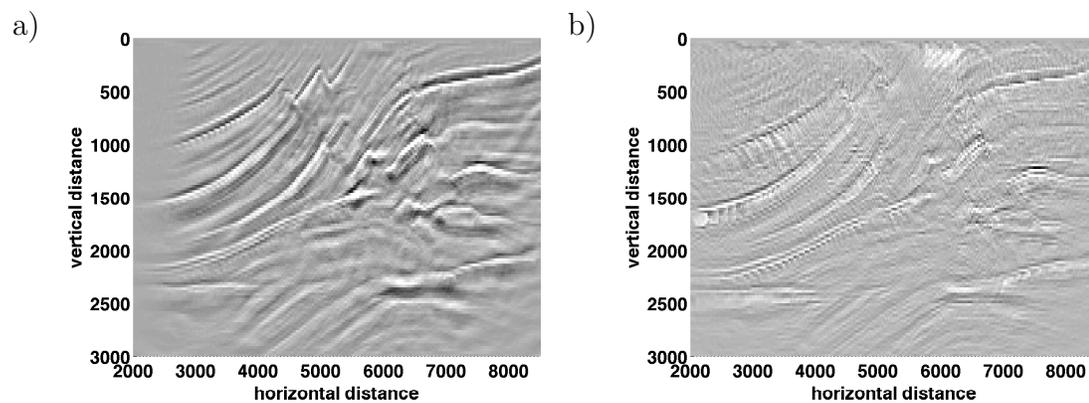


Figure 6.9: Marmousi migration with a) 41 plane waves and b) 41 shot records.

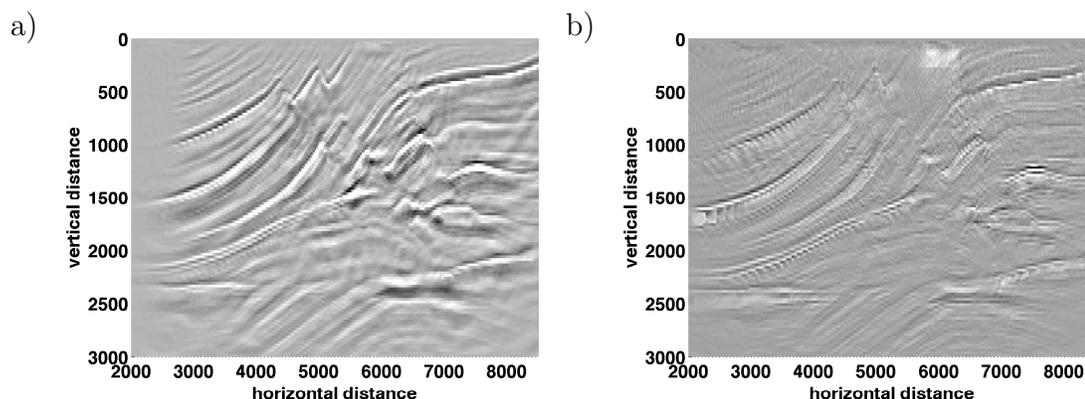


Figure 6.10: Marmousi migration with a) 81 plane waves and b) 81 shot records.

profile migration calculated with 240 shots (Figure 6.8) is clearly superior to the final plane-wave image calculated with 81 plane waves (Figure 6.10). We then calculate the residual between an intermediate calculation and this 240-shot image. These residuals are shown in Figure 6.11. Although the final plane-wave image does not have the fine detail of the final shot-profile image, it is clear that the plane-wave image approaches the final shot-profile image much faster than the shot-profile image itself converges. That is, the 81 plane-wave image has a significantly lower residual than the 81 shot image. In fact, the full numerical results reveal that the shot-profile image requires 175 shots to reach the same level of residual as the 81 plane-wave image.

6.4 Discussion

The plane-wave migration resolves the image efficiently, requiring only a few plane waves to adequately resolve the structure. The single horizontal plane-wave image suggested the placement of the horizontal contact, and hinted at the location of the top of the circle feature. With the addition of two plane waves at approximately

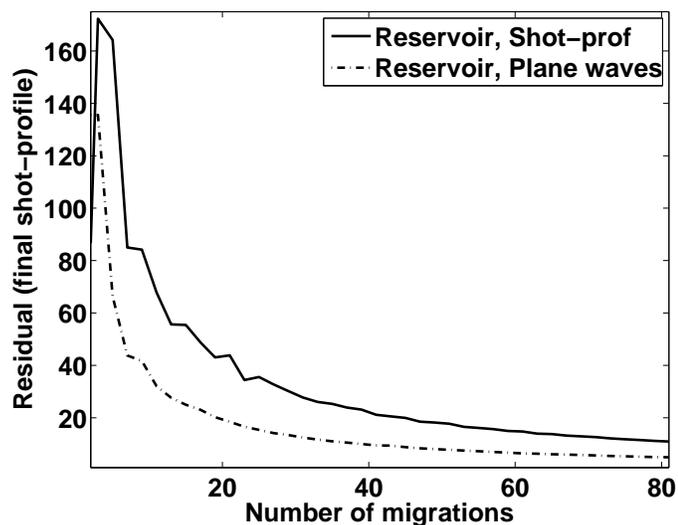


Figure 6.11: Residual calculated between an intermediate image and the final “best” image calculated with 240 shots.

+30° and −30°, however, the horizontal contact and a great deal of the shape of the circle were clearly revealed. Optimal results for this particular example appear to manifest at 7 plane waves. The qualitative difference between using 7 plane waves and 51 plane waves is small, so with plane-wave migration a comparable image can be calculated in this case in something like 15% of the original shot-profile migration calculation time.

The final process using 51 plane waves did not result in an image of the same quality as the shot-profile migration. This suggests that the algorithm as implemented is somehow suboptimal. One possible shortcoming is the method for choice of plane waves to use in migration. This issue was explored for real data by Stork and Kapoor (2004), who specifically wondered about how many plane waves may be required for reliable imaging. In this experiment, plane waves were chosen in an ad hoc fashion, simply based on a constant increment of the time-delays used to

create the effective plane waves. Perhaps a more careful algorithm for the selection of plane wave distribution could be developed in order to optimize the imaging.

The Marmousi migration yielded similar results. The plane-wave image very quickly converged to a final image that was not as clear as the final shot-profile migration. However, the plane-wave method revealed a useful image with significantly fewer migrations than the shot-profile image. It is expected that the quality of the final plane-wave image is strongly dependent on the number of shot and receiver locations. It is in situations of very fine spacing that plane-wave migration is expected to yield its best value in the sense that the number of plane-waves required to produce an image stays nearly constant, but the quality improves with finer sampling. This contrasts with shot-profile migration, in which many more shots would directly require a similar increase in the number of migrations.

6.5 Chapter summary

Plane-wave migration is useful for efficient prestack depth wave-equation migration. By generating effective plane waves and using a plane wave source model, a standard wave-equation algorithm can be easily adapted to plane-wave migration. Though it is relatively straight-forward to implement such an algorithm, other details, such as the number of plane waves to use and specific orientation of these plane waves, remain unsolved problems in many cases.

The residual measure introduced here may be useful in many situations. In addition to giving a numerical measure of when performing more calculation (in the form of adding more plane waves to the image) leads to diminishing returns, there are other possible uses. This residual may be used to guide the selection of the actual

plane waves used in migration. In the case of asymmetric geology, for example, it may be desirable to use an angular range of, say, -30° to $+5^\circ$. The residual may also be used to guide focused illumination of a particular geologic feature. For example, the residual may be used to determine which plane waves enhance the imaging of salt flanks, subsalt features, or other poorly-illuminated regions.

Chapter 7

Locally WKBJ Operators

7.1 Introduction

Chapter 4 considered the case where the locally-homogeneous GPSPI operator was extended to include a horizontal velocity gradient. This chapter will describe an operator that treats the case of a vertical velocity gradient, $v(z) = v_0 + mz$. The approach will be similar to that of Margrave (2001) to extend the GPSPI operator to a local vertical gradient.

7.1.1 The design of the lWKBJ operator

The composition of many GPSPI operators

Consider decomposing the operator $\mathbf{T}_{\alpha(0:\Delta z)}$ from equation 4.1 into a cascade of operators taking the wavefield from a depth $z = 0$ to $z = \Delta z$ in N discrete steps. Suppressing ω , $k(x)$, and ξ dependence in α we have an operator

$$\mathbf{T}_{\alpha(0:\Delta z)} = \left(\mathbf{T}_{\alpha((N-1)\frac{\Delta z}{N}:\Delta z)} \circ \cdots \circ \left(\mathbf{T}_{\alpha(\frac{\Delta z}{N}:2\frac{\Delta z}{N})} \circ \mathbf{T}_{\alpha(0:\frac{\Delta z}{N})} \right) \right), \quad (7.1)$$

which takes the wavefield from depth $z = 0$ to $z = \Delta z$ in N steps as required, where $\left(\mathbf{T}_{\alpha(\frac{\Delta z}{N}:2\frac{\Delta z}{N})} \circ \mathbf{T}_{\alpha(0:\frac{\Delta z}{N})} \right)$ represents the composition of $\mathbf{T}_{\alpha(\frac{\Delta z}{N}:2\frac{\Delta z}{N})}$ with $\mathbf{T}_{\alpha(0:\frac{\Delta z}{N})}$. If we treat the symbols of these operators as elliptic PSDO symbols (Saint Raymond, 1991), then the standard asymptotic expansion may be calculated,

$$\alpha(0 : \Delta z) \sim \alpha \left(0 : \frac{\Delta z}{N} \right) \alpha \left(\frac{\Delta z}{N} : 2\frac{\Delta z}{N} \right) \cdots \alpha \left((N-1)\frac{\Delta z}{N} : \Delta z \right), \quad (7.2)$$

truncating all but the first term in the expansion.

Since each symbol in equation 7.2 is an exponential like equation 1.46, the N -fold product becomes an exponential with an N -fold sum in the exponent,

$$\alpha(k(x_0), \xi, \omega, 0 : \Delta z) = \begin{cases} \exp\left(i \frac{\Delta z}{N} \sum_{j=1}^N \sqrt{\frac{\omega^2}{v_j^2 - \xi^2}}\right), & |\xi| \leq \frac{\omega}{v_j} \\ \exp\left(-\frac{\Delta z}{N} \sum_{j=1}^N \left| \sqrt{\frac{\omega^2}{v_j^2 - \xi^2}} \right|\right), & |\xi| > \frac{\omega}{v_j} \end{cases}. \quad (7.3)$$

Where $v_j = v(x_0, j\Delta z/N)$, and the symbol is valid at a given output point $(x_0, \Delta z)$. In the limit as $N \rightarrow \infty$ the summation becomes an integral reminiscent of a WKBJ approximation (see e.g. Aki and Richards, 2002), leading us to the “locally WKBJ operator” (hereafter lWKBJ). Practically, this limit is not taken, but rather a finite number of terms is used. In this work, we have used $N = 10$, as greater values show no discernable improvement. As the correction is for kinematic reasons, a zeroth order WKBJ-type correction that addresses only the phase of the symbol is sufficient.

Although real-earth velocity models may contain vertical velocity gradients, this is not the reason for this approach. The advantage is that a velocity with positive gradient in z has the effect of limiting the aperture of the migration operator on the line $(x, 0) \ x \in \mathbb{R}$. Consider a point source at depth in a medium with a positive vertical velocity gradient. The upward-traveling raypaths from the point source will all intersect the surface within a finite lateral aperture from the source. Figure 7.1 demonstrates this effect. The effect of the lWKBJ process may be seen in amplitude (Figure 7.2) and phase (Figure 7.3), as it effectively smoothes out the sharp corner found at the evanescent boundary of the locally homogeneous symbol amplitude. The stability of this operator was motivated from an intuitive raypath argument suggesting the limitation of the support of the operator on the input data, however this symbol amplitude reveals the mathematical source of the new-found stability: the sharp corner in the original locally-homogeneous symbol amplitude causes trouble via Gibbs phenomenon upon truncation (Morse and Feshbach, 2005, §6.3). If

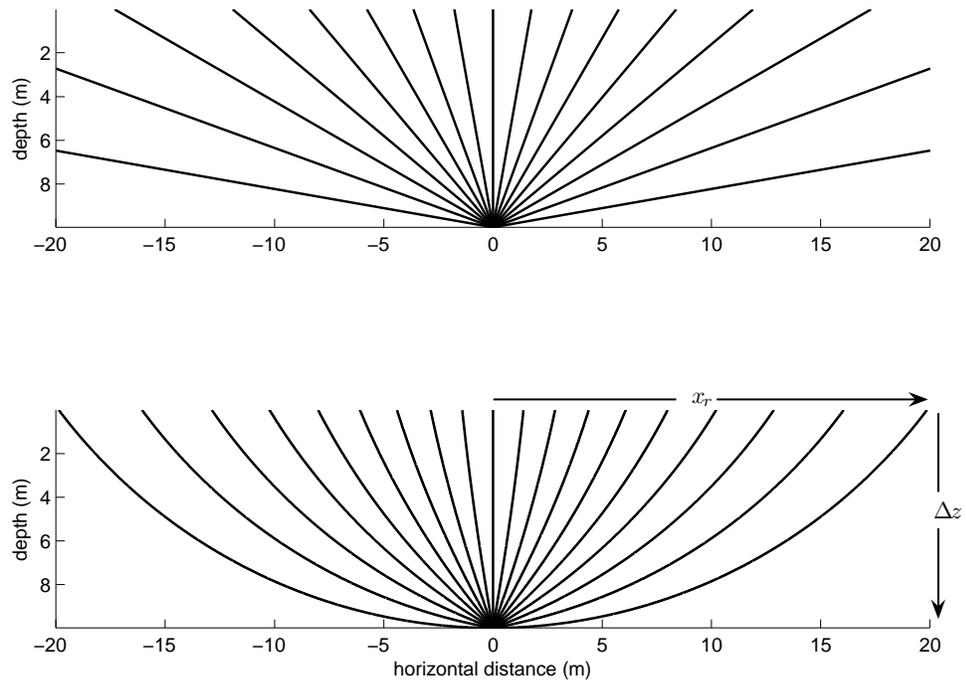


Figure 7.1: In the top panel, raytracing through a homogeneous medium is shown. Note that rays leaving horizontally will travel to infinity. In the bottom panel, rays are traced through a LWKB medium. Aperture radius was limited to 20 meters, and takeoff angles from -90° to 90° are displayed in ten meter increments

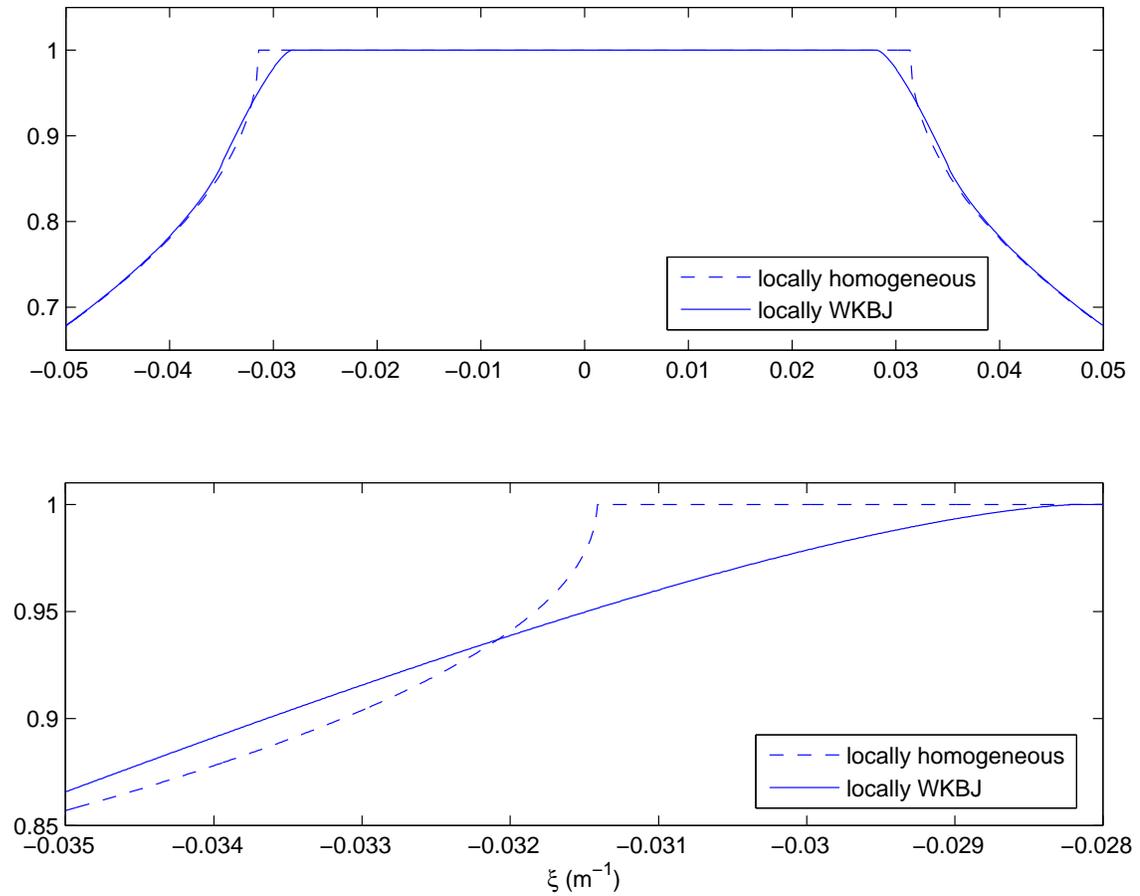


Figure 7.2: Symbol amplitudes for a locally homogeneous medium and for a lwkBJ medium are compared. $x_r = 30m$, $v_{loc} = 2000m/s$, $v_0 = 1785m/s$, $m = 45s^{-1}$, $dz = 10m$

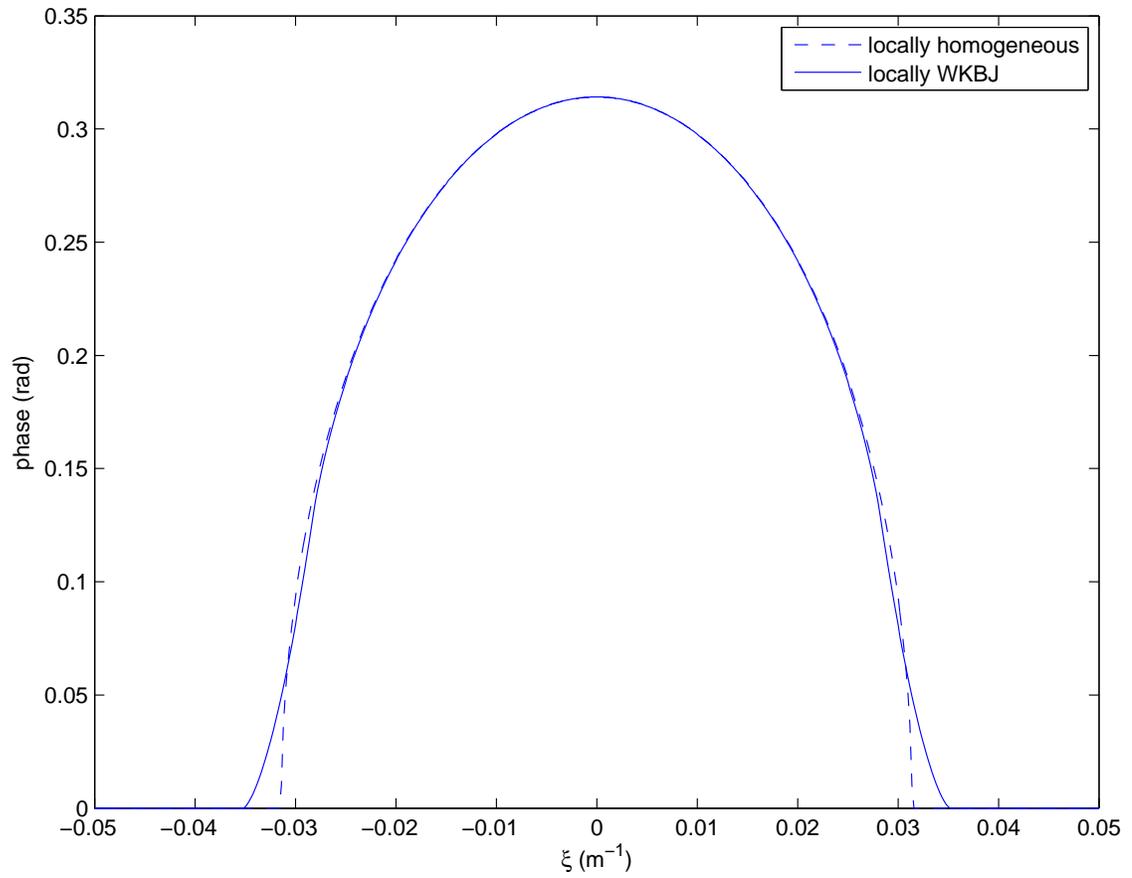


Figure 7.3: Symbol phases for a locally homogeneous medium and for a lwkBJ medium are compared. $x_r = 30m$, $v_{loc} = 2000m/s$, $v_0 = 1785m/s$, $m = 45s^{-1}$, $dz = 10m$

this sharp corner is removed, the Gibbs phenomenon is significantly reduced and so stability is enhanced. The lwkBJ method provides a fast and simple method for rounding the corner without losing operator fidelity. The series of terms as in equation 7.3 shows the origin of the rounding. Effectively there are N terms of successively narrower wavelike regions multiplying together. This blurs the evanescent boundary somewhat, giving the desired “soft” corner.

7.1.2 The $\omega - x$ operator

Given the symbol for a PSDO operator such as equation 1.46 or 7.3, we can develop an equivalent operation that takes place in the $\omega - x$ domain as

$$\Psi(x, \Delta z, \omega) = \int_{\mathbb{R}} \Psi(x', 0, \omega) W(k(x'), x - x', 0 : \Delta z) dx' \quad (7.4)$$

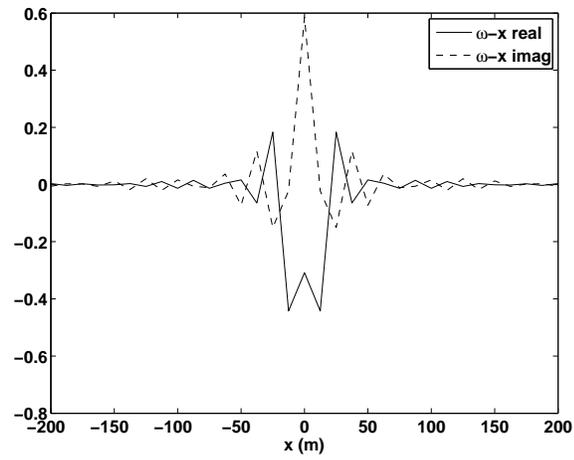
where

$$W(k(x'), x - x', 0 : \Delta z) = \int_{\mathbb{R}} \alpha(k(x'), \xi, \omega, 0 : \Delta z) e^{i\xi(x-x')} d\xi \quad (7.5)$$

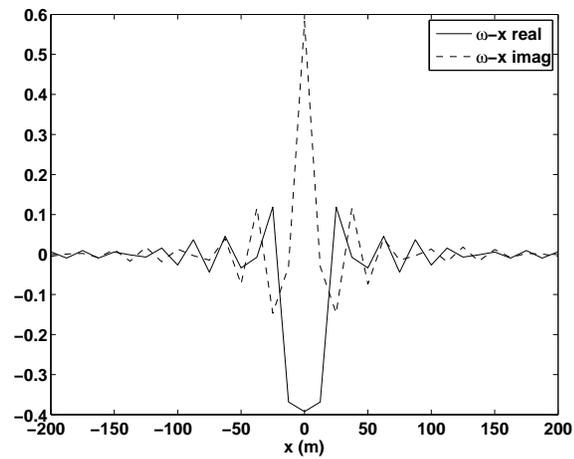
is the nonstationary convolution kernel of the lwkBJ operator in the $\omega - x$ domain and α is defined in equation 7.3.

Convolution kernels are shown in figure 7.4, with a FOCI-stabilized kernel compared to a typical lwkBJ kernel and an “overstable” kernel (i.e., its aperture radius was severely limited to exaggerate the effect). The convolution kernels have a large spatial extent, but the figures have been limited to show a width of 400m. These kernels may be transformed back to symbol form using the inverse of equation 7.5, and so we may examine the amplitude and phase spectra to understand their effects. The amplitude spectra are compared in Figure 7.5. The FOCI-stabilized kernel shows excellent stability in the wavelike region, with some instability near the evanescent region that must be corrected in the algorithm. The lwkBJ-stabilized kernels show

a)



b)



c)

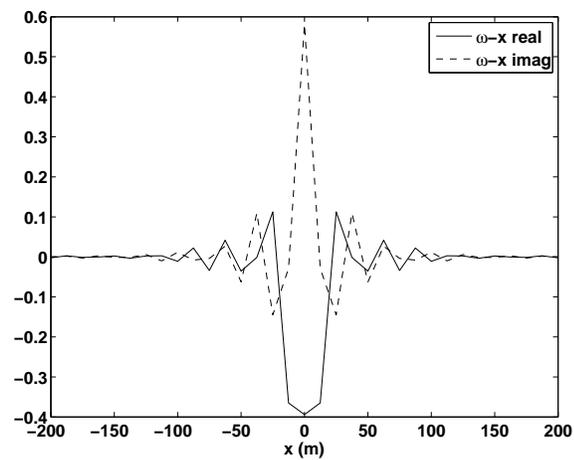


Figure 7.4: $\omega - x$ operator kernel for FOCI(a), lwKBJ with normal stabilization (b), and lwKBJ with overstabilization (c). The kernel is not compactly supported, only a 400m extent is shown here.

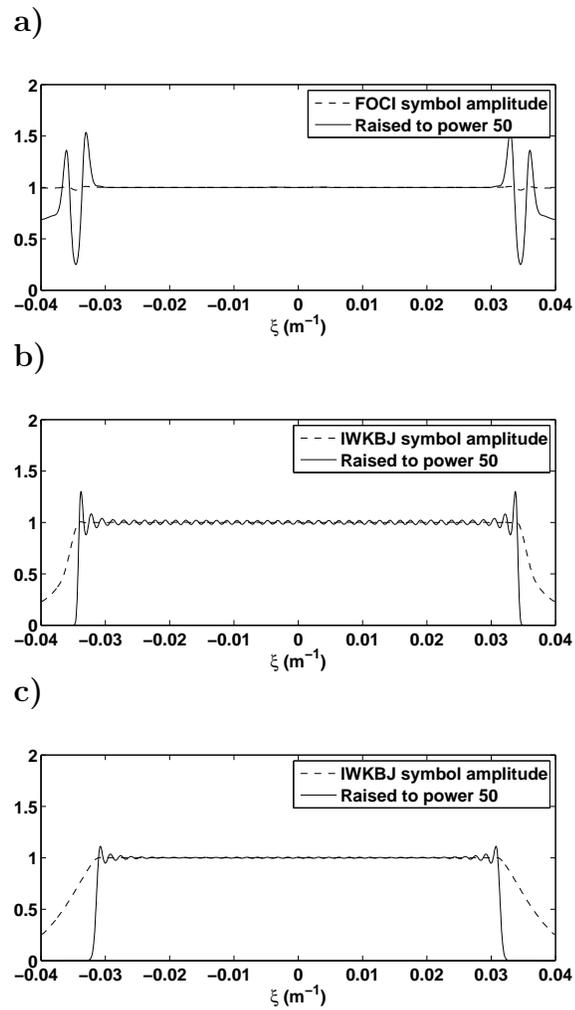


Figure 7.5: Equivalent truncated $\omega - \xi$ symbol amplitude for FOCI(a), IWKBJ with normal stabilization (b), and IWKBJ with overstabilization (c).

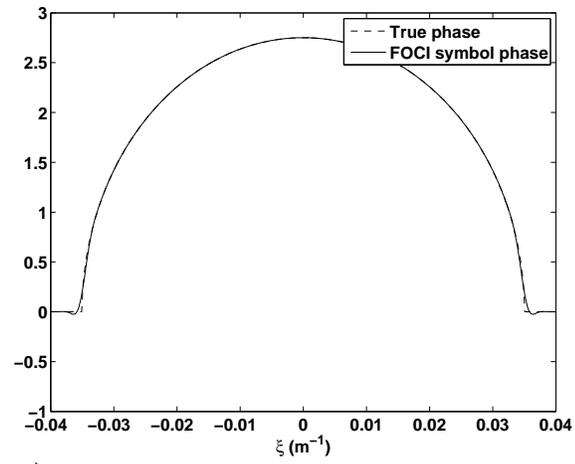
slight instabilities throughout the amplitude spectrum, but with far better control of the instabilities near the evanescent region. The exaggerated LWKBJ kernel shows the expected improved stability, though the width of the propagating wavelike region is slightly reduced. Thus the LWKBJ operator has a natural damping response near the evanescent boundary.

The respective phase spectra are compared in Figure 7.6. We expect that the phase spectra should closely match the phase spectrum of the exact symbol. The FOCI-stabilized kernel matches extremely well. The typical LWKBJ-stabilized kernel is comparable to the FOCI-stabilized kernel. The exaggerated LWKBJ-stabilized kernel matches well throughout most of the spectrum, but has significant phase error near the evanescent boundary. From Figure 7.5, however, the phase error lies exactly within the newly-damped region. This self-censoring property is due to the variation in the width of the wavelike region of the amplitude spectra as a function of depth from $z_0 \rightarrow z_0 + \Delta z$. That is, at the end of the interval $[z_0, z_0 + \Delta z]$ the $v(x, z)$ takes its maximum value which is greater than v_{loc} . This leads to a narrower evanescent region and therefore a damping of this region and this self-censoring. This corresponds to an effective loss of extrapolation of extremely high-angle components of the wavefield. Typically, these high horizontal wavenumbers do not contribute much to the final image, so it is not a significant loss.

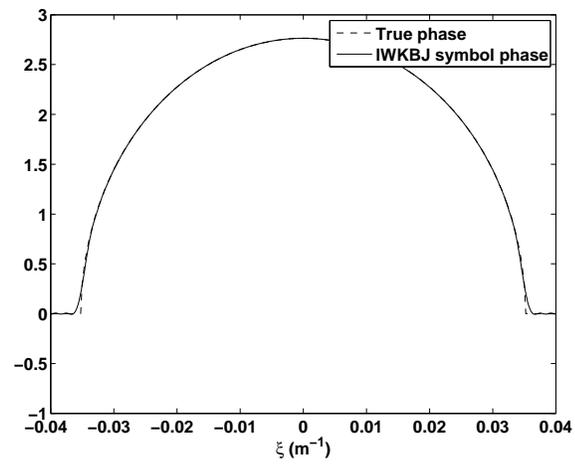
7.2 Marmousi migration

Full prestack migrations of the Marmousi model were calculated using both FOCI and our LWKBJ algorithm. The two migration codes were identical in all respects except for the different operator kernels.

a)



b)



c)

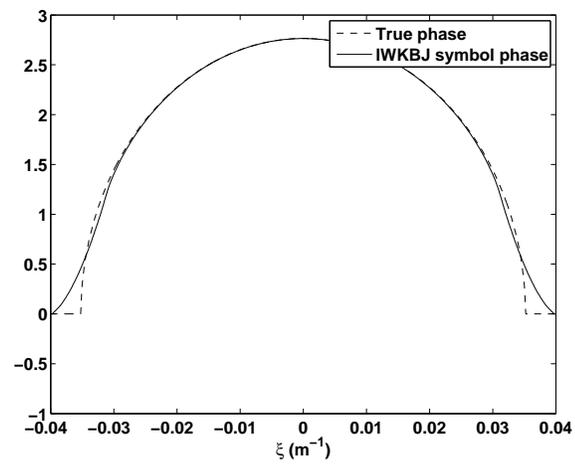


Figure 7.6: Equivalent truncated $\omega - \xi$ symbol phase for FOCI(a), lWKBJ with normal stabilization (b), and lWKBJ with overstabilization (c).

A prestack migration image of the Marmousi model was generated using both standard FOCI (Figure 7.7) and lwkBJ (Figure 7.8) operator kernels.

Several subsections of the image are compared more closely in Figures 7.9–7.12. Each figure contains three panels: the velocity model for the region, and the FOCI and lwkBJ images for that region. On the top, the Marmousi velocity model for the region is displayed, in the middle, the FOCI image is shown, and on the bottom the lwkBJ image is displayed.

In region 1 (Figure 7.9) the complex layering is clearly imaged by both algorithms. It is difficult to choose a “better” image, though in several cases (e.g., near (3600, 1300) and (4400, 1000)) the lwkBJ image seems to have better continuity of the reflectors. In region 2 (Figure 7.10) the images are slightly different. The lwkBJ image seems to display better reflector continuity, but in some cases loses some of the sharpness of the FOCI image. In region 3 (Figure 7.11) again, the lwkBJ image seems to have better continuity. Also, the fault line appears more obviously in the lwkBJ image. Finally, in region 4 (Figure 7.12), the reservoir target area is displayed. This reservoir target is well-imaged by both algorithms. The lwkBJ image appears to have less noise contamination than the FOCI image. This is especially evident in the top-right corner of each region, around (7500, 2250), though it may be observed throughout the region.

7.2.1 Parameter choice and performance

For the choice of $v(z) = v_0 + mz$, two constraints are required to uniquely determine the velocity function. Since the $v(z)$ velocity function is purely conceptual, designed to stabilize the operator and not to represent any physical gradient, for our first constraint we require that the overall travelttime through the small depth step Δz

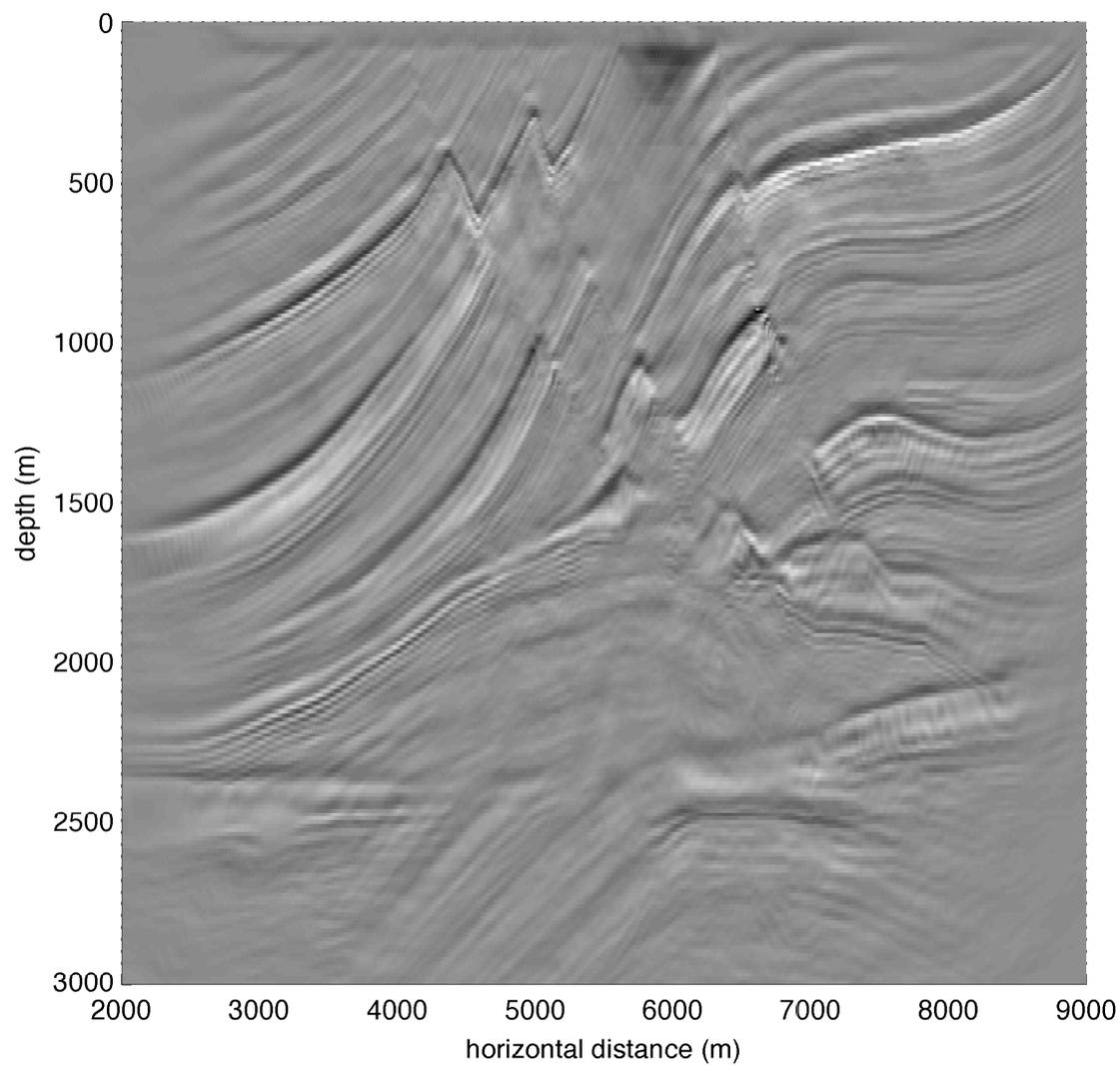


Figure 7.7: Marmousi image, FOCI migration with full stabilization. (31 point forward operator, 41 point inverse operator, 31 point final window, 187.5m effective operator length.)

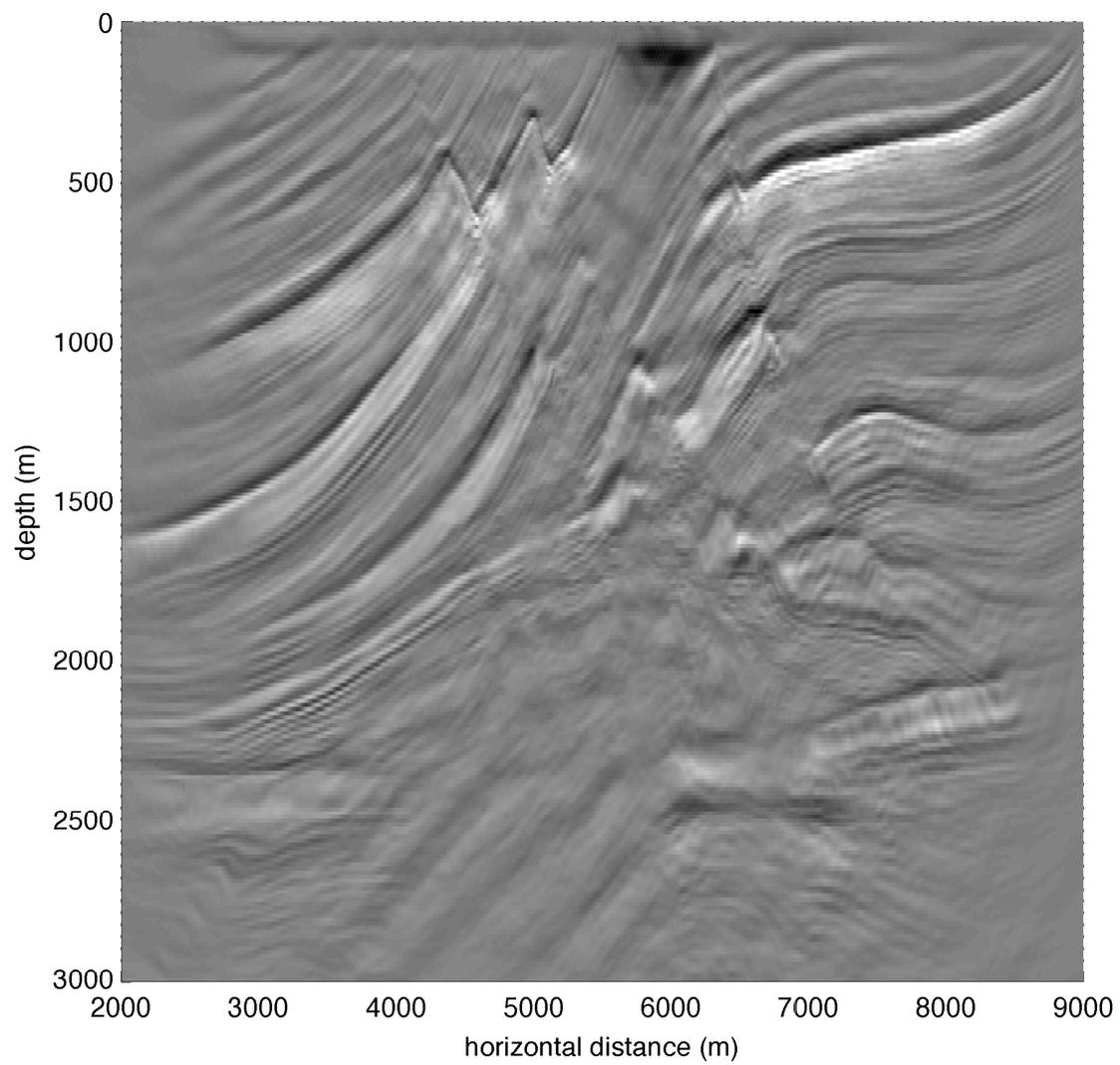


Figure 7.8: Marmousi image, lwkBJ migration. (31 point (187.5m) operator, 40m aperture radius.)

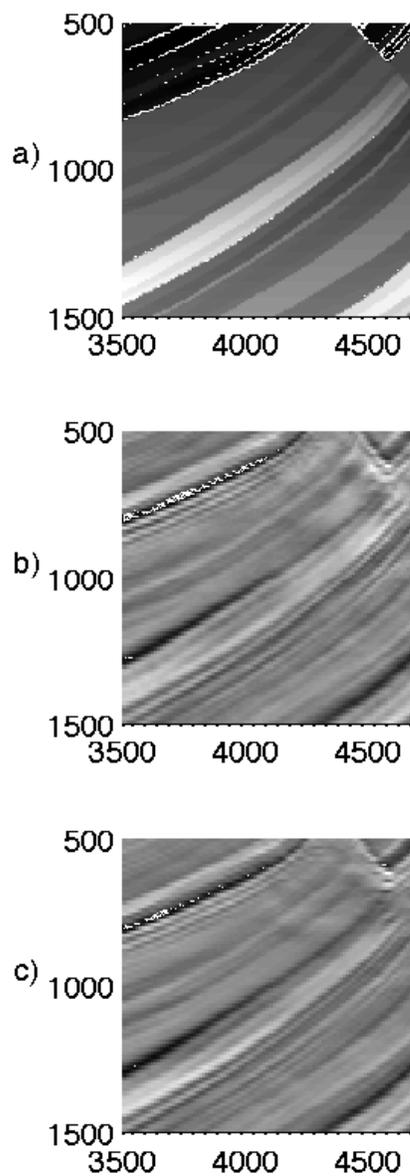


Figure 7.9: Marmousi region 1. a) Velocity model of the region b) FOCI image c) IWKBJ image. Note the improved continuity in reflectors near (3600, 1300) and (4400, 1000)

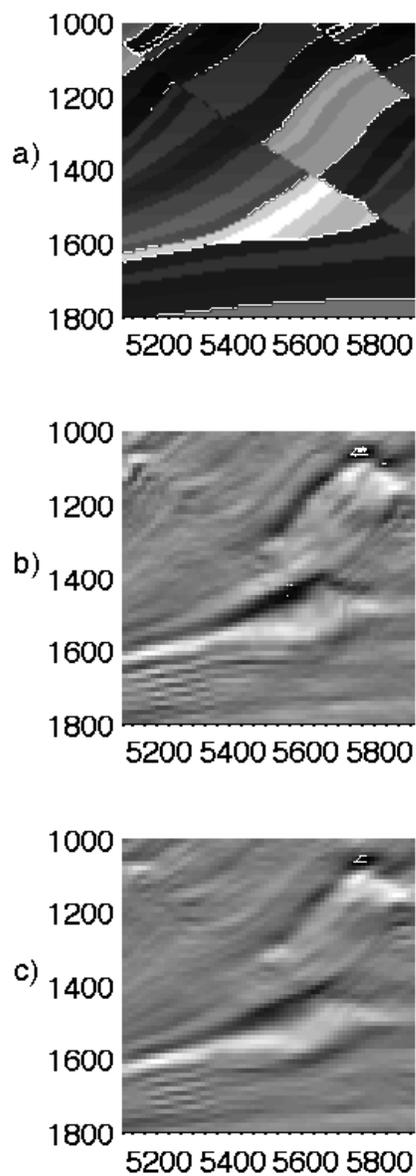


Figure 7.10: Marmousi region 2. a) Velocity model of the region b) FOCI image c) LWKBJ image. The LWKBJ may have better reflector continuity, but seems to lose some focus compared to the FOCI image.

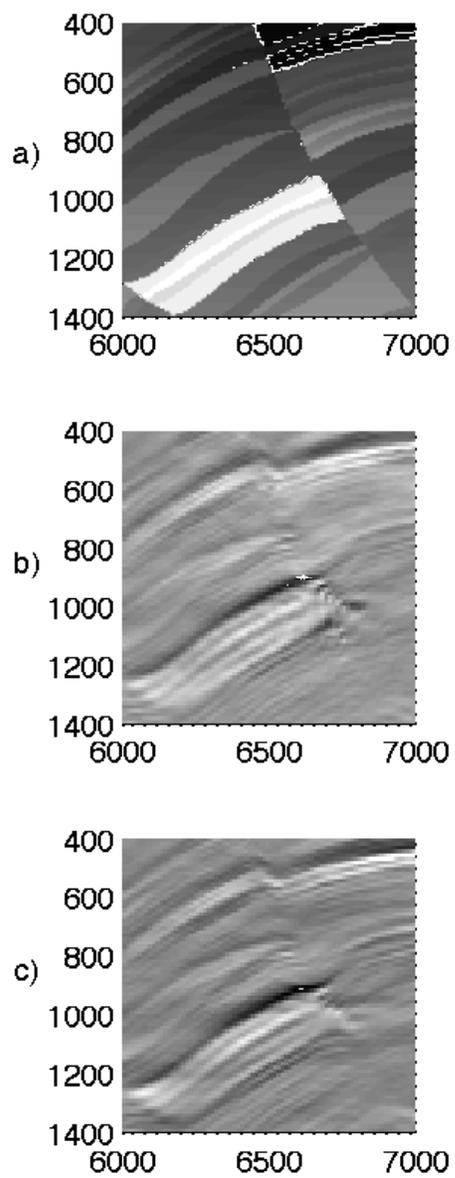


Figure 7.11: Marmousi region 3. a) Velocity model of the region b) FOCI image c) LWKBJ image. The LWKBJ image appears to have better continuity, with a more prominent fault line.

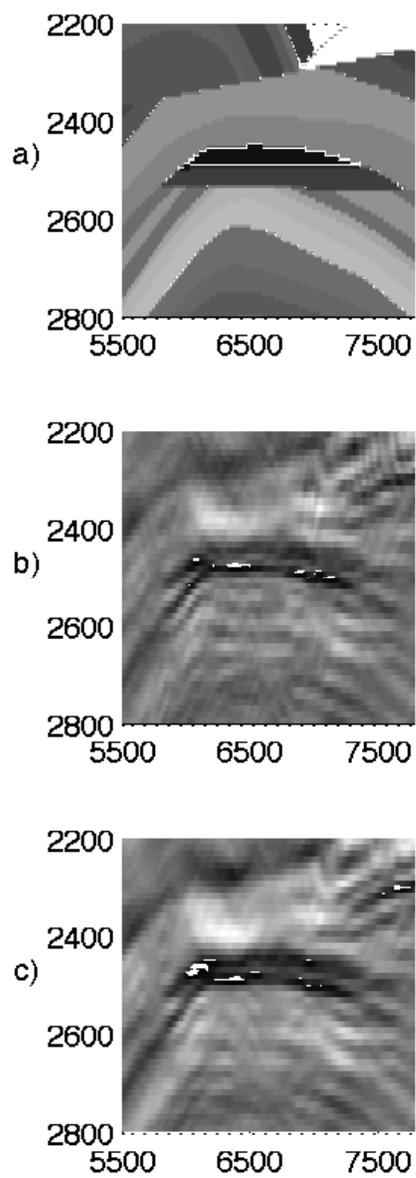


Figure 7.12: Marmousi region 4. a) Velocity model of the region b) FOCI image c) LWKBJ image. The reservoir target is well imaged by both algorithms.

through the $v(z)$ medium match the traveltimes through the same depth step at the velocity model's local value v_{ref} . For our second constraint, we choose the gradient such that a ray with a takeoff angle of 90° at the output point is normally incident at the input depth at precisely the desired operator's aperture radius.

This leads us to two equations for the two unknowns. From raypath arguments (see e.g. Aki and Richards, 2002) it can be shown that, for an aperture radius x_r , a depth step of Δz , and an initial velocity v_0 , the required gradient (accelerator) m is

$$m = \frac{2v_0\Delta z}{x_r^2 - \Delta z^2}. \quad (7.6)$$

By equating the traveltimes for a linear $v(z)$ with a constant velocity medium at v_{ref} , the second constraint is given in equation 7.7,

$$\log \left(1 + \frac{m\Delta z}{v_0} \right) = \frac{m\Delta z}{v_{ref}}. \quad (7.7)$$

From equations 7.6 and 7.7 it is obvious that the aperture x_r must not approach Δz too closely, or a singularity results.

Combining equations 7.6 and 7.7 and solving for v_0 yields equation 7.8:

$$v_0 = v_{ref} \log \left(1 + \frac{2\Delta z^2}{x_r^2 - \Delta z^2} \right) \frac{(x_r^2 - \Delta z^2)}{2\Delta z^2}. \quad (7.8)$$

So with equations 7.6 and 7.8, the required $v(z)$ is expressed purely in terms of reference velocity v_{ref} , aperture radius x_r and depth-step size Δz .

A factor in choosing the operator length parameters for the operation of the lwkBJ operator is time-to-compute vs. image quality. Simply put, longer operators usually mean a better image, but also a slower calculation. In figure 7.13, another lwkBJ Marmousi image with a 101 point (625m) operator and an 80m aperture is shown. In figure 7.14, a final lwkBJ Marmousi image with a 15 point (87.5m)

operator and a $23m$ aperture is shown. All calculations were performed using MATLAB 7.1 on Linux-based PC computers with 3.06 GHz Intel Pentium 4 CPUs. The 31 point ($187.5m$) operator requires approximately 4 minutes per shot record to migrate, resulting in a total time of about 16 hours. The 101 point operator requires approximately 4.6 minutes to calculate each shot record, with a total time of slightly more than 18 hours. The 15 point ($87.5m$) operator requires approximately 2 minutes per shot record to migrate, resulting in a total time of about 8 hours. Obviously the calculation time is not directly proportional to the operator length as there is significant overhead and other numerical complications within the code, but operator length does have an impact on the total time taken. As can be seen from a careful examination of figures 7.8 and 7.13, the 101 point operator does result in small improvements in the quality of the final image. These improvements were gained at the cost of an extra 15% calculation time. The 15 point operator does give an interpretable image compared to the 30 point operator, and saves about half the total computation time. However, the 15 point image does show incipient instability artifacts around coordinates ($5000-8000, 2500+$), demonstrating that this stabilization method does have some limitations.

Aperture radius does not directly affect calculation time, since a convolution kernel designed with a narrower aperture radius but the same operator length will simply be a padded version of the operator with the shorter aperture radius. That is, a 31 point operator with a $20m$ aperture radius will compute in the same time as a 31 point operator with a $30m$ aperture radius, the only difference will be in stability. Aperture radius can affect the quality of the image if the radius is not significantly larger than the depth step of each extrapolation. Also, a smaller aperture radius does allow a shorter operator length, which will result in a faster calculation. The

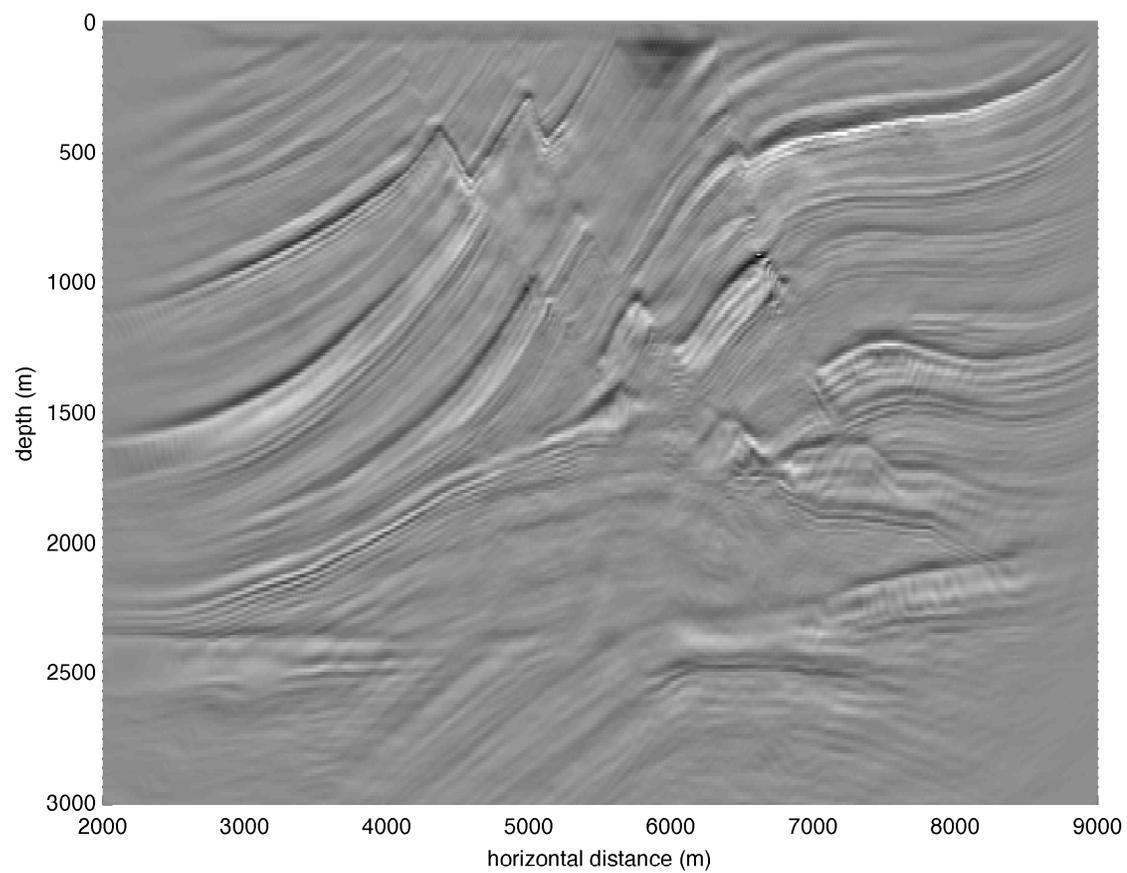


Figure 7.13: Marmousi image, lwKBJ migration. (101 point (625m) operator, 80m aperture radius.)

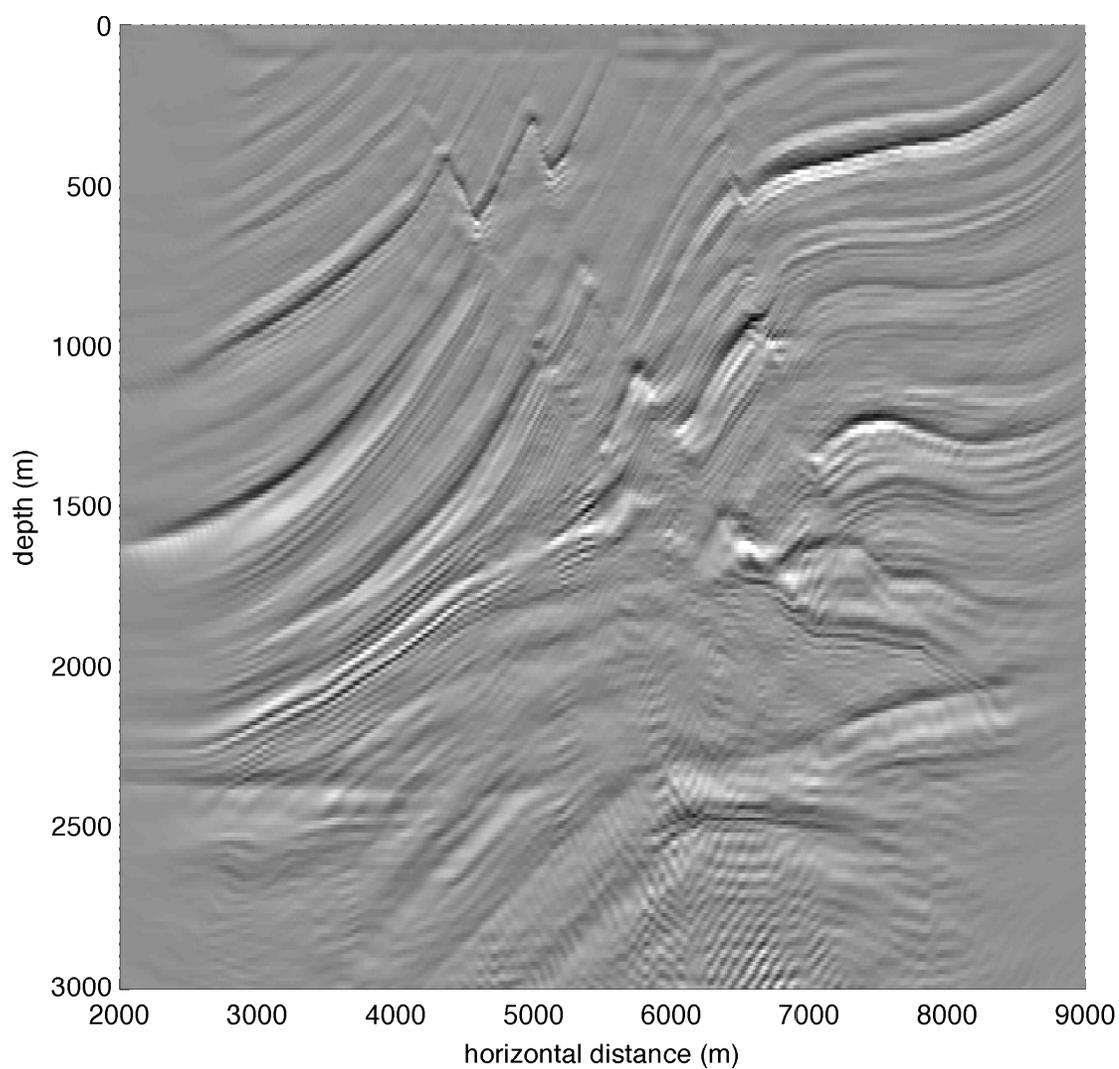


Figure 7.14: Marmousi image, lwKBJ migration. (15 point (87.5m) operator, 23m aperture radius.) Instability artifacts contaminate the image near the reservoir target, (5000 – 8000, 2500+).

stabilization of the operator is ineffective if the radius is not significantly less than the radial extent of the operator. In the Marmousi data set, traces are $12.5m$ apart. Thus, a 31 point operator has an effective radial extent of $15 \times 12.5m = 187.5m$. Therefore, the choice of a $40m$ aperture radius is smaller than the radial extent of the operator ($187.5m$), but greater than the depth step of the extrapolation ($12.5m$).

Testing on migrations of the Marmousi model suggests that a useful aperture radius measure is such that the ratio of operator radial extent OL to aperture radius x_r is somewhat greater than the ratio of aperture radius to depth-step size Δz . Also, we require that $OL > x_r > \Delta z$. Dividing the two ratios and calling the quotient ρ :

$$\left(\frac{\Delta x(OL - 1)/2}{x_r} \right) \left(\frac{\Delta z}{x_r} \right) = \rho \quad (7.9)$$

allows a numerical description of the aperture radius x_r where OL is the operator length in points and Δx is trace spacing in meters. Therefore, aperture radius x_r may be suggested as follows, given that useful values of this ratio are in the range $\rho = (1, 2.5]$. Solving equation 7.9 for aperture:

$$x_r = \sqrt{\frac{\Delta x \Delta z(OL - 1)}{2\rho}}. \quad (7.10)$$

This aperture radius balances the requirement for effective natural truncation of the operator to avoid Gibbs phenomenon instability, while maximizing the quality of the final image.

7.3 Conclusions and chapter summary

The extrapolation of Fourier plane waves by the GPSPI operator is effectively a stable operation in the $\omega - \xi$ for seismic imaging purposes, but it becomes unstable when performed in the $\omega - x$ domain due to the necessary truncation of an operator

that is infinitely long. By introducing a local positive vertical velocity gradient into the design of each operator that preserves vertical traveltime but limits the spatial extent of the operator to a chosen horizontal distance, it is possible to design an effectively stable $\omega - x$ operator which can produce high-quality migration images of complicated models like the Marmousi model.

This operator is constructed as the composition of many operators which have pseudodifferential operator symbols in the form of exponentials. Since we approximate the composition of these symbols as simply the multiplication of the symbols, the resulting symbol is also an exponential where the argument is the sum of all arguments of each composed symbol. In the limit where an infinite number of symbols are composed, this sum may be represented as an integral, leading to a WKBJ-style integrated phase approximation. Although other methods exist to stabilize explicit $\omega - x$ wavefield extrapolation, the main benefit of this approach is its simplicity: it allows for relatively easy and efficient implementation of a stable version of this operator.

Chapter 8

Conclusions

In Chapter 1, a one dimensional wave equation was derived to describe the propagation of the deflection of a string under tension. This may be the simplest wave equation imaginable, and yet the derivation required a dramatic approximation: we assumed that there was *no finite deflection of the string!* This means that any real, physical situation that may conceivably be adequately described by this equation, such as the plucking of a guitar string, is strictly incorrect. It is a theoretical wave equation, that cannot describe exactly any real physical situation.

Virtually every wave equation applied to seismic imaging has an analogous approximation: we assume that the medium does not change in any finite sense as a wave passes through it. We have *linearized* the problem, in that we assume that the passage of a wave does not appreciably impact the passage of other waves.

Of course, in reality this is strictly *not* the case, as the material must compress, expand, or shear some small but non-zero amount in order for a measurable wave-field to exist. This deformation necessarily alters the density and other physical parameters of the material in some finite way, and so the passage of the wave *must* alter the passage of waves in general, in a non-zero fashion. Presumably all serious students of the wave equation recognize this approximation, yet we are comfortable with it in varying degrees. We realize that without this approximation, solutions to the wave equations derived would be hopelessly intractable for even the simplest imaginable circumstances.

We also have a fundamental intuition for what this approximation says: the

passage of a wave *does* alter the supporting medium, but only in a very small way that essentially does not impact our investigations. We also have an intuition for when this assumption is invalid: *e.g.* when the medium is catastrophically altered by the deformation. A hammer striking the metal of an anvil may only temporarily and trivially deform the anvil. A subterranean detonation of a high-yield nuclear weapon, on the other hand, may be expected to dramatically and permanently alter the rocks surrounding it¹. The wave propagation properties of this rock will be dramatically non-linear nearby the detonation in this case. Yet we may also expect that the rock will behave approximately linearly when we are far away from this source. A linearized wave equation will be hopelessly inadequate in the region near the explosion, yet may give useful results farther away. The validity of an approximation depends entirely upon its context.

This linearity approximation is emphasized in order to underscore the fact that in seismic imaging, we are always making assumptions, and we must continuously question the validity of these assumptions. We must always keep in mind that we are always finding the wrong answers to not-quite the right questions!

In my research, I have addressed the issue of approximate wave equation solutions in several different ways. I have not directly considered the linearity of the wave equation as used in the examples in preceding paragraphs, however – this issue remains hidden underneath the details, lurking. Instead, I have addressed more subtle approximations that do not exist in the “standard” (acoustic) wave equations as derived in this dissertation, but rather exist within the approximate solutions to the “exact” acoustic wave equation made under the assumption that the domi-

¹There is always a small region in the neighbourhood of a seismic source where non-linear effects dominate.

nant wavelength of the propagating wavefield is significantly smaller than the scale of variation within the medium. These are commonly known as “high-frequency” approximations. We assume that in the limit where the wavelength becomes much smaller than the scale of variation in the medium, the approximate solutions we find converge to the “exact” solutions to the wave equation. Of course, a diminishing wavelength is equivalent to an increasing frequency.

High-frequency approximations are common in many fields of wave equation studies. In many of the traditional wave-equation problems, ranging from underwater acoustic sound propagation to visible-light optics, high-frequency approximations are frequently justified. In exploration seismology, however, this validity is worth questioning. Seismic investigations are often conducted with sources that produce negligible energy at frequencies higher than 100 Hz. Wave propagation speeds within rocks are rarely slower than, say, 1500 m/s (the nominal acoustic wavespeed of water). This produces wavelengths of 15 m as a *minimum* wavelength. The bulk of exploration seismic energy is more commonly in the 20-30 Hz range, propagating in rocks with wavespeeds perhaps 3000 m/s or more. This gives more usual wavelengths of 100 m or larger. Meanwhile, geologic layers frequently have appreciable variations in wavespeed and other relevant properties on scales of meters or even finer. Even the coarsest determination of geologic layers seldom defines layers thicker than a few tens of meters. Yet to have a “high frequency” medium for even the most optimistic case, we require geologic layering to vary on spatial scales such that 15 m is a negligible distance in comparison. This is grossly unrealistic, and it behooves us to consider the validity of the results that are derived based on our questionable assumptions.

In Chapter 2, I showed the results of my investigations into one of the most com-

mon high-frequency approximate solutions to the wave equation, the eikonal equation. Here I demonstrated that eikonal solutions in proxy velocity models that are essentially smoothed versions of the “true” velocity model can yield traveltimes that more closely match the traveltimes derived from solutions to the wave equation made without a high-frequency approximation. In a heuristic sense, low-frequency wavefields appear to propagate approximately as high-frequency wavefields in a smoothed version of the original medium. This is directly and economically useful in virtually any wave equation problem that requires efficient approximate solutions when the medium is variable on length scales that are comparable to the dominant wavelengths of the wavefield: simply solve an eikonal equation in a smoothed version of the medium, and the solution will be useful at lower frequencies.

In Chapter 3, I investigated a very different sort of high-frequency approximation, but with some shared characteristics with the results in Chapter 2. In this chapter, I worked with pseudodifferential solutions to the “exact” wave equation. These solutions, in principle, can yield exact, full-frequency validity. Unfortunately, they are computationally excruciating to calculate. To bring these algorithms into useability, we approximate the operator symbol using a high-frequency approximation. That is, the operator symbol is exact in the limit $\omega \rightarrow \infty$, and approximate for any finite ω . This is a much lower-level high-frequency approximation, that will yield much more valid results at a wider range of frequencies than a high-frequency approximation to the wavefield directly. In this case, as with the eikonal solutions, I have shown that performing the high-frequency approximate solution on a smoothed version of the full velocity model can give an operator symbol that more closely matches the full-frequency operator symbol. Again, this is directly and economically useful in that it allows a higher fidelity solution without introducing any

significant extra computational or analytical complexity.

Chapter 4 shows the results from my work to incorporate a more physically descriptive operator symbol based on mathematical theory. This theory effectively introduces a dependence on a derivative in horizontal slowness of the medium into the symbol. Since the locally-homogeneous approximation is effectively a high-frequency approximate symbol, this work extended the symbol beyond this. Although the overall economic benefits of this extended symbol were slight, I believe it was certainly worth investigating whether the standard elliptic pseudodifferential calculus could be employed directly to improve imaging with GPSPI and related algorithms. The shortest way to state the result of this investigation is, “not easily”. I believe that this result underscores the overall effectiveness of the locally homogeneous operator, though it does leave open a possibility for the improvement of GPSPI-type algorithms with the elliptic calculus, if significant technological challenges are adequately addressed.

Chapter 5 approaches the high-frequency approximation from a different angle. Rather than introducing an idea which extends the high-frequency approximation to a lower frequency regime, instead I adapt an existing full-frequency algorithm to a new application. With this work, I show that it is feasible to consider using Early Arrival Waveform Tomography for time-lapse seismic monitoring. Its development was motivated by seismic imaging and migration, but I believe that it could see significant use in the detection of a time-dependent change in an underlying velocity model. Practical acquisition constraints are considerable, but not insurmountable.

In Chapter 6, I continue in a similar direction. Instead of introducing a new algorithm to directly extend the range of validity of approximate solutions, this work on the measurement of the convergence of plane wave migrations was intended to

assist in the adoption of existing algorithms which make fewer (*i.e.* less strict) high-frequency assumptions. Although the GPSPI algorithm contains a high-frequency approximation in the design of the operator symbol, it is a much milder approximation than found in *e.g.* the eikonal equation, or other approximations that apply a high-frequency assumption directly to the wavefield. In this work, I am intending to make a higher fidelity algorithm like GPSPI more palatable for general commercial use by allowing for the efficient measurement of the convergence of a plane wave GPSPI migration algorithm. This imaging algorithm can help to reduce the computational cost of seismic imaging, and a useful measurement of this improvement can definitely assist in the widespread use of these sorts of algorithms.

Finally, in Chapter 7, I extend the operator symbol in GPSPI again, this time in order to incorporate a derivative in the vertical velocity model. Although in a sense this does incorporate a more interesting local model of velocity than the simple locally-homogeneous symbol, the true benefit is not the inclusion of more complicated physics. In fact, the actual useful part of this vertical gradient ends up being the simplification of the implementation of GPSPI-type algorithms. Algorithms of this type frequently end up being most efficiently applied in the space-frequency domain. The vertical gradient has a net effect of stabilizing the Gibbs phenomenon which develops in a naïve implementation of GPSPI in this domain. This is a similar overall benefit to that shown in Chapter 6: although we are not introducing a new broader-frequency validity, we are making an existing algorithm that has attractive low-frequency features much more practical and attractive to use in standard migration.

In conclusion, practical solutions to the wave equation invariably involve at least some level of approximation. Simply put, they're wrong. My overall accomplishment

during the completion of my dissertation has been to carefully look at especially the high-frequency assumptions within these solutions that underlie modern exploration seismic imaging, and to extend the range of validity of those solutions at least somewhat towards the range of frequencies/wavelengths that are common in seismic imaging.

In short, I have made seismic imaging less wrong than ever.

Bibliography

- Keiiti Aki and Paul G. Richards. *Quantitative Seismology*. University Science Books, second edition, 2002.
- Richard C. Aster, Brian Borchers, and Clifford H. Thurber. *Parameter Estimation and Inverse Problems*, volume 90 of *International Geophysics Series*. Elsevier Academic Press, 2005.
- A. J. Berkhout. *Seismic migration: Imaging of acoustic energy by wave field extrapolation*. Elsevier, 1984.
- Biondo Biondi. Solving the frequency-dependent eikonal equation. Technical Report 73, Stanford Exploration Project, 1997.
- N. Bleistein, J. W. Stockwell Jr., and J.K. Cohen. *Mathematics of multidimensional seismic imaging, migration, and inversion*. Springer-Verlag, 2001.
- A. Bourgeois, M. Bourget, P. Lailly, M. Poulet, P. Ricarte, and R. Versteeg. Marmousi, model and data. In *Proceedings of the 1990 EAGE Workshop on Practical Aspects of Seismic Data Inversion*, 1991.
- John P. Boyd. *Chebyshev and Fourier spectral methods*. Dover Publications Inc., 2001.
- A. J. Brenders and R. G. Pratt. Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model. *Geophysical Journal International*, 168:133–151, 2006.

- A. J. Brenders and R. G. Pratt. Efficient waveform tomography for lithospheric imaging: implications for realistic, two-dimensional acquisition geometries and low-frequency data. *Geophysical Journal International*, 168:152–170, 2007.
- Richard L. Burden and Douglas J. Faires. *Numerical Analysis*. Brooks Cole, December 2004. ISBN 0534392008.
- David Cho, Chad Hogan, and Gary Margrave. Acoustic finite difference parameter analysis and modelling in MATLAB. Technical report, CREWES Research Report, 2007.
- Jon Claerbout. Toward a unified theory of reflector mapping. *Geophysics*, 36:467–481, 1971.
- Jon F. Claerbout. *Imaging the Earth's Interior*, chapter 5. Blackwell Scientific Publishers Inc., 1985.
- CREWES. CREWES Matlab Library. 2010. available at <http://www.crewes.org>.
- M. V. de Hoop, J. H. Le Rousseau, and R. S. Wu. Generalization of the phase-screen approximation for the scattering of acoustic waves. *Wave Motion*, 31:43–70, 2000.
- B. Duquet, P. Lailly, and A. Ehinger. 3D plane wave migration of streamer data. In *Expanded abstracts of the 71st SEG International Convention*, 2001.
- J.T.E. Etgen. Stability analysis of explicit depth extrapolation through laterally-varying media. In *Expanded abstracts of the 64th SEG International Convention*, 1994.
- R. J. Ferguson and G. F. Margrave. Explicit Fourier wavefield operators. *Geophysical Journal International*, 165(1):259–271, 2006.

- Louis Fishman. Applications of directional wavefield decomposition, phase space, and path integral methods to seismic wave propagation and inversion. *Pure and Applied Geophysics*, 159:1637–1679, 2002.
- Louis Fishman, A. K. Gautesen, and Zhiming Sun. Uniform high-frequency approximations of the square root Helmholtz operator symbol. *Wave Motion*, 26:127–161, 1997.
- Jenö Gazdag. Wave equation migration with the phase-shift method. *Geophysics*, 43:1342–1556, 1978.
- Jenö Gazdag. Wave equation migration with the accurate space derivative method. *Geophysical Prospecting*, 28:60–70, 1980.
- Jenö Gazdag and Piero Sguazzero. Migration of seismic data by phase shift plus interpolation. *Geophysics*, 49:124–131, 1984.
- Alain Grigis and Johannes Sjöstrand. *Microlocal Analysis for Differential Operators*, volume 196 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 1994.
- D. Hale. Stable explicit depth extrapolation of seismic wavefields. *Geophysics*, 56:1770–1777, 1991.
- R. H. Hardin and F. D. Tappert. Applications of the split-step Fourier method to the numerical solution of nonlinear and variable coefficient wave equations. *SIAM Review*, 15:423, 1973.
- Eugene Hecht. *Optics*, chapter 5.2. Addison Wesley, 3rd edition, 1997.

- Chad M. Hogan and Gary F. Margrave. Frequency-dependent velocity smoothing in GPSPI migration. Technical report, CREWES Research Report, 2006.
- O. Holberg. Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave phenomena. *Geophysical prospecting*, 35:629–655, 1987.
- O. Holberg. Towards optimum one-way wave propagation. *Geophysical Prospecting*, 36:99–114, 1988.
- Lars Hörmander. *The Analysis of Linear Partial Differential Operators*, volume 1-4. Springer, 1985.
- P. Lailly. The seismic inverse problem as a sequence of before stack migrations. In J. B. Bednar, R. Redner, E. Robinson, and A. Weglein, editors, *SIAM Philadelphia, Conference on Inverse Scattering: Theory and Application*, 1983.
- Rod Lakes. *Viscoelastic Solids*. CRC Press, 1998.
- J.H. Le Rousseau and M.V. de Hoop. Modeling and imaging with the scalar generalized-screen algorithms in isotropic media. *Geophysics*, 66:1551–1568, 2001.
- Laurence R. Lines, Raphael Slawinski, and R. Phillip Bording. A recipe for stability analysis of finite-difference wave equation computations. Technical report, CREWES Research Report, 1998.
- F. Liu, R. Stolt, D. Hanson, and R. Day. Plane wave source composition: an accurate phase encoding scheme for prestack migration. In *Expanded abstracts of the 72nd SEG International Convention*.

- F. Liu, D. Hanson, N. Whitmore, R. Day, and R. Stolt. Toward a unified analysis for source plane-wave migration. *Geophysics*, 71(4):S129–S139, 2006.
- Gary F. Margrave. Direct Fourier migration for vertical velocity variations. *Geophysics*, 66:1504–1514, 2001.
- Gary F. Margrave. Theory of nonstationary linear filtering in the Fourier domain with application to time-variant filtering. *Geophysics*, 63:244–259, 1998.
- Gary F. Margrave and Robert J. Ferguson. Wavefield extrapolation by nonstationary phase shift. *Geophysics*, 64:1067–1078, 1999.
- Gary F. Margrave, Saleh M. Al-Saleh, Hugh D. Geiger, and Michael P. Lamoureux. The FOCI algorithm for seismic depth migration. Technical report, CREWES Research Report, 2004.
- Gary F. Margrave, Hugh D. Geiger, Saleh M. Al-Saleh, and Michael P. Lamoureux. Optimizing explicit depth migration with a stabilizing Wiener filter and spatial resampling: 75th Annual International Meeting. *SEG, Expanded Abstracts*, 2005.
- Gary F. Margrave, Hugh D. Geiger, Saleh M. Al-Saleh, and Michael P. Lamoureux. Improving explicit depth migration with a stabilizing Wiener filter and spatial resampling. *Geophysics*, 71:S111–S120, 2006.
- André Martinez. *An Introduction to Semiclassical and Microlocal Analysis*. Universitext. Springer, 2002.
- P. R. Mora. Nonlinear two-dimensional elastic inversion of multioffset seismic data. *Geophysics*, 52:1211–1228, 1987.

- Philip M. Morse and Herman Feshbach. *Methods of Theoretical Physics*. Feshbach Publishing LLC, 2005.
- C. Mosher and D. Foster. Offset plane wave propagation in laterally varying media. In *Mathematical Methods in Geophysical Imaging V: Proceedings of SPIE 3453*, 1998.
- Isaac Newton. *Philosophiæ Naturalis Principia Mathematica*. Royal Society, London, 1687.
- David M. Pai. Generalized f-k (frequency-wavenumber) migration in arbitrarily varying media. *Geophysics*, 53:1547–1555, 1988.
- R. G. Pratt. Frequency domain elastic wave modeling by finite differences: A tool for cross-hole seismic imaging. *Geophysics*, 55:626–632, 1990.
- R. G. Pratt. Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model. *Geophysics*, 64(3):888–901, 1999.
- R. G. Pratt and R. M. Shipp. Seismic waveform inversion in the frequency domain, part 2: Fault delineation in sediments using crosshole data. *Geophysics*, 64(3):902–914, 1999.
- R. G. Pratt and M. H. Worthington. Inverse theory applied to multi-source cross-hole tomography. part I: Acoustic wave-equation method. *Geophysical Prospecting*, 38:287–310, 1990.
- R. G. Pratt, C. Shin, and G. Hicks. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophysical Journal International*, 133:341–362, 1998.

- R. Gerhard Pratt. *Waveform Tomography: An introduction to theory and practice*, April 2008. Course notes from: Waveform Tomography 2008: Beyond the First Arrival.
- José Pujol. *Elastic wave propagation and generation in seismology*. Cambridge University Press, 2003.
- Lennart Råde and Bertil Westergren. *Mathematics handbook for science and engineering*. Springer-Verlag, Berlin, 2004.
- W. Rietveld. *Controlled Illumination in Prestack Seismic Migration*, chapter 5.4. Delft University of Technology, 1995.
- W. Rietveld, A. Berkhout, and C. Wapenaar. Optimum seismic illumination of hydrocarbon reservoirs. *Geophysics*, 57:1334–1345, 1992.
- Etienne Robein. *Velocities, Time-imaging and Depth-imaging: Principles and Methods*. EAGE Publications BV, 2004.
- L. Romero, D. Ghiglia, C. Ober, and S. Morton. Phase encoding of shot records in prestack migration. *Geophysics*, 65:426–436, 2000.
- Xavier Saint Raymond. *Elementary Introduction to the Theory of Pseudodifferential Operators*. Studies in Advanced Mathematics. CRC Press, Inc., 1991.
- J. A. Sethian and A. M. Popovici. Three dimensional travelttime computation using the fast marching method. *Geophysics*, 64:516–523, 1999.
- L. Sirgue and R. G. Pratt. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies. *Geophysics*, 69:231–248, 2004.

- P.L. Stoffa, J.T. Fokkema, R.M. de Luna Freire, and W.P. Kessinger. Split-step Fourier migration. *Geophysics*, 55:410–421, 1990.
- Christiaan C. Stolk. A pseudodifferential equation with damping for one-way wave equation propagation in inhomogeneous acoustic media. *Wave Motion*, 40(2):111–121, 2004.
- C. Stork and J. Kapoor. How many p values do you want to migrate for delayed-shot wave equation migration? In *Expanded abstracts of the 74th SEG International Convention*, 2004.
- A. Tarantola. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49:1259–1266, 1984.
- A. Tarantola. *Inverse Problem Theory and Model Parameter Estimation*. Society for Industrial and Applied Mathematics, 2005.
- V. Červený. *Seismic Ray Theory*. Cambridge University Press, 2001.
- C.P.A. Wapenaar and J.L.T. Grimbergen. A discussion on stability analysis of wave-field depth extrapolation. In *Expanded abstracts of the 68th SEG International Convention*, pages 1716–1719, 1998.
- N. Whitmore. *An imaging hierarchy for common angle plane wave seismograms*. PhD thesis, University of Tulsa, 1995.
- Marta Jo Woodward. Wave-equation tomography. *Geophysics*, 57(1):15–26, January 1992.

Ru-Shan Wu. Wide-angle elastic wave one-way propagation in heterogenous media and an elastic wave complex-screen method. *Journal of Geophysical Research*, 99: 751–766, 1994.

Öz Yilmaz. *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data*. Society of Exploration Geophysicists, 2001.

UNIVERSITY OF CALGARY
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a dissertation entitled “Extending high-frequency asymptotic solutions to wave equations to lower-frequency regimes” submitted by Chad M. Hogan in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Dr. Gary F. Margrave
Department of Geoscience

Dr. Laurence Lines
Department of Geoscience

Dr. Michael P. Lamoureux
Department of
Mathematics and
Statistics, Department of
Geoscience

Dr. Robert Ferguson
Department of Geoscience

Dr. Elise Fear
Department of Electrical
and Computer
Engineering

Dr. Matthew Yedlin
University of British
Columbia

Date