CWP-517
May 2005

Data Continuation for Data Regularization and Internal Multiples

Alison E. Malcolm

— Doctoral Thesis —
Geophysics

Defended on April 1, 2005

Committee Chair:	Prof. James A. McNeil
Advisor:	Prof. Maarten V. de Hoop
Co-Advisor:	Prof. John A. Scales
Committee members:	Prof. Michael L. Batzle
	Prof. Yaoguo Li
	Prof. Paul A. Martin

Center for Wave Phenomena
Colorado School of Mines
Golden, Colorado 80401
(1) 303 273-3557

Abstract

Seismic data collected in the field are often not ideal for processing. The process known as data continuation computes data not recorded from those that are recorded so that data requirements for processing techniques can be met. Although there are many techniques of data continuation currently used in seismic processing, the majority of these assume that the seismic wave velocity is either constant or varying only with depth. A notable exception is the downward continuation of data, often referred to as survey sinking, for which techniques applicable in most velocity models exist.

We extend data continuation techniques used to fill in missing data to velocity models in which caustics are generated in the wavefield. To do this, we use a method based on the composition of Fourier integral operators. To demonstrate that this method doesn't introduce false reflections, we show that the composite operator is also a Fourier integral operator. We illustrate the utility of this theory with a synthetic example, with caustics, in which we fill in missing traces in a shot record. This method is computationally more expensive than similar methods that assume simple velocity models.

First order internal multiples are a source of errors seismic imaging. Artifacts caused by internal multiples are often similar to true reflectors and thus can be difficult to attenuate. Typically multiples are estimated in the data and then subtracted from the data before an image is created. We propose a method by which artifacts in the image are estimated as part of the imaging process; an integral part of this method is the downward continution of data.

Table of Contents

Acknowledgments

First and foremost I want to express my gratitude to my advisor Martijn without whom this thesis would not have happened. Doing science with Martijn has been a wild ride and though I often felt that I was in over my head he always seemed to know when I really needed help and when to leave me to figure things out on my own. I appreciate not only the direct guidance he gave me for my thesis work, but also the breadth of interests and topics he shared with me in many more general scientific discussions. I am also grateful (and impressed) that he never lost patience with me despite my giving him plenty of reasons to.

For most of my time at Mines I have worked with Henri Calandra at Total. I am grateful to him for all of his help and advice with the practical side of the work in this thesis, organizing the financial support for much of this work, the opportunity to visit and work in France and the French lessons.

I was fortunate to have gone through a comprehensive exam process that required two projects with two different faculty members. The work I did with John Scales in the Physical Acoustics Lab was some of the most exciting times I spent on campus. I feel that sharing in John's enthusiasm, his physical insights and his sense of fun has made me a better scientist.

I am grateful for the assistance I have received at every step of the way from the rest of my committee, Mike Batzle, Paul Martin, Jim McNeil and Yaoguo Li as well as the department head Terry Young. I also wish to thank them for their understanding of the somewhat chaotic way in which my thesis was distributed. I have also appreciated the gallons of red ink from Ken Larner, the helpful advice (and jokes) from Roel Snieder and the coding assistance from Dave Hale.

From the first day I arrived at Mines Jérôme Le Rousseau has helped to smooth the way for me. He, along with Günther Hörmann, really got me started in the theory of microlocal analysis and taught me where to find the information I needed and how to learn it. Jérôme has also been a constant source of support and encouragement to me throughout my time here; he introduced me to Henri, encouraged me to work with Martijn and taught me many of his vast supply of computer tricks.

The comraderie amongst the students in the department and in particular in CWP has helped me to really enjoy my time at Mines. I am grateful to all of the students I have shared space with for their individual insights and for sharing with me a little of their own culture (and to Alex and Nicole for keeping me in touch with my own). There are a few students whose presence really smoothed the way for this thesis, to whom I am particularly grateful: Alex–for great car conversations, David–for his always unique perspectives, Ivan–for sharing his intelligence and sense of humour, Matt R.–for pretending I knew more math than him, Sverre and Stig–for bits and pieces of code, and Tamara–for all the gal stuff. I

Chapter 1

Introduction

The subject of this thesis is data continuation through operator composition, in the general context of seismic imaging. Two applications are investigated, seismic data regularization and the attenuation of imaging artifacts caused by internal multiples. Each chapter is written to be published independently. As such, each contains its own summary and introduction describing the material contained in it. The purpose of this introduction is to tie the four papers together and to give the reader a general overview of the thesis contents. Rather than explain the basics of seismic processing here, I refer the reader not familiar with exploration seismology to two excellent books on the subject: Claerbout (1985) and Yilmaz (1987). Similarly, the reader not familiar with the techniques of microlocal analysis is referred either to the book by Sjöstrand & Grigis (1994), or Appendix A of Le Rousseau (2001). Some of the theory used for multiple prediction is described in detail in Appendix G.

Because this thesis is a complitation of papers, it is not designed to be read cover-to-cover. Chapters 2 and 3 deal primarily with data regularization, whereas Chapters 4 and 5 deal with estimating imaging artifacts caused by internal multiples. The majority of the appendices are associated to particular chapters, as noted in the footnotes at the beginning of each chapter. Appendices A and B give some constant velocity examples of the theory described in Chapter 2. Appendices C-F give some of the more detailed derivations associated with Chapter 3. Appendix G attempts to describe some of the theory of the Lippmann-Schwinger and Bremmer series, which are used extensively in Chapter 4. Appendix H describes some of the artifacts that caused problems in the algorithm development phase leading up to the examples shown in Chapter 5. In Appendix I the amplitude factors necessary for the imaging procedure performed in Chapter 5 are discussed to bridge the gap between the theoretical development and the implementation of these amplitude factors. The final two appendices, J and K are associated with Chapter 4. Appendix J gives a particularly long proof of one of the results in Chapter 4 and Appendix K compares the approach for suppressing internal multiples discussed here with that of Weglein *et al.* (1997) and ten Kroode (2002).

Data continuation is the process by which data are computed at a location different from that at which they were collected. A typical example from seismic imaging is so called survey-sinking *wave equation*[1] migration (Bevc *et al.*, 2003). In this example, data collected at the surface are used to estimate data that would have been recorded at some depth, say

[1]This type of migration goes by many names, the most common of which is wave equation migration. This name is somewhat misleading, however, as all migration techniques are based on the wave equation.

(the distance between the source and receiver) were compensated for (i.e., normal-moveout corrected) with the goal of producing a flat section that could be averaged to give a one-dimensional (1D) estimate of subsurface properties. This is in essence a data regularization as the correction for offset is designed to mimic zero-offset data, which are more easily interpreted than are nonzero-offset data. These first approaches made strong assumptions on Earth structure [see Green (1938) for a list], most of which are not at all valid in areas of current interest for oil exploration. Two of the most fundamental assumptions were that the underlying structure consists of a stack of horizontal layers and that the background velocity is constant or at most varies consistently (e.g., linearly) as a function of depth only.

Accounting for dip came first, before computers were readily available, by adjusting the velocity used in common-midpoint (CMP) stacking[3] to stack to account for dipping layers. This is discussed by Levin (1971), drawing on the derivation of the reflection response to a dipping layer made by Brown (1969). Again the goal is to reduce a common-depth-point (CDP)[4] gather to a set of identical traces to be stacked to determine Earth structure at the shared depth point. Further research along these lines introduced the process known as *Dip moveout* (DMO), described in detail by Hale (1983, 1991). Dip moveout is a method by which zero-offset data are estimated from nonzero-offset data extending normal moveout to dipping reflectors. The strengths of this method are in its relatively weak dependence on the precise subsurface velocity and the speed with which it can be computed. A weakness of the method is its inability to account for lateral velocity variations. This is addressed as part of this thesis.

The data regularization method described in this thesis involves the composition of two (or three) operators. The first is an imaging operator from which an estimate of subsurface properties is obtained. The second is a modeling operator that models output data with the desired acquisition geometry. (The third is a restriction operator, used to set up the desired output geometry; it is discussed in both Chapter 2 and 3.) A similar vein of research has been followed by others, a few of which are described in the following paragraphs.

A recently developed method, similar to DMO is azimuth moveout (AMO), first described by Biondi *et al.* (1998). This technique allows the regularization of a data set from multi-azimuth to a single-azimuth. It has proved particularly useful to do this in marine seismology where data are typically collected with a streamer running behind a boat. In this situation there are not enough data to make a true 3D image, but the data are spread over too large an area in the direction perpendicular to the streamer for 2D processing techniques. The technique works by using DMO to go from the original data set to a zero-offset data set and then inverse DMO to go from this zero-offset data set to a nonzero-offset data set at a chosen azimuth. In a constant-background velocity, these imaging and modeling operators can be composed to form a single operator mapping from one azimuth to another. The constant background assumption is often sufficient provided the true subsurface velocity is not overly complex.

[3]The midpoint is halfway between the source and receiver; stacking is an average over offset.

[4]A CDP gather is a set of traces that are assumed to have scattered from the same point at depth. A CDP gather is often approximated by a common midpoint gather (CMP) in which all of the traces share the same midpoint between the source and receiver; the two are equivalent in horizontally layered media.

continuation of data from one source-receiver configuration to another without introducing false reflections. (For a detailed explanation see the discussion around Equation 25.22 of Hörmander (1985a).)

Chapter 2 treats the two special cases of DMO and AMO discussed previously. In Appendix A a number of these results are derived for a constant-background velocity, including the derivation of a closed form expression for the DMO impulse response. Appendix A also includes an explanation as to why not all the phase variables[6] can necessarily be integrated out when applying the method of stationary phase. In Appendix B, we give the AMO impulse response.

1.1.2 Examples and Illustrations

The purpose of Chapter 3 is to explain and illustrate the theory described in Chapter 2 in the context of exploration geophysics rather than mathematics. There is little material directly repeated from the previous chapter, although the derivations in Appendix D are contained also in Appendix A and Appendix B in a slightly different form.

The first sections of this chapter give a physical overview of the theory developed in Chapter 2. As part of this development a number of impulse responses are shown to illustrate the salient points of the theory. These impulse responses are not sufficient to demonstrate the applicability of the theory but they do illustrate the construction of the composite operators and show the extent, both in time and in space, of the resultant operators. Following this, the method is illustrated by estimating missing traces in shot records. A description of the algorithm used to do this is also given in this section.

The algorithm discussed in Chapter 3 works but is not the most computationally efficient algorithm possible. Examples of a more efficient algorithm are the fast marching techniques of Sethian (2002). With these techniques, it is possible to construct the table, which will be described in Chapter 3, relating subsurface and surface parameters, with a single pass through the velocity model. Once the table is constructed, data continuation, velocity model building, and imaging can be performed through direct operations on the table. Another way to get a faster algorithm is to use the curvelet techniques currently being developed by my fellow student Huub Douma.

The data regularization theory described in this thesis extends the family of velocity models in which data regularization techniques can be applied. The cost of this extension is in additional time requirements for computations and the requirement that the velocity model be known.

---

[6]Phase variables are the variables typically integrated out through the method of stationary phase. They are often, though not always, Fourier variables associated with the space/time variables of the equation. For an operator to be an FIO its phase is required to be homogeneous of degree one in these variables, i.e., $\phi(x, \lambda\theta) = \lambda\phi(x, \theta)$, where θ is the phase variable and the FIO, F, is defined through its action on a test function a by

$$(Fa) = \int\int a(x, \theta) u(x) e^{\mathrm{i}\phi(x,\theta)} \mathrm{d}\theta \mathrm{d}x \,.$$

Fokkema *et al.* (1994); Berkhout & Verschuur (1997); Verschuur & Berkhout (1997) and van Borselen (2002). In some sense, the work presented here is an extension of the surface-related multiples attenuation theory proposed by Aminzadeh & Mendel (1981), who used the Bremmer series to suppress surface multiples. The closest approach to that discussed here, however, is that of Jakubowicz (1998) in which he implicitly uses the Generalized Bremmer series to estimate internal multiples, under the traveltime monotonicity assumption of ten Kroode (2002) (see Figure 1.2). The work of Weglein *et al.* (1997), the mathematical fundamentals of which are presented by ten Kroode (2002), is another series expansion of the wavefield used to estimate internal multiples. Their approach uses a scattering series derived from the Lippmann-Schwinger equation, which will be referred to in what follows as the Lippmann-Schwinger series.

1.2.1 Theoretical description

Chapter 4 describes a theory for attenuating imaging artifacts caused by multiples. This theory involves the development of a hybrid series between the Lippmann-Schwinger and Bremmer series. These two series are discussed in more detail in Appendix G.

The Lippmann-Schwinger series, first proposed by Lippmann (1956) [see also Lippmann & Schwinger (1950); Lippmann (1950)], has been investigated extensively for attenuating internal multiples by Weglein *et al.* (1997, 2003) and ten Kroode (2002). This series is an expansion in the difference between the wave operators in a known reference model and of the unknown true velocity model. Each successive term in the series is of higher order in the difference between the wave operators in these two models. Internal multiples can be estimated from this series, and attenuated through the associated inverse series, which is constructed by assuming that the medium contrast can be expanded in a series with each successive term of higher order in the data. Currently, there are no estimates for convergence on either the forward or inverse seres. With two assumptions on the velocity model, however, this series can be used to attenuate internal multiples without knowledge of the subsurface velocity. The two assumptions are that the wavefield does not contain caustics, and the traveltime monotonicity assumption of ten Kroode (2002), illustrated in Figure 1.2.

The Bremmer series was first developed by Bremmer (1951) for modeling atmospheric phenomena. This series builds up a wavefield by adding together the contributions from successive reflections from different layers in the medium. As constructed by Bremmer, this series is applicable only in a velocity model consisting of a stack of horizontal layers. This series was generalized by de Hoop (1996) to arbitrary reflector geometry and variable background; it is this generalization that is used throughout the thesis. The Bremmer series separates the wavefield into its up- and down-going constituents and as such fits into the framework of a *wave-equation* migration in which an image is formed by propagating the wavefield down into the Earth and extracting the part of the wavefield at zero offset and zero time to be added to the image at each depth. This imaging method works well in the presence of caustics. de Hoop (1996) gives convergence estimates for the Generalized Bremmer series. Drawbacks to the use of this series are that it is not straightforward to

involved in the multiple scattering. The computational cost of the algorithm is not prohibitive however; the multiples are estimated with roughly twice the computational cost of a standard migration.

1.3 Acknowledgment

Each chapter of this thesis was written to be published with various co-authors; a full citation of each paper may be found in the footnotes on the title page of the associated chapter. Although each of them was written with the goal of being part of this thesis, much of the work presented was done in close collaboration with Maarten de Hoop, Jérôme Le Rousseau and Henri Calandra. The author gratefully acknowledges their contributions here.

Chapter 2

Seismic wavefield 'continuation' in the single scattering approximation: A framework for Dip and Azimuth MoveOut[1]

2.1 Summary

Seismic data are commonly modeled by a high-frequency single scattering approximation. In this paper we use methods from microlocal analysis and the theory of Fourier integral operators, to study continuation of the seismic wavefield in this single scattering approximation. This amounts to a linearization in the medium coefficient about a smooth background. The discontinuities are contained in the medium perturbation. We use the smooth background to derive the continuation as the composition of imaging, modeling and restriction operators.

2.2 Introduction

In reflection seismology one places point sources and point receivers on the Earth's surface. The source generates acoustic waves in the subsurface, that are reflected where the medium properties vary discontinuously. The recorded reflections that can be observed in the data are used to reconstruct these discontinuities. In principle, the recordings are taken on an acquisition manifold, made up of all source and receiver positions and a time interval. In practice, however, certain subsets in the acquisition manifold are not covered. In this paper, we discuss how, and conditions when, data can be continued from an acquisition submanifold to the complete acquisition manifold.

The data are commonly modeled by a high-frequency single scattering approximation. This amounts to a linearization in the medium coefficient about a smooth background. The discontinuities are contained in the medium perturbation (Beylkin, 1985). Thus a linear operator, the modeling operator, depending on the background, that maps the perturbation to the data is obtained. The smooth background (C^∞) is associated with a computational medium (which can be chosen) rather than a physical one, the distributional ($\mathcal{E}'$) perturba-

[1]This chapter has been published, along with appendices A and B as:
de Hoop, M. V., Malcolm, A. E. and Le Rousseau, J. H. 2003. Seismic wavefield 'continuation' in the single scattering approximation: A framework for dip and azimuth moveout *Can. Appl. Math. Q.*, **10**, 199-238.

smooth background. (iii) AMO can be employed to carry out approximate (based on a linearized scattering model) seismic data 'regularization'.

The basic idea of investigating the composition of imaging and modeling operators dates back in particular to the work of Goldin (1994).

2.3 High-frequency Born modeling and imaging

We consider the scalar wave equation for acoustic waves in a constant density medium in $\mathbb{R}^n$. We introduce coordinates $x \in \mathbb{R}^n$. The scalar acoustic wave equation is given by

$$Pu = f\,, \qquad P = c(x)^{-2}\frac{\partial^2}{\partial t^2} + \sum_{j=1}^{n} D_{x_j}^2, \tag{2.1}$$

where $D_x = -\mathrm{i}\frac{\partial}{\partial x}$. The equation is considered in a time interval $]0,T[$.

If $c \in C^\infty$ the solution operator of (2.1) propagates singularities along bicharacteristics. These are the solutions of a Hamilton system with Hamiltonian given by the principal symbol of P,

$$P(x,\xi,\tau) = -c(x)^{-2}\tau^2 + \|\xi\|^2.$$

The Hamilton system is given by

$$\frac{\partial(x,t)}{\partial\lambda} = \frac{\partial P}{\partial(\xi,\tau)}, \quad \frac{\partial(\xi,\tau)}{\partial\lambda} = -\frac{\partial P}{\partial(x,t)}. \tag{2.2}$$

Its solutions will be parameterized by initial position (x_0), take-off direction ($\alpha \in S^{n-1}$), frequency (τ) and time (t),

$$\boldsymbol{x} = \boldsymbol{x}(x_0,\alpha,\tau,t)$$

and similarly for $\boldsymbol{t},\boldsymbol{\xi}$; τ is invariant along the Hamilton flow. The evolution parameter λ is the time t.

By Duhamel's principle, a causal solution operator for the inhomogeneous equation (2.1) is given by

$$u(x,t) = \int_0^t \int G(x,t-t_0,x_0) f(x_0,t_0)\,\mathrm{d}x_0\mathrm{d}t_0, \tag{2.3}$$

where G defines a Fourier integral operator (FIO) with canonical relation, Λ_G, that is essentially a union of bicharacteristics,

$$\Lambda_G = C_+ \cup C_-,\ C_\pm = \{(\boldsymbol{x}(x_0,\alpha,\tau,\pm t), t, \boldsymbol{\xi}(x_0,\alpha,\tau,\pm t), \mp\tau\ ;\ x_0, \underbrace{-\mp(\tau/c(x_0))\alpha)}_{\xi_0}\}.$$

Let

$$(x_I, x_0, \underbrace{\xi_J, \tau}_{\theta}) \quad \text{with} \quad I \cup J = \{1,\dots,n\}\,,\ N := |J|+1\,,$$

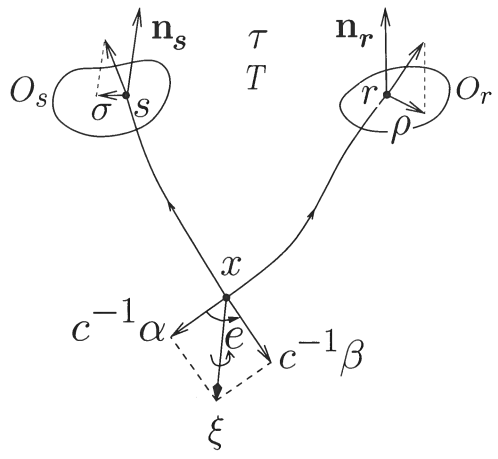


Figure 2.1. Source-receiver bicharacteristics.

Theorem 2.3.1. *(Rakesh, 1988) With Assumption 1 the map F is a Fourier integral operator $\mathcal{D}'(X) \to \mathcal{D}'(Y)$ of order $(n-1)/4$ with canonical relation*

$$\Lambda_F = \{(\boldsymbol{s}(x_0,\beta), \boldsymbol{r}(x_0,\alpha), \overbrace{T(x_0,\alpha)+T(x_0,\beta)}^{T(x_0,\alpha,\beta)}, \boldsymbol{\sigma}(x_0,\beta), \boldsymbol{\rho}(x_0,\alpha), \tau; \\ x_0, -\underbrace{(\tau/c(x_0))\,(\alpha+\beta)}_{\boldsymbol{\xi}(x_0,\alpha,\beta,\tau)}) \mid (x_0,\alpha,\beta) \in K,\ \tau \in \mathbb{R}\backslash 0\} \subset T^*Y\backslash 0 \times T^*X\backslash 0\,, \tag{2.8}$$

where $K \subset \mathbb{R}^n \times \{(\alpha,\beta) \in S^{n-1} \times S^{n-1} \mid \alpha+\beta \neq 0\}$. Here,

$$\underbrace{\boldsymbol{x}(x_0,\alpha,\tau,T(x_0,\alpha))}_{=:\boldsymbol{r}(x_0,\alpha)} \in O_r\,, \quad \underbrace{\boldsymbol{x}(x_0,\beta,\tau,T(x_0,\beta))}_{=:\boldsymbol{s}(x_0,\beta)} \in O_s,$$

which expresses that the time T is locally solved from the equation describing the intersection of the rays with the acquisition manifold, while

$$\boldsymbol{\rho}(x_0,\alpha) = (I - \mathbf{n_r} \otimes \mathbf{n_r})\boldsymbol{\xi}(x_0,\alpha,\tau,T(x_0,\alpha)) \tag{2.9}$$

where $\mathbf{n_r}$ is the unit normal to O_r at $\boldsymbol{r}(x_0,\alpha)$. A similar expression holds for $\boldsymbol{\sigma}(x_0,\beta)$.

The parametrization of Λ_F is illustrated in Figure 2.1. The cotangent vectors σ, ρ can be identified with acquisition 'slopes' p_s, p_r, in accordance with $\boldsymbol{\sigma}(x_0,\beta) = -\tau p_s(x_0,\beta)$ and $\boldsymbol{\rho}(x_0,\alpha) = -\tau p_r(x_0,\alpha)$.

Assumption 1 is microlocal. One can identify the conic set of points $(s,r,t,\sigma,\rho,\tau) \in T^*Y\backslash 0$ where this assumption is violated. If the symbol $\psi = \psi(s,r,t,\sigma,\rho,\tau)$ vanishes on a neighborhood of this set, then the composition ψF of the pseudodifferential cutoff

minimum number of phase variables is given by the corank of the projection

$$D\pi \ : \ T\Lambda \to T(Y \times X)$$

at $(x_0, \alpha, \beta, \tau)$, which is here given by

$$\text{corank } D\pi = 1 + \text{corank } \frac{\partial \boldsymbol{s}}{\partial \beta}(x_0, \beta) + \text{corank } \frac{\partial \boldsymbol{r}}{\partial \alpha}(x_0, \alpha).$$

This corank is > 1 when s or r is in a caustic point relative to x_0. Let

$$\Lambda'_F = \Lambda_F \backslash \{\text{closed neighborhood of } \{\lambda \in \Lambda_F \mid \text{corank } D\pi > 1\}\} \ . \tag{2.10}$$

Λ'_F can be described by phase functions of the 'traveltime' form $\tau\,(t - T^{(m)})$ with the only phase variable being τ. Here, $T^{(m)}$ is the value of the time variable in (2.8). The index m labels the branches of the multi-valued traveltime function. Thus the set $\{T^{(m)}\}_{m \in M}$ describes the canonical relation (2.8) except for a neighborhood of the subset of the canonical relation where the mentioned projection is degenerate. Each $T^{(m)}$ can be viewed as a function defined on a subset $D^{(m)}$ of $X \times O_s \times O_r$. We define $F^{(m)}$ to be a contribution to F with phase function given by $\tau(t - T^{(m)}(x, s, r))$, and symbol $A^{(m)}$ in a suitable class such that on the subset Λ'_F of the canonical relation F is given microlocally by $\sum_{m \in M} F^{(m)}$.

2.4 Generalized Radon transform

We can use $(x, \xi) \in T^*X \backslash 0$ as local coordinates on the canonical relation (2.8) (cf. (Hörmander, 1985b, Prop. C.3.3)). In addition, we need to parameterize the subsets (these are characteristic strips) of the canonical relation given by $(x, \xi) =$ constant; we denote such parameters by e. The canonical relation (2.8) was parameterized by (x, α, β, τ). We relate (x, ξ, e) by a coordinate transformation to (x, α, β, τ): A suitable choice when $\alpha \neq \beta$ is the scattering angles given by de Hoop *et al.* (1999)

$$\boldsymbol{e}(x, \alpha, \beta) = \left(\arccos(\alpha \cdot \beta), \frac{-\alpha + \beta}{2 \sin(\arccos(\alpha \cdot \beta)/2)} \right) \in]0, \pi[\times S^{n-2}. \tag{2.11}$$

On $D^{(m)}$ there is a map $(x, \alpha, \beta) \mapsto (x, s, r)$. We define $\boldsymbol{e}^{(m)} = \boldsymbol{e}^{(m)}(x, s, r)$ as the composition of $\boldsymbol{e}$ with the inverse of this map.

In preparation for the generalized Radon transform (GRT) we define the 'angle' transform, L, via a restriction in F^* of the mapping $\boldsymbol{e}^{(m)}$ to a prescribed value e, i.e. the distribution kernel of each contribution $F^{(m)^*}$ is multiplied by $\delta(e - \boldsymbol{e}^{(m)}(x, s, r))$ (which is justified by (Hörmander, 1983, Thm. 8.2.10)). Invoking the Fourier representation of this δ, the kernel of L follows as

$$L(x, e, r, s, t) = \sum_{m \in M} (2\pi)^{-(n-1)} \int \overline{A^{(m)}(x, s, r, \tau)} \exp[\mathrm{i}\Phi^{(m)}(x, e, s, r, t, \varepsilon, \tau)] \, \mathrm{d}\tau \mathrm{d}\varepsilon, \tag{2.12}$$

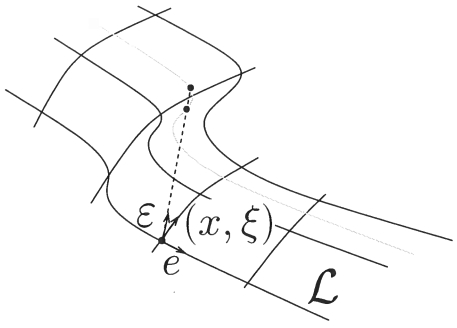


Figure 2.3. The origin of artifacts generated by the GRT. (Inside the $T^*Y\backslash 0$ box of Figure 2.2.)

with F can be evaluated through solving the system of equations

$$\begin{aligned} r &= \boldsymbol{r}(x,\alpha), && (2.15)\\ s &= \boldsymbol{s}(x,\beta), && (2.16)\\ T^{(m)}(z,s,r) &= T(x,\alpha)+T(x,\beta), && (2.17)\\ \boldsymbol{\rho}^{(m)}(z,s,r,\tau,\varepsilon) &= -\tau p_r(x,\alpha), && (2.18)\\ \boldsymbol{\sigma}^{(m)}(z,s,r,\tau,\varepsilon) &= -\tau p_s(x,\beta). && (2.19) \end{aligned}$$

(The frequency is preserved.) Equations (2.15)-(2.17) imply that the image point z must lie on the isochrone determined by (x,s,r). Equations (2.18)-(2.19) enforce a match of slopes (apparent in the appropriate 'slant stacks') in the measurement process,

$$\begin{aligned} -\tau\partial_r T^{(m)}(z,s,r)+\langle\varepsilon,\partial_r \boldsymbol{e}^{(m)}(z,s,r)\rangle &= -\tau p_r(x,\alpha), && (2.20)\\ -\tau\partial_s T^{(m)}(z,s,r)+\langle\varepsilon,\partial_s \boldsymbol{e}^{(m)}(z,s,r)\rangle &= -\tau p_s(x,\beta). && (2.21) \end{aligned}$$

For $\varepsilon\neq 0$ the take-off angles of the pairs of rays at (r,s) following from the right-hand sides of (2.15)-(2.19) may be distinct from those following from the left-hand sides. Equations (2.20)-(2.21) imply the matrix compatibility relation (upon eliminating ε/τ)

$$\begin{aligned} &[\partial_r \boldsymbol{e}^{(m)}(z,s,r)]^{-1}[p_r(x,\alpha)-\partial_r T^{(m)}(z,s,r)]\\ &\quad=[\partial_s \boldsymbol{e}^{(m)}(z,s,r)]^{-1}[p_s(x,\beta)-\partial_s T^{(m)}(z,s,r)]. \end{aligned} \tag{2.22}$$

The geometrical composition equations determining the artifacts are solved as follows: For each $(x,\alpha,\beta)\in K$ solve the $(3n-2)$ equations (2.15)-(2.17), (2.22) for the $(3n-2)$ unknowns (z,s,r). (From (2.20) we then obtain ε/τ hence ε.)

The GRT reconstructs a distribution in $\mathcal{E}'(X)$ smoothly indexed by $e\in E$. We can trivially extend the domain of modeling operator F from $\mathcal{E}'(X)$ to $\mathcal{E}'(X\times E)$, its action

has full rank. In other words

$$\frac{\partial y''}{\partial(x,\alpha,\beta,\tau)} \quad \textit{has maximal rank.}$$

Applying (Duistermaat, 1996, Thm. 4.2.2) to the pair F and the restriction $\mathcal{R}^c$ from $O_s \times O_r \to \Sigma^c$ with Assumption 4 implies that $\mathcal{R}^c F$ is an FIO of order $(n+c-1)/4$ with canonical relation

$$\begin{aligned}\Lambda_F^c = \{(y',t,\eta',\tau;x,\xi) \mid \exists\ (y',y'',\eta',\eta'') \text{ such that} \\ y''=0 \text{ and } (y,\eta;x,\xi)\in\Lambda_F\} \\ \subset T^*Y^c\backslash 0 \times T^*X\backslash 0.\end{aligned} \tag{2.23}$$

We will encounter two examples: Zero offset (ZO), where $c=n-1$ and $\Sigma^c := \Sigma_0 \subset \mathrm{diag}(\partial X)$ (subject to the $n-1$ constraints $r=s$ when $\arccos(\alpha\cdot\beta)=0$ and e_o at x follows from (2.11)), and common azimuth (CA), where $c=1$ and $\Sigma^c := \Sigma_A$ subject to one constraint typically of the form that the $(n-1)$st coordinate in $r-s$ is set to zero, while $E_o \ni e$ at x follows from the mapping $\boldsymbol{e}^{(m)}$. We set $Y_0 = \Sigma_0 \times]0,T[$ and $Y_A = \Sigma_A \times]0,T[$.

The restriction to acquisition submanifolds is placed in the context of inversion in Nolan & Symes (1997).

2.6 Exploding reflector modeling

In this section we introduce a procedure to model zero-offset (ZO) data: data with coinciding sources and receivers. To ensure that the zero-offset experiment can be modeled by an FIO we invoke Assumption 4 with $\Sigma^c := \Sigma_0$. We denote its canonical relation by Λ_0.

For the zero-offset reduction to be 'image preserving', i.e. for the associated normal operator to be pseudodifferential, we mention

Assumption 5. *The projection*

$$\pi_{Y_0} \ : \ \Lambda_0 \to T^*Y_0\backslash 0$$

is an embedding.

(In fact, Assumption 4 with $\Sigma^c = \Sigma_0$ implies that π_{Y_0} is an immersion.) This assumption is most easily verified by checking whether an element (y_0,η_0) in $T^*Y_0\backslash 0$ uniquely determines an element $(x_0, \xi_0 = \partial_{x_0} T_0)$ in $T^*X\backslash 0$ smoothly; here, T_0 is the zero-offset traveltime. (In fact, Assumption 5 implies that the projection π_{Y_0} is a diffeomorphism, which coincides with Beylkin's condition (Beylkin, 1985).)

Remark 2.6.1. Assumptions 4 and 5 precisely allow the introduction of so-called map migration-demigration between the wavefront set of zero-offset data and the wavefront set of the singular medium perturbation.

In the absence of Assumption 5 we introduce the notion of the exploding reflector (ER) model in the following

Remark 2.6.4. Subjecting the configuration to Assumption 5, the exploding reflector modeling, F_0, is up to leading order singularities equivalent to restricting the multiple-offset modeling to zero offset, $\mathcal{R}_0\delta G$, where $\mathcal{R}_0$ is the restriction of $X \times X \times]0,T[$ to Y_0. Otherwise, the exploding reflector models only part of the zero-offset data.

2.7 Transformation to zero offset: Dip MoveOut

In applications, the data at zero offset is usually missing: Receivers cannot be placed on top of sources. Hence, as a first example, we analyze the continuation of multiple finite-offset seismic data to zero-offset seismic data. Dip MoveOut is the process following upon composing ER modeling with L_U, the imaging GRT for a neighborhood of a given value of e (conventionally for given value of offset $r-s$); the sing supp of the Lagrangian-distribution kernel of the resulting operator is what seismologists call the DMO 'impulse response'. The compose, F_0L, is a well-defined operator $\mathcal{D}'(Y) \to \mathcal{D}'(Y_0)$. Its wavefront set is contained in the composition of the wavefront sets of F_0 and L (Duistermaat, 1996, Thm. 1.3.7), hence in the composition of canonical relations,

$$\begin{aligned}\Lambda_E \circ \Lambda'_L = \{(z,t_0,\zeta,\tau_0;s,r,t,\sigma,\rho,\tau) \mid \exists\ (x,\xi,\varepsilon) \text{ such that} \\ (z,t_0,\zeta,\tau_0;x,\xi) \in \Lambda_E \text{ and } (x,e,\xi,\varepsilon;s,r,t,\sigma,\rho,\tau) \in \Lambda_L\} \\ \subset T^*Y_0\backslash 0 \times T^*Y\backslash 0\,. \end{aligned} \tag{2.25}$$

with $\Lambda'_L = \{(x,\xi;s,r,t,\sigma,\rho,\tau) \mid \exists\ \varepsilon \text{ such that } (x,e,\xi,\varepsilon;s,r,t,\sigma,\rho,\tau) \in \Lambda_L\}$. Whether the compose is an FIO is yet to be investigated.

Using the parametrization of Λ_E in (2.24) and the parametrization of Λ_L in (2.13), the compose (2.25) can be evaluated through solving a system of equations, the first n being trivial fixing the scattering point $x_0 = x$, the second n equating the cotangent vectors

$$\underbrace{2\tau_0\partial_x T(x_0,\alpha_0)}_{\boldsymbol{\xi}_0(x_0,\alpha_0,\tau_0)} = \underbrace{\tau\partial_x T^{(m)}(x_0,s,r) - \langle\varepsilon,\partial_x \boldsymbol{e}^{(m)}(x_0,s,r)\rangle}_{\boldsymbol{\xi}^{(m)}(x_0,s,r,\tau,\varepsilon)}\,. \tag{2.26}$$

Given x_0, these constitute n equations with the n unknowns (α_0,τ_0). Thus for each (s,r,τ,ε) we need to solve these equations.

Note that, given $e = \boldsymbol{e}^{(m)}(x_0,s,r)$, we can obtain r from s (cf. (2.11)). Thus we can parametrize the composition $\Lambda_E \circ \Lambda'_L$ by (x_0,s,τ,ε). We can interpret the computation of the composition as follows: (i) Given (x_0,s) we compute r and then $T^{(m)}$; (ii) then, given (τ,ε) we compute $\sigma = \boldsymbol{\sigma}^{(m)}$ and $\rho = \boldsymbol{\rho}^{(m)}$; (iii) we solve (2.26) for (α_0,τ_0); (iv) with these initial values, we solve the Hamiltonian flow (with (2.2) in the exploding reflector model) up to its intersection with the acquisition manifold Y_0, from which we deduce t_0 and z, as well as ζ.

Theorem 2.7.1. *With Assumptions 2 and 3 the composition F_0L_U yields a smooth family of FIOs parametrized by e. The compose is called Dip MoveOut. Its canonical relation is given by (2.25)*

$$\Lambda_D = \Lambda_E \circ U_L = \{(z,t_0,\zeta,\tau_0;s,r,t,\sigma,\rho,\tau)\}$$

$\int \mathrm{d}e F_0\langle N^{-1}\rangle L$ *is an FIO,* $\mathcal{D}'(Y) \to \mathcal{D}'(Y_0)$. *With Assumption 5 the reduced dataset is image preserving.*

The proof follows that of Theorem 2.3.2 closely (see (Stolk & de Hoop, 2002, Thm. 4.5)).

Remark 2.7.4. The adjoint $(F_0 L_U)^*$ is by Theorem 2.7.1 also an FIO. This operator is called 'inverse' DMO.

2.8 Continuation, transformation to common azimuth

Continuation

We analyze the 'continuation' of multiple finite-offset seismic data.

The compose FL is a well-defined operator $\mathcal{D}'(Y) \to \mathcal{D}'(Y)$. Its wavefront set is contained in the composition of the wavefront sets of F and L (Duistermaat, 1996, Thm. 1.3.7), hence in the composition of canonical relations,

$$\begin{aligned}\Lambda_F \circ \Lambda_L' = \{&(s_2, r_2, t_2, \sigma_2, \rho_2, \tau_2; s_1, r_1, t_1, \sigma_1, \rho_1, \tau_1) \mid \exists\ (x, \xi, \varepsilon) \text{ such that} \\ &(s_2, r_2, t_2, \sigma_2, \rho_2, \tau_2; x, \xi) \in \Lambda_F \text{ and } (x, e, \xi, \varepsilon; s_1, r_1, t_1, \sigma_1, \rho_1, \tau_1) \in \Lambda_L\} \\ &\subset T^*Y\backslash 0 \times T^*Y\backslash 0. \end{aligned} \tag{2.28}$$

Whether the compose is an FIO is yet to be investigated.

Using the parametrizations of Λ_F in (2.8) and Λ_L in (2.13), the compose (2.28) can be evaluated through solving a system of equations, the first n being trivial fixing the scattering point $x_0 = x$, the second n equating the cotangent vectors

$$\underbrace{\tau_2 \partial_x T(x_0, \alpha, \beta)}_{\boldsymbol{\xi}(x,\alpha,\beta,\tau_2)} = \underbrace{\tau_1 \partial_x T^{(m)}(x, s, r) - \langle \varepsilon, \partial_x \boldsymbol{e}^{(m)}(x, s, r)\rangle}_{\boldsymbol{\xi}^{(m)}(x,s,r,\tau_1,\varepsilon)}. \tag{2.29}$$

Given $\boldsymbol{e}(x, \alpha, \beta) = e$ ($n-1$ constraints) these constitute n equations with the $2n-1$ unknowns (α, β, τ_2). (On $D^{(m)}$ the constraints on e can be invoked on s, r instead, viz. via the inverse of the map $(x, \alpha, \beta) \mapsto (x, s, r)$ as before.)

Lemma 2.8.1. *With Assumptions 2 and 3 the composition* FL_U *yields a smooth family of FIOs parametrized by* e. *Their canonical relations are given by*

$$\Lambda_C = \Lambda_F \circ U_L = \{(s_2, r_2, t_2, \sigma_2, \rho_2, \tau_2; s_1, r_1, t_1, \sigma_1, \rho_1, \tau_1)\}$$

parameterized by $(x_0, \alpha, s_1, \tau_1, \varepsilon)$, *where upon substituting* $x = x_0$ *and once* r_1 *is obtained from* s_1 *through the value* e *of* $\boldsymbol{e}^{(m)}$ *(which mapping is defined below equation (2.11)),* $(s_1, r_1, t_1, \sigma_1, \rho_1)$ *are given in (2.13), and, given* (α, ε), $(s_2, r_2, t_2, \sigma_2, \rho_2)$ *are given in Theorem 2.3.1 in which* (β, τ_2) *are obtained by solving (2.29).*

Proof. First we extend the operator F to act on distributions in $\mathcal{E}'(X \times E)$ by assuming that the action does not depend on $e \in E$. The calculus of FIOs gives sufficient conditions that the composition of two FIOs, here F and L_U, is again an FIO. The essential condition

resulting operator is what seismologists call the AMO 'impulse response'. The composition $F\,L_U$ has been addressed in Lemma 2.8.1. The general restriction has been addressed in Section 2.5. Here we combine these results in the following

Theorem 2.8.4. *With Assumptions 2, 3 and 4 with $Y^c = Y_A$, the composition $\mathcal{R}^1_A F\,L_U$ yields a smooth family of FIOs parametrized by e. The resulting operator is called Azimuth MoveOut.*

The following Bolker-like condition ensures that the restriction to common azimuth is 'image preserving'. Let Λ_A denote the canonical relation of $\mathcal{R}^1_A F$ in accordance with the analysis of Section 2.5,

Assumption 7. *The projection*

$$\pi_{Y_A} \;:\; \Lambda_A \to T^*Y_A\backslash 0$$

is an embedding.

This assumption is most easily verified whether an element in $T^*Y_A\backslash 0$ uniquely determines an element in $T^*X\backslash 0$ smoothly given the medium c_0.

Using 'all' the data (when available), integration over the $(n-1)$ dimensional e removes the artifacts under the Bolker condition, Assumption 2: We obtain the transformation to common azimuth (TCA)

Corollary 2.8.5. *Let $\langle N^{-1}\rangle$ denote the regularized inverse of the normal operator in Theorem 2.3.2. With Assumptions 1, 2 and 4 (with $\Sigma^c = \Sigma_A$), the composition $\mathcal{R}^1_A F\langle N^{-1}\rangle F^* = \int \mathrm{d}e \mathcal{R}^1_A F\langle N^{-1}\rangle L$ is an FIO, $\mathcal{D}'(Y) \to \mathcal{D}'(Y_A)$. With Assumption 7 the reduced dataset is image preserving.*

The proof follows that of Theorem 2.3.2 closely (see (Stolk & de Hoop, 2002, Thm. 4.5)).

2.9 Examples

We give a seismologists' perspective on Dip and Azimuth MoveOut. We illustrate their perspective in the constant coefficient c_0 case. This is the common case where the associated transformations are applied. In this paper, however, we have established the methodology to honor the heterogeneity in the subsurface.

The smooth background coefficient function c_0 is called the (seismic) velocity model and characterizes the speed at which waves travel through the medium. Invoking Cartesian coordinates, the acquisition manifold is obtained by setting the nth coordinate of s and r to zero. Then O_s and O_r are open subsets of a plane hypersurface. In seismology, the midpoint in this hypersurface is defined as $y = \frac{1}{2}(s+r)$ and the offset is defined as $h = \frac{1}{2}(r-s)$. In some sense, the midpoint is associated with the direction of ξ while the offset is a particular choice for e. Here, we assume that c_0 is constant. We will illustrate both DMO and AMO, i.e. the singular supports of their respective kernels. In this section, we will highlight the transition from a parametrization including (y,h,t) to a parametrization including (s,e,t) where e relates to the scattering angles.

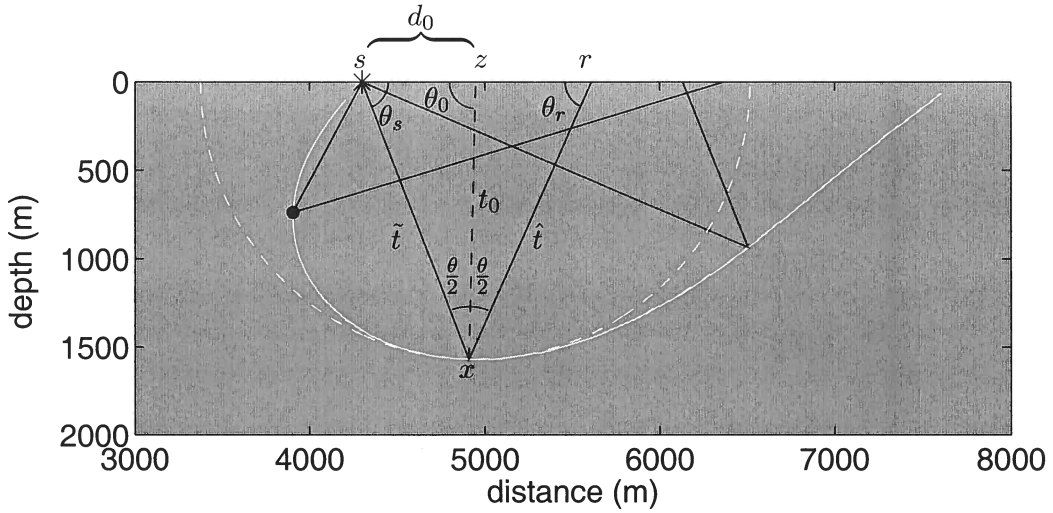


Figure 2.6. Notation for the derivation of the constant medium impulse response (cf. Appendix A). Black lines are rays, the dashed white curve is the zero-offset isochron, the solid white curve is the e 'isochron' and the black dot is the location of the scattering point for the rays shown in Figure 2.5.

that the shape of the finite e 'isochron' differs from the one of the finite h isochron, but that the shapes of exploding reflector model isochrons are the same.

Figure 2.7 shows the 'isochron' in the (s, e, t) parametrization for different values of e. All isochrons, except the exploding reflector one, have two points in common. One of these is the point at which the source ray travels for one time sample less than the full (fixed) time before the ray is scattered and returns to the acquisition surface; the other is the source point.

The impulse response of the DMO operator is the zero-offset traveltime t_0 and the distance d_0 from the source (s) to the exploding reflector source/receiver position (z) both as a function of the direction (θ_s) of the ray at the source (related to σ); all other parameters are fixed. In Figure 2.8 we plot these functions parameterically against one another. They are derived in Appendix A.

Azimuth MoveOut

The azimuth in Azimuth MoveOut Biondi *et al.* (1998) is the polar angle associated with the two-dimensional offset ($n = 3$) in the acquisition manifold. As the key parameter, we will employ the azimuthal angle in e rather than azimuth in h.

The composition of canonical relations that determines the AMO canonical relation, implies the 'matching' in (x, ξ) of two 'isochrons', one associated with the imaging operator L_U and one associated with the modeling operator F. The points at which these two 'isochrons' touch and share the same (co)tangent plane are the points which contribute to

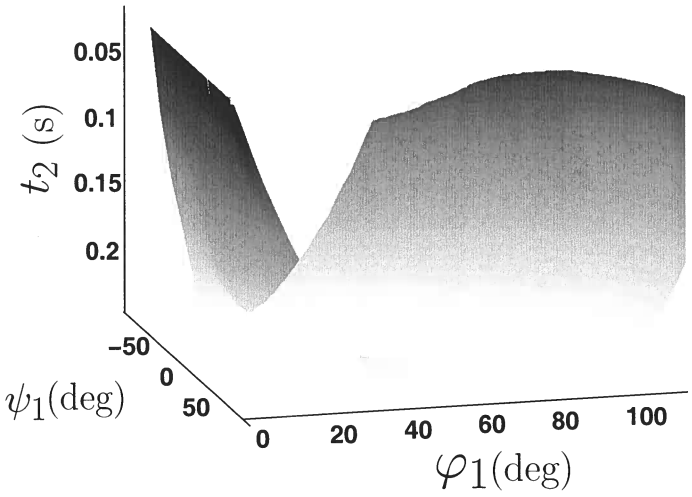


Figure 2.9. Surface plot of the AMO impulse response as a function of input ray directions.

the canonical relation of the AMO operator.

The impulse response of the AMO operator is the traveltime t_2 as a function of the direction $((\theta_s, \psi_s))$ of the ray at the source (related to σ); all other parameters are fixed. In Figure 2.9 we plot this function. The expression is derived in Appendix B.

Chapter 3

The applicability of DMO/AMO in the presence of caustics[1]

3.1 Summary

Reflection seismic data continuation is the computation of data at source and receiver locations that differ from those in the original data, using what data are available. We develop a general theory of data continuation in the presence of caustics and illustrate it with three examples: dip moveout (DMO), azimuth moveout (AMO), and offset continuation. This theory does not require knowledge of the reflector positions. We construct the output data set from the input through the composition of three operators: an imaging operator, a modeling operator, and a restriction operator. This results in a single operator that maps directly from the input data to the desired output data. We use the calculus of Fourier integral operators to develop this theory in the presence of caustics. For both DMO and AMO, we compute impulse responses in a constant-velocity model and in a more complicated model in which caustics arise. This analysis reveals errors that can be introduced by assuming, for example, a model with a constant vertical velocity-gradient when the true model is laterally heterogeneous. Data continuation uses as input a subset (common-offset, common-angle) of the available data, which may introduce artifacts in the continued data. One could suppress these artifacts by stacking over a neighborhood of input data (using a small range of offsets, or angles, for example). We test data continuation on synthetic data from a model known to generate imaging artifacts. We show that stacking over input scattering angles suppresses artifacts in the continued data.

3.2 Introduction

Data collected in the field are often not ideal for processing. For example, zero-offset data are important in seismic data processing but limitations preclude collecting such data in the field. In general, we refer to methods to remedy this problem as *data continuation* or *data mapping*. Stolt (2002) gives an excellent description of why data continuation is necessary, as well as a theory for performing data mapping with a constant-background-velocity model. Patch (2002) gives an example of data continuation in medical imaging.

[1]This chapter has been published, along with Appendices C, D, E, and F as:
Malcolm, A. E., de Hoop, M. V. and Le Rousseau, J. H. 2005. The applicability of DMO/AMO in the presence of caustics. *Geophysics* **70 S1-S17.**

Data continuation is not a stand-alone process but rather exists within a larger framework. Imaging-inversion, migration-velocity-analysis and 'offset' continuation are closely dependent on one another; the operators that connect the three processes are the *annihilators* of the seismic data (Stolk & de Hoop, 2002; Brandsberg-Dahl *et al.*, 2003b).

We first give an outline of how the continuation operators are constructed and then show three examples: DMO, AMO and offset continuation in shot records. To illustrate the properties of the first two operators we compute their impulse responses in both a constant-velocity model and a gas lens model. For the offset continuation section, we demonstrate the applicability of our theory by filling in missing offsets in synthetic shot records. We include three appendices with derivations of certain parts of this theory in constant-velocity media. Appendix A contains the derivation of the impulse responses in closed form. In Appendix B we derive an expression for the amplitudes of the data continuation operator. This result could also be applied in varying velocity models by using the root-mean-square velocity as a local approximation to the true velocity. In Appendix C we show that if the input and output data configurations are the same, the operator does not change the positions of reflections in the data.

3.3 Imaging-Modeling-Restriction

Data continuation is developed from the composition of three processes: imaging/migration, modeling, and restriction. By *composition* we mean applying one operator after another and simplifying the result, by integrating over intermediate phase variables, into a single operator. The continuation process is a single operation that computes data for the desired acquisition geometry from the original data. This is similar to the derivation of AMO (Biondi *et al.*, 1998), in which the observed data are first migrated to form an image of the subsurface (imaging), from which data with another acquisition geometry are modeled (modeling). The output data are assumed to have a certain acquisition azimuth, which is accomplished with the use of a *restriction operator*.

Composing the operators, given the acquisition surface, requires that we impose certain conditions on the velocity model (e.g., removing *grazing caustics*, caustics that are both at and parallel to the acquisition surface). In our scheme, the first operator, the imaging operator, is a Generalized Radon Transform [GRT; Miller *et al.* (1987) and de Hoop & Brandsberg-Dahl (2000)]. The second operator models the data for all possible (continuous) source and receiver positions at the surface, using an image of the subsurface as input. The final operator restricts this modeling operator so that only the desired output data are modeled, rather than the complete data set. Composing the three operators yields a single operator that computes the desired output data directly from the input data. This composition is done in Appendix B, for a general continuation operator, in a constant-velocity (or mildly depth varying) model. All three operators are Fourier integral operators (FIO)[2].

[2]Fourier integral operators, which generalize Fourier integrals, can be used to solve wave equations. These operators are characterized by an amplitude and a phase, but the phase is not necessarily linear in the space variables as it is in Fourier transforms. Stationary phase analysis yields a leading-order asymptotic

the final operator. In some cases, for example 3D DMO in constant media (Bleistein *et al.*, 2000, p. 326), additional phase variables must remain in the final operator, as explained in Appendix A of de Hoop *et al.* (2003a).

3.4 DMO

For DMO, finite-offset data are used in the imaging step, after which (normal-incidence) zero-offset data are modeled. When the zero-offset modeling is done by restricting a multiple-offset modeling operator to zero-offset, data from non-normal incidence data will also be modeled. We construct an exploding-reflector (Loewenthal *et al.*, 1976; Claerbout, 1985) modeling operator and use it in place of the restricted modeling operator. The exploding-reflector modeling operator models only a single ray from the scattering point, to the surface point. (From this point onward we will refer to this, single, ray as the zero-offset ray and the associated data as zero-offset data.) This means that for DMO the composition of three operators (imaging, modeling, and restriction) is reduced to a composition of only two (imaging and exploding-reflector modeling).

Using the exploding-reflector modeling operator rather than the full zero-offset modeling operator results in a different amplitude for the final transformation to zero-offset (TZO). A partial explanation for this is that the exploding-reflector operator models the data along a single ray as if the wave speed of the medium were half its true value, while the zero-offset modeling operator models two rays (up- and down-going) in the true medium, with coincident surface and subsurface positions. The amplitude of the exploding-reflector modeling operator differs from that of the restricted multiple-offset modeling operator (RMO); the RMO models a ray that travels from the surface to the subsurface applying geometrical spreading, followed by modeling a ray from the subsurface to the surface applying geometrical spreading once again (thus resulting in the geometrical spreading squared). In contrast, the exploding-reflector modeling operator considers the geometrical spreading for just one ray with half the wave speed along the path.

3.4.1 Homogeneous model

For any operator, we can construct a table relating the input parameters to the output parameters. In migration, for example, this table would relate a point in the subsurface (the output of migration) to the source position, receiver position and two-way travel time (the input to migration). The table can be parameterized in different ways. For example, midpoint and offset, or source position and scattering angle can be used in place of the source and receiver positions. Fixing different sets of these parameters allows us to plot cross-sections of this table. (We use the term cross-section in analogy with the cross-sections of a function; a cross-section of a function of two variables is the function values as a function of one variable with the other variable held fixed.) As a first example, we plot the standard migration ellipse in Figure 3.1; in this case the midpoint, offset and traveltime are fixed giving us a particular cross-section of the table. Similarly we can fix the source point,s, scattering angle, θ, and traveltime giving what we will call an *angle isochron.* An example

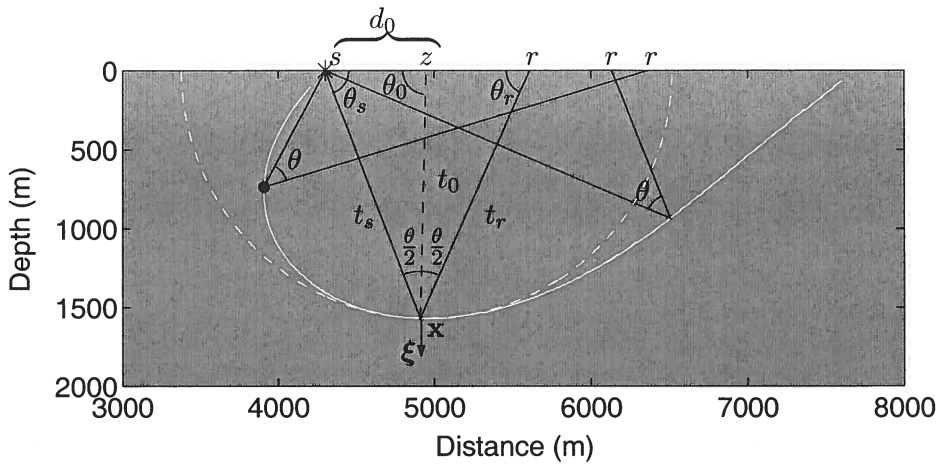


Figure 3.2. Angle DMO, in a constant-velocity medium ($v = 1.7$ km/s). The black lines are rays; the dashed white line is the exploding-reflector isochron; the solid white line is the angle isochron. The black dot is the scattering point from Figure 3.1; this point will have a different contribution to the operator illustrated in this figure than it does to the operator illustrated in Figure 3.1. The angle θ_0, between the zero-offset ray and the surface is variable; it depends on the particular value of the scattering angle. This illustration shows θ_0 close to 90°, which is not generally the case. The notation shown in this figure is used throughout the text and in Appendix A, where the impulse response is computed.

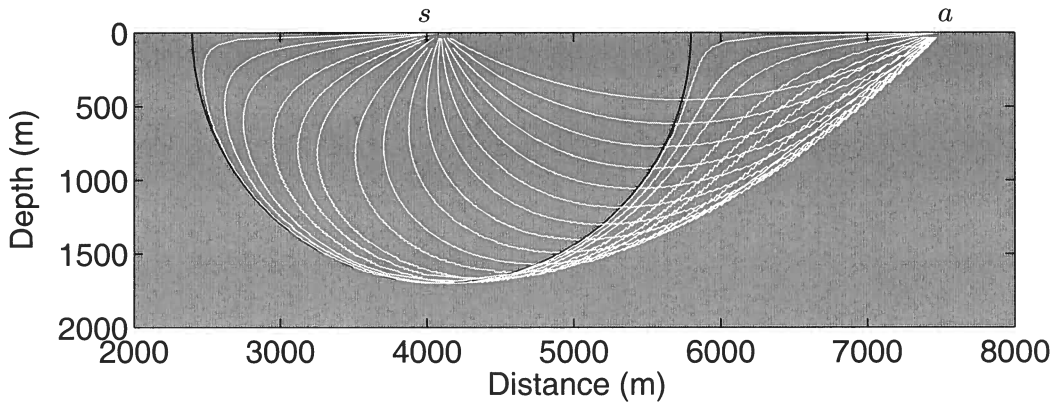


Figure 3.3. Angle isochrons in a constant-velocity medium ($v = 1.7$ km/s), for various scattering angles. All other parameters are the same as in Figure 3.2. The smaller the scattering angle, the deeper the isochron penetrates into the model. The angle isochron for the smallest scattering angle shown, 2.86°, is the closest to the exploding-reflector isochron (black circle), and the angle isochron for the largest scattering angle, 148.97°, is the shallowest line. This plot shows several cross-sections of the migration table described in the text.

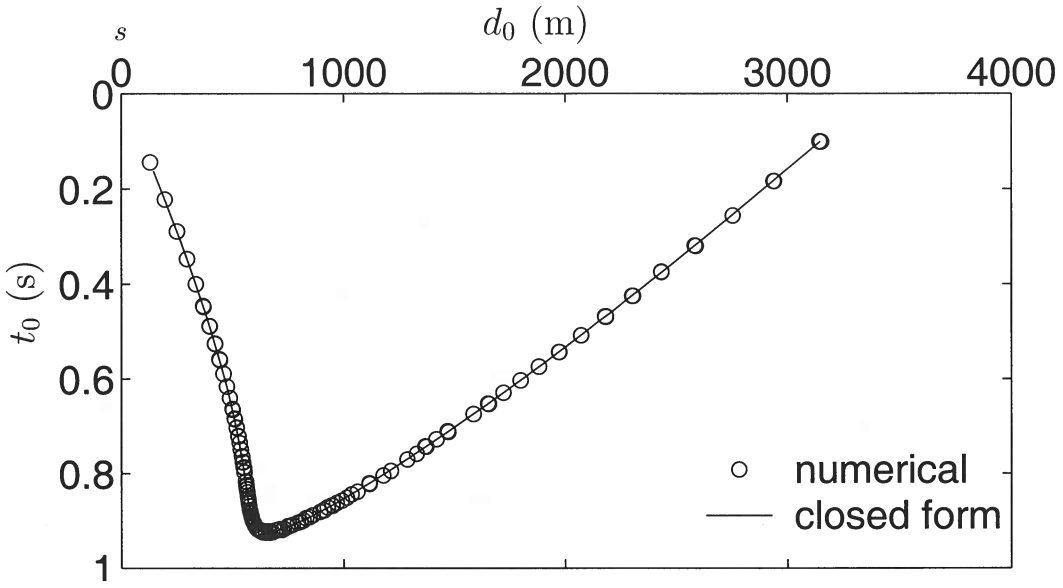


Figure 3.5. DMO impulse response for a constant-velocity medium, for scattering angle 45°, velocity 1.7 km/s, $T = 2$ s. The spacing of the points is constant in take-off angle at the source ray (increments of 5.73°). Areas of the impulse response with denser sampling, which are regions of high curvature, can be expected to have relatively high associated amplitude.

To compute the impulse response in the angle domain (using the notation of Figure 3.2), we fix the scattering angle θ, the source position s, and the traveltime $T = t_s + t_r$, and compute the one-way zero-offset traveltime t_0 as a function of the distance, d_0, from the source position s to the zero-offset source/receiver position z. In a constant-velocity medium, this impulse response can be computed in closed form. This is done in Appendix A. Figure 3.5 shows the impulse response for angle DMO, as derived in Appendix A. This impulse response is the zero-offset traveltime as a function of the distance from the source position to the zero-offset position (the distance d_0 in Figure 3.2). The solid line is the closed-form solution, and the points are computed numerically with ray tracing; the spacing between the points is constant (increment 5.73°) in the take-off angle θ_s (defined in Figure 3.2).

3.4.2 Gas-lens model

The gas-lens model consists of a vertical velocity gradient (0.45 s^{-1}) beginning at 1600 m/s, with a low-velocity circular lens with Gaussian parameter variations (maximum velocity contrast 800 m/s) located at lateral position 4600 m, and depth 600 m, with a diameter (Gaussian standard deviation) of 600 m. This model, introduced by Brandsberg-Dahl *et al.* (2003a), is based on a feature in the BP Valhall field. Throughout this subsection, we compare results for this model with those of a constant-velocity-gradient model (the same as the lens model but without the lens). The background shading of Figures 3.6 and 3.7 depicts the velocity model, with darker shading indicating higher velocity.

In Figure 3.6, we show the relationship between the shape of the angle isochron in the lens model (solid line) and that in the constant-gradient model (dashed line). The

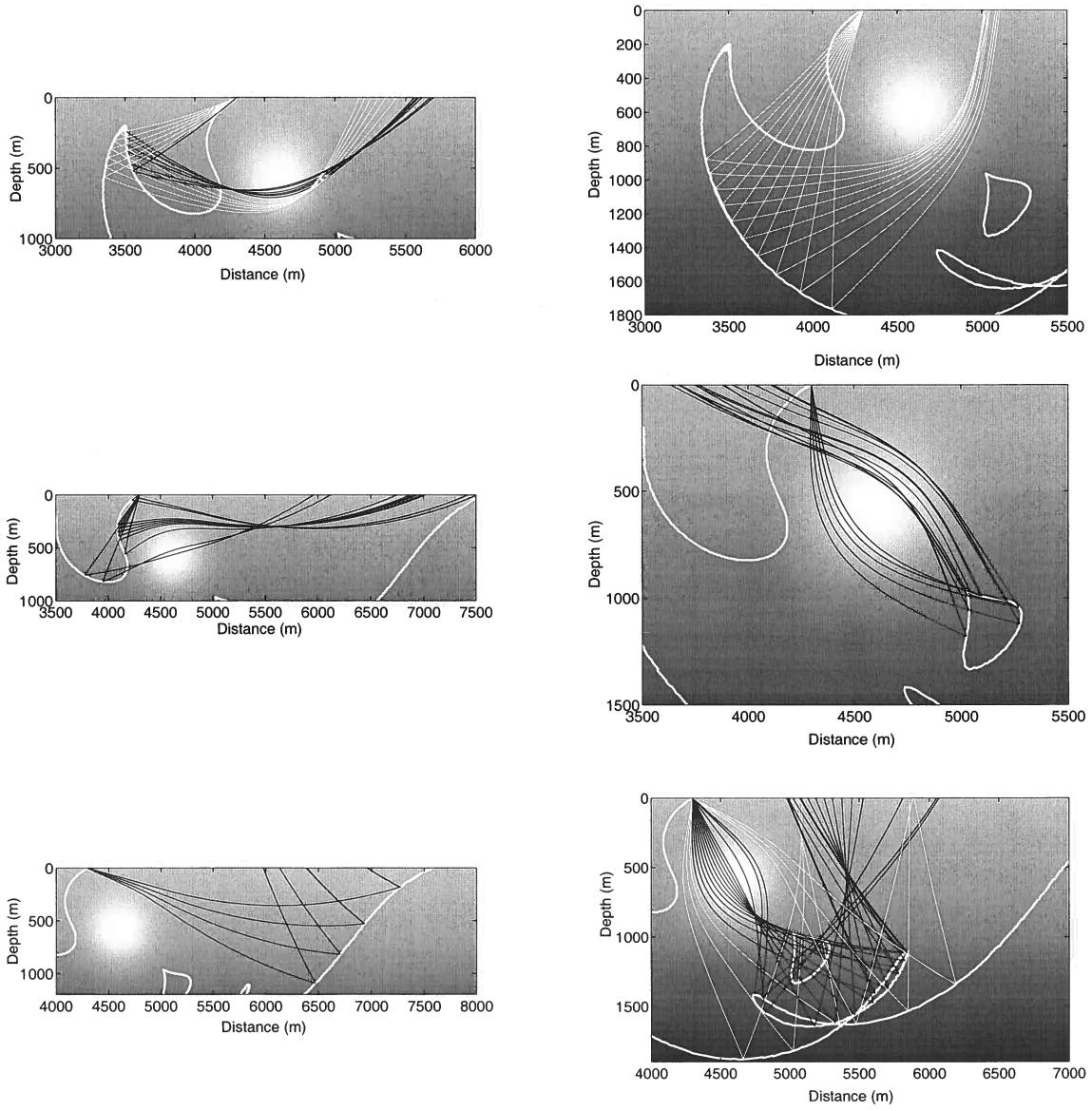


Figure 3.7. Raypaths in the lens model for scattering angle fixed at 45°. This figure illustrates the origin of the complicated structure of the angle isochron by showing the rays that were used to compute it. Where both white and black rays are present, white rays come from the main branch while black rays come from inner (closer to lens) branches. Each panel shows a different region of the same angle isochron.

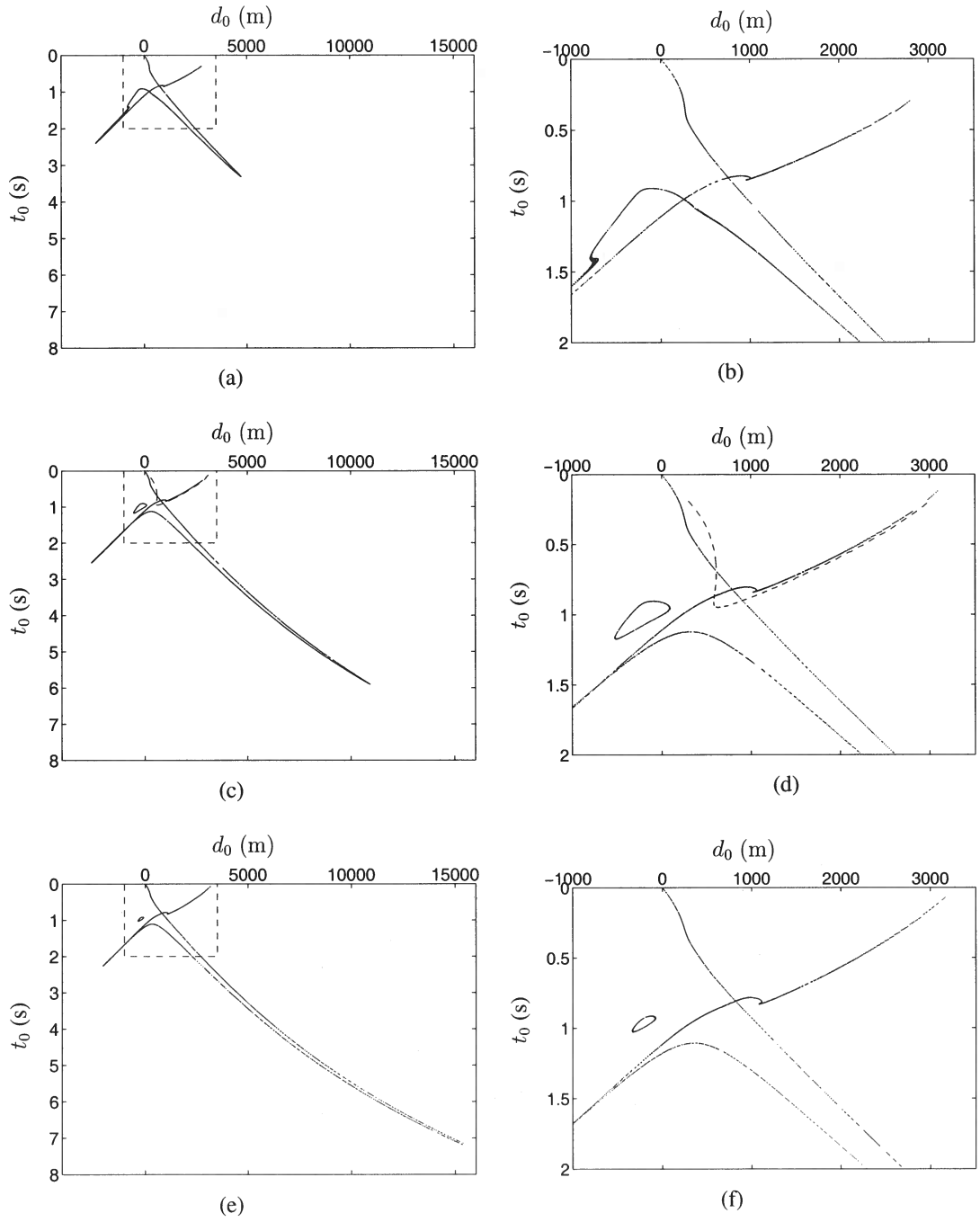


Figure 3.8. Angle DMO impulse responses for the lens model for scattering angles of 40° (a and b), 45° (c and d), and 50° (e and f). The dashed box in a, c, and e outlines the region shown in b, d, and f. The dashed curves in c and d are the impulse responses computed in the constant vertical velocity gradient that makes up the background of the lens model.

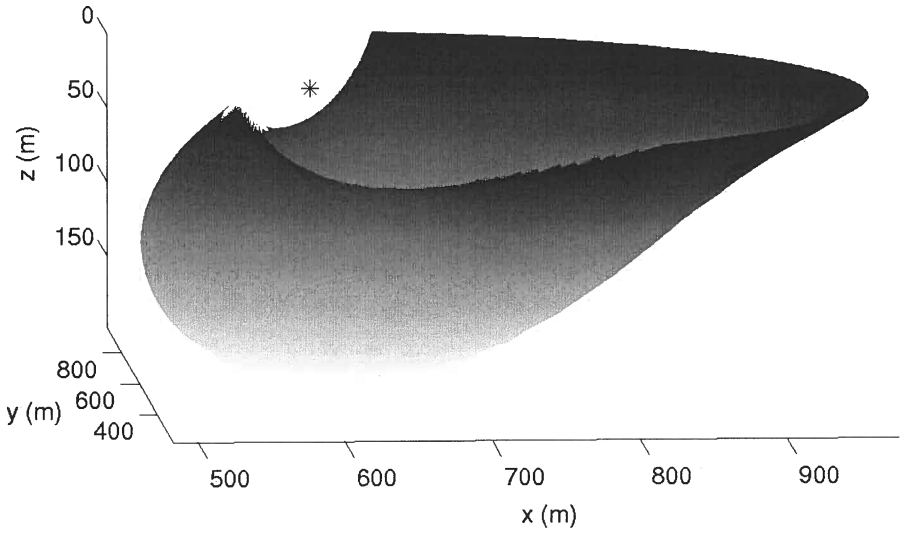


Figure 3.10. 3D angle isochron in a constant-velocity medium, the asterisk indicates the position of the source. This extends to three dimensions the angle isochron shown in 2D in Figure 3.2.

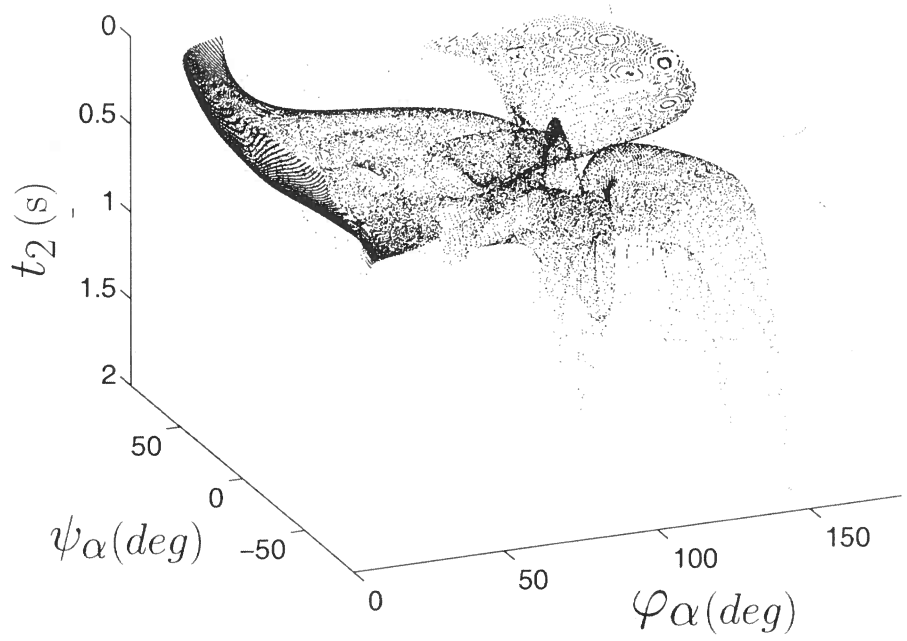


Figure 3.13. AMO impulse response for the lens model. Compared with the impulse response for DMO (2D case) in Figure 3.8, the 3D nature of AMO introduces further complications in this impulse response.

spherical lens with Gaussian parameter variations of 100 m diameter (standard deviation of Gaussian) with center at $x_1 = 600$ m, $x_2 = 600$ m, $x_3 = 100$ m and a maximum velocity contrast of 800 m/s. The model extends to $x_1 = 2400$ m, $x_2 = 2400$ m, $x_3 = 800$ m with an initial velocity of 1600 m/s and a vertical gradient of 0.45 s^{-1}.

The angle isochron, shown in Figure 3.12, exhibits a gross shape similar to that in the constant-velocity case (Figure 3.10). Just as with the 2D angle isochron (Figure 3.6), the 3D version has a complicated shape, in which we recognize the presence of caustics. Figure 3.13 shows the AMO impulse response, which bears little resemblance to that for the constant-velocity model (Figure 3.11). The complications in the AMO impulse response and angle isochrons for the lens model are significantly greater for the 3D problem than for the 2D one.

3.6 Synthetic Data Example

We show an example of data continuation. The example uses the same 2D lens model used for the DMO impulse response calculations, with the addition of a reflector from

is based on the escape equations developed by Sethian (2002).

Figure 3.15 shows the results of this procedure applied to the data in three synthetic shot records for the model in Figure 3.14. The top panels show a shot record reconstructed away from the influence of the lens. At this position the event is filled in correctly. Because the event is so simple at this point, interpolation would work just as well. The second panels down show the reconstruction of a shot record on the edge of the lens. Traces with events from both the dipping and flat portions of the reflector are missing, making the reconstruction more difficult. The reconstruction is reasonably successful although there are some amplitude errors. The next panel down shows a shot record for the shot right over the lens. We have removed the traces containing the caustic and are attempting to fill them in with the algorithm. Again the reconstruction is not perfect but it is considerably better than what would have been achieved using conventional interpolation schemes. The data shown have been bandpass filtered to match the frequency content of the original data. We expect some amplitude inaccuracies because the amplitudes are obtained from equation (E.27) which are for a constant velocity model, whereas this model is clearly more complicated. In these examples we have corrected for only the obliquity factor ($|\cos\theta_i|^2$).

The single shot record in Figure 3.16 shows the same computation as above it with a range of only 2° in scattering angle for the input data (compared with the section above it which uses all the available data). A particularly high amplitude artifact appears just below the latest true event in this section. This example can be generated much more quickly than the other images because so few input angles are used. The other panels most likely use more angles than necessary and so the optimal data quality and computational time trade-off is probably somewhere in between the two.

Finally, to mimic the DMO discussion above, we show in Figure 3.17 the exploding reflector data computed from the original data set (with offsets from -6 to 7 km) while removing offsets from -100 m to 100 m. The smallest offset data (offset of 6 m) from the true data set are shown for comparison. This illustrates the difference between zero-offset modeling and exploding reflector modeling.

3.7 Conclusions

We have described and illustrated a method for source-receiver continuation of seismic data in the presence of caustics. In the absence of caustics the method reduces to a form of offset continuation. The computational complexity of this method depends on the geologic complexity and varies in space as the geology changes. In the most complex situation a table must be constructed relating the subsurface and surface parameters, and a search performed in this table; the computational complexity depends on the algorithms used to perform these steps. Both DMO and AMO are examples of data continuation that can be obtained from the continuation framework presented here. It is possible to continue data only when the subsurface point and reflector orientation generating the data point are sampled in the original data set. Thus, the issue of illumination is directly related to data continuation. Similarly, data continuation provides a framework for estimating the smallest necessary data set, by using a combination of acquisition and data continuation

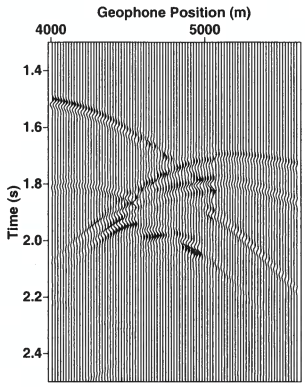


Figure 3.16. This extra section shown at the bottom illustrates the presence of artifacts when insufficient stacking is done over the input scattering angle. It is the shot record at $s = 4560$ m, with only a 2° range of input scattering angle.

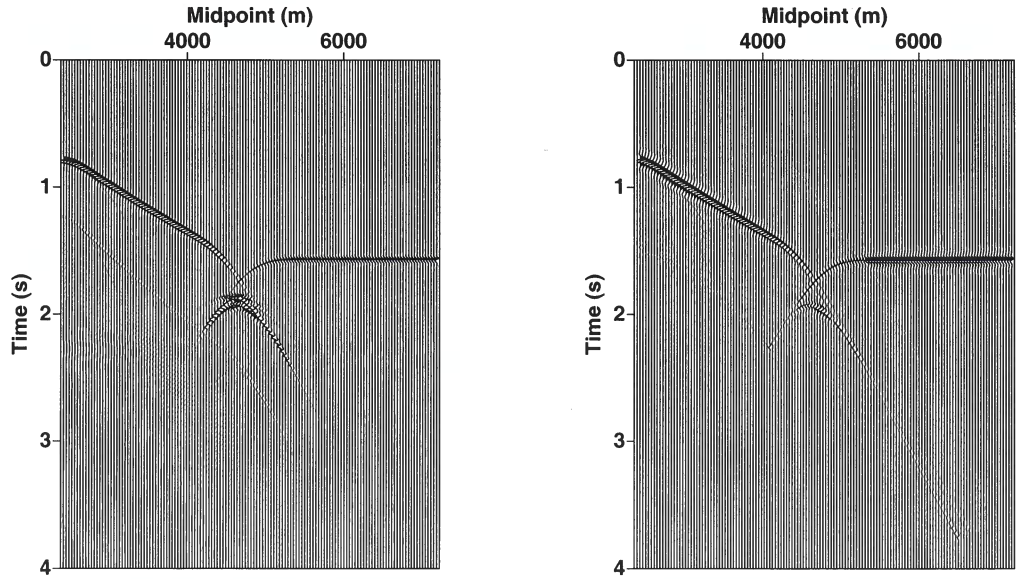


Figure 3.17. The left panel shows the smallest offset available in the synthetic data set (6 m offset). The right panel is the exploding reflector data computed from the portion of the synthetic data set with offsets between -100 m and 100 m removed.

Chapter 4

A method for inverse scattering based on the generalized Bremmer coupling series[1]

4.1 Summary

Imaging with seismic data is typically done under the assumption of single scattering. Here we formulate a theory that includes multiply scattered waves in the imaging process. We develop both a forward and an inverse scattering series derived from the Lippmann-Schwinger equation and the Bremmer coupling series. We estimate leading-order internal multiples explicitly using the third term of the forward series. From the inverse series, two images are constructed, one formed with all the data, the other with the estimated leading-order internal multiples; the final image is formed from the difference of these two images. We combine the modeling of the leading-order internal multiples with the construction of the second image resulting in one two-part imaging step.

4.2 Introduction

A seismic experiment is typically modeled as a set of sources at the Earth's surface that generate waves that are reflected once from medium discontinuities in the subsurface and recorded at a set of receivers again located on the surface. The goal of this paper is to move beyond the single-reflection assumption to allow for multiply scattered waves. We consider only scalar waves and assume that the sources and receivers are on the same horizontal surface. A finite collection of scatterers with a separation large compared to the wavelength is also assumed, so that each reflection may be treated separately.

Fokkema & van den Berg (1993) developed a rigorous theory for the suppression of surface-related multiples. A surface-related multiple is a wave that has been reflected at least three times, with at least one reflection at the surface. Their analysis is derived from the reciprocity theorem in integral form and results in a Neumann series representation to predict surface-related multiples. If assumptions allowing the construction of data at zero-offset, such as those given by de Hoop *et al.* (2003a) are satisfied, then, in theory, Fokkema and van den Berg's theory solves the surface-related multiple attenuation problem. This

[1]This chapter has been accepted, along with Appendices J and K to Inverse Problems as:
Malcolm, A. E. and de Hoop, M. V. A method for inverse scattering based on the generalized Bremmer coupling series.

multiples (Fokkema *et al.*, 1994; Berkhout & Verschuur, 1997; Verschuur & Berkhout, 1997; van Borselen, 2002). In these methods, a particular layer is identified as the multiple generator (i.e. the layer where the second reflection occurs) and the surface-related multiple attenuation is adapted to be applied at that layer. Dragoset and Jeričević give a practical algorithm for attenuating surface-related multiples in (Dragoset & Jericević, 1998); an algorithm such as that discussed by Dragoset and Jeričević could be used for internal multiples in any of the mentioned extensions. Weglein and others (Weglein *et al.*, 1997) have used the Lippmann-Schwinger series to model and process seismic data, including the suppression of both surface-related and internal multiples, without knowledge of the velocity model. In ten Kroode (2002) the mathematical theory behind that approach is given in both one and two dimensions. He shows that internal multiples can be estimated without knowledge of the velocity model if the velocity model satisfies two conditions: ten Kroode's traveltime monotonicity assumption (this condition is described in Appendix K), and the condition that the wavefield contains no caustics. When the two assumptions of ten Kroode are satisfied, our method can be rewritten in a form consistent with the method of Weglein *et al.* (1997); this is discussed further in Appendix K. Jakubowicz (1998) proposes a method for modeling internal multiples by correlating one primary reflection with the convolution of two other primary reflections; his approach implicitly uses the Bremmer series and is similar to the work presented here under ten Kroode's traveltime monotonicity assumption. Kelamis *et al.* (2002) use an approach similar to that of Jakubowicz, in which the multiples are constructed from a combination of different data sets, both at the surface and in the subsurface. In any method that predicts internal multiples and subtracts them, an adaptive subtraction technique such as that suggested by Guitton & Verschuur (2004) must be used.

Aside from reflection seismology, there are other applications in which multiply scattered waves are important. In earthquake seismology, Burdick & Orcutt (1979) investigate the truncation of the generalized ray sum, from which they find earth models in which the inclusion of internal multiples becomes important. Revenaugh & Jordan (1987) observe both internal and surface-related multiples and use them to estimate the attenuation quality factor, Q, of the mantle. In (Revenaugh & Jordan, 1989, 1991), the same authors use multiples to investigate layering in the mantle. Bostock *et al.* (2001) use incident teleseismic P-waves scattered from a free surface and then subsurface structure before being recorded in an inversion scheme in which the teleseismic P-wave coda is used to invert for subsurface structure. For synthetic aperture radar (SAR) data, Cheney & Borden (2002) derive a theory to relate the singular structure (wavefront set) of the object to the singular structure of the multiply scattered data.

In the next section we describe the techniques of the directional decomposition used in the Bremmer series. In the third section, we describe some of the details of the construction of one-way Green functions. This is followed by a description of the contrast-source method used for the Lippmann-Schwinger series. In the fifth section, we construct the hybrid series. In the sixth section we use the hybrid series to model data, giving the first of our three main results in (4.84). The proof of this result is given in Appendix J. Following this, we summarize a method of constructing an inverse to the modeling operator. We then describe, through a series of results in Section 4.9, a method to estimate artifacts in the

of pseudodifferential operator matrices $\mathsf{Q}(z)$ such that microlocally,

$$U = \begin{pmatrix} u_+ \\ u_- \end{pmatrix} = \mathsf{Q}(z)D\,, \qquad X = \begin{pmatrix} f_+ \\ f_- \end{pmatrix} = \mathsf{Q}(z)M\,. \tag{4.5}$$

The diagonalization procedure requires that cut-offs be applied to U to remove constituents of the wavefield that propagate anywhere horizontal; these cut-offs are described in the following section. We have omitted any indication that these cut-offs are not present in this section to keep the notation in this section consistent with the notation in the remainder of the paper, in which the cut-offs are assumed to have been applied. In this notation, $u_\pm$ satisfy the one-way wave equations

$$\left(I\partial_z + \mathsf{Q}(z)\partial_z\mathsf{Q}^{-1}(z) - \mathsf{B}\right)U = X\,, \tag{4.6}$$

where I is the identity matrix, introducing

$$\mathsf{B} = \mathsf{Q}(z)\mathsf{A}\mathsf{Q}^{-1}(z) = \begin{pmatrix} \mathrm{i}B_+(z,x,D_x,D_t) & 0 \\ 0 & \mathrm{i}B_-(z,x,D_x,D_t) \end{pmatrix}, \tag{4.7}$$

where $B_\pm$ has principal symbol $b_\pm(z,x,\xi,\tau) = \pm\tau\sqrt{c(z,x)^{-2} - \tau^{-2}\|\xi\|^2} = \pm b(z,x,\xi,\tau)$, which corresponds with k_z in the seismological notation.

With the conventions used here, u_+ represents downward propagating waves and u_- represents upward propagating waves. (As is standard in geophysics, we have chosen the positive z-axis downward.) The columns of the Q operator matrix are an operator generalization of eigenvectors and we are free to choose their normalization in the operator sense. We choose the vertical power flux normalization of de Hoop (1996) so as to make $B_\pm$ in (4.6) self-adjoint (the normalization changes the sub-principal part of the operator). In this normalization, the composition and decomposition operators are

$$\mathsf{Q} = \frac{1}{2}\begin{pmatrix} (Q_+^*)^{-1} & -\mathcal{H}Q_+ \\ (Q_-^*)^{-1} & \mathcal{H}Q_- \end{pmatrix}, \qquad \mathsf{Q}^{-1} = \begin{pmatrix} Q_+^* & Q_-^* \\ \mathcal{H}Q_+^{-1} & -\mathcal{H}Q_-^{-1} \end{pmatrix}, \tag{4.8}$$

where * denotes the adjoint of the operator, $\mathcal{H}$ is the Hilbert transform in time, and the principal symbol of both $Q_\pm$ is given by $\left(\frac{\tau^2}{c(z,x)^2} - \|\xi\|^2\right)^{-1/4}$. From these expressions we find that

$$u = Q_+^*u_+ + Q_-^*u_-, \text{ and } f_\pm = \pm\tfrac{1}{2}\mathcal{H}Q_\pm f. \tag{4.9}$$

The $Q_\pm$ operators act in the time variable as time convolutions.

In the flux normalization, the term $\mathsf{Q}^{-1}\partial_z\mathsf{Q}$ in (4.6) is of lower-order in the singularities (i.e. the operator is smoothing in comparison with other terms), thus we omit it. If required, its contribution can be accounted for by including it in the B matrix. We introduce the propagators for the one-way wave equations (4.6) as

$$(\partial_z + \mathsf{B})\,\mathsf{L} = I\delta, \quad \mathsf{L} = \begin{pmatrix} G_+ & 0 \\ 0 & G_- \end{pmatrix}. \tag{4.10}$$

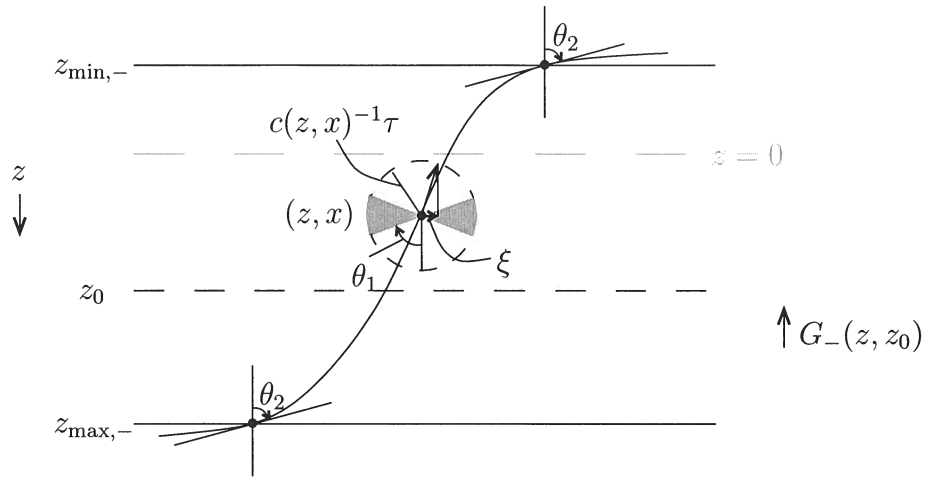


Figure 4.1. Removing horizontal propagations. The symbol of the cut-off operator ψ is one up to an angle θ_1 and then decays smoothly to zero at the angle θ_2. This removes all propagation at angles larger than θ_2, i.e., the region within the gray wedges.

In phase space, we introduce the set

$$I_\theta = \{(z,x,t,\zeta,\xi,\tau) \,|\, \arcsin(c(z,x)\|\tau^{-1}\xi\|) < \theta, |\zeta| < C|\tau|\}, \tag{4.14}$$

illustrated in Figure 4.2, where C is the maximum slowness. Finally, we construct the sets

$$J_-(z_0,\theta) = \{(z,x,t,\zeta,\xi,\tau) \in I_\theta \,|\, \tau^{-1}\zeta < 0 \text{ and } z_{\max-}(z,x,\xi,\tau,\theta) \geq z_0\}, \tag{4.15}$$

and

$$J_+(z_0,\theta) = \{(z,x,t,\zeta,\xi,\tau) \in I_\theta \,|\, \tau^{-1}\zeta > 0 \text{ and } z_{\min+}(z,x,\xi,\tau,\theta) \leq z_0\}. \tag{4.16}$$

Figure 4.1 illustrates the set $J_-(z_0,\theta_2)$, considering the shaded region as excluded from the set.

The sets $J_\pm$ encompass the regions of phase space that must be excluded in order to remove horizontally propagating singularities while analyzing $G_\pm(z,z_0)$. To actually remove singularities from these regions, we define a pseudodifferential cutoff $\psi_- = \psi_-(z,z_0,x,D_x,D_t)$ with symbol satisfying

$$\psi_-(z,x,\xi,\tau) \sim 1 \text{ on } J_-(z_0,\theta_1), \tag{4.17}$$

$$\psi_-(z,x,\xi,\tau) \in S^\infty \text{ outside } J_-(z_0,\theta_2), \text{ if } z - z_0 > \delta > 0; \tag{4.18}$$

here $0 < \theta_1 < \theta_2$. Singularities propagating at an angle less than θ_1 are unaffected by the

$G_-^*(z_0, x_0, t_0 - t, z, x)$ follows from

$$\begin{aligned} & \int \mathrm{d}s_0\, \mathrm{d}t_0\, v(z_0, s_0, t_0) \left(\int \mathrm{d}s\, \mathrm{d}t\, G_-(z_0, s_0, t_0 - t, z, s) u(z, s, t) \right) \\ = & \int \mathrm{d}s\, \mathrm{d}t \left(\int \mathrm{d}s_0\, \mathrm{d}t_0\, v(z_0, s_0, t_0) G_-(z_0, s_0, t_0 - t, z, s) \right) u(z, s, t) \\ = & \int \mathrm{d}s\, \mathrm{d}t \left(\int \mathrm{d}s_0\, \mathrm{d}t_0 (G_-(z_0, z))^*(s, t - t_0, s_0) v(z_0, s_0, t_0) \right) u(z, s, t) \end{aligned} \tag{4.21}$$

Using the self-adjoint property of B, $G_-(z_0, s_0, t_0 - t, z, s) = G_+(z, s, t - t_0, z_0, s_0)$, microlocally so that $(G_-(z_0, z))^* = G_+(z, z_0)$. A similar result holds with $+$ and $-$ interchanged. Note that $G_\pm$ are causal.

Remark 4.4.2. The $G_\pm$ propagators obey the reciprocity relation (of the time convolution type)

$$Q_+^*(z) G_+(z, z_0) Q_+(z_0) = -Q_-^*(z_0) G_-(z_0, z) Q_-(z)\,. \tag{4.22}$$

This reciprocity relation is derived from the reciprocity of the full-wave Green propagator.

Remark 4.4.3. We have

$$G_-(z, z') G_-(z', z'') = G_-(z, z'')\,, \tag{4.23}$$

for $z < z' < z''$; this property is known as the semi-group property. The same property holds for G_+.

In the above, we have nowhere assumed the absence of caustics in the wavefield. This section has addressed the necessary assumption that rays are nowhere horizontal: the double-square-root DSR assumption (Stolk & de Hoop, 2004a, Assumption 2).

4.5 Scattering: Contrast source formulation

The Bremmer formulation assumes a degree of smoothness in the velocity model. In the contrast formulation of the Lippmann-Schwinger approach, the velocity, c, is split into a background, c_0, which is here assumed to be smooth (C^∞) and a singular contrast, δc, which is here assumed to be a superposition of conormal distributions. A series is then constructed with terms of increasing order in δc. We use a hybrid of the two approaches; the contrast-source integral equation (Lippmann-Schwinger) subjected to a directional decomposition (Bremmer). We begin with the wave equation in the smooth background and in the true medium respectively

$$(\partial_z + \mathsf{A}_0) D_0 = M, \qquad (\partial_z + \mathsf{A}) D = M, \tag{4.24}$$

where the subscript 0 indicates that an operator is using the smooth background parameters and no subscript indicates an operator acting on the full medium. Subtracting the equation

operators of de Hoop (1996) viz,

$$V = \mathsf{Q}\delta\mathsf{A}D_t^2\mathsf{Q}^{-1} = \tfrac{1}{2}\mathcal{H}\begin{pmatrix} Q_+\ a\ Q_+^* & Q_+\ a\ Q_-^* \\ -Q_-\ a\ Q_+^* & -Q_-\ a\ Q_-^* \end{pmatrix} D_t^2 = \begin{pmatrix} S_{++} & S_{-+} \\ S_{+-} & S_{--} \end{pmatrix} D_t^2 . \quad (4.33)$$

Here, S_{++} and S_{--} are interpreted as transmission operators since they govern scatterings between singularities traveling in the same principal direction before and after scattering. In contrast, S_{+-} and S_{-+} are interpreted as reflection operators because they govern scatterings that result in a change of principal direction; from up-going to down-going and down-going to up-going, respectively.

To simplify the notation, we define

$$P_0 = \partial_z + \mathsf{B}_0 , \quad (4.34)$$

and its forward parametrix,

$$\mathsf{L}_0 = \begin{pmatrix} G_+ & 0 \\ 0 & G_- \end{pmatrix} . \quad (4.35)$$

Recall also that $V = \mathsf{Q}(z)\ \delta\mathsf{A}\ \mathsf{Q}^{-1}(z)$ from (4.33). In this notation, (4.31) reduces to

$$P_0\ \delta U = -VU , \quad (4.36)$$

or

$$\delta U = -\mathsf{L}_0(VU) . \quad (4.37)$$

The V operator is a distributional multiplication along with a second time derivative, whereas L_0 is the forward parametrix of a partial differential operator. Writing $U = U_0 + \delta U$ gives

$$\delta U = -\mathsf{L}_0(VU_0) - \mathsf{L}_0(V\delta U) , \quad (4.38)$$

or equivalently,

$$(I + \mathsf{L}_0\ V)\ \delta U = -\mathsf{L}_0(VU_0) . \quad (4.39)$$

As was done in (4.26), we take out the time derivative in the formulation of V. Thus we introduce $\widehat{V}$, the matrix of $S_{\pm\pm}$ operators (cf (4.33)), viz,

$$V(z, x, D_t) = \widehat{V}(z, x)D_t^2, \quad (4.40)$$

which results in

$$(I + D_t^2\mathsf{L}_0\ \widehat{V})\ \delta U = -D_t^2\mathsf{L}_0(\widehat{V}U_0) , \quad (4.41)$$

where $\widehat{V}\delta U$ and $\widehat{V}U_0$ are products of distributions (subject to the condition that their wavefront are favorably oriented (Friedlander & Joshi, 1998, proposition 11.2.3), (Hörmander, 1983, Theorem 8.2.10)). This is the resolvent equation in our hybrid Lippmann-Schwinger-Bremmer formulation for scattered waves. (See Yosida (1995) for details on resolvent equations.)

where

$$\delta U_1(\widehat{V}) = D_t^2 \mathsf{L}_0(\widehat{V} U_0), \quad \text{and} \quad \delta U_m(\widehat{V}) = D_t^2 \mathsf{L}_0(\widehat{V} \delta U_{m-1}(\widehat{V})), \quad m = 2, 3, \ldots . \tag{4.43}$$

Each subsequent term in the series is a multilinear operator of higher order than previous terms.

To compare the Bremmer series formulation ((VII.1)-(VII.22) of de Hoop (1996)) to the recursion in (4.42) we first make the following identifications. From (VII.1) and (VII.12) we note that W_0 of de Hoop (1996) corresponds to δU_1. From this formulation we note that $-D_t^2 \mathsf{L}_0 \widehat{V}$ corresponds with K of equation (VII.15) in de Hoop (1996) and (4.42) corresponds to equation (VII.22). The major difference between this hybrid series and the Bremmer series is in that the coupling of the different components is different. In the Bremmer series the reflection and transmission operators come from derivatives of the medium contrast whereas in the hybrid series they come from difference between the reference and true model. We use this hybrid formulation to derive operators that model both 'singly' and 'triply' scattered waves.

The expressions in (4.42) and (4.43) are not quite in the form of observables, however; data are acquired only at the Earth's surface, but the L_0 operator models data at all depths. We therefore define a restriction operator, R, which restricts a distribution to the acquisition surface, $z = 0$. This operator does not account for the free-surface boundary condition, thus we assume a homogeneous medium (with no reflectors) above the acquisition surface. In this we way have excluded incoming waves from above the acquisition surface. We assume that there are no reflectors at or near this surface, i.e., we assume that the support of the medium contrast, a, does not contain source or receiver points. The composition $R\mathsf{L}_0$ is well-defined provided there are no grazing rays (Stolk & de Hoop, 2002), which have been excluded already by the ψ cut-off from section 4.4. The composition with Q^{-1} also does not change the properties of the composite operator provided we satisfy the assumptions in the generalized Bremmer series (de Hoop, 1996). Observable quantities are obtained by applying Q^{-1} to δU, as in (4.30). We thus rewrite (4.42)

$$R\mathsf{Q}^{-1}\delta U(\widehat{V}) = \sum_{m=1}^{M} (-1)^m R\mathsf{Q}^{-1}\delta U_m(\widehat{V}) . \tag{4.44}$$

With

$$R\mathsf{Q}^{-1}\delta U_1(\widehat{V}) = D_t^2 R\mathsf{Q}^{-1}\mathsf{L}_0(\widehat{V} U_0) , \tag{4.45}$$

we introduce the operator

$$\mathsf{M}_0 = R\mathsf{Q}^{-1}\mathsf{L}_0 . \tag{4.46}$$

We then have,

$$R\mathsf{Q}^{-1}\delta U(\widehat{V}) \;=\; -D_t^2 \mathsf{M}_0 \left(\widehat{V} \left(U_0 + \sum_{m=1}^{M} (-1)^m \delta U_m(\widehat{V}) \right) \right) , \tag{4.47}$$

These equations are assumed to hold anywhere in the interior of the scattering region. Restricting δU to the surface and transforming it into observables by applying $R\mathsf{Q}^{-1}$ to (4.52-4.54) yields a recursion for $\widehat{V}_m$ in terms of the data d,

$$\begin{aligned}
\delta D &= -D_t^2\mathsf{M}_0(\widehat{V}_1U_0) && (4.55)\\
0 &= -D_t^2\mathsf{M}_0(\widehat{V}_2U_0) + D_t^4\mathsf{M}_0(\widehat{V}_1\mathsf{L}_0(\widehat{V}_1U_0)) && (4.56)\\
0 &= -D_t^2\mathsf{M}_0(\widehat{V}_3U_0) + D_t^4\mathsf{M}_0(\widehat{V}_2\mathsf{L}_0(\widehat{V}_1U_0)) + D_t^4\mathsf{M}_0(\widehat{V}_1\mathsf{L}_0(\widehat{V}_2(U_0))) && (4.57)\\
&\quad -D_t^6\mathsf{M}_0(\widehat{V}_1\mathsf{L}_0(\widehat{V}_1\mathsf{L}_0(\widehat{V}_1U_0)))
\end{aligned}$$

etc.

These equations hold on the acquisition surface, $z = 0$. In general, $\partial_z d$ (the second component of δD) is not recorded. We assume that we record only the up-going field, d_-, from $z > 0$. With this assumption, $d = Q_-^* d_-$ and $\partial_z d = -\mathcal{H}Q_-^{-1}d_-$ allowing $\partial_z d$ to be estimated directly from d.

The first term in the series, given in (4.55), models singly scattered data. The third term, in (4.57), models leading-order internal multiples as well as other primary events such as the one shown on the right in Figure 4.3. (The second term, given in (4.56), models events which have scattered twice, including primary events with one transmission and one reflection.)

Equation (4.57) can be simplified using (4.53). This is done by noting that the distributions $D_t^2\mathsf{L}_0(\widehat{V}_2U_0)$ from the second term of (4.57) and $D_t^4\mathsf{L}_0(\widehat{V}_1\mathsf{L}_0(\widehat{V}_1U_0))$ from the third term are identical by (4.53) and $D_t^2\mathsf{M}_0\widehat{V}_1(\cdot)$, which acts on these distributions (again in the second and third terms) is a linear operator. With this simplification we have, for (4.57)

$$D_t^2\mathsf{M}_0(\widehat{V}_3U_0) = D_t^4\mathsf{M}_0(\widehat{V}_2\mathsf{L}_0(\widehat{V}_1U_0))\,. \tag{4.58}$$

The general equation in the recursion follows from the fact that higher order terms are built from lower-order terms through the application of $D_t^2\mathsf{M}_0\widehat{V}_i$ to $(j-i)^{\text{th}}$-order terms to form terms of order j. For example, terms of order 4 are formed by subtracting $D_t^2\mathsf{M}_0\widehat{V}_1$ applied to (4.54), $D_t^2\mathsf{M}_0\widehat{V}_2$ applied to (4.53), and $D_t^2\mathsf{M}_0\widehat{V}_3$ applied to the right side of (4.52), from $D_t^8\mathsf{M}_0\widehat{V}_4U_0$. In general terms of order j will contain sub-series of the form

$$D_t^2\mathsf{M}_0\widehat{V}_1(\text{sum of terms of order } j-1 \text{ from (4.52)-(4.54)}), \tag{4.59}$$

$$D_t^2\mathsf{M}_0\widehat{V}_2(\text{sum of terms of order } j-2 \text{ from (4.52)-(4.54)}), \tag{4.60}$$

etc. For $j \geq 2$ the sub-series in parentheses sum to zero because of the zero on the left-hand side of (4.53).

We obtain the final form of the recursion,

$$\begin{aligned}
&D_t^2\mathsf{M}_0(\widehat{V}_jU_0) = D_t^4\mathsf{M}_0(\widehat{V}_{j-1}\mathsf{L}_0(\widehat{V}_1U_0))\,, \quad j \geq 2\,,\\
&\text{while } D_t^2\mathsf{M}_0(\widehat{V}_1U_0) = -\delta D
\end{aligned} \tag{4.61}$$

We give here an alternate derivation of this equation, resulting in our equation (4.76). We formulate the solution only for the upward propagating constituent of δU_1, which we denote by $\delta u_{-,1}$. We first determine the form of the down-going constituent of U_0, denoted by $u_{+,0}$, which is the down-going wave excited at the surface and arriving at the scattering point. With the expression for the source f_+ in (4.9) and that for u_+ in (4.11) we find that

$$u_{+,0}(z_1,x_1,t_1,z_0,s_0) = \tfrac{1}{2}\int_{-\infty}^{z_1} \mathrm{d}\tilde{z}_0 \int \mathrm{d}\tilde{s}_0 \int_{\mathbb{R}} \mathrm{d}\tilde{t}_{s_0} G_+(z_1,x_1,t_1-\tilde{t}_{s_0},\tilde{z}_0,\tilde{s}_0) \\ \mathcal{H}Q_{+,\tilde{s}_0}(\tilde{z}_0) f(\tilde{z}_0,\tilde{s}_0,\tilde{t}_{s_0},z_0,s_0)\,, \quad (4.66)$$

where we will adopt the convention that an integral without limits is assumed to be an integration over $\mathbb{R}^{n-1}$. In general, s represents a source position, r represents a receiver position, t is a time variable and z is depth, regardless of subscripts and superscripts. The notation $Q_{-,s}(z)$ is short for $Q_-(z,s,D_s,D_t)$. The t integrations are limited implicitly by the causality of the Green operator, G, discussed in Remark 4.4.1. The operator G_+ in (4.66) propagates between the levels z_0 and z_1, with its action being in the lateral variables $\tilde{s}_0$, and $\tilde{t}_{s_0}$; we will also use the notation $G_+(z_1,z_0)$ for the propagator G_+ when the lateral positions in which it acts are unambiguous. We adopt the standard kernel notation that the input variables to an operator are written to the right of the output variables. We are justified in writing the time dependence of $G_\pm$ as the difference of elapsed time and initial source time as the wave equation is time translation invariant. Expression (4.66) is valid for $z_1 > z_0$. The parameters z_0, s_0 are assumed to be known.

Next, we derive an expression for c_-, the up-going constituent in the contrast source given by,

$$\begin{pmatrix} c_+ \\ c_- \end{pmatrix} = VU_0 = V \begin{pmatrix} u_{+,0} \\ u_{-,0} \end{pmatrix}. \quad (4.67)$$

Using the expression for V in (4.33), and recalling that $u_{-,0} = 0$ for $z > 0$, we obtain an expression for c_-,

$$c_-(z_1,x_1,t_1) = -\tfrac{1}{2}\mathcal{H}D_{t_1}^2 Q_{-,x_1}(z_1) a(z_1,x_1) Q^*_{+,x_1}(z_1) u_{+,0}(z_1,x_1,t_1,z_0,s_0)\,. \quad (4.68)$$

Substituting c_- from (4.68) for f_- in (4.11) gives

$$\delta u_{-,1}(z_0,r_0,t_{r_0},z_0,s_0) = -\tfrac{1}{2}\mathcal{H}D_{t_{r_0}}^2 \int_{z_0}^{\infty} \mathrm{d}z_1 \int \mathrm{d}x_1 \int_{\mathbb{R}} \mathrm{d}t_1 \\ G_-(z_0,r_0,t_{r_0}-t_1,z_1,x_1) \underbrace{Q_{-,x_1}(z_1)\; a(z_1,x_1) Q^*_{+,x_1}(z_1)}_{S_{+-}} u_{+,0}(z_1,x_1,t_1,z_0,s_0) \quad (4.69)$$

in the diagonal system without the restriction to the Earth's surface, $z_0 = \tilde{z}_0 = 0$. This is the first term in the series in (4.42)-(4.43). Because a is compactly supported in z_1, the integral over z_1 is actually over a compact set. As in the previous section, we assume that the medium contrast, a, has its support away from $z = 0$. To obtain modeled data, we

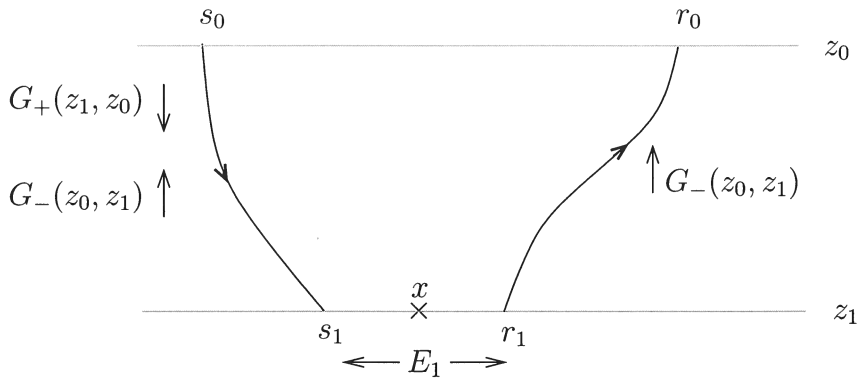


Figure 4.4. Notation for single scattering modeling.

To obtain a more compact expression, we return to operator notation, first introducing

$$(H(z_0, z_1))(s_0, r_0, t - t_0, s_1, r_1) = \int_{\mathbb{R}} (G_-(z_0, z_1))(r_0, t - t' - t_0, r_1)(G_-(z_0, z_1))(s_0, t', s_1)\,\mathrm{d}t', \quad (4.75)$$

the kernel of the propagator $H(z_0, z_1)$ associated with the so-called double-square-root (DSR) equation (Claerbout, 1985), which propagates data from the depth z_1 to the depth z_0. Substituting this expression for the two Green functions in (4.74) gives equation (3.10) of Stolk & de Hoop (2004a, Theorem 5.1),

$$d_1(s_0, r_0, t_{r_0}) = -\tfrac{1}{4} D_t^2 Q^*_{-,r_0}(0) Q^*_{-,s_0}(0) \int_0^\infty \mathrm{d}z_1 \left(H(0, z_1) Q_{-,r_1}(z_1) Q_{-,s_1}(z_1)(E_2 E_1 a) \right)(s_0, r_0, t_{r_0}). \quad (4.76)$$

4.7.2 Leading-order internal multiple scattering

In (4.76), we showed how singly scattered data can be constructed given the medium perturbation. Our ultimate goal is to construct the medium contrast given data containing both primaries and leading-order internal multiples. In this section we establish a relation between the modeling of primaries and internal multiples.

Following the diagram in Figure 4.5, the first scattering of the internal multiple, from s_0 through s_2, r_2 to m_r is nearly identical to the single scattering case. We cannot use the H operator however, because the second leg (from r_2 to m_r) does not reach the surface,

which acts as a contrast source for the final wave, propagating from (z_3, r_3) to $(0, r_0)$,

$$d_3(s_0, r_0, t_4) = -\tfrac{1}{2}\mathcal{H}D_{t_4}^2 Q^*_{-,r_0}(0)\int_0^\infty \mathrm{d}z_3 \int \mathrm{d}x_3 \int_{\mathbb{R}} \mathrm{d}t_3$$
$$G_-(0, r_0, t_4 - t_3, z_3, x_3) Q_{-,x_3}(z_3) a(z_3, x_3) Q^*_{+,x_3}(z_3) \delta u_{+,2}(z_3, x_3, t_3, 0, s_0)\,, \quad (4.79)$$

where we have returned to observables through the operator $R\mathsf{Q}^{-1}$, introduced in (4.47). For the above construction to be valid, (z_1, x_1), (z_2, x_2) and (z_3, x_3) cannot be arbitrarily close to one another.

We now apply reciprocity (4.22) to the G_+ occurring in the expression for $\delta u_{+,2}$ in (4.78). We do this by substituting the expression for $\delta u_{+,2}$ in (4.78) into (4.79) to use the Q_+ operators from both expressions combined and introduce the extension operators E_1, E_2. This gives

$$d_3(s_0, r_0, t_4) = -\tfrac{1}{4}D_{t_4}^4 \int_0^\infty \mathrm{d}z_3 \int \mathrm{d}s_3 \int \mathrm{d}r_3 \int_{\mathbb{R}} \mathrm{d}t_3 \int_0^{z_3} \mathrm{d}z_1 \int \mathrm{d}m_s \int \mathrm{d}m_r \int_{\mathbb{R}} \mathrm{d}t_a$$
$$Q^*_{-,r_0}(0) G_-(0, r_0, t_4 - t_3, z_3, r_3) Q_{-,r_3}(z_3)$$
$$Q^*_{-,m_s}(z_1) G_-(z_1, m_s, t_3 - t_a, z_3, s_3) Q_{-,s_3}(z_3)$$
$$(E_1 a)(z_3, s_3, r_3)(E_1 a)(z_1, m_s, m_r) Q^*_{-,m_r}(z_1) \delta u_{-,1}(z_1, m_r, t_a, 0, s_0)\,; \quad (4.80)$$

we have also introduced the extension operator E_1, to split each of the m and x_3 integrations into two.

Identifying the propagator $Q_{-,x_a}(z_a) G_-(z_a, z_b) Q_{-,z_b}(z_b)$ in (4.80) with the function $G(z_a, x_a, t, z_b, x_b)$ in equation (8) of ten Kroode (2002) along with the substitution of the expression for $\delta u_{-,1}$ in (4.77) shows the correspondence of (4.80) with expression (8) in ten Kroode (2002). (Note that $V(x)$ in ten Kroode (2002) is $a(z, x)$ here.)

We interchange the order of integration in t_3 and t_a, and change integration variables from t_3 to $t_3' = t_3 - t_a$, introducing the E_2 operator at the third scatter. This results in

$$d_3(s_0, r_0, t_4) = -\tfrac{1}{4}D_t^4 \int_0^\infty \mathrm{d}z_3 \int \mathrm{d}s_3 \int \mathrm{d}r_3 \int_{\mathbb{R}} \mathrm{d}t_{30} \int_{\mathbb{R}} \mathrm{d}t_a \int_0^{z_3} \mathrm{d}z_1 \int \mathrm{d}m_s \int \mathrm{d}m_r$$
$$\int_{\mathbb{R}} \mathrm{d}t_3' Q^*_{-,r_0}(0) G_-(0, r_0, t_4 - t_a - t_3' - t_{30}, z_3, r_3) Q_{-,r_3}(z_3)$$
$$Q^*_{-,m_s}(z_1) G_-(z_1, m_s, t_3', z_3, s_3) Q_{-,s_3}(z_3)$$
$$(E_2 E_1 a)(z_3, s_3, r_3, t_{30})(E_1 a)(z_1, m_s, m_r) Q^*_{-,m_r}(z_1) \delta u_{-,1}(z_1, m_r, t_a, 0, s_0)\,, \quad (4.81)$$

which is a modeling operator for triply scattered waves. We need not introduce E_2 at the m_s, m_r scattering point here, but it will be required later. Equations (4.81) and (4.78) are expressed entirely in terms of up-going propagators (G_-); they comprise the $m = 2$ term of the forward series, given in the summation in (4.47).

The recursion in equation (4.61) demonstrates that it is possible to express the triply scattered data, d_3, in terms of the singly scattered data d_1. The first step to writing d_3 in

thus far. To this end, we now discuss an inverse scattering theory. From the inverse series, constructed in section 4.6.2, we note that only a single-scattering inverse is required, because for each term in the series we estimate $\widehat{V}_j$ from $\mathsf{M}_0(\widehat{V}_j U_0)$ based on the recursion in (4.61).

The Born modeled data, d_1, as given in (4.48), is measured in the field but its vertical derivative is not. A left inverse to the Born modeling operator, the inverse scattering operator, can be constructed under the double-square root (DSR) assumption (see section 4.4 and Stolk & de Hoop (2004a)). This operator gives an estimate of $a(z,x)$ via

$$\langle a_1 \rangle = T d_1 \,, \tag{4.85}$$

where T is the inverse scattering operator and $\langle a_1 \rangle$ is the single scattering estimate of a. Once an estimate of a is obtained, it can be substituted into (4.33) to obtain an estimate of $\widehat{V}$.

Stolk & de Hoop (2004b) give a detailed method for inverse scattering from singly scattered data; here we give a brief summary. The construction involves the depth-to-time conversion operator, $\bar{K}$, defined as

$$\bar{K} : a \mapsto -\int_0^\infty H(0,z)(E_2 a)(z,\cdot,\cdot,\cdot)(s,r,t)\mathrm{d}z \,. \tag{4.86}$$

Stolk and de Hoop show that this operator is an invertible Fourier integral operator. Upon substitution of a point source in (4.76), we obtain

$$d_1 = \tfrac{1}{4} D_t^2 Q^*_{-,s}(0) Q^*_{-,r}(0) \bar{K} J(E_1 a) \,. \tag{4.87}$$

The operator J (denoted V by Stolk & de Hoop (2004b)), has symbol

$$\begin{aligned} &J(z,s,r,\zeta,\sigma,\rho) = \\ &\quad |\tau|^{-1}(c_0(z,s)^{-2} - \tau^{-2}\|\sigma\|^2)^{-1/4}(c_0(z,r)^{-2} - \tau^{-2}\|\rho\|^2)^{-1/4}|_{\tau=\Theta^{-1}(z,s,r,\zeta,\sigma,\rho)} \,. \end{aligned} \tag{4.88}$$

This operator is related to the $Q_{-,s}(z)Q_{-,r}(z)$ appearing in (4.76); the difference is that J is applied before the E_2 extension operator whereas $Q_{-,s}(z)Q_{-,r}(z)$ is applied after (note that Q and E_2 do not commute). The map Θ is defined by

$$\Theta(z,s,r,\sigma,\rho,\tau) = -b(z,s,\sigma,\tau) - b(z,r,\rho,\tau) \,. \tag{4.89}$$

Stolk & de Hoop (2004a, Lemma 4.1) show that $\tau \mapsto \zeta = \Theta(z,s,r,\sigma,\rho,\tau)$ is a diffeomorphism. The mapping from frequency to vertical wavenumber described by this map is required for J to be applied before the E_2 extension operator.

After defining the adjoint operator in space (restriction to $h = 0$) by $R_1 = E_1^*$, the adjoint operator in time (restriction to $t = 0$) by $R_2 = E_2^*$, and the normal operator $\bar{\Xi}$ which is equal to the principal part of $\bar{K}^*\bar{K}$ we have

$$\bar{\Phi}(z,x,D_z,D_x)a = R_1 J^{-1} \bar{\Xi}^{-1} \bar{K}^* Q^*_{-,s}(0)^{-1} Q^*_{-,r}(0)^{-1} D_t^{-2} d_1 \,, \tag{4.90}$$

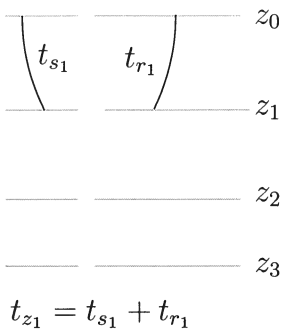


Figure 4.6. Time notations used to estimate d_3 at z_1.

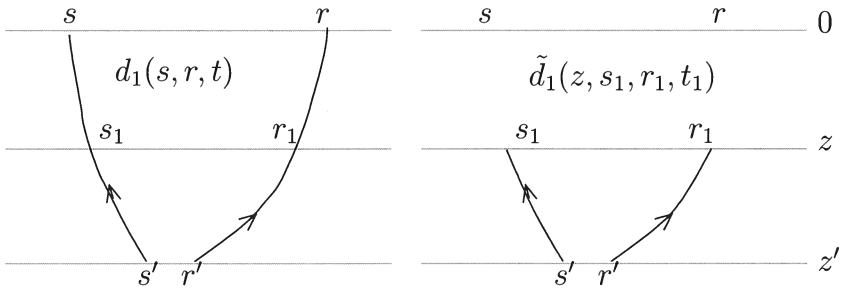


Figure 4.7. Illustration of Lemma 4.9.1; the construction of $\tilde{d}_1$ from d_1.

Lemma 4.9.1. *We define*

$$\tilde{d}_1(z,s,r,t) = -\tfrac{1}{4}D_t^2\int_z^\infty \mathrm{d}z' \left(H(z,z')Q_{-,r'}(z')Q_{-,s'}(z')(E_2E_1a)(z',\overset{s'}{\cdot},\overset{r'}{\cdot},\overset{t'}{\cdot})\right)(s,r,t)\,. \quad (4.94)$$

For $t > 0$

$$\left(H(0,z)^*Q^*_{-,s}(0)^{-1}Q^*_{-,r}(0)^{-1}d_1\right)(s,r,t) = \tilde{d}_1(z,s,r,t)\,, \quad (4.95)$$

where d_1 *is modeled by (4.76).*

This Lemma is illustrated in Figure 4.7.

Proof. We first define $\bar{a} = \chi a$ where χ is the characteristic function of (z,∞). With this definition we write $\tilde{d}_1$ as

$$\tilde{d}_1(s,r,t) = -\tfrac{1}{4}D_t^2\int_0^\infty \mathrm{d}z' \left(H(z,z')Q_{-,r'}(z')Q_{-,s'}(z')(E_2E_1\bar{a})(z',\overset{s'}{\cdot},\overset{r'}{\cdot},\overset{t'}{\cdot})\right)(s,r,t)\,. \quad (4.96)$$

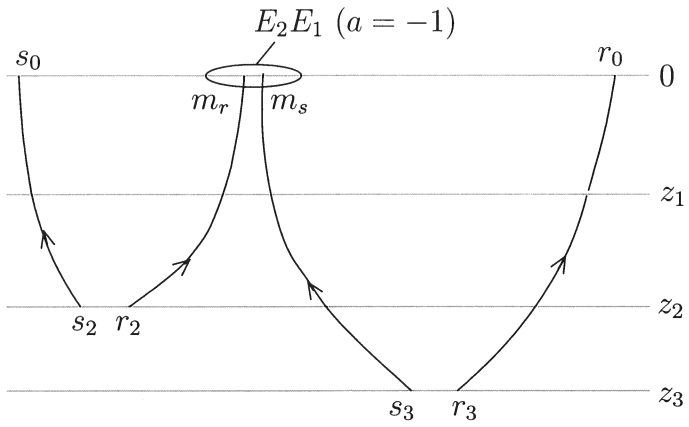


Figure 4.8. Illustration of the surface-related multiple case (SRME).

relating our method to the surface-related multiple elimination (SRME) procedure of Fokkema & van den Berg (1993, chapter 12). This is illustrated in Figure 4.8.

The following theorem describes the relation between the internal multiple estimated at the surface through (4.84) given in Theorem 4.7.1 and the estimate of $\tilde{d}_3$ defined in (4.99).

Theorem 4.9.3. *Let the data be modeled by the forward scattering series (4.47) for $M = 2$. Then there is the following correspondance between the leading-order internal multiple modeled at the surface and $\tilde{d}_3$*

$$d_3(s_0, r_0, t_0) = Q^*_{-,r_0}(0) Q^*_{-,s_0}(0) \int_0^\infty \mathrm{d}z_1 \left(H(0, z_1) \tilde{d}_3(z_1, \overset{s}{\cdot}, \overset{r}{\cdot}, \overset{t}{\cdot}) \right) (s_0, r_0, t_0) \,. \tag{4.102}$$

The theorem is illustrated in Figure 4.9.

Proof. We begin by returning to (4.77),

$$\begin{aligned} \delta u_{-,1}(z_1, m, t_a, 0, s_0) = \\ -\tfrac{1}{4} D^2_{t_a} Q^*_{-,s_0}(0) \int_{z_1}^\infty \mathrm{d}z_2 \int \mathrm{d}s_2 \int \mathrm{d}r_2 \int_{\mathbb{R}} \mathrm{d}t_0 \int_{\mathbb{R}} \mathrm{d}t' G_-(z_1, m, t_a - t' - t_0, z_2, r_2) \\ \int \mathrm{d}s_1 \int_{\mathbb{R}} \mathrm{d}t_{s_1} G_-(0, s_0, t_{s_1}, z_1, s_1) G_-(z_1, s_1, t' - t_{s_1}, z_2, s_2) \\ Q_{-,r_2}(z_2) Q_{-,s_2}(z_2) (E_2 E_1 a)(z_2, s_2, r_2, t_0) \,, \end{aligned} \tag{4.103}$$

The same sequence of steps applied to (4.81) gives

$$d_3(s_0,r_0,t_4) = -\tfrac{1}{4}D_t^2\int_0^\infty \mathrm{d}z_1\int \mathrm{d}r_1\int \mathrm{d}t_{r_1} Q^*_{-,r_0}(0)G_-(0,r_0,t_{r_1},z_1,r_1)$$
$$\int \mathrm{d}m_s \int \mathrm{d}m_r Q^*_{-,m_s}(z_1)\int_{\mathbb{R}} \mathrm{d}t_a \tilde{d}_1(z_1,m_s,r_1,t_4-t_a-t_{r_1})$$
$$(E_1 a)(z_1,m_s,m_r)Q^*_{-,m_r}(z_1)\delta u_{-,1}(z_1,m_r,t_a,0,s_0)\,, \quad (4.106)$$

where we have also interchanged the order of integration. Substituting the expression for $\delta u_{-,1}$ from (4.105) into (4.106) and re-ordering the Q operators and the G_- propagators results in

$$d_3(s_0,r_0,t_4) =$$
$$D_t^2 Q^*_{-,r_0}(0)Q^*_{-,s_0}(0)\int_0^\infty \mathrm{d}z_1 \int \mathrm{d}m_s \int \mathrm{d}m_r Q^*_{-,m_s}(z_1)(E_1 a)(z_1,m_s,m_r)$$
$$Q^*_{-,m_r}(z_1)\int_{\mathbb{R}} \mathrm{d}t_a G_-(0,z_1)\tilde{d}_1(z_1,m_s,{}^{r_1}_{\cdot},t_4-t_a-{}^{t_{r_1}}_{\cdot})G_-(0,z_1)\tilde{d}_1(z_1,{}^{s_1}_{\cdot},m_r,t_a-{}^{t_{s_1}}_{\cdot})\,. \quad (4.107)$$

Combining the two G_- propagators into a single H operator gives the result.

□

Equation (4.102) is equivalent to (4.76) with $\tilde{d}_3$ taking the place of the contrast source.

Theorem 4.9.4. *Assume the inverse scattering series (4.62) for $M = 2$. If we replace d_1 in (4.95) in Lemma 4.9.1 by d and a in equation (4.99) for $\tilde{d}_3$ by a_1 then*

$$\langle a_3(z,x)\rangle = (\Phi^{-1}R_1J^{-1}\bar{\Xi}^{-1}R_2D_t^{-2}\tilde{d}_3(z,{}^{s}_{\cdot},{}^{r}_{\cdot},{}^{t}_{\cdot}))(x)\,. \quad (4.108)$$

Proof. Recall from the recursion in (4.61) that

$$D_t^2\mathsf{M}_0(\widehat{V}_3U_0) = D_t^6\mathsf{M}_0(\widehat{V}_1\mathsf{L}_0(\widehat{V}_1\mathsf{L}_0(\widehat{V}_1U_0)))\,. \quad (4.109)$$

Theorem 4.7.1 shows that d_3 is third order in $\widehat{V}_1$ and thus third order in d. We then estimate $\widehat{V}_3$ directly from d_3 using (4.92).

$$\begin{aligned}\langle a_3\rangle &= \Phi^{-1}R_1J^{-1}\bar{\Xi}^{-1}K^*D_t^2d_3\\ &= \Phi^{-1}R_1J^{-1}\bar{\Xi}^{-1}R_2H(0,z)^*Q^*_{-,r}(0)^{-1}Q^*_{-,s}(0)^{-1}D_t^{-2}d_3\,. \quad (4.110)\end{aligned}$$

The argument in the proof of Lemma 4.9.1 can be repeated for the expression for d_3 in (4.102), recalling that $\tilde{d}_3$ is defined for $t > 0$ giving

$$\tilde{d}_3(z,s,r,t) = (H(0,z)^*Q^*_{-,s}(0)^{-1}Q^*_{-,r}(0)^{-1}d_3)(z,s,r,t)\,, \quad (4.111)$$

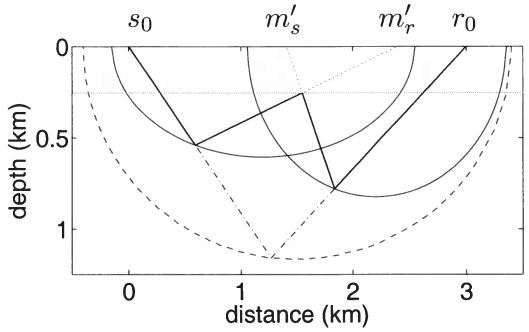
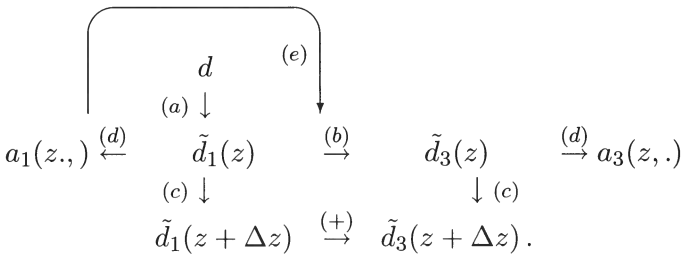


Figure 4.10. A contribution to $\langle a_3 \rangle$. The solid rays are the triply scattered rays. The dash-dot line is the singly scattered contribution with the same source and receiver positions as well as slopes. The dashed curve is the single scattering isochron, for the time t_4 corresponding to the amount of time required to travel along the triply scattered path. The shaded region extends to the depth level z_1 to which the entire wavefield is propagated before generating the image correction via $\langle a_3 \rangle$.

4.10 Discussion

We propose a method for attenuating artifacts in the image generated by leading-order internal multiples. We give two main results: a structure for modeling leading-order internal multiples in (J.8) and (4.84), and a system to estimate leading-order internal multiples as part of the imaging procedure in (4.112). Our suggested algorithm is illustrated by the following flowchart



In (a) the data are downward continued to the depth z, through Lemma 4.9.1. Following this, in (b) leading-order internal multiples are estimated via (4.99). In (c), both the data and the estimated multiple are propagated to the next depth, again through (4.95) in Lemma 4.9.1. An image is formed, in (d), at this depth via (4.112). The image is also used to obtain an estimate of a used in the estimate of d_3 from (4.99). The theory discussed requires knowledge of the velocity model to the depth z_1 of the up-to-down scatter at which the image is formed. In addition, an adaptive subtraction technique is necessary

Chapter 5

Identification of image artifacts due to leading-order internal multiples [1]

5.1 Summary

First order internal multiples are a source of coherent noise in seismic images. There are a number of techniques to estimate internal multiples in the data, but few methods exist that estimate imaging artifacts caused by internal multiples. We propose a method to do this in which the artifacts are estimated as part of the imaging process. Our technique is based on a hybrid of the Lippmann-Schwinger scattering series and the generalized Bremmer coupling series. Although we require knowledge of the velocity model this allows us to estimate internal multiples without assumptions inherent to other methods.

5.2 Introduction

Internal multiples have been recognized in seismic experiments for a long time (Sloat, 1948). A lot more is known about the attenuation of surface-related multiples (Aminzadeh & Mendel, 1980; Fokkema *et al.*, 1994; Berkhout & Verschuur, 1997; Verschuur & Berkhout, 1997; Weglein *et al.*, 1997), than is known about the attenuation of internal multiples (Weglein *et al.*, 1997; Jakubowicz, 1998; Kelamis *et al.*, 2002; ten Kroode, 2002; van Borselen, 2002). It is still not possible, to estimate multiples in data with sufficient accuracy to remove all the artifacts they introduce in seismic images. In this paper, we propose a technique for estimating imaging artifacts caused by internal multiples as part of the imaging process.

Fokkema & van den Berg (1993) use reciprocity to show the possibility of predicting surface-related multiples through a Neumann series expansion. As a point of departure we use the generalized Bremmer coupling series to model internal multiples, because its behavior and convergence are known (de Hoop, 1996). We then construct a hybrid series, using the contrast source formulation from the Lippmann-Schwinger scattering series (Lippmann, 1956; Weglein *et al.*, 1997), to estimate imaging artifacts caused by leading-order internal multiples. This hybrid series is amenable to the downward continuation approach unlike the Kirchhoff approximation (de Hoop, 2004). Using this method requires knowledge of the velocity model. Technically this knowledge is necessary only to the depth of the shallowest reflector involved in the generation of internal multiples (the depth of the up-to-down reflec-

[1] This chapter is expected to grow into a published paper with co-authors M. V. de Hoop and H. Calandra

many problems (see van Stralen (1997) for an overview) and the convergence of various generalizations of the original series has also been a subject of interest (Atkinson, 1960; Corones, 1975; Gray, 1983; McMaken, 1986). Aminzadeh & Mendel (1980, 1981) were the first to propose a method using the Bremmer series to attenuate surface-related multiples in a horizontally layered medium.

To estimate artifacts in the image caused by first-order internal multiples (FOIM), we proceed in two steps. We first develop a method to model FOIM, using the hybrid series; this is described in Section 5.3. We circumvent the traveltime monotonicity condition through downward continuation. Following this, we predict internal multiples from downward continued data, to estimate the artifacts they cause in the image, using ideas from the inverse series in Section 5.4. In Section 5.5 we describe an algorithm to perform these two steps at the same time. We illustrate the application of this algorithm to synthetic and field data in Section 5.6.

5.3 The scattering series

We begin by decomposing the pressure wavefield, u, into its up- and down-going constituents, $u_\pm$ ($-$ denotes an up-going constituent and $+$ a down-going constituent). Following Stolk & de Hoop (2004a), we find that

$$\begin{pmatrix} u \\ \partial_z u \end{pmatrix} = \mathsf{Q}^{-1} \begin{pmatrix} u_+ \\ u_- \end{pmatrix}, \tag{5.1}$$

where

$$\mathsf{Q} = \frac{1}{2} \begin{pmatrix} (Q_+^*)^{-1} & -\mathcal{H} Q_+ \\ (Q_-^*)^{-1} & \mathcal{H} Q_- \end{pmatrix}; \tag{5.2}$$

$\mathcal{H}$ denotes the Hilbert transform in time and * denotes adjoint. From this,

$$Q_-^* G_- \mathcal{H} Q_-,$$

generates the upgoing constituents of the full-wave Green's function. The Q matrix, along with its inverse, diagonalizes the wave operator written as a first order system and thus splits the wavefield into its up- and down-going constituents.

The hybrid series uses the decomposition discussed above along with the contrast source from the Lippmann-Schwinger series. We use a subscript 0 to indicate the field in the background model and δ to represent a contrast, thus the field U in the unknown true medium is related to that in the known background medium by $U = U_0 + \delta U$. We denote by δU_j the vector of up- and down-going wave constituents scattered j times. The terms in the hybrid forward scattering series are related by

$$\delta U_1(\widehat{V}) = D_t^2 \mathsf{L}_0(\widehat{V} U_0), \quad \text{and} \quad \delta U_m(\widehat{V}) = D_t^2 \mathsf{L}_0(\widehat{V} \delta U_{m-1}(\widehat{V})), \; m = 2, 3, \ldots \tag{5.3}$$

5.4 Inverse scattering

In inverse scattering the goal is to solve for $\widehat{V}$ in terms of the data d. To this end, we assume that the contrast operator $\widehat{V}$ can be written as a series

$$\widehat{V} = \sum_{m\in\mathbb{N}} \widehat{V}_m(d)\,, \tag{5.9}$$

where $\widehat{V}_m$ is of order m in the data. Substituting this equation into (5.6) leads to the following relation between the $\widehat{V}_m(d)$,

$$D_t^2\mathsf{M}_0(\widehat{V}_m U_0) = D_t^4\mathsf{M}_0(\widehat{V}_{m-1}\mathsf{L}_0(\widehat{V}_1 U_0))\ \ , m\geq 2\,, \tag{5.10}$$

where

$$\delta D = -D_t^2\mathsf{M}_0(\widehat{V}_1 U_0)\,. \tag{5.11}$$

From this it follows that

$$-D_t^2\mathsf{M}_0(\widehat{V}U_0) = \delta D - \left(\sum_{m\in\mathbb{N}} D_t^2\mathsf{M}_0(\widehat{V}_m\delta U)\right)\,. \tag{5.12}$$

If we ignore the second term on the right-hand side, the problem of expressing $\widehat{V}$ in terms of the data reduces to inverse scattering in the Born approximation (Stolk & de Hoop, 2004b). In the context of wave-equation migration, the inverse scattering procedure is split into two parts: downward continuation and imaging, which are discussed in the next section.

5.5 Artifacts due to internal multiples in imaging

The first step in imaging is downward continuation. To do this, we apply the adjoint propagator $H(0,z)^*$ to the modeled data in (5.8) yielding the downward continued data at depth z,

$$\tilde{d}_1(z) = H(0,z)^* Q^*_{-,s}(0)^{-1} Q^*_{-,r}(0)^{-1} d\,, \tag{5.13}$$

for $t\geq 0$. This downward continuation uses the usual background velocity model to estimate the data that would have been recorded at the depth z. The downward continuated data, $\tilde{d}_1(z)$ is found directly from the medium contrast via

$$\tilde{d}_1(z,s,r,t) = -\tfrac{1}{4}D_t^2\int_z^\infty \mathrm{d}z' \left(H(z,z')Q_{-,r'}(z')Q_{-,s'}(z')(E_2E_1a)(z',\overset{s'}{\cdot},\overset{r'}{\cdot},\overset{t'}{\cdot})\right)(s,r,t)\,. \tag{5.14}$$

Next, we apply the 'true-amplitude' imaging condition, M, to the downward continued data

$$a_1(z,.) = R_1 M R_2 D_t^{-2}\tilde{d}_1(z) \tag{5.15}$$

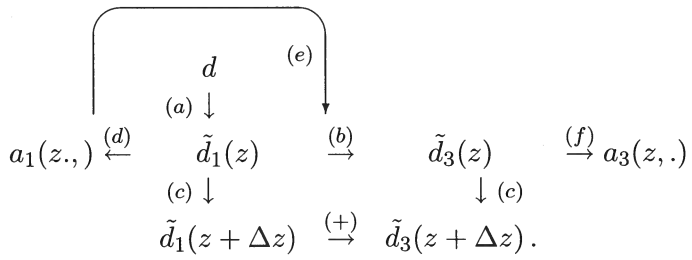
where $R_1 = E_1^*$ and $R_2 = E_2^*$ are the usual imaging conditions ($t=0$ and $h=0$); this procedure is described in Stolk & de Hoop (2004b). We apply (5.13)-(5.15) to (5.12), from

Using relation (5.17) along with (5.13) we find that artifact in the image caused by FOIM at the depth z can be estimated by

$$a_3(z,.) = R_1 M R_2 D_t^{-2} \tilde{d}_3(z)\,, \tag{5.19}$$

the analog of (5.15).

We propose an algorithm summarized by the following flow chart



The algorithm can be divided into several steps. First, in (a), we downward continue the data to the depth z. Then in (b), the multiples are estimating using (5.18); this also requires an estimate of the image in (d). An estimate of the artifacts in the image is made in (f). The data and multiples are then downward continued to the next depth in (c).

5.6 Examples

Techniques like the angle-domain filtering proposed by Sava & Guitton (2005) are promising because they attenuate multiples directly in the image as opposed to in the data. In this way, even though the multiples are still not completely removed their location in the image is known. Thus, they are less likely to be misinterpreted as primary reflection energy.

In this section we describe both synthetic and field data results of the technique described above to estimate artifacts from FOIM directly in the image space. Three different synthetic models are presented. A simple flat model illustrates the steps of the algorithm. Following this a more complicated model is used to test the ability of the method to estimate imaging artifacts caused by FOIM in the presence of caustics and to test the sensitvity of the method to the velocity model. The third model shows the techniques applicability in a model with more complicated reflector shapes. The field data example is a 2D line extracted from a 3D survey in the Gulf of Mexico.

The algorithm used to generate the examples shown here falls into the category of a Generalized Screen Propagator. It is implemented as a split step propagator along with an implicit finite difference residual wide-angle correction. The propagator works in the midpoint-offset coordinates, requiring us to use a subset of the available data to obtain an uniformly sampled data set in these coordinates. This algorithm was proposed by Jin *et al.* (1998).

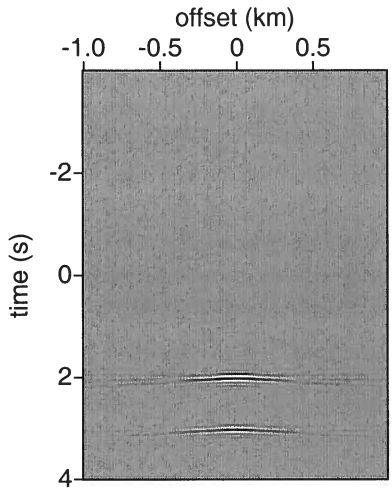


Figure 5.4. The estimated multiple, in the data, at a depth of 1.5 km; note the agreement with the true multiple in Figure 5.3.

sampling.

In our method, the data are first downward continued as part of a standard wave-equation migration technique ((a) in the flowchart). In Figure 5.3 we show $\tilde{d}_1(z = 1.5 \text{ km})$, a single common-midpoint gather (cmp) downward continued to the depth $z = 1.5$ km of the top of the layer. The primary reflected from the top of the layer is located around $t = 0$, the reflection from the bottom of the layer at about $t = 1$ s and the first order internal multiple at about $t = 2$ s.

We now estimate the multiples at depth using equation (5.16). This requires restricting $\tilde{d}_1$ to time $t > 0$. The procedure removes the primary reflection from the current depth (which theoretically arrives at $t = 0$), in this case 1.5 km, before doing the convolution. If this process is not done correctly and energy remains at $t \leq 0$, all primary reflections from deeper depths will be duplicated in the estimated multiples section. In this model a simple time-windowing procedure is sufficient, because the reflections are far apart in time. In some situations, we find a τ-p filter to be more effective. This is because we typically see tails at small positive and negative times caused by the band-limited signal. A τ-p filter is more effective at removing these tails when they are mixed with later reflections. Figure 5.3 show the results of applying the τ-p filter to the data.

Once the negative time contributions to the data have been removed, the multiple is estimated with (5.16), through a convolution. The convolved wavefield is multiplied by an

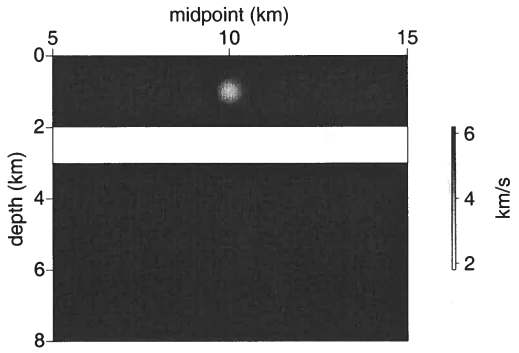


Figure 5.6. Velocity model, similar to the flat layered example discussed previously, with the addition of a low-velocity lens to demonstrate that the method works in laterally heterogeneous velocity models.

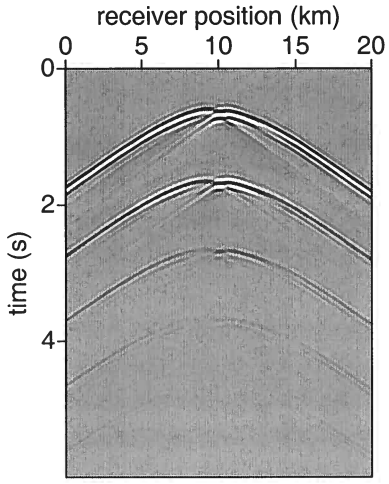


Figure 5.7. Shot record from $s = 9.8$ km, 200 m to the left of the center of the lens. Note the caustic introduced by the lens around zero-offset. The ringing on the second primary (at about 2 s) and the multiple (at about 3 s) is numerical dispersion from the modeling of the data.

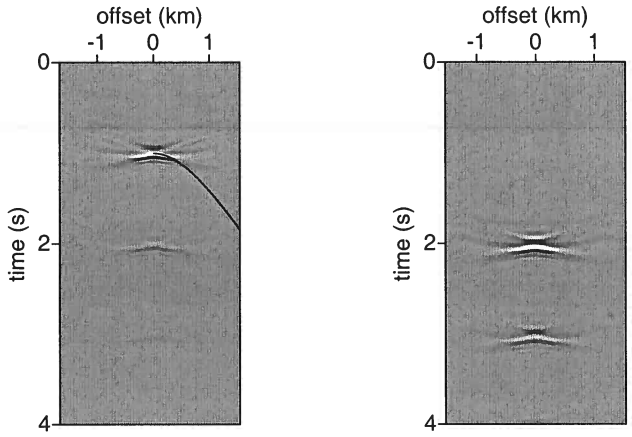


Figure 5.9. Left: Common midpoint gather at 9.8 km and 2 km depth, after the $t \leq 0$ times have been removed. Note the disappearance of the multi-pathing as the data are now below the lens. The solid line shows the expected moveout curve for the reflection from the bottom of the reflector. Right: estimated multiples at this depth.

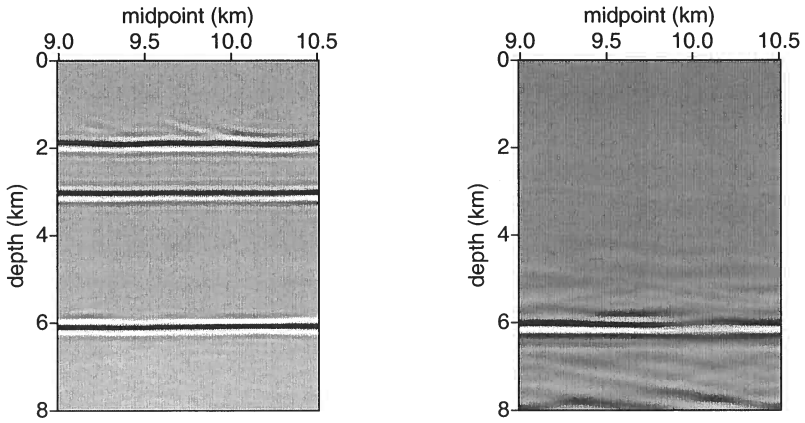


Figure 5.10. Left: Image with an artifact from the first-order internal multiple at approximately 6 km depth. Right: Estimated artifacts from first-order internal multiples.

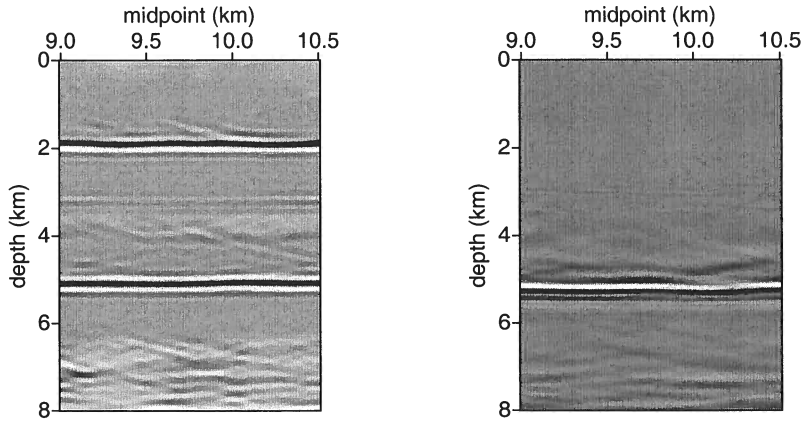


Figure 5.12. In these images, the reflector has been extended to 3.5 km from 3 km to test the sensitivity of the method to the velocity model. Left: Image with artifacts from internal multiples. Right: Estimated artifacts from first-order internal multiples.

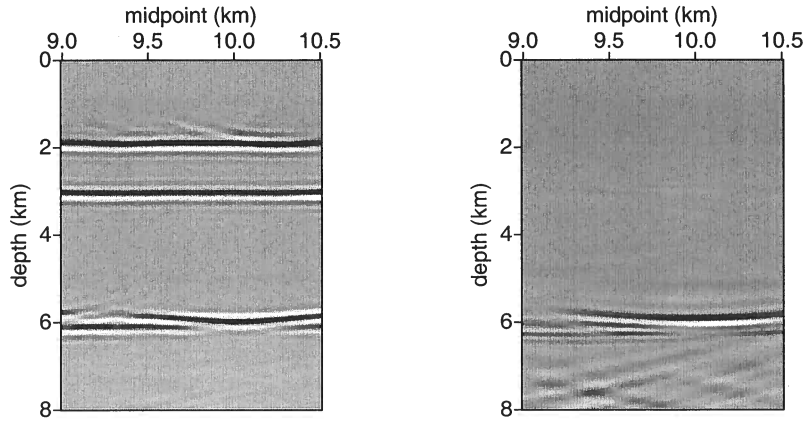


Figure 5.13. In these images, a second lens has been added beneath the layer to introduce a laterally varying velocity perturbation. Left: Image with artifacts from internal multiples. Right: Estimated artifacts from first-order internal multiples.

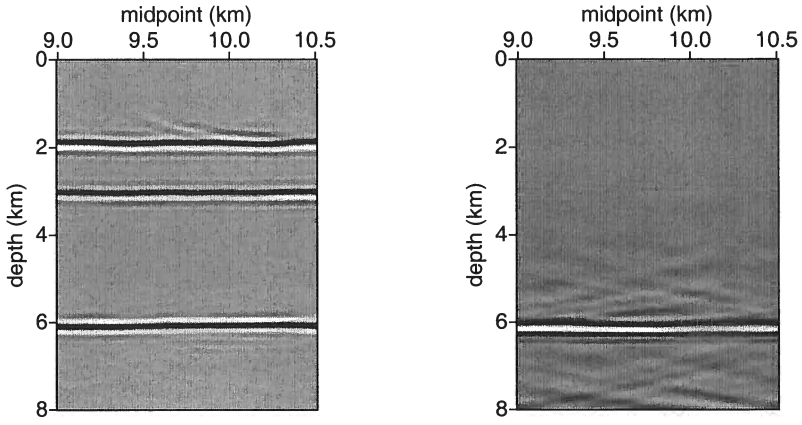


Figure 5.15. In this model the lens was moved 0.2 km deeper than in the correct velocity model. Because this perturbation is above the top of the layer, we expect this to have an impact on the estimated multiple. Note the phase difference between the estimated artifact and the image. Left: Image with artifacts from first-order internal multiples. Right: Estimated artifacts from first-order internal multiples.

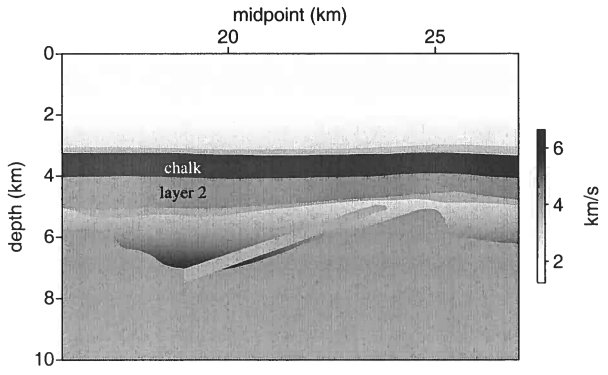


Figure 5.16. Velocity model for the North Sea example.

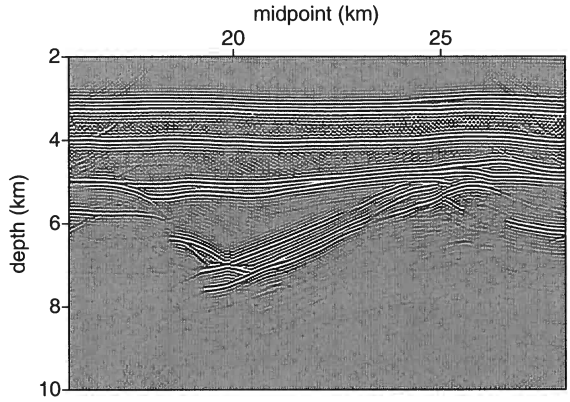


Figure 5.18. The image in the chalk model, with artifacts from a first-order internal multiple at about 4.5 km depth.

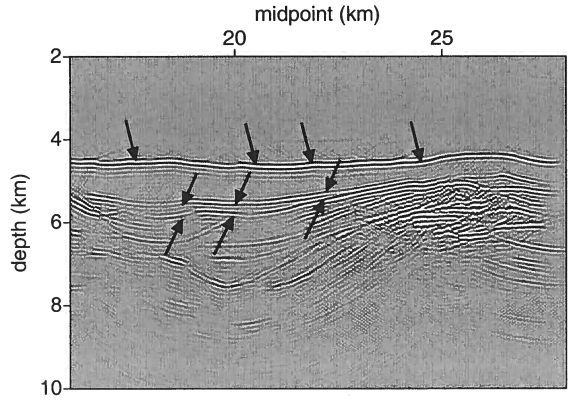


Figure 5.19. This figure repeats the estimated artifacts from internal multiples shown in Figure 5.18, with arrows indicating locations at which multiples can be observed in the results of ten Kroode (2005).

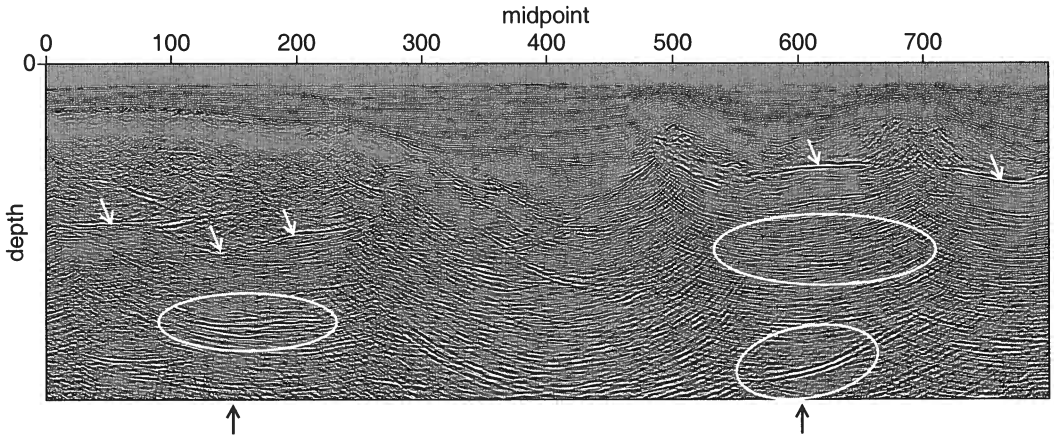


Figure 5.20. Image for the real data example. The base of salt is marked with white arrows and three areas of the image are marked, two of which contain artifacts from internal multiples (left and top right) and one of which does not (bottom right). The locations of the CIGs shown in Figures 5.22 and 5.23 are marked with the black arrows below the image.

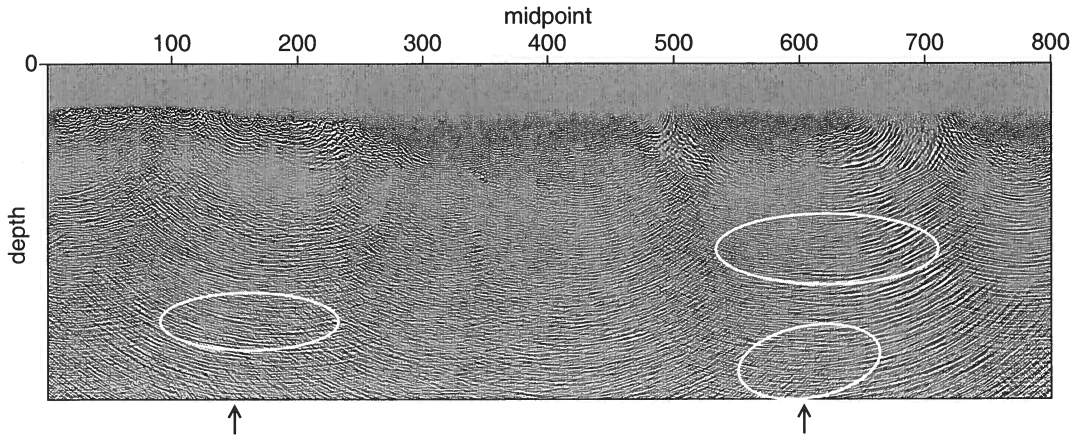


Figure 5.21. Estimated artifacts for the real data example. The three areas marked are the same as those marked on the image above. The locations of the CIGs shown in Figures 5.22 and 5.23 are marked with the arrows below the image.

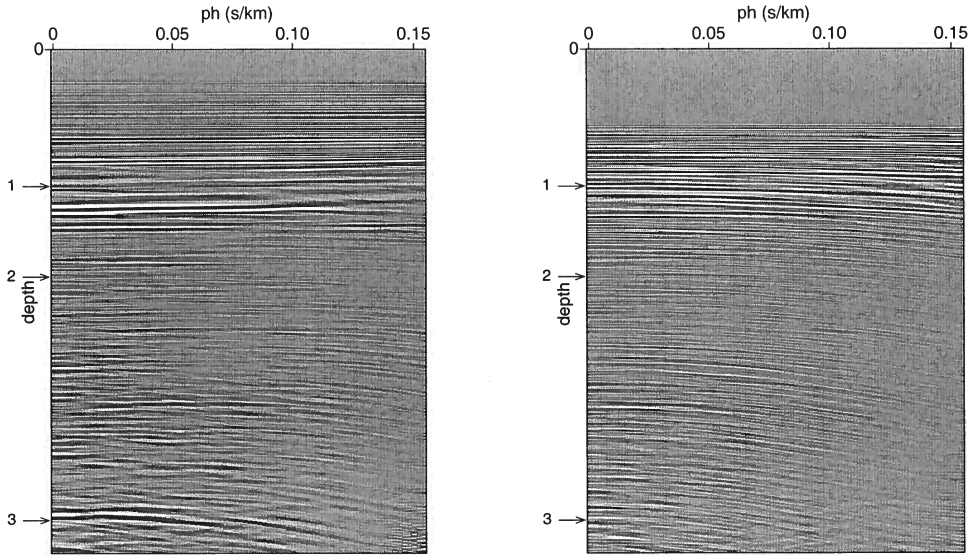


Figure 5.23. Real data example, common image gather at cmp 600. On the left is the standard image gather and on the right are the estimated artifacts from internal multiples. Arrows 1 and 2 indicate the locations of accurately estimated artifacts. Arrow number 2 is within the upper-right highlighted region of the image. Arrow number 3 marks the artifact in the lower right highlighted region on the image shown in Figure 5.20. That this artifact does not appear in the estimated artifacts indicates that this energy does not come from an internal multiple.

to be done to have a robust algorithm for real data applications. This example illustrates that at the least this method can be used to ascertain whether artifacts in the image come from internal multiples or not.

5.7 Discussion

We have described a method to estimate imaging artifacts caused by first-order internal multiples. This method requires knowledge of the velocity model down to the top of the layer that generates the multiple (the depth of the up-to-down reflection). The main computational cost of the algorithm comes from the downward continuation of the data and the internal multiples. Because two data sets are downward continued (the data themselves and the estimated multiples), the cost of the algorithm described here is about twice that of a usual pre-stack depth migration, plus the cost of the removal of negative times. The removal of negatives times can be compuationally expensive because it is necessary to

Chapter 6

Conclusions and Future Work

6.1 Conclusions

Data continuation has many application in seismology, two of which are discussed in this thesis. By employing the mathematics of Fourier Integral Operators, we are able to extend data regularization or data mapping techniques to situations in which the wavefield contains caustics. In addition, we have shown that techniques for the downward continuation of data allow us to dispence with two commonly made assumptions in multiple attenuation: that the wavefield does not contain caustics and the traveltime monotonicity assumption. We have developed a technique that is able to estimate imaging artifacts caused by first-order internal multiples without these two assumption. Both of these applications also involve operator composition, in which several operators are combined to make a final, single, operator. For data regularization, this composition comes with conditions under which the resulting operator does not introduce false reflections in the data. For multiples the application of successive Green's operators illustrates the need for the traveltime monotonicity assumption used in other multiple attenuation techniques.

For data regularization we have shown that, even in the presence of caustics, it is possible to fill in missing data. We have illustrated this theory by computing impulse responses and by filling in missing data in synthetic shot records. The major limitation of this technique is its computational cost. Although we believe it to be possible to fill in missing data with the cost on the order of a prestack migration that is still expensive as a pre-conditioning step. On the other hand, this technique is designed to be applied in the most complicated geologic structures where traditional, faster techniques fail.

We have developed a theory for multiple attenuation in the presence of caustics, when the traveltime monotoncity condition does not hold. We have begun to illustrate the practical value of the process but a lot of work remains in that direction. First of all, artifacts caused by internal multiples have significantly smaller amplitudes than do primary reflections. This presents problems for accurately modeling data containing multiples, but not containing other artifacts such as aliasing, numerical dispersion, refractions and boundary reflections. Even if such modeled data are available it is also difficult to see the artifacts from internal multiples on synthetic data because of errors in the migration itself, such as wrap-around, aliasing and wide-angle effects that interfere with the low amplitude multiples. The problem of modeling data is removed by application to real data but this comes with its own set of problems as the data contain noise, have been pre-processed introducing other errors that are difficult to assess.

analysis because they do not satisfy the assumptions of the method; they are triply scattered not singly scattered. Multiples are sensitive to the velocity model, however, in particular they are sensitive to the structure in which they scatter three times rather than the single time primaries are reflected. By estimating the multiples in the image, or preferably in the image gather, the flatness criteria often used in velocity analysis can be replaced with a joint criteria on the flatness of the primaries and the successful prediction of artifacts from internal multiples. In addition the matching between the predicted and true multiple is influenced by the velocity above the shallowest scattering point only, providing an indicator of which parts of the velocity model need to be updated.

Multiples are also more sensitive to reflector positions in the Earth because they scatter three times rather than once. When they are mixed with primaries it is difficult to use this information because whether the recorded energy is a primary scattered deep in the Earth or a multiple scattered shallower is not easily assessed. Seperating the multiples from the primaries opens up the possibility of imaging with multiples as is currently under investigation for surface-related multiples by Brown (2004).

References

Alkhalifah, T., & de Hoop, M. V. 2004. Residual Dip moveout in VTI media. *Geophysical Prospecting, in print.*

Aminzadeh, F., & Mendel, J. M. 1980. On the Bremmer series decomposition: Equivalence between two different approaches. *Geophysical Prospecting*, **28**, 71–84.

Aminzadeh, F., & Mendel, J. M. 1981. Filter Design for suppression of surface multiples in a non-normal incidence seismogram. *Geophysical Prospecting*, **29**, 835–852.

Artley, C., & Hale, D. 1994. Dip-moveout processing for depth-variable velocity. *Geophysics*, **59**, 610–622.

Atkinson, F. V. 1960. Wave propagation and the Bremmer series. *J. Math. Anal. Appl.*, **1**, 225–276.

Berkhout, A. J., & Verschuur, D. J. 1997. Estimation of multiple scattering by iterative inversion, Part I: Theoretical considerations. *Geophysics*, **62**(5), 1586–1595.

Bevc, D., Fliedner, M., Crawley, S., & Biondi, B. 2003. Wave equation imaging comparisons: Survey sinking vs. shot profile methods. *SEG Technical Program Expanded Abstracts*, **22**(1), 885–888.

Beylkin, G. 1985. Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform. *J. of Math. Phys.*, **26**, 99–108.

Biondi, B., Fomel, S., & Chemingui, N. 1998. Azimuth moveout for 3-D prestack imaging. *Geophysics*, **63**, 574–588.

Black, J. L., Schleicher, K. L., & Zhang, L. 1993. True-amplitude imaging and dip moveout. *Geophysics*, **58**, 47–66.

Bleistein, N. 1990. Born DMO revisited. *60th Annual International Meeting, SEG, Expanded Abstracts*, 1366–1369.

Bleistein, N., & Jaramillo, H. 2000. A platform for data mapping in scalar models of data acquisition. *Geophysical Prospecting*, **48**, 135–161.

Bleistein, N., Cohen, J. K., & Jaramillo, H. 1999. True amplitude transformation to zero offset of data from curved reflectors. *Geophysics*, **64**, 112–129.

Bleistein, N., Cohen, J. K., & Stockwell, J. W. Jr. 2000. *Mathematics of multidimensional seismic imaging, migration and inversion.* New York: Springer-Verlag.

de Hoop, M. V., Burridge, R., Spencer, C., & Miller, D. 1994. Generalized radon transform amplitude versus angle (GRT/AVA) migration/inversion in anisotropic media. *Pages 15–27 of: Proc SPIE 2301.* SPIE.

de Hoop, M. V., Spencer, C., & Burridge, R. 1999. The resolving power of seismic amplitude data: An anisotropic inversion/migration approach. *Geophysics*, **64**(3), 852–873.

de Hoop, M. V., Malcolm, A. E., & Le Rousseau, J. H. 2003a. Seismic wavefield 'continuation' in the single scattering approximation: A framework for dip and azimuth moveout. *Can. Appl. Math. Q.*, **10**, 199–238.

de Hoop, M. V., Le Rousseau, J. H., & Biondi, B. 2003b. Symplectic structure of wave-equation imaging: A path-integral approach based on the double-square-root equation. *Geophys. J. Int.*, **153**, 52–74.

Deregowski, S. G., & Rocca, F. 1981. Geometrical optics and wave theory of constant offset sections in layered media. *Geoph. Prosp.*, **29**, 374–406.

Dragoset, W. H., & Jericević, Z. 1998. Some remarks on surface multiple attenuation. *Geophysics*, **63**, 772–789.

Duistermaat, J. J. 1996. *Fourier integral operators.* Boston: Birkhäuser.

Fokkema, J. T., & van den Berg, P. M. 1993. *Seismic applications of acoustic reciprocity.* Amsterdam: Elsevier.

Fokkema, J.T., Van Borselen, R. G., & Van den Berg, P.M. 1994. *Removal of inhomogeneous internal multiples.* Eur. Assn. of Expl. Geophys. Page Session:H039.

Fomel, S. 1995. Amplitude preserving offset continuation in theory, Part 1: the offset continuation equation. *Pages 179–196 of: Stanford Exploration Project Preprint*, vol. SEP-84.

Fomel, S. 2003. Theory of differential offset continuation. *Geophysics*, **68**(2), 718–732.

Fomel, S., & Bleistein, N. 2001. Amplitude preservation for offset continuation: Confirmation for Kirchhoff data. *Journal of Seismic Exploration*, **10**, 121–130.

Fomel, S., & Prucha, M. 1999. Angle-gather time migration. *SEP-100*, 141–150.

Foss, S-K., De Hoop, M. V., & Ursin, B. 2003. Linearized 2.5-D parameter imaging-inversion in anisotropic elastic media. *CWP-431-P.*

Fowler, P. J. 1998. A comparative overview of dip moveout methods. *68th Annual International Meeting, SEG, Expanded Abstracts*, 1744–1747.

Friedlander, F. G., & Joshi, M. 1998. *Introduction to the theory of distributions.* Cambridge university press.

Jaramillo, Herman, & Bleistein, Norman. 1998. Seismic data mapping. *SEG Technical Program Expanded Abstracts*, **17**(1), 1991–1994.

Jin, Shengwen, Wu, Ru-Shan, & Peng, Chengbin. 1998. Prestack depth migration using a hybrid pseudo-screen propagator. *SEG Technical Program Expanded Abstracts*, **17**(1), 1819–1822.

Kelamis, P., Erickson, K., Burnstad, R., Clark, R., & Verschuur, D. 2002. *Data-driven internal multiple attenuation - Applications and issues on land data.* Soc. of Expl. Geophys. Pages 2035–2038.

Kennett, B. L. N. 1974. Reflections, Rays and Reverberations. *Bull. Seism. Soc. Am.*, **64**, 1685–1696.

Kennett, B. L. N. 1979a. The suppression of surface multiples on seismic records. *Geoph. Prosp.*, **27**, 584–600.

Kennett, B. L. N. 1979b. Theoretical Reflection Seismograms for Elastic Media. *Geoph. Prosp.*, **27**, 301–321.

Kennett, B. L. N. 1983. *Seismic Wave Propagation in Stratified Media.* Cambridge: Cambridge University Press.

Le Rousseau, J. H. 2001. *Microlocal analysis of wave-equation imaging and generalized-screen propagators.* Ph.D. thesis, Colorado School of Mines.

Levin, F. K. 1971. Apparent velocity from dipping interface reflections. *Geophysics*, **36**, 510–516.

Liner, C. L. 1991. Born theory of wave-equation dip moveout. *Geophysics*, **56**(2), 182–189.

Lippmann, B. A. 1950. Variational principles for scattering processes II Scattering of slow Neutrons by Para-Hydrogen. *Phys. Rev.*, **79**(3), 481–486.

Lippmann, B. A. 1956. Rearrangement Collisions. *Phys. Rev.*, **102**(1), 264–268.

Lippmann, B. A., & Schwinger, J. 1950. Variational principles for scattering processes I. *Phys. Rev.*, **79**(3), 469–480.

Loewenthal, D., Lu, L., Roberson, R., & Sherwood, J. 1976. The wave equation applied to migration. *Geophysical Prospecting*, **24**(2), 380–399.

McMaken, H. 1986. On the convergence of the Bremmer series for the Helmholtz equation in 2d. *Wave Motion*, **8**, 277–283.

Miller, D., & Burridge, R. 1992. *Geophysical inversion.* SIAM. Chap. Multiparameter inversion, dip-moveout, and the generalized Radon transform, pages 46–58.

Sorin, V., & Ronen, S. 1989. Ray-geometrical analysis of dip moveout amplitude distribution. *Geophysics*, **54**, 1333–1335.

Stolk, C. C. 2001 (July). *Microlocal analysis of the scattering angle transform.* Preprint, Dept. of Computational and Applied Mathematics, Rice University, http://www.caam.rice.edu/~cstolk/angle.ps.

Stolk, C. C., & de Hoop, M. V. 2002. Microlocal analysis of seismic inverse scattering in anisotropic elastic media. *Comm. Pure Appl. Math.*, **55**(3), 261–301.

Stolk, C. C., & de Hoop, M. V. 2004a. Modeling of seismic data in the downward continuation approach. *submitted to SIAM J. Appl. Math.* CWP468P.

Stolk, C. C., & de Hoop, M. V. 2004b. Seismic inverse scattering in the downward continuation approach. *submitted to SIAM J. Appl. Math.* CWP469P.

Stolk, C. C., & Symes, W. W. 2004. Kinematic artifacts in prestack depth migration. *Geophysics*, **69**(2), 562–575.

Stolk, C. C., de Hoop, M. V., & Symes, W. W. 2005. Kinematics of shot-geophone migration. *in preparation.*

Stolt, R. H. 1978. Migration by Fourier Transform. *Geophysics*, **43**(1), 23–48.

Stolt, R.H. 2002. Seismic data mapping and reconstruction. *Geophysics*, **67**, 890–908.

ten Kroode, A. P. E. 2002. Prediction of internal multiples. *Wave Motion*, **35**, 315–338.

ten Kroode, A. P. E. 2005. personal communication.

ten Kroode, A. P. E., Smit, D.-J., & Verdel, A. R. 1998. A microlocal analysis of migration. *Wave Motion*, **28**, 149–172.

Treves, F. 1980a. *Introduction to pseudodifferential and Fourier integral operators.* Vol. 2. New York: Plenum Press.

Treves, F. 1980b. *Introduction to pseudodifferential and Fourier integral operators.* Vol. 1. New York: Plenum Press.

Tygel, M., Schleicher, J., Hubral, P., & Santos, L. T. 1998. 2.5-D true-amplitude Kirchhoff migration to zero offset in laterally inhomogeneous media. *Geophysics*, **63**(2), 557–573.

van Borselen, R. 2002. Data-driven interbed multiple removal: Strategies and examples. *In: Expanded Abstracts.* Society of Exploration Geophysicists.

van Stralen, M. J. N. 1997. *Directional decomposition of electromagnetic and acoustic wave-fields.* Ph.D. thesis, Delft University of Technology.

Verschuur, D. J., & Berkhout, A. 1997. Estimation of multiple scattering by iterative inversion, Part II: Practical aspects and examples. *Geophysics*, **62**(5), 1596–1611.

Appendix A

Dip MoveOut: $n = 3$ and constant coefficient[1]

The aim of this appendix is two-fold: (i) to show that the analysis presented in the main text encompasses the usual DMO analysis in the absence of caustics as practiced in seismology, and (ii) to clarify the issue of number of phase variables needed in the OI representation of the DMO kernel for $n = 3$ in the constant coefficient case.

It is noted, that in the case of constant c, e can be chosen to be offset $h = \frac{1}{2}(r - s)$ in the acquisition manifold. We define an acquisition submanifold, Y', by prescribing the value of h. Throughout the analysis, in particular of the operator L, the manifold Y can be replaced by the submanifold Y' and the cotangent bundle $T^*Y\backslash 0$ by $T^*Y'\backslash 0$.

A.1 Modeling and imaging operators

In the case of a medium with constant velocity c, the generating function S in (2.4) is simply given by $-\tau T(x, x_0)$ with $T(x, x_0)$ the traveltime function along the ray connecting x with x_0, viz.

$$T(x, x_0) = \frac{|x - x_0|}{c}. \tag{A.1}$$

Since, away from the point source, no caustics occur, the traveltime function is single valued and only one phase variable, namely τ, is required in the phase function. We will use Cartesian coordinates.

The Green's function, G, is given by the OI (cf. (2.6))

$$G(x, t, x_0) = \int \frac{1}{8\pi^2 c^2 |x - x_0|} \exp[\mathrm{i}\tau(t - T(x, x_0))]\, \mathrm{d}\tau,$$

from which the modeling operator kernel of F is derived,

$$F(s_1, s_2, r_1, r_2, t, x_0) = \int \frac{-\tau^2}{16c^7\pi^3 |r - x_0||s - x_0|} \delta c(x_0) \exp[\mathrm{i}\tau(t - T(x_0, s_1, s_2, 0, r_1, r_2, 0))]\, \mathrm{d}\tau,$$

[1]This appendix has been published, along with Chapter 2 and Appendix B in:
de Hoop, M. V., Malcolm, A. E. and Le Rousseau, J. H. 2003. Seismic wavefield 'continuation' in the single scattering approximation: A framework for dip and azimuth moveout *Can. Appl. Math. Q.*, **10**, 199-238.

is transversal.

Proof. The partial derivatives of Ψ with respect to the phase variables are given by

$$\frac{\partial\Psi}{\partial(x_0)_i} = -\frac{2\tau_0}{c}\frac{((x_0)_i - z_i)}{|x_0 - z|} + \frac{\tau}{c}\left(\frac{((x_0)_i - y_i - h_i)}{|x_0 - y - h|} + \frac{((x_0)_i - y_i + h_i)}{|x_0 - y + h|}\right) , \quad (A.3)$$

$$\frac{\partial\Psi}{\partial\tau} = -t + \frac{1}{c}\left(|x_0 - y - h| + |x_0 - y + h|\right) , \quad (A.4)$$

$$\frac{\partial\Psi}{\partial\tau_0} = t_0 - \frac{2|x_0 - z|}{c} , \quad (A.5)$$

$i = 1, 2, 3$. The form of the differentials with respect to all the variables is:

		$\mathrm{d}(\partial_{(x_0)_i}\Psi)$ $(i=1,2,3)$	$\mathrm{d}(\partial_\tau\Psi)$	$\mathrm{d}(\partial_{\tau_0}\Psi)$
(i)	y_j $(j=1,2)$	$*$	$*$	$*$
(ii)	t	$\boxed{0}$	$\boxed{-1}$	$\boxed{0}$
(iii)	z_j $(j=1,2)$	$-\frac{2\tau_0}{c}\left(-\frac{\delta_{ij}}{\lvert x_0-z\rvert} + \frac{((x_0)_i-z_i)((x_0)_j-z_j)}{\lvert x_0-z\rvert^3}\right)$	$*$	$*$
(iv)	t_0	$\boxed{0}$	$\boxed{0}$	$\boxed{1}$
(v)	$(x_0)_j$ $(j=1,2,3)$	$*$	$*$	$*$
(vi)	τ	$*$	0	0
(vii)	τ_0	$-\frac{2}{c}\frac{((x_0)_i-z_i)}{\lvert x_0-z\rvert}$	0	0

Because of the entries related to t and t_0 [rows (ii) and (iv)] the rank of the matrix is $2 + \mathrm{rank}(\mathrm{d}(\frac{\partial\Psi}{\partial x_0}))$. Now, $\frac{c}{2\tau_0} \times (iii)_j - \frac{c}{2}\frac{((x_0)_j - z_j)}{|x_0 - z|^2} \times (vii)$, for $j = 1, 2$ yields the form of rows (iii) and (vii):

$$\begin{array}{c|ccc} (iii)_1 & \frac{1}{|x_0-z|} & 0 & 0 \\ (iii)_2 & 0 & \frac{1}{|x_0-z|} & 0 \\ (vii) & * & * & -\frac{2}{c}\frac{((x_0)_3-z_3)}{|x_0-z|} \end{array} .$$

Since $|x_0 - z| > 0$, $(0 = z_3 < (x_0)_3)$, $\mathrm{rank}(\mathrm{d}(\frac{\partial\Psi}{\partial x_0})) = 3$. The rank of the differentials is therefore maximal; the phase is nondegenerate. It follows that the composition of the two canonical relations is transversal (Hörmander, 1985b, Thm. 21.2.19). □

Parametrization of the canonical relation

As already mentioned above in the case $e = h$ one can restrict the DMO operation to a constant offset one, replacing Y by Y'. Comparing to the main text, observe that one does not need the cotangent variable ε to parameterize the canonical relation: we make use of only s, x_0, τ here. In this case Λ_D then follows as

$$\Lambda_D = \{(\boldsymbol{z}(s, x_0), \boldsymbol{t}_0(s, x_0), \boldsymbol{\zeta}(s, x_0, \tau), \boldsymbol{\tau}_0(s, x_0, \tau); \boldsymbol{y}(s), \boldsymbol{t}(s, x_0), \boldsymbol{\eta}(s, x_0, \tau), \tau)\} .$$

and a similar expression for $\partial_{r_1}\theta$.

(A.6)

From (x_0, s) we determine the direction of the ray at the source,

$$-\beta = \frac{s - x_0}{|s - x_0|} , \tag{A.7}$$

and the traveltime

$$\tilde{t} = \frac{|s - x_0|}{c} . \tag{A.8}$$

The angle θ_s is defined through

$$\beta = (\cos\theta_s, \sin\theta_s) . \tag{A.9}$$

Using the relation $\theta_r = \theta_s + \theta$ (cf. Figure 3.2) we find the angle θ_r which defines the direction of the ray at the receiver

$$-\alpha = -(\cos\theta_r, \sin\theta_r) . \tag{A.10}$$

The receiver ray traveltime then follows from

$$\sin\theta_r c\hat{t} = \frac{(x_0)_3}{c} . \tag{A.11}$$

Then the receiver position is found to be

$$r = -\hat{t}c\alpha + x_0 . \tag{A.12}$$

The total traveltime is simply given by

$$T = \tilde{t} + \hat{t} . \tag{A.13}$$

The cotangent variables σ and ρ are then given by

$$\sigma = -\frac{\tau}{c}\cos\theta_s - \varepsilon\partial_{s_1}\theta(x_0, s, r) , \tag{A.14}$$

$$\rho = -\frac{\tau}{c}\cos\theta_r - \varepsilon\partial_{r_1}\theta(x_0, s, r) . \tag{A.15}$$

We determine (α_0, τ_0) from the equality

$$\frac{\tau}{c}(\alpha + \beta) - \varepsilon\partial_{x_0}\theta(x_0, s, r) = \xi = 2\left(\frac{\tau_0}{c}\right)\alpha_0 . \tag{A.16}$$

The zero-offset traveltime, t_0, thus follows as

$$t_0 = (x_0)_3/(c\,\sin\theta_0) ,$$

Proof. The structure of the differentials of f with respect to (x_0, τ, τ_0) is:

	(a) $\mathrm{d}(\partial_{(x_0)_j}\Psi)$ $(j=1,2,3)$	(b) $\mathrm{d}(\partial_\tau \Psi)$	(c) $\mathrm{d}(\partial_{\tau_0}\Psi)$
(i) $(x_0)_i$	$-\frac{2\tau_0}{c}\left(\frac{\delta_{ij}}{\lvert x_0-z\rvert}-\frac{((x_0)_i-z_i)((x_0)_j-z_j)}{\lvert x_0-z\rvert^3}\right)$ $+\frac{\tau}{c}\left(\frac{\delta_{ij}}{\lvert x_0-y-h\rvert}-\frac{((x_0)_i-y_i-h_i)((x_0)_j-y_j-h_j)}{\lvert x_0-y-h\rvert^3}\right)$ $+\frac{\tau}{c}\left(\frac{\delta_{ij}}{\lvert x_0-y+h\rvert}-\frac{((x_0)_i-y_i+h_i)((x_0)_j-y_j+h_j)}{\lvert x_0-y+h\rvert^3}\right)$	$\frac{1}{c}\left(\frac{((x_0)_i-y_i-h_i)}{\lvert x_0-y-h\rvert}+\frac{((x_0)_i-y_i+h_i)}{\lvert x_0-y+h\rvert}\right)$	$\frac{2}{c}\frac{((x_0)_i-z_i)}{\lvert x_0-z\rvert}$
(ii) τ	$\frac{1}{c}\left(\frac{((x_0)_j-y_j-h_j)}{\lvert x_0-y-h\rvert}+\frac{((x_0)_j-y_j+h_j)}{\lvert x_0-y+h\rvert}\right)$	0	0
(iii) τ_0	$-\frac{2}{c}\frac{((x_0)_j-z_j)}{\lvert x_0-z\rvert}$	0	0

with $i=1,2,3$. On the stationary point set, S_Ψ, in view of equation (A.3), rows (ii) and (iii) are linearly dependent and the same holds for columns (b) and (c). Equation (A.3) for $(x_0)_3$ $((x_0)_3 \neq 0)$ gives $-\frac{2\tau_0}{c}\frac{1}{\lvert x_0-z\rvert}+\frac{\tau}{c}\left(\frac{1}{\lvert x_0-y-h\rvert}+\frac{1}{\lvert x_0-y+h\rvert}\right)=0$ and hence a simplification of the upper left 3×3 matrix. Therefore $\mathrm{rank}(\partial_{x_0}f, \partial_\tau f, \partial_{\tau_0} f)$ is that of:

	(a)	(b)
(i)	$2\tau_0\frac{((x_0)_i-z_i)((x_0)_j-z_j)}{\lvert x_0-z\rvert^3}$ $-\tau\frac{((x_0)_i-y_i-h_i)((x_0)_j-y_j-h_j)}{\lvert x_0-y-h\rvert^3}$ $-\tau\frac{((x_0)_i-y_i+h_i)((x_0)_j-y_j+h_j)}{\lvert x_0-y+h\rvert^3}$	$((x_0)_i-z_i)$
(ii)	$((x_0)_j-z_j)$	0

with $i,j=1,2,3$. By subtracting $[2\tau_0\frac{((x_0)_j-z_j)}{\lvert x_0-z\rvert^3}-\tau(\frac{((x_0)_j-y_j-h_j)}{\lvert x_0-y-h\rvert^3}+\frac{((x_0)_j-y_j+h_j)}{\lvert x_0-y+h\rvert^3})]\times(b)$ from $(a)_j$, $j=1,2,3$, and using the fact that $h_2=h_3=0$, $z_3=y_3=0$, and $z_2=y_2$ on S_Φ, the matrix further simplifies to

	(a)	(b)
$(i)_1$	$-\tau\frac{(z_1-y_1-h_1)((x_0)_j-y_j-h_j)}{\lvert x_0-y-h\rvert^3}-\tau\frac{(z_1-y_1+h_1)((x_0)_j-y_j+h_j)}{\lvert x_0-y+h\rvert^3}$	$((x_0)_1-z_1)$
$(i)_2$	0	$((x_0)_2-y_2)$
$(i)_3$	0	$(x_0)_3$
(ii)	$((x_0)_j-z_j)$	0

Performing a similar operation with rows instead yields that the rank is that of $((x_0)_3>0)$

	$(a)_1$	$(a)_3$	(b)
$(i)_1$	$-\tau\frac{(z_1-y_1-h_1)^2}{\lvert x_0-y-h\rvert^3}-\tau\frac{(z_1-y_1+h_1)^2}{\lvert x_0-y+h\rvert^3}$	0	$((x_0)_1-z_1)$
$(i)_3$	0	0	$(x_0)_3$
(ii)	$((x_0)_1-z_1)$	$(x_0)_3$	0

(A.17)

The signature of the second order differential $(\partial^2\Psi/\tau)/(\partial^2((x_0)_1,(x_0)_3,\tau_0))$ is constant. It is easy to compute it at a point x_0 half-way between source and receiver. The signature is then -1.

At stationarity the phase function simplifies to

$$\tau(\frac{|x_0-y-h|}{c}+\frac{|x_0-y+h|}{c}-t)\,,$$

and the entire amplitude of the associated oscillatory integral representatio becomes

$$\frac{\tau^{1/2}\tau_0^2}{32c^{14}(2\pi)^{3/2}|z-x_0|^2|y+h-x_0||y-h-x_0|}\exp(-\mathrm{i}\pi/4)\frac{1}{|H|^{1/2}}\,.$$

Appendix B

Azimuth MoveOut: $n = 3$ and constant coefficient[1]

In a constant velocity medium, it is possible to derive an expression for the impulse response in closed form. In Biondi *et al.* (1998) the impulse response of AMO was derived as the time t_2 as a function of translation in midpoint location, $\frac{1}{2}(r_2 + s_2) - \frac{1}{2}(r_1 + s_1)$ for given offsets $\frac{1}{2}(r_1 - s_1)$ and $\frac{1}{2}(r_2 - s_2)$. Here, we determine time t_2 as a function of ray direction at s_1 (associated with σ_1) for given e_1 and (subsurface) scattering angle from e_2 and (acquisition surface) azimuth direction, i.e. direction of $\frac{1}{2}(r_2 - s_2)$.

We derive the impulse response in two steps. First, we determine the three-dimensional DMO zero-offset traveltime (t_0) from (s_1, r_1, t_1), and then we determine the AMO time (t_2) by performing inverse DMO from the zero-offset ray to (s_2, r_2, t_2). Since the zero-offset ray will always be in the plane defined by the source and receiver rays, we need only compute the scattering angle in that plane, as a function of the initial ray angles at the source s_1. We apply the DMO formula derived in Appendix A.

In Figure B.1 we introduce the unit vectors

$$\tilde{\alpha}_1 = (\cos(\varphi_1)\cos(\psi_1), \cos(\varphi_1)\sin(\psi_1), \sin(\varphi_1)) , \tag{B.1}$$

$$\Xi = (\cos(\varphi_2)\cos(\psi_2), \cos(\varphi_2)\sin(\psi_2), \sin(\varphi_2)) . \tag{B.2}$$

($\tilde{\alpha}_1$ determines σ_1 and Ξ determines ξ.) We observe that $w = \tilde{\alpha}_1 - \lambda\Xi$, while also w lies in the $x_3 = 0$ plane. We evaluate λ by setting $w_3 = 0$,

$$\lambda = \frac{\sin(\varphi_1)}{\sin(\varphi_2)} . \tag{B.3}$$

Then

$$w = \begin{pmatrix} \cos(\varphi_1)\cos(\psi_1) - \frac{\sin(\varphi_1)}{\sin(\varphi_2)}\cos(\varphi_2)\cos(\psi_2) \\ \cos(\varphi_1)\sin(\psi_1) - \frac{\sin(\varphi_1)}{\sin(\varphi_2)}\cos(\varphi_2)\sin(\psi_2) \\ 0 \end{pmatrix} \tag{B.4}$$

with

$$\|w\|^2 = \cos^2(\varphi_1) + \frac{\sin^2(\varphi_1)}{\tan^2(\varphi_2)} - \frac{\sin(2\varphi_1)}{\tan(\varphi_2)}\cos(\psi_1 - \psi_2) . \tag{B.5}$$

[1]This appendix has been published, along with Chapter 2 and Appendix A in:
de Hoop, M. V., Malcolm, A. E. and Le Rousseau, J. H. 2003. Seismic wavefield 'continuation' in the single scattering approximation: A framework for dip and azimuth moveout *Can. Appl. Math. Q.*, **10**, 199-238.

The angle $\tilde{\theta_1}$ is defined in Figure B.1 and is given by

$$\tilde{\theta_1} = \text{acos}\left(\frac{\tilde{\alpha}_1 \cdot w}{\|w\|}\right) . \tag{B.6}$$

With $\tilde{\theta_1}$ we derive the zero-offset time t_0 using (A.17),

$$t_0 = \frac{t_1 \sin(\tilde{\theta_1}) \sin(\hat{\theta_1})}{(\sin(\tilde{\theta_1}) + \sin(\hat{\theta_1})) \sin(\theta_{0;1})} . \tag{B.7}$$

We now rotate this DMO ray geometry about the Ξ axis to obtain the desired azimuthal orientation. We have chosen our coordinates such that this orientation coincides with the 1 axis, which implies that $w := (1, 0, 0)$. We determine the time t_2 by applying inverse DMO to the rotated geometry. Thus,

$$\theta_{0;2} = \text{acos}(\Xi \cdot (1, 0, 0)) = \text{acos}(\cos(\varphi_2) \cos(\psi_2)) , \tag{B.8}$$

and it follows that

$$t_2 = \tilde{t}_2 + \hat{t}_2 = \frac{t_0(\sin(\tilde{\theta_2}) + \sin(\hat{\theta_2})) \sin(\theta_{0;2})}{\sin(\tilde{\theta_2}) \sin(\hat{\theta_2})} . \tag{B.9}$$

Appendix C

Table of Symbols for Appendix D, E, F and Chapter 3[1]

Table C.1. Table of symbols. In general quantities with a sub or super-script s refer to source quantities, with r to receiver quantities. Bold symbols are vectors and an underlined symbol is a function of other variables.

symbol	in or near	meaning
θ	p. 37	scattering angle
s	p. 38	source position
a	p. 38	rightmost surface position in Figure 3.3
d_0	p. 40, Figure 3.2	distance from source to zero offset surface position
$\mathbf{x}$	p. 40, Figure 3.2	subsurface position
$\boldsymbol{\xi}$	p. 40, Figure 3.2	migration dip
z, $\mathbf{z}$	p. 40, Figure 3.2	surface position of zero-offset ray
T	p. 41	two-way time
t_s	p. 41 Figure 3.2	time along source ray
t_0	p. 41 Figure 3.2	one-way zero-offset time
θ_s	p. 41, Figure 3.2	angle between surface and source ray
φ_α	p. 46	spherical angle measured clockwise from x
ψ_α	p. 46	spherical angle measured downward from $x-y$ plane
θ_0	Eqn D.2, Figure 3.2	angle from zero-offset ray to surface
y	p. 140	midpoint
h, $\mathbf{h}$	Eqn. D.8	half-offset
F	Eqn. E.1	modeling operator
δc	Eqn. E.1	velocity contrast
c	Eqn. E.1	smooth background velocity
$G(\mathbf{r}, t, \mathbf{x})$	Eqn. E.1	Green's function for singularities from $\mathbf{x}$ to $\mathbf{r}$ in time t
X	Eqn. E.1	set of subsurface scattering points

[1]This appendix has been published, along with Chapter 3 and Appendices D, E, and F as: Malcolm, A. E., de Hoop, M. V. and Le Rousseau, J. H. 2005. The applicability of DMO/AMO in the presence of caustics. *Geophysics* **70 S1-S17.**

Appendix D

Impulse Responses[1]

D.1 DMO

For a constant-velocity model one can derive, in closed form, the shape of the angle-domain DMO impulse response. To do this, we fix the source position, s, the scattering angle, θ, and the two-way traveltime T, and compute the output time t_0 and the source to zero-offset distance, d_0. We use the notation defined in Figure 3.2.

Since we assume the traveltime to be fixed and the medium to have a constant velocity (straight rays), we have the following three relations (see Figure 3.2):

$$\theta_r = \pi - \theta - \theta_s \tag{D.1}$$

$$\theta_0 = \pi - \frac{\theta}{2} - \theta_s \tag{D.2}$$

$$T = t_s + t_r. \tag{D.3}$$

From the law of sines we derive the following relationships:

$$t_0 = \frac{T \sin\theta_s \sin\theta_r}{(\sin\theta_s + \sin\theta_r)\sin\theta_0} \tag{D.4}$$

$$d_0 = \frac{vT \sin\theta_r \sin(\frac{\theta}{2})}{(\sin\theta_s + \sin\theta_r)\sin\theta_0}, \tag{D.5}$$

which gives both the one-way zero-offset travel time t_0 and the distance d_0 between the source and the zero-offset source/receiver as a function of the source angle θ_s. Figure 3.5 shows this impulse response (t_0 as a function of d_0). In these expressions, θ and T are fixed, while the other angles vary.

Although equations (D.4), (D.5) are given in terms of θ_s, θ_r and θ_0, it is possible to fully determine both t_0 and d_0 in terms of only two angles. Using equation (D.1) to substitute for θ_r and (D.2) for θ_s, equation (D.4) simplifies to

$$t_0 = \frac{\sin^2\theta_0 - \sin^2(\frac{\theta}{2})}{2\sin^2\theta_0 \cos(\frac{\theta}{2})}. \tag{D.6}$$

[1]This appendix has been published, along with Chapter 3 and Appendices C, E, and F as:
Malcolm, A. E., de Hoop, M. V. and Le Rousseau, J. H. 2005. The applicability of DMO/AMO in the presence of caustics. *Geophysics* **70 S1-S17.**

Appendix E

On Amplitudes[1]

E.1 Modeling

We begin with the kernel of the Born modeling operator (de Hoop *et al.* (2003a); equation 7),

$$F[\delta c](\mathbf{r},\mathbf{s},t) = \int_X \int_0^t G(\mathbf{r},t-t',\mathbf{x})\, 2c^{-3}(\mathbf{x})\partial_{t'}^2 \delta c(\mathbf{x}) G(\mathbf{x},t',\mathbf{s})\, \mathrm{d}t'\mathrm{d}\mathbf{x}. \tag{E.1}$$

In the above equation, $G(\mathbf{r},t-t',\mathbf{x})$ is the Green's function for the ray from the scattering point to the receiver, and $2\,c^{-3}(\mathbf{x})\,\partial_{t'}^2\,\delta c(\mathbf{x})\,G(\mathbf{x},t',\mathbf{s})$ is the contrast source at the scattering point, generated by the true source at the surface. The background velocity is denoted by $c(\mathbf{x})$, $\delta c(\mathbf{x})$ is the velocity perturbation that contains the locations of the reflectors, and X is the set of scattering points in the subsurface. The notation $F[\delta c](\mathbf{r},\mathbf{s},t)$ indicates that the operator F acts on the perturbation δc, with the result being dependent on the variables $(\mathbf{r},\mathbf{s},t)$. We use the full Born theory; the amplitudes will change if the Kirchhoff approximation is used in place of Born. For example Biondi *et al.* (1998) use the Zhang-Black Black *et al.* (1993) amplitudes, which are different from those used in this paper. The amplitude in equation (7.2.1) of Bleistein *et al.* (2000) should be compared with that of equation (E.1). Equation (20) of Black *et al.* (1993) gives the comparison between Black's DMO amplitudes and Hale's DMO amplitudes. Equation (7.6.36) of Bleistein *et al.* (2000) compares the Bleistein amplitudes with those of Hale (1991).

From this point onward we will assume that the background velocity, c, of the medium is constant. Thus, the Green's function is given by

$$G(\mathbf{y},t,\mathbf{x}) = \frac{1}{2\pi}\int \frac{1}{4\pi|\mathbf{y}-\mathbf{x}|} e^{-i\omega(t-T(\mathbf{x},\mathbf{y}))}\mathrm{d}\omega, \tag{E.2}$$

where $T(\mathbf{x},\mathbf{y}) = \frac{|\mathbf{x}-\mathbf{y}|}{c}$ is the traveltime between $\mathbf{x}$ and $\mathbf{y}$. The Green's function, as written in equation (E.2), is the kernel of a Fourier integral operator Stolk & de Hoop (2002). To

[1]This appendix has been published, along with Chapter 3 and Appendices C, D, and F as:
Malcolm, A. E., de Hoop, M. V. and Le Rousseau, J. H. 2005. The applicability of DMO/AMO in the presence of caustics. *Geophysics* **70 S1-S17.**

where $A(\mathbf{s},\mathbf{x},\mathbf{r}) = \frac{1}{16\pi^2|\mathbf{x}-\mathbf{s}||\mathbf{x}-\mathbf{r}|}$ and $T(\mathbf{r},\mathbf{x},\mathbf{s}) = T(\mathbf{x},\mathbf{r}) + T(\mathbf{s},\mathbf{x})$. The integration in t results in $2\pi\delta(\omega-\omega')$. Performing both this integration and that in ω' results in

$$N[\delta c](\mathbf{y}) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{d}\omega \int_{O_r\times O_s}\int_X \left[\omega^4\, 4c^{-6} A(\mathbf{s},\mathbf{x},\mathbf{r})A(\mathbf{s},\mathbf{y},\mathbf{r}) \right. \\ \left. e^{i\omega(T(\mathbf{r},\mathbf{x},\mathbf{s})-T(\mathbf{r},\mathbf{y},\mathbf{s}))}\right] \delta c(\mathbf{x})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{s}\, \mathrm{d}\mathbf{r}. \tag{E.7}$$

We now change variables to the ray directions at the scattering point ($\boldsymbol{\alpha}^s$ and $\boldsymbol{\alpha}^r$). We will use subscripts $1,2,3$ to denote the components of a vector in the x_1, x_2, x_3 direction respectively. To change from $(\mathbf{s},\mathbf{r})$ to $(\boldsymbol{\alpha}^s,\boldsymbol{\alpha}^r)$, we compute the Jacobian

$$\frac{\partial(\alpha_1^s,\alpha_2^s)}{\partial(s_1,s_2)}\frac{\partial(\alpha_1^r,\alpha_2^r)}{\partial(r_1,r_2)} = \frac{y_3^2}{|\mathbf{y}-\mathbf{s}|^4}\frac{y_3^2}{|\mathbf{y}-\mathbf{r}|^4} \tag{E.8}$$

since

$$\alpha_{1,2}^s = \frac{y_{1,2}-s_{1,2}}{|\mathbf{y}-\mathbf{s}|} \quad \text{and} \quad \alpha_{1,2}^r = \frac{y_{1,2}-r_{1,2}}{|\mathbf{y}-\mathbf{r}|}.$$

Performing this change of variables we obtain

$$N[\delta c](\mathbf{y}) \simeq \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{d}\omega \int_{S_s^2\times S_r^2}\int_X \left[\omega^4\, 4c^{-6} A(\underline{\mathbf{s}},\mathbf{y},\underline{\mathbf{r}})^2 \frac{|\mathbf{y}-\underline{\mathbf{s}}|^4|\mathbf{y}-\underline{\mathbf{r}}|^4}{y_3^4} \right. \\ \left. e^{i\omega(T(\underline{\mathbf{r}},\mathbf{x},\underline{\mathbf{s}})-T(\underline{\mathbf{r}},\mathbf{y},\underline{\mathbf{s}}))}\right] \delta c(\mathbf{x})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\boldsymbol{\alpha}^s\, \mathrm{d}\boldsymbol{\alpha}^r. \tag{E.9}$$

Underlined symbols are used to indicate that a variable is a function of other variables and S_s^2 (S_r^2) is the unit sphere on which the direction of the source (receiver) ray lies. We have expanded $A(\underline{\mathbf{s}},\mathbf{x},\underline{\mathbf{r}})$ about $\mathbf{y}$, using that $\mathbf{x}\approx\mathbf{y}$ at stationarity in $\boldsymbol{\alpha}^r$ and $\boldsymbol{\alpha}^s$ (see de Hoop & Brandsberg-Dahl (2000, p. 553) and Beylkin (1985)).

Expanding $T(\underline{\mathbf{r}},\mathbf{x},\underline{\mathbf{s}})$ to first order in a Taylor series about $\mathbf{y}$, we can write the phase as

$$\boldsymbol{\Gamma}\cdot(\mathbf{x}-\mathbf{y}), \tag{E.10}$$

where $\boldsymbol{\Gamma} = \nabla_{\mathbf{y}} T(\underline{\mathbf{r}},\mathbf{y},\underline{\mathbf{s}})$. We will also scale the ω variable by $|\boldsymbol{\Gamma}|^{-1}$ and introduce $\boldsymbol{\nu} = \boldsymbol{\Gamma}/|\boldsymbol{\Gamma}|$. Note that $\boldsymbol{\nu}$ corresponds to the migration dip. The expression thus becomes

$$N[\delta c](\mathbf{y}) \simeq \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{d}\omega \int_{S_s^2\times S_r^2}\int_X \left[4c^{-6}\omega^4 \mu_{LS}(\underline{\mathbf{r}},\boldsymbol{\alpha}^r,\underline{\mathbf{s}},\boldsymbol{\alpha}^s,\mathbf{y}) e^{i\omega\boldsymbol{\nu}\cdot(\mathbf{x}-\mathbf{y})}\right] \delta c(\mathbf{x})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\boldsymbol{\alpha}^s\, \mathrm{d}\boldsymbol{\alpha}^r, \tag{E.11}$$

letting $\mu_{LS}(\underline{\mathbf{r}},\boldsymbol{\alpha}^r,\underline{\mathbf{s}},\boldsymbol{\alpha}^s,\mathbf{y}) = A(\underline{\mathbf{s}},\mathbf{y},\underline{\mathbf{r}})^2 \frac{\partial(\mathbf{s},\mathbf{r})}{\partial(\boldsymbol{\alpha}^s,\boldsymbol{\alpha}^r)}|\boldsymbol{\Gamma}|^{-5}$.

We now change variables again, to the scattering angle θ, scattering azimuth ψ and migration dip $\boldsymbol{\nu}$ (these quantities are defined in terms of $\boldsymbol{\alpha}^s$ and $\boldsymbol{\alpha}^r$ by de Hoop & Brandsberg-Dahl (2000, eqn (127))). This introduces a Jacobian, $\frac{\partial(\boldsymbol{\alpha}^s,\boldsymbol{\alpha}^r)}{\partial(\boldsymbol{\nu},\theta,\psi)}$, which is computed for the general case by Burridge & Beylkin (1988) and for the homogeneous, isotropic case by Bur-

in a Taylor series about $\mathbf{x} = \mathbf{y}$, we can write the phase of the above operator as

$$i\omega\boldsymbol{\nu}\cdot(\mathbf{x}-\mathbf{y}) - i\omega'\nabla_y T(\mathbf{r},\mathbf{y},\mathbf{s})\cdot(\mathbf{x}-\mathbf{y}) - i\omega' T(\mathbf{r},\mathbf{y},\mathbf{s}).$$

We can now perform the integration in $\mathbf{x}$ (noting that $\mathbf{x}\approx\mathbf{y}$ at stationarity), giving

$$(2\pi)^3\delta(\omega\boldsymbol{\nu} - \omega'\nabla_y T(\mathbf{r},\mathbf{y},\mathbf{s})) = \frac{(2\pi)^3}{\omega'^2|\boldsymbol{\Gamma}|^2}\delta(\omega-\omega'|\Gamma|)\delta(\boldsymbol{\nu}-\mathbf{v}), \tag{E.16}$$

where $\mathbf{v} = \nabla_y T(\mathbf{r},\mathbf{y},\mathbf{s})/|\nabla_y T(\mathbf{r},\mathbf{y},\mathbf{s})|$. This allows us to integrate out both the $\boldsymbol{\nu}$ and ω variables resulting in

$$\begin{aligned} PF^*[u](\mathbf{y}) &= \frac{1}{8\pi^2}\int_{O_r\times O_s}\frac{1}{|\boldsymbol{\Gamma}|^2}(\Lambda(\mathbf{y},\mathbf{v}))^{-1}\frac{1}{2}c^3A(\mathbf{s},\mathbf{y},\mathbf{r}) \\ &\quad \frac{1}{\pi}\mathrm{Re}\left[\int_{\mathbb{R}_+} u(\mathbf{s},\mathbf{r},\omega')e^{-i\omega' T(\mathbf{r},\mathbf{y},\mathbf{s})}\,\mathrm{d}\omega'\right]\mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{r}. \end{aligned} \tag{E.17}$$

Changing variables from $(\mathbf{s},\mathbf{r})$ to $(\boldsymbol{\alpha}^s,\boldsymbol{\alpha}^r)$ as before and substituting $\mu_{LS}(\underline{\mathbf{r}},\boldsymbol{\alpha}^r,\underline{\mathbf{s}},\boldsymbol{\alpha}^s,\mathbf{y})$ gives

$$\begin{aligned} I[u](\mathbf{y}) = PF^*[u](\mathbf{y}) &= \frac{1}{16\pi^2}\int_{S_s^2\times S_r^2}\frac{c^3|\boldsymbol{\Gamma}|^3(\Lambda(\mathbf{y},\mathbf{v}))^{-1}\mu_{LS}(\underline{\mathbf{r}},\boldsymbol{\alpha}^r,\underline{\mathbf{s}},\boldsymbol{\alpha}^s,\mathbf{y})}{A(\underline{\mathbf{s}},\mathbf{y},\underline{\mathbf{r}})} \\ &\quad \frac{1}{\pi}\mathrm{Re}\left[\int_{\mathbb{R}_+} e^{-i\omega' T(\underline{\mathbf{r}},\mathbf{y},\underline{\mathbf{s}})}u(\underline{\mathbf{s}},\underline{\mathbf{r}},\omega')\,\mathrm{d}\omega'\right]\mathrm{d}\boldsymbol{\alpha}^s\,\mathrm{d}\boldsymbol{\alpha}^r. \end{aligned} \tag{E.18}$$

We can then perform the integration with respect to ω', obtaining

$$I[u](\mathbf{y}) = \frac{1}{16\pi^2}\int_{S_s^2\times S_r^2}\frac{c^3|\boldsymbol{\Gamma}|^3(\Lambda(\mathbf{y},\mathbf{v}))^{-1}\mu_{LS}(\underline{\mathbf{r}},\boldsymbol{\alpha}^r,\underline{\mathbf{s}},\boldsymbol{\alpha}^s,\mathbf{y})}{A(\underline{\mathbf{s}},\mathbf{y},\underline{\mathbf{r}})}u(\underline{\mathbf{s}},\underline{\mathbf{r}},T(\underline{\mathbf{s}},\mathbf{y},\underline{\mathbf{r}}))\,\mathrm{d}\boldsymbol{\alpha}^s\,\mathrm{d}\boldsymbol{\alpha}^r. \tag{E.19}$$

In equation (E.18), we recognize the form of equation (27) of Miller *et al.* (1987), by noting that

$$|\boldsymbol{\Gamma}| = \frac{2\cos\left(\frac{\theta}{2}\right)}{c},$$

and setting $\mu_{LS} = 1$. The two equations are not exactly the same however; our equation contains a factor, $\frac{\sin\theta}{\Lambda}$. The $\sin\theta$ is introduced by changing to the same coordinates as used by Miller et al. The factor Λ^{-1} is present since we construct the least squares inverse and is equal to $\int_{E_\theta}\sin\theta\,\mathrm{d}\theta$.

We now continue with the stationary phase analysis in (σ_1', σ_2'), following Bleistein *et al.* (2000). The stationary phase condition

$$\partial_{\sigma_{1,2}'}\Phi = \nabla_{\underline{x}}T \cdot \frac{\partial \underline{\mathbf{x}}}{\partial \sigma_{1,2}'} = 0,$$

gives that $\nabla_{\underline{x}}T$ is parallel to $\nabla_{\underline{x}}T'$ or equivalently, the unit normal to the input angle isochron $\hat{\mathbf{n}}'$ is equal to the unit normal, $\hat{\mathbf{n}}$, to the output standard isochron. The Hessian of second derivatives is

$$\mathbf{H} = \det\left[\frac{\partial^2 T}{\partial\sigma_1'\partial\sigma_2'}\right] = |\nabla_{\underline{x}}T|^2 det[\mathbf{B}_{kl}' - \mathbf{B}_{kl}]$$

where we have used the relation

$$\frac{\partial^2 T}{\partial\sigma_l'\partial\sigma_k'} = \frac{\partial^2 T}{\partial\underline{x}_i\partial\underline{x}_j}\frac{\partial\underline{x}_i}{\partial\sigma_k'}\frac{\partial\underline{x}_j}{\partial\sigma_l'} + \frac{\partial T}{\partial\underline{x}_i}\frac{\partial^2\underline{x}_i}{\partial\sigma_k'\partial\sigma_l'} \tag{E.25}$$

and

$$\frac{\partial T}{\partial\underline{x}_i}\frac{\partial^2\underline{x}_i}{\partial\sigma_k'\partial\sigma_l'} = \hat{\mathbf{n}}\cdot\frac{\partial^2\underline{x}}{\partial\sigma_k'\partial\sigma_l'}|\nabla_{\underline{x}}T| = \mathbf{B}_{kl}'|\nabla_{\underline{x}}T(\underline{\mathbf{r}}', \underline{\mathbf{x}}, \mathbf{s}')|, \tag{E.26}$$

using the fact that $\hat{\mathbf{n}}' = \hat{\mathbf{n}}$ at stationarity. $\mathbf{B}_{kl}$ is the second fundamental tensor, which is a rotation of the matrix of principle curvatures for the level sets (standard isochrons) of the function T. $\mathbf{B}_{kl}'$ denotes $\mathbf{B}_{kl}$ in the input parameters (i.e., a rotation of the matrix of principle curvatures for the level sets (angle isochrons) of T'). Equation (E.26) gives $\mathbf{B}_{kl}'$ because the (σ_1', σ_2') are coordinates on the angle isochrons (level sets of T'). The transformation of the first term in equation (E.25) to $\mathbf{B}_{kl}$ is explained in section 7.7 of Bleistein *et al.* (2000).

We note at this point that the preceding stationary phase analysis is applicable only when the Hessian, $\mathbf{H}$, is nonzero. One example of when $\mathbf{H} = 0$ is when the input and output configurations are the same, in which case $\mathbf{B}_{kl} = \mathbf{B}_{kl}'$. Assuming that $\mathbf{H} \neq 0$ we may write the final continuation operator as

$$\begin{aligned} FI[u](\mathbf{s}, \mathbf{r}, t) \quad = & \int_{E_{\theta'}\times E_{\psi'}} \frac{-1}{2\pi}\int_{\mathbb{R}_+}\int_{\mathbb{R}}\frac{1}{4\sqrt{2\pi}}\int_{S'}\omega\sin\theta'|\mathbf{\Gamma}|^3(\Lambda(\underline{\mathbf{x}}_0, \mathbf{v}))^{-1} \\ & \mu_{LS}(\underline{\mathbf{r}}', \underline{\boldsymbol{\alpha}}^{r'}, \mathbf{s}', \underline{\boldsymbol{\alpha}}^{s'}, \underline{\mathbf{x}}_0)|\partial_{\sigma_1}\underline{\mathbf{x}}\wedge\partial_{\sigma_2}\underline{\mathbf{x}}|\left|\frac{\partial(\boldsymbol{\nu}', \psi', \theta')}{\partial(\mathbf{s}', \psi', \theta')}\right|\frac{A(\mathbf{s}, \underline{\mathbf{x}}_0, \mathbf{r})}{A(\mathbf{s}', \underline{\mathbf{x}}_0, \underline{\mathbf{r}}')\sqrt{\mathbf{H}}|\nabla_{\underline{x}_0}T'|} \\ & u(\mathbf{s}', \underline{\mathbf{r}}', T(\underline{\mathbf{r}}', \underline{\mathbf{x}}_0, \mathbf{s}'))e^{i\omega(T(\mathbf{r}, \underline{\mathbf{x}}_0, \mathbf{s})-t)+i\frac{\pi}{4}sig(\mathbf{H})}\mathrm{d}\mathbf{s}'\,\mathrm{d}\omega\,\mathrm{d}\tau'\,\mathrm{d}\theta'\,\mathrm{d}\psi'. \end{aligned} \tag{E.27}$$

In this expression, $\underline{\mathbf{x}}_0$ is the stationary point (in σ_1' and σ_2', i.e., $\underline{\mathbf{x}}_0 = \underline{x}(\sigma_1'^0, \sigma_2'^0, \tau')$, $\nabla_{\underline{x}_0}T' = \nabla_{\underline{x}}T(\underline{\mathbf{r}}', \underline{\mathbf{x}}, \mathbf{s}')|_{\underline{x}=\underline{x}_0}$, and $sig(\mathbf{H})$ is not evaluated as this will depend on the exact input and output configuration.

Appendix F

Properties of FF^* for Common Offset[1]

A formal definition of a pseudodifferential operator is beyond the scope of this paper; we refer the reader to either Treves (1980b) or Appendix A of de Hoop *et al.* (2003b) for the mathematical details. Roughly speaking, however, an operator is pseudodifferential if it can be written in the form of a forward and inverse Fourier transform along with a multiplication by an amplitude in the Fourier domain. In general FF^* is not a pseudodifferential operator. In certain cases, however, when the input and output offsets are constrained to be equal for example, FF^* does have this property. In this particular case, we write the kernel of FF^* as that of a pseudodifferential operator and extract its amplitude behavior.

The phase of FF^* is given by (cf. equation (E.20))

$$\Phi = -\omega(t - T(\mathbf{r}, \mathbf{x}, \mathbf{s})) + \omega'(t' - T(\mathbf{r}', \mathbf{x}, \mathbf{s}')), \tag{F.1}$$

in which $\mathbf{x}$, ω and ω' are identified as phase variables. Subjecting equation (F.1) to the common offset condition, the travel time, T, and phase function, Φ, can be rewritten in terms of midpoints, $\mathbf{y} = \frac{\mathbf{s}+\mathbf{r}}{2}$, $\mathbf{y}' = \frac{\mathbf{s}'+\mathbf{r}'}{2}$ and offset, $\mathbf{h} = \frac{\mathbf{s}-\mathbf{r}}{2} = \frac{\mathbf{s}'-\mathbf{r}'}{2}$, as

$$\Phi = -\omega(t - T(\mathbf{x}, \mathbf{h}, \mathbf{y})) + \omega'(t' - T(\mathbf{x}, \mathbf{h}, \mathbf{y}')). \tag{F.2}$$

Treatment of phase variables

The representation of the kernel of the operator FF^* will contain integrations over $\mathbf{x}$, ω and ω'. The application of this operator to input data will result in integrations over $\mathbf{y}'$ and t', the input midpoint and time (but not over $\mathbf{h}$ as this is fixed). For this phase to be that of a pseudodifferential operator, we must be able to write it in the form

$$\omega'(t' - t) + \boldsymbol{\eta}' \cdot (\mathbf{y} - \mathbf{y}'), \tag{F.3}$$

where $\omega' t'$ and $-\boldsymbol{\eta}' \cdot \mathbf{y}'$ are associated with Fourier transforms and $-\omega' t$ and $\boldsymbol{\eta}' \cdot \mathbf{y}$ are associated with inverse Fourier transforms. Thus the oscillatory integral representation of

[1]This appendix has been published, along with Chapter 3 and Appendices C, D, and E as:
Malcolm, A. E., de Hoop, M. V. and Le Rousseau, J. H. 2005. The applicability of DMO/AMO in the presence of caustics. *Geophysics* **70 S1-S17.**

for given $(\mathbf{y}, t)$ and $\mathbf{h}$ still fixed (note that the $\underline{x}_3^0$ dependence contains implicitly a dependence on x_1, x_2). With this definition, equation (F.13) gives (F.3). To compute the Jacobian associated with the transformation $(x_1, x_2) \to (\eta_1', \eta_2')$, we compute

$$\mathbf{J} = \frac{\partial(\eta_1', \eta_2')}{\partial(x_1, x_2)} = \frac{\omega'^2}{c^2} \begin{vmatrix} \partial_{x_1}\eta_1' + \partial_{x_3}\eta_1' \dfrac{\partial \underline{x}_3^0}{\partial x_1} & \partial_{x_2}\eta_1' + \partial_{x_3}\eta_1' \dfrac{\partial \underline{x}_3^0}{\partial x_2} \\ \partial_{x_1}\eta_2' + \partial_{x_3}\eta_2' \dfrac{\partial \underline{x}_3^0}{\partial x_1} & \partial_{x_2}\eta_2' + \partial_{x_3}\eta_2' \dfrac{\partial \underline{x}_3^0}{\partial x_2} \end{vmatrix}. \tag{F.15}$$

Introducing the vectors $\mathbf{p} = (\mathbf{x} - \mathbf{y} - \mathbf{h})$ and $\mathbf{q} = (\mathbf{x} - \mathbf{y} + \mathbf{h})$, and orienting the system of coordinates so that $\mathbf{h} = (h_1, 0, 0)$, we have, upon substituting $T = \frac{|\mathbf{q}|+|\mathbf{p}|}{c}$,

$$J = \frac{\omega'^2}{c^2} \begin{vmatrix} \frac{-|\mathbf{q}|^2+q_1^2}{|\mathbf{q}|^3} + \frac{-|\mathbf{p}|^2+p_1^2}{|\mathbf{p}|^3} + \left(\frac{q_1q_3}{|\mathbf{q}|^3} + \frac{p_1p_3}{|\mathbf{p}|^3}\right)\frac{\partial \underline{x}_3^0}{\partial x_1} & \frac{q_1q_2}{|\mathbf{q}|^3} + \frac{p_1p_2}{|\mathbf{p}|^3} + \left(\frac{q_1q_3}{|\mathbf{q}|^3} + \frac{p_1p_3}{|\mathbf{p}|^3}\right)\frac{\partial \underline{x}_3^0}{\partial x_2} \\ \frac{q_1q_2}{|\mathbf{q}|^3} + \frac{p_1p_2}{|\mathbf{p}|^3} + \left(\frac{q_1q_3}{|\mathbf{q}|^3} + \frac{p_1p_3}{|\mathbf{p}|^3}\right)\frac{\partial \underline{x}_3^0}{\partial x_1} & \frac{-|\mathbf{q}|^2+q_2^2}{|\mathbf{q}|^3} + \frac{-|\mathbf{p}|^2+p_2^2}{|\mathbf{p}|^3} + \left(\frac{q_2q_3}{|\mathbf{q}|^3} + \frac{p_2p_3}{|\mathbf{p}|^3}\right)\frac{\partial \underline{x}_3^0}{\partial x_2} \end{vmatrix}. \tag{F.16}$$

We solve equation (F.4) for $\underline{x}_3^0$,

$$\underline{x}_3^0 = \frac{1}{2ct}\sqrt{-2p_1^2c^2t^2 - 4q_2^2c^2t^2 + p_1^4 + q_1^4 - 2p_1^2q_1^2 - 2q_1^2c^2t^2 + c^4t^4}, \tag{F.17}$$

noting that $q_2 = p_2$ and $q_3 = p_3 = \underline{x}_3^0$ and taking the positive square-root as x_3 is constrained to be greater than zero. We then compute

$$\frac{\partial \underline{x}_3^0}{\partial x_1} = \frac{k_1}{p_3}, \qquad \text{where} \quad k_1 = \frac{-(p_1+q_1)(q_2^2 + x_3^2 + q_1q_2 + |\mathbf{q}||\mathbf{p}|)}{c^2t^2} \tag{F.18}$$

$$\frac{\partial \underline{x}_3^0}{\partial x_2} = \frac{-q_2}{p_3}, \tag{F.19}$$

where we have again used the stationarity condition (F.4) in the numerator of (F.18). Substituting (F.18) and (F.19) into (F.16) reduces the Jacobian matrix to

$$\begin{vmatrix} -\frac{1}{|\mathbf{q}|} + \frac{q_1^2+q_1k_1}{|\mathbf{q}|^3} - \frac{1}{|\mathbf{p}|} + \frac{p_1^2+p_1k_1}{|\mathbf{p}|^3} & 0 \\ * & -\frac{1}{|\mathbf{q}|} - \frac{1}{|\mathbf{p}|}, \end{vmatrix} \tag{F.20}$$

where the $*$ represents a term that we do not need to compute as it is multiplied by 0 in the Jacobian. Thus we have for the Jacobian

$$\mathbf{J} = \frac{\omega'^2}{c^2}\left(\frac{1}{|\mathbf{p}|^2} + \frac{1}{|\mathbf{q}|^2} - \frac{p_1^2 + k_1p_1}{|\mathbf{p}|^3}\left(\frac{1}{|\mathbf{p}|} + \frac{1}{|\mathbf{q}|}\right) - \frac{q_1^2 + k_1q_1}{|\mathbf{q}|^3}\left(\frac{1}{|\mathbf{p}|} + \frac{1}{|\mathbf{q}|}\right) + \frac{2}{|\mathbf{q}||\mathbf{p}|}\right). \tag{F.21}$$

After some algebra, it can be shown that

$$-\frac{p_1^2 + k_1p_1}{|\mathbf{p}|^3}\left(\frac{1}{|\mathbf{p}|} + \frac{1}{|\mathbf{q}|}\right) + \frac{1}{|\mathbf{q}||\mathbf{p}|} = \frac{\mathbf{p}\cdot\mathbf{q}}{|\mathbf{p}|^3|\mathbf{q}|}, \tag{F.22}$$

where $\underline{x}_1$ and $\underline{x}_2$ are underlined as they are now functions of (η_1', η_2'). It can be shown that the amplitude of this operator is the reciprocal of that of Miller *et al.* (1987) equation (27), upon changing variables in that equation from $\boldsymbol{\xi}$ to $\mathbf{y}$, using the Jacobian given in de Hoop *et al.* (1999). This amplitude construction can be used to correctly account for the amplitudes in offset continuation. For that case, one would simply fix both the output offset, $\mathbf{h} = \frac{\mathbf{s}-\mathbf{r}}{2}$, and the input offset, $\mathbf{h}' = \frac{\mathbf{s}'-\mathbf{r}'}{2}$ and follow the same procedure as for the case $\mathbf{h} = \mathbf{h}'$.

Appendix G

A Comparison of the Lippmann-Schwinger and Bremmer Series

The reason that internal multiples cause artifacts in imaging is that they violate the single scattering assumption, also known as the Born approximation. Two series to move beyond this assumption have been proposed in the literature, the Bremmer series (Bremmer, 1951) and the Lippmann-Schwinger series (Lippmann, 1956). Both of these series are designed to approximate the Green's function or Green's operator solving the wave equation. The wave equation is given by

$$c(\mathbf{x})^{-2}\partial_t^2 u(\mathbf{x},t) - \partial_x^2 u(\mathbf{x},t) - \partial_y^2 u(\mathbf{x},t) - \partial_z^2 u(\mathbf{x},t) = f(\mathbf{x},t)\,, \tag{G.1}$$

where c is the velocity, $\mathbf{x} = (x, y, z)$ is a spatial position vector and f is the source of waves. The Green's function, g, solves the wave equation in the sense that

$$\begin{aligned} c(\mathbf{x})^{-2}\partial_t^2 g(\mathbf{x},t-t',\mathbf{x}') - \partial_x^2 g(\mathbf{x},t-t',\mathbf{x}') - \partial_y^2 g(\mathbf{x},t-t',\mathbf{x}') - \partial_z^2 g(\mathbf{x},t-t',\mathbf{x}') \\ = \delta(\mathbf{x}-\mathbf{x}')\delta(t-t')\,, \end{aligned} \tag{G.2}$$

where δ is the dirac-delta function. Equation (G.1) can also be written in operator form, defining the wave operator, L, as

$$Lu = f\,. \tag{G.3}$$

This allows the definition of the Green's operator, G, via

$$LG = I\,, \tag{G.4}$$

where I is the identity operator. The Green's function is the kernel of the Green's operator so

$$[G\phi](\mathbf{x},t) = \int \mathrm{d}\mathbf{x}' \int \mathrm{d}t' g(\mathbf{x},t-t',\mathbf{x}')\phi(\mathbf{x}',t')\,, \tag{G.5}$$

where ϕ is the test function on which the operator acts. Thus the Green's operator propagates the wavefield, ϕ, from the position $(\mathbf{x}', t')$ to the position $(\mathbf{x}, t)$, which is why it is also referred to as a propagator. From the result that

$$g * f = u\,, \tag{G.6}$$

Lippmann-Schwinger series

$$G = \left(\sum_{n=0}^{\infty}(-G_0V)^n\right)G_0 = G_0 - G_0VG_0 + G_0VG_0VG_0 - \dots . \tag{G.13}$$

To my knowledge there are no references that investigate the choice of background model to ascertain whether or not $\|G_0V\| < 1$, meaning that this series is not, in general, convergent.

An advantage of the Lippmann-Schwinger series is that, unlike the Bremmer series, it has a corresponding inverse series that expresses difference operator, V, in terms of the Green's operator, which is directly related to the data. To express the medium contrast in terms of the data, assume the medium contrast can be decomposed into a series,

$$V = \sum_{j=1}^{\infty} V_j\,, \tag{G.14}$$

where V_j is of order j in the data. So, while the forward series expresses the Green's operator as a series with each term being of higher order in the medium contrast, the inverse series expresses the medium contrast in terms of a series in which each term is of higher order in the data.

This assumption is then substituted into (G.13), and terms of equal order in the data are equated, giving the following relations:

$$\begin{aligned} d &= G - G_0 = -G_0V_1G_0 && \text{(G.15)}\\ 0 &= -G_0V_2G_0 + G_0V_1G_0V_1G_0 && \text{(G.16)}\\ 0 &= -G_0V_3G_0 + G_0V_1G_0V_2G_0 + G_0V_2G_0V_1G_0 - G_0V_1G_0V_1G_0V_1G_0 && \text{(G.17)} \end{aligned}$$

etc..

Equation (G.15) serves to define the data (or the scattered field as in Weglein *et al.* (1997)). This difference is directly related to seismic data because the difference between G and G_0 is in the reflections from differences between the true and reference media. The terms in (G.15) are first-order in the data, those in (G.16) are second-order in the data, and so on. By substituting (G.16) into (G.17) and continuing to higher order in the series it is found that

$$G_0V_jG_0 = \underbrace{G_0V_1G_0V_1\cdots G_0V_1}_{j \text{ times}}G_0\,. \tag{G.18}$$

Since the Green's operators, G_0, are in the background medium, they are known and V_j can be estimated in this manner.

To estimate V_j from equation (G.18) requires an operator to estimate V_j from $G_0V_jG_0$, but this is the only inverse operator needed to estimate all of the V_j. The single-scattering case,

$$d = -G_0V_1G_0 \tag{G.19}$$

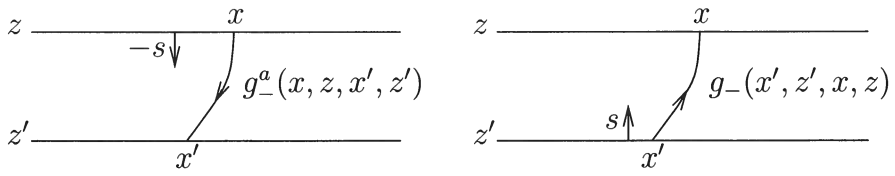


Figure G.1. Illustration of the reciprocity relation given in (G.26).

source constituents.

The Bremmer series is derived by considering the following derivative

$$\partial_z(g^a U) = \partial_z(g^a)U + g^a \partial_z U\,, \tag{G.23}$$

using the notation

$$g^a = \begin{pmatrix} g^a_+ & 0 \\ 0 & g^a_- \end{pmatrix} \tag{G.24}$$

for the matrix of adjoint Green's functions. The adjoint Green's functions are defined by taking the adjoint of the diagonalized wave equation

$$I\partial_z g^a + g^a \mathsf{B}^T = \delta(\mathbf{x}-\mathbf{x_0})\delta(z-z_0)\,; \tag{G.25}$$

where I is the identity operator. In this section $\mathbf{x} = (x, y)$ because the z dependence has been separated due to the up-down decomposition of the wavefield. The adjoint Green's function is related to the Green's function by

$$g^a(x,z,x',z') = -g(x',z',x,z)\,. \tag{G.26}$$

Recall that G_+ is the down-going propagator and G_- is the up-going propagator (with g_+ and g_- as their associated kernels). The adjoint operator propagates backward in time. Thus, G^a_+ propagates waves upward and G^a_- downward. The minus sign in (G.26) follows from reciprocity as in Remark 4.4.2. This is illustrated in Figure G.1.

Equation (G.23) can be expanded, by substituting (G.25) for the first term and (G.22) for the second, to

$$\begin{aligned} \partial_z(g^a U) &= (\delta(\mathbf{x}-\mathbf{x_0})\delta(z-z_0) + g^a\mathsf{B}^T)U + g^a(X - \mathsf{Q}(z)\partial_z\mathsf{Q}(z)^{-1}U - \mathsf{B}U) \\ &= \delta(\mathbf{x}-\mathbf{x_0})\delta(z-z_0)U + g^a X - g^a\mathsf{Q}(z)\partial_z\mathsf{Q}(z)^{-1}U\,, \end{aligned} \tag{G.27}$$

recalling that $\mathsf{B}^T = \mathsf{B}$ because the $B_+ = -B_-$ in the flux normalization. Defining G^a as the matrix of Green's operators associated with the matrix, g^a, of Green's functions and

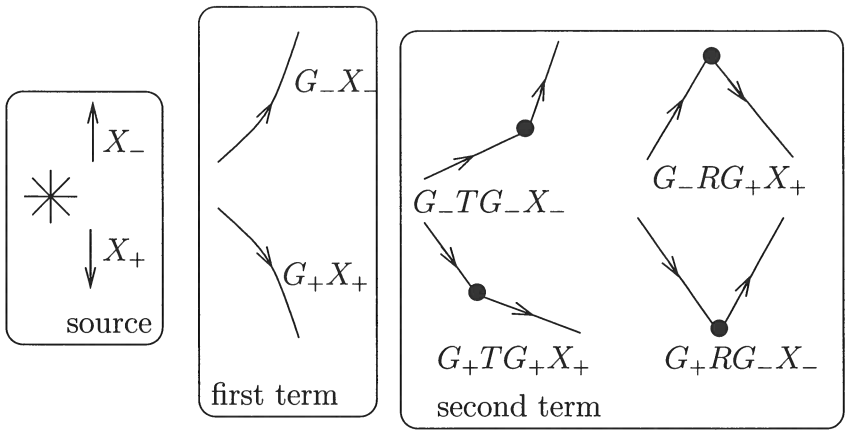


Figure G.2. Illustration of the first two terms of the Bremmer series.

G.3 The Hybrid Series

Chapter 4 describes a hybrid series between the Lippmann-Schwinger and Bremmer series. This is done in an attempt to capture the advantages of both series. The Lippmann-Schwinger series has the advantage that its form admits both a forward and inverse series. The advantage of the Bremmer series is that it is convergent and the splitting of the wavefield into its up- and down-going constituents fits naturally in a downward continuation or wavefield extrapolation migration. The hybrid series first separates the wavefield into its up- and down-going constituents as in the Bremmer series. This is done in a known background model. The diagonalization operators computed in the known background are then applied to the wavefield in the true medium, whose difference from the background medium is unknown, creating two equations in two different media, as in the Lippmann-Schwinger series.

Perhaps the most fundamental difference is in the writing of the series as a set of matrix equations rather than scalar operator equations. The Lippmann-Schwinger series is a series of operators applied to a single wavefield; each term of the series is of higher order in the contrast between the true and reference media, parameterized by the operator V. The hybrid series is a series of matrix operators applied to two wavefields; the up- and down-going constituents of the wavefield on which the operators of the Lippmann-Schwinger series acts. Each term of the series is of higher-order in the difference operator, which is now a matrix, and each term of the series also contains higher order scattering than does the previous term. The reflection and transmission operators of the hybrid series are different from those of the Bremmer series because the reflection operators in the Bremmer series are vertical derivatives of the medium whereas in the hybrid series these operators are based on differences between the true and background media.

The inverse hybrid series is constructed in exact parallel to the Lippmann-Schwinger

Appendix H

A comparison of different boundary conditions and edge effects in migration

The goal of this appendix is to show that boundary effects can be a source of significant artifacts in imaging. An understanding of the basics of wave-equation migration algorithms is assumed, although brief descriptions are given of the most important points.

Estimating multiples in the image space requires a better estimate of the image than is required for ordinary imaging. The reason for this is twofold. First, multiples have smaller amplitude than do primaries because they have reflected three times rather than once. Second, the process of migration focusses energy from primaires in time at $t = 0$ and in offset at $h = 0$, but energy from multiples is not correctly focussed, and thus does not make as strong a contribution to the image. The purpose of this appendix is to illustrate, primarily through examples, some of the imaging problems that need to be taken into account when imaging with multiples that may not be as important for standard imaging.

The migration algorithm used in Chapter 5 works by extrapolating the wavefield, sampled in midpoint, offset and time, in depth using a propagator provided by Total. The propagator begins with a standard Stolt phase shift (Stolt, 1978), followed by a lens correction to correct for lateral velocity variations and a wide angle correction using a tridiagonal matrix solve. The precise details of the propagator are not important. The essential point is that the propagator estimates the data at depth $z + \Delta z$ from the data at depth z. The zero time, zero offset imaging condition is then applied resulting in the final image. (The energy recorded at $h = 0$ and $t = 0$ is the reflected energy from a reflector at the current depth.) The entire algorithm works in the temporal frequency domain. The phase shift is applied in the spatial frequency or wavenumber domain. Each propagation step therefore requires Fourier transforms in both offset and wavenumber. To use the prime factor Fast Fourier Transform (FFT) the input data are padded with zeros to a length that is a product of the factors available to the FFT routine. (In this case, these factors are: 2,3,4,5,7,8,9,11,13 and 16.)

The model used throughout this appendix consists of a homogeneous velocity of 6 km/s with a low-velocity layer from 1.5 to 2.5 km depth with velocity 2 km/s. The model is sampled every 10 m, with 40 km lateral extent and 12 km depth. This model was chosen because the strong velocity contrast results in a strong internal multiple. Data were generated for this model using standard explicit second-order finite differences, using the program *sufdmod2*. A single shot at the center of the model was simulated with receivers spaced every 10 m; the source and receivers were at 300 m depth in the model. Four seconds

A Robin boundary condition is a mix of a Neumann and Dirichlet boundary condition; van Stralen (1997) found this particular boundary condition to be accurate in wave propagation problems. Because of the structure of the propagator used here, I chose to implement this boundary condition by propagating the wavefield from one depth to the next and then replacing the boundary elements of the wavefield with the solution of the boundary condition equations. The affects of the boundary conditions are shown in Figure H.5.

Finally, I came to the conclusion that tapering and boundary conditions is simply not sufficient and I simply padded the wavefield with zeros to approximately two and a half times its original size. This seems to attenuate the artifacts to the point where the image is sufficiently clean to analyze the contribution to the image from internal multiples. Padding in different variables is shown in Figures H.9,H.10 and H.11.

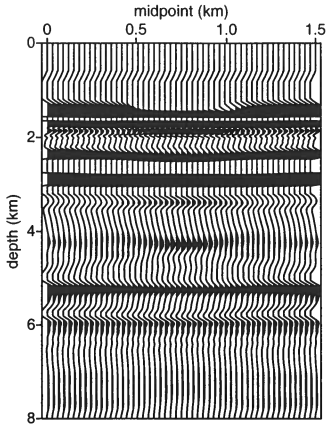


Figure H.5. Corner frequencies of bandpass filter: 1,3,20,30; padding of 25 points in both offset and midpoint; padding of 2079 points in time; no tapering; boundary conditions applied in both midpoint and offset.

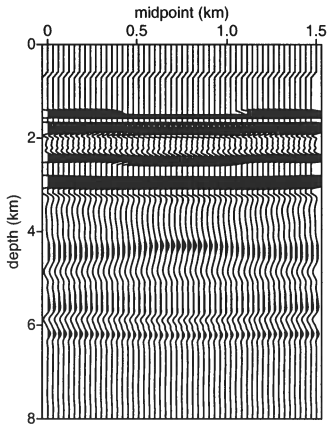


Figure H.6. Corner frequencies of bandpass filter: 1,3,20,30; padding of 25 points in both offset and midpoint; padding of 2079 points in time; taper in offset; no boundary conditions.

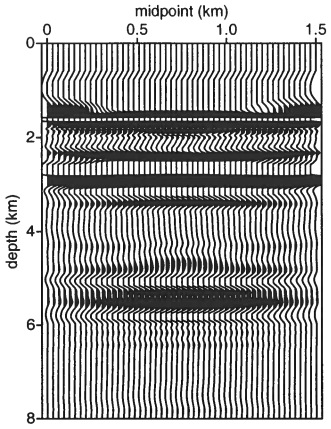


Figure H.7. Corner frequencies of bandpass filter: 1,3,20,30; padding of 25 points in both offset and midpoint; padding of 2079 points in time; taper in midpoint; no boundary conditions.

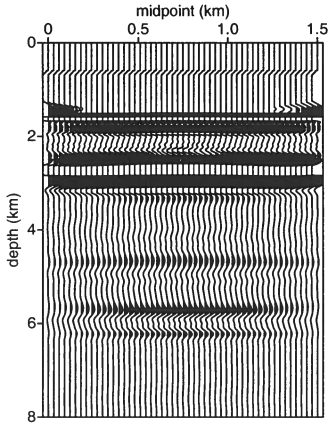


Figure H.8. Corner frequencies of bandpass filter: 1,3,20,30; padding of 25 points in both offset and midpoint; padding of 2079 points in time; taper in both midpoint and offset; no boundary conditions.

Appendix I

Amplitude correction factors used in imaging

Chapter 4 gives a theory for predicting artifacts caused by internal multiples in imaging. In order to estimate these artifacts accurately, the various amplitude factors discussed must be implemented and applied to the data. In this appendix, I discuss the implementation of these factors and illustrate their application to two synthetic data sets. I describe their application as part of the standard imaging procedure rather than in the context of multiple estimation because this is the simplest application of these operators.

I.1 Algorithm

The goal of this appendix is to bridge the gap between the theory described by Stolk & de Hoop (2004a) for amplitude preserved wave-equation migration and the implementation of these ideas. I thus begin directly with Stolk & de Hoop (2004a, proposition 3.2), in which the following expression for the wave-equation angle transform is given,

$$(A_{\mathrm{WE}}d)(z,x,p) = j^{-1}R_3\bar{\Xi}^{-1}Q_{-,s}(z)^{-1}Q_{-,r}(z)^{-1}H^*(z,0) \\ Q_{-,s_0}(0)^{-1}Q_{-,r_0}(0)^{-1}D_t^{-2}d(s_0,r_0,t)\,. \quad \text{(I.1)}$$

In the above expression, d is the data, as a function of the source position, s_0, receiver position, r_0 and time, t. The Q operators are the elements of generalized operator eigenvectors which arise in the diagonalization of the wave-equation written as a first order system of differential equations (see Section 4.3 for more details). The propagator H is the downward continuation operator which is the Green operator of the double-square-root (DSR) equation (the DSR equation was introduced by Claerbout (1985)). The pseudo-differential operator $\bar{\Xi} = \bar{K}^*\bar{K}$ is the normal operator, where

$$\bar{K} = LE_2\,, \quad \text{(I.2)}$$

$$E_2 : b(z,r,s) \mapsto \delta(t)b(z,r,s)\,, \quad \text{(I.3)}$$

and

$$\bar{L}g = \int_0^Z H(0,z)g(z,\cdot,\cdot,\cdot)\,\mathrm{d}z. \quad \text{(I.4)}$$

The restriction operator R_3 is the beam-forming operator which gives the wavefield as a function of angle. The j^{-1} operator is a jacobian applied to the final image gather.

equation in (I.8) gives

$$\left(\frac{1}{2\pi}\right)^2 \int \mathrm{d}\rho_0 \int \mathrm{d}\omega \int \mathrm{d}r_0' \int \mathrm{d}t' (c(0,r_0')^{-2}\omega^2 - \rho_0^2)^{\frac{1}{4}} \mathcal{F}_t^{-1}[\omega^{-2}\widehat{d}(s_0,r_0,\omega)](s_0,r_0',t')e^{\mathrm{i}\omega(t-t')+\mathrm{i}\rho_0(r_0-r_0')}, \quad \text{(I.9)}$$

where the subscript t to $\mathcal{F}$ to indicates that it is a Fourier transform only in time. This is equivalent to

$$\begin{aligned} &\left(\frac{1}{2\pi}\right)^2 \int \mathrm{d}\rho_0 \int \mathrm{d}\omega \int \mathrm{d}r_0' (c(0,r_0')^{-2}\omega^2 - \rho_0^2)^{\frac{1}{4}} e^{\mathrm{i}\rho_0(r_0-r_0')+e^{\mathrm{i}\omega t}} \int \mathrm{d}t' \qquad \text{(I.10)} \\ &\qquad \mathcal{F}_t^{-1}[\omega^{-2}\widehat{d}(s_0,r_0,\omega)](s_0,r_0',t')e^{-\mathrm{i}\omega t'} \\ = &\left(\frac{1}{2\pi}\right)^2 \int \mathrm{d}\rho_0 \int \mathrm{d}\omega \int \mathrm{d}r_0' (c(0,r_0')^{-2}\omega^2 - \rho_0^2)^{\frac{1}{4}} \omega^{-2}\widehat{d}(s_0,r_0,\omega)e^{\mathrm{i}\rho_0(r_0-r_0')+\mathrm{i}\omega t}. \end{aligned}$$

There are several ways to approximate the above form, although its direct application is quite costly. One possible approximation scheme is discussed in the following section.

Screen approximation In (I.10) the forward and inverse Fourier transforms in the (r_0, ρ_0) variables cannot be separated because q depends on both variables. Because it is more efficient to apply the operator if this can be done, I approximate the symbol of Q by a phase screen. The idea of a phase screen is to approximate the symbol of an operator by a product of two symbols, one depending only on r_0 and the other only on ρ_0. Further details on screens and generalized screens can be found in de Hoop *et al.* (2003b). The slowness contrast is

$$u(z,x) = c(x,z)^{-2} - c_0(z)^{-2}, \quad \text{(I.11)}$$

where c_0 is the laterally invariant reference velocity. Using this define

$$q_0(z,\rho,\omega) = \left(\left(\frac{\omega}{c_0(z)}\right)^2 - \rho^2\right)^{\frac{1}{4}}, \quad \text{(I.12)}$$

as the symbol of Q in a background velocity model. Substituting this into the expression above for q gives

$$\begin{aligned} q(z,\rho,\omega) &= (q_0(z,\rho,\omega)^4 + u(z,x))^{\frac{1}{4}} \qquad \text{(I.13)} \\ &= q_0(z,\rho,\omega)\left(1 + \frac{u(z,x)}{q_0(z,\rho,\omega)^4}\right)^{\frac{1}{4}}. \end{aligned}$$

Performing a first order Taylor expansion about $u = 0$ of the term in brackets gives

$$q(z,\rho,\omega) \approx q_0(z,\rho,\omega)\left(1 + \tfrac{1}{4}\frac{u(z,x)}{q_0(z,\rho,\omega)^4}\right). \quad \text{(I.14)}$$

The symbol of H also depends on both the space and phase variables, making its application computationally expensive. There are many methods of approximating this operator, a few of which are given in de Hoop *et al.* (2003b) and references therein. This completes the data propagatation to depth z, denoted in the following by $d_p(z,s,r,t)$.

I.1.3 Imaging condition

To generate a CIG at depth z, the second set of amplitude corrections and the restriction operator R_3 are applied to the data. The first two operators that must be applied to the data are again Q^{-1} operators whose approximate symbols are given in (I.16). Thus, they are applied via

$$-\left(\frac{1}{2\pi}\right)^2 \mathcal{F}_\rho^{-1}\mathcal{F}_\sigma^{-1}\mathcal{F}_\omega^{-1}\left[q_0(z,\rho,\omega)q_0(z,\sigma,\omega)\right.$$
$$\left.\mathcal{F}_r\mathcal{F}_s\left[(1+\tfrac{1}{4}c_0(z)^2u(z,r))(1+\tfrac{1}{4}c_0(z)^2u(z,s))\widehat{d}_p(z,s,r,\omega)\right]\right](z,s,r,t)\,. \quad \text{(I.18)}$$

Next $\bar{\Xi}$ must be applied to the data. It's symbol is given by

$$c(z,s)^{-2}(c(z,s)^{-2}-\omega^{-2}\sigma^2)^{-\frac{1}{2}}+c(z,r)^{-2}(c(z,r)^{-2}-\omega^{-2}\rho^2)^{-\frac{1}{2}} \quad \text{(I.19)}$$

the phase screen approximation of this symbol, computed as in section I.1.1, is

$$c_0(z)^{-2}(c_0(z)^{-2}-\omega^{-2}\sigma^2)^{-\frac{1}{2}}(1+u(s,z)c_0(z)^2)(1-\tfrac{1}{2}c_0(z)^2u(s,z))$$
$$+\,c_0(z)^{-2}(c_0(z)^{-2}-\omega^{-2}\rho^2)^{-\frac{1}{2}}(1+u(r,z)c_0(z)^2)(1-\tfrac{1}{2}c_0(z)^2u(r,z))\,. \quad \text{(I.20)}$$

To simplify notation define

$$\xi_p(z,\omega,\alpha)=c_0(z)^{-2}(c_0(z)^{-2}-\omega^{-2}\alpha^2)^{-\frac{1}{2}}$$

and

$$\xi_s(z,x)=(1+u(x,z)c_0(z)^2)(1-\tfrac{1}{2}c_0(z)^2u(x,z))\,.$$

The application of $\bar{\Xi}$ to the data can then be written as

$$d_o(z,s,r,t)=\mathcal{F}_t^{-1}\left[\mathcal{F}_s^{-1}\left[\xi_p(z,\omega,\sigma)\mathcal{F}_s\left[\xi_s(z,s)\widehat{d}_i(z,s,r,\omega)\right]\right]\right.$$
$$\left.+\mathcal{F}_r^{-1}\left[\xi_p(z,\omega,\rho)\mathcal{F}_r\left[\xi_s(z,r)\widehat{d}_i(z,s,r,\omega)\right]\right]\right]\,, \quad \text{(I.21)}$$

which gives the final corrected data at depth z. The CIG is now formed through the restriction operator R_3,

$$a(z,x,p)=\int_{\mathbb{R}} d_o(z,x-\tfrac{h}{2},x+\tfrac{h}{2},ph)\chi(z,x,h)\,\mathrm{d}h, \quad \text{(I.22)}$$

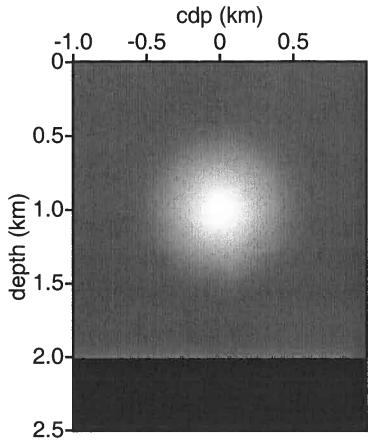


Figure I.1. Velocity model for the simple lens example. The low-velocity lens is described by $c(x,z) = 1 - 0.4e^{-9[x^2+(x-1)^2]}$; the velocity increases from 1 km/s to 1.15 km/s at a depth of 2 km to create a reflector.

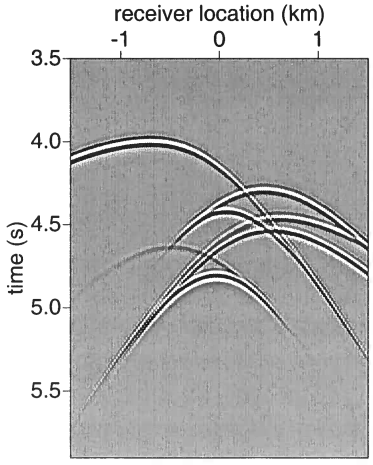


Figure I.2. Shot record from $x = -500m$ for the simple lens model.

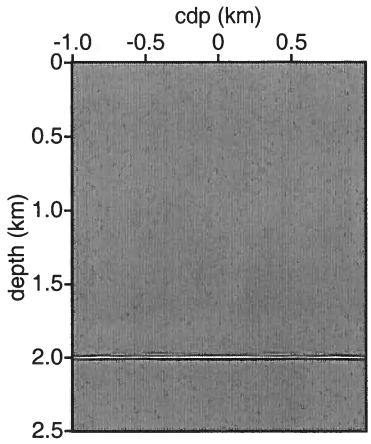


Figure I.5. Image in the simple lens model with Q operators applied.

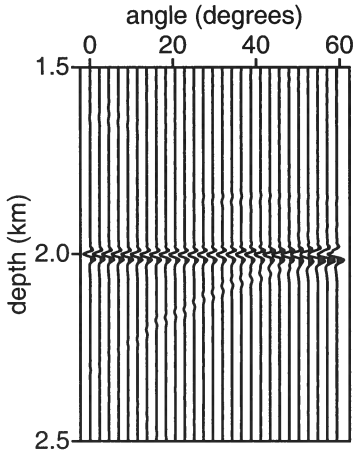


Figure I.6. Common image gather at $x = 0$ in the simple lens model with Q operators applied.

I.2.2 Valhall toy lens model

This gas-lens model, shown in Figure I.9 consists of a vertical velocity gradient ($0.45\ s^{-1}$) beginning at 1600 m/s, with a low-velocity circular lens with Gaussian parameter variations (maximum velocity contrast 800 m/s) located at lateral position 4600 m, and depth 600 m, with a diameter (Gaussian standard deviation) of 600 m. This model, introduced by Brandsberg-Dahl *et al.* (2003a), is based on a feature in the BP Valhall field; it is the same model used in Chapter 3. A typical shot record is shown in Figure I.10.

An image in this model is shown in Figure I.11 with no Q or Ξ operators applied. There are some imaging artifacts in this image, caused by edge reflections in the propagation (these show up as smiles near the edges). The dipping portion of the reflector is also spatially aliased. An image gather is shown in Figure I.12, at a position beneath the lens. As expected this image gather is flat with contributions only from the correct depth location.

Figures I.13 and I.14 show the image and CIG respectively with the Q operators applied both at the surface and the image point. The CIG in particular is sharper than without the Q operators, although once again high frequency noise is introduced. Figures I.15 and I.16 show the results of applying the Ξ operator. Again Ξ seems to have a small influence on the relative amplitudes, although the amplitudes at large angles are somewhat stronger.

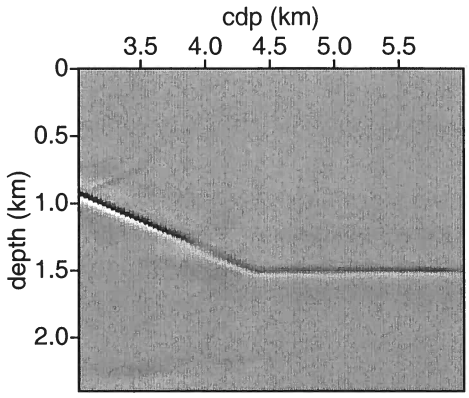


Figure I.11. Image in the Valhall model with no Q or Ξ operators applied.

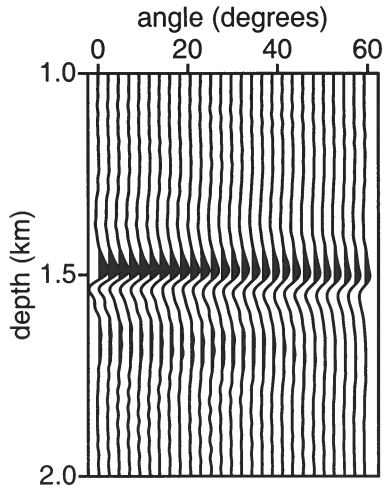


Figure I.12. Common image gather at $x = 4689$ m in the Valhall model with no Q or Ξ operators applied.

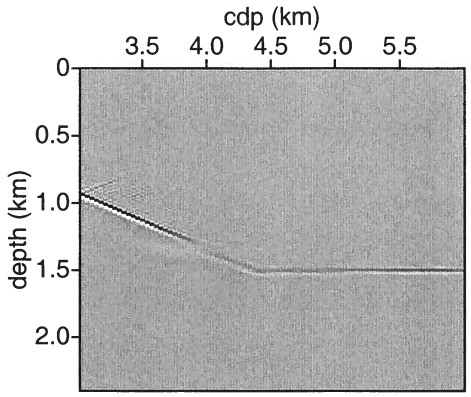


Figure I.15. Image in the Valhall model with both Q and Ξ operators applied.

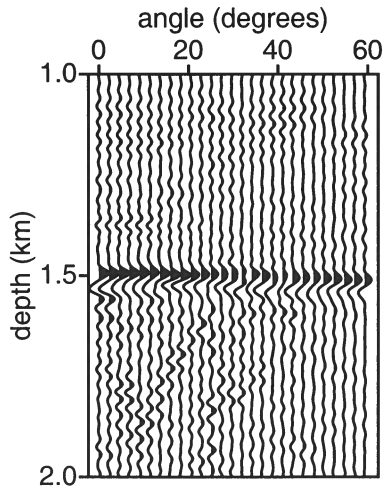


Figure I.16. Common image gather at $x = 4689$ m in the Valhall model with both Q and Ξ operators applied.

Appendix J

Proof of Theorem 4.7.1[1]

The proof rest on the semi-group property (4.23), discussed previously. The idea is to use this property to extend the two Green functions in (4.81 meeting at (z_1, m_s, m_r) to the surface (see Figure J.1). The resulting operators are then rearranged to pair the G_- operators to substitute the double-square-root Green function, H. We go through this procedure twice, once for $\delta u_{1,-}$ and once for the other elements of (4.81).

We start by applying the procedure outlined above to $\delta u_{1,-}$, beginning with the semi-group property applied to (4.77),

$$\begin{aligned}\delta u_{-,1}(z_1, m_r, t_a, 0, s_0) = -\tfrac{1}{4} D_t^2 Q^*_{-,s_0}(0) \int_{z_1}^{\infty} \mathrm{d}z_2 \int \mathrm{d}s_2 \int \mathrm{d}r_2 \int_{\mathbb{R}} \mathrm{d}t_0 \int_{\mathbb{R}} \mathrm{d}t' \int \mathrm{d}m'_r \int_{\mathbb{R}} \mathrm{d}t_{m'_r} \\ G^*_-(z_1, m_r, t_{m'_r}, 0, m'_r) G_-(0, m'_r, t_a + t_{m'_r} - t' - t_0, z_2, r_2) \\ G_-(0, s_0, t', z_2, s_2)\; Q_{-,r_2}(z_2) Q_{-,s_2}(z_2) (E_2 E_1 a)(z_2, s_2, r_2, t_0)\,, \quad \text{(J.1)}\end{aligned}$$

where $t_a + t_{m'_r}$ is the time required to travel from the source at s_0 to the pseudo-receiver at m'_r, as illustrated in Figure J.2. We now begin to rearrange the terms in preparation for the H substitution.

Since t' is independent of the other variables we bring this integration to the inside to replace the two G_- operators by the H operator

$$\begin{aligned}\delta u_{-,1}(z_1, m_r, t_a, 0, s_0) = -\tfrac{1}{4} D_{t_a}^2 Q^*_{-,s_0}(0) \int \mathrm{d}m'_r \int_{\mathbb{R}} \mathrm{d}t_{m'_r} G^*_-(z_1, m_r, t_{m'_r}, 0, m'_r) \\ \int_{z_1}^{\infty} \mathrm{d}z_2 \int \mathrm{d}s_2 \int \mathrm{d}r_2 \int_{\mathbb{R}} \mathrm{d}t_0 H(0, s_0, m'_r, t_a + t_{m'_r} - t_0, z_2, s_2, r_2) Q_{-,r_2}(z_2) Q_{-,s_2}(z_2) \\ (E_2 E_1 a)(z_2, s_2, r_2, t_0)\,. \quad \text{(J.2)}\end{aligned}$$

This completes the manipulations of $\delta u_{1,-}$.

[1]This appendix has been accepted, along with Chapter 4 and Appendix K, to Inverse Problems as: Malcolm, A. E. and de Hoop, M. V. A method for inverse scattering based on the generalized Bremmer coupling series.

Next, we apply the same procedure to the second Green function in (4.81),

$$d_3(s_0,r_0,t_4) = -\tfrac{1}{4}D_t^4\int_0^\infty \mathrm{d}z_3\int \mathrm{d}s_3\int \mathrm{d}r_3\int_{\mathbb{R}}\mathrm{d}t_{30}\int_{\mathbb{R}}\mathrm{d}t_a\int_0^{z_3}\mathrm{d}z_1\int \mathrm{d}m_s\int \mathrm{d}m_r\int_{\mathbb{R}}\mathrm{d}t_3'$$
$$Q^*_{-,r_0}(0)G_-(0,r_0,t_4-t_a-t_3'-t_{30},z_3,r_3)Q_{-,r_3}(z_3)$$
$$Q^*_{-,m_s}(z_1)\int \mathrm{d}m_s'\int_{\mathbb{R}}\mathrm{d}t_{m_s'}G^*_-(z_1,m_s,t_{m_s'},0,m_s')G_-(0,m_s',t_3'+t_{m_s'},z_3,s_3)$$
$$Q_{-,s_3}(z_3)(E_2E_1a)(z_3,s_3,r_3,t_{30})(E_1a)(z_1,m_s,m_r)Q^*_{-,m_r}(z_1)$$
$$\delta u_{-,1}(z_1,m_r,t_a,0,s_0)\,, \quad \text{(J.3)}$$

where $t_{m_s'}$ is defined by analogy to $t_{m_r'}$ (see Figure J.2). We now begin to rearrange terms in (J.3) in preparation of the H substitution.

Since G^*_- and the propagator proceeding it do not have variables in common, we interchange their order. We also change variables from t_3' to $t_3'' = t_3' + t_{m_s'}$, interchanging the t_3' and $t_{m_s'}$ integrations. This results in

$$d_3(s_0,r_0,t_4) =$$
$$-\tfrac{1}{4}D_t^4\int_0^\infty \mathrm{d}z_3\int \mathrm{d}s_3\int \mathrm{d}r_3\int_{\mathbb{R}}\mathrm{d}t_{30}\int_{\mathbb{R}}\mathrm{d}t_a\int_0^{z_3}\mathrm{d}z_1\int \mathrm{d}m_s\int \mathrm{d}m_r\int_{\mathbb{R}}\mathrm{d}t_a\int \mathrm{d}m_s'\int_{\mathbb{R}}\mathrm{d}t_{m_s'}\int_{\mathbb{R}}\mathrm{d}t_3''$$
$$Q^*_{-,r_0}(0)Q^*_{-,m_s}(z_1)G^*_-(z_1,m_s,t_{m_s'},0,m_s')G_-(0,r_0,t_4-t_a-t_3''+t_{m_s'}-t_{30},z_3,r_3)$$
$$Q_{-,r_3}(z_3)G_-(0,m_s',t_3'',z_3,s_3)Q_{-,s_3}(z_3)$$
$$(E_2E_1a)(z_3,s_3,r_3,t_{30})(E_1a)(z_1,m_s,m_r)Q^*_{-,m_r}(z_1)\delta u_{-,1}(z_1,m_r,t_a,0,s_0)\,. \quad \text{(J.4)}$$

We now substitute H from (4.75) for the time convolution of the two G_- operators above, interchanging the order of integration, to obtain

$$d_3(s_0,r_0,t_4) = -\tfrac{1}{4}D_{t_4}^4Q^*_{-,r_0}(0)\int_0^\infty \mathrm{d}z_3\int_0^{z_3}\mathrm{d}z_1\int \mathrm{d}m_s\int \mathrm{d}m_r\int_{\mathbb{R}}\mathrm{d}t_a\int \mathrm{d}m_s'\int_{\mathbb{R}}\mathrm{d}t_{m_s'}$$
$$Q^*_{-,m_s}(z_1)G^*_-(z_1,m_s,t_{m_s'},0,m_s')\int \mathrm{d}s_3\int \mathrm{d}r_3\int_{\mathbb{R}}\mathrm{d}t_{30}$$
$$H(0,m_s',r_0,t_4-t_a+t_{m_s'}-t_{30},z_3,s_3,r_3)Q_{-,s_3}(z_3)Q_{-,r_3}(z_3)(E_1a)(z_1,m_s,m_r)$$
$$(E_2E_1a)(z_3,s_3,r_3,t_{30})Q^*_{-,m_r}(z_1)\delta u_{-,1}(z_1,m_r,t_a,0,s_0)\,. \quad \text{(J.5)}$$

We have now extended both Green operators to the surface, what remains is the combining of the G^*_- operators in (J.5) and (J.2) into an H^* operator.

To do this, we substitute (J.2) into (J.5). We then interchange operators to combine the two G^*_- terms, as well as changing the order of integration to move the t_a integral inside

translation invariant) and changing time variables from t_a to $t_b = t_a + t_{m'_r}$ we arrive at a structure into which the distribution W defined in the theorem statement can be inserted. This W distribution is a new field constituent generated through the convolution of the two data sets on which the two Green functions in (J.8) act. The $\mathbf{d}_1$ data constituents cannot be extracted directly from the data unless ten Kroode's traveltime monotonicity assumption is satisfied. If this assumption is not satisfied one could generate $\mathbf{d}_1$ as $d_1 - \mathbf{D}\langle a\rangle$, where

$$(\mathbf{D}\langle a\rangle)(z_1, s_0, r_0, t) = -\tfrac{1}{4} D_t^2 Q^*_{-,r}(0) Q^*_{-,s}(0) \int_0^{z_1} \mathrm{d}z \int \mathrm{d}s \int \mathrm{d}r \int_{\mathbb{R}} \mathrm{d}t_0 \\ H(0, s_0, r_0, t - t_0, z, s, r) Q_{-,r}(z) Q_{-,s}(z) (E_2 E_1 \langle a\rangle)(z, s, r, t_0)\,, \quad \text{(J.9)}$$

is the data modeled from an estimate, $\langle a\rangle$, of the medium contrast down to the depth z_1.

Two changes are required to insert the W distribution into (J.8). First, the lower bound on the t_b integral is extended to 0, rather than $t_{m'_r}$ because $t_b > t_{m'_r}$ by definition. Second, to overlay the distribution W with the expression in braces in (J.8) we need only make the identification $t = t_4 + t_{m'_r} + t_{m'_s}$.

In the definition of W, we identify a new time variable $t_{m'} = t_{m'_r} + t_{m'_s}$ in the expression for t. To introduce this variable we change variables from $t_{m'_r}$ to $t_{m'}$, substituting the expression for W from (4.83) into (J.8)

$$d_3(s_0, r_0, t_4) = D_t^2 \int_0^{\infty} \mathrm{d}z_1 \int \mathrm{d}m_s \int \mathrm{d}m_r \int_{\mathbb{R}} \mathrm{d}t_{m_0} \\ Q^*_{-,m_s}(z_1)(E_2 E_1 a)(z_1, m_s, m_r, t_{m_0}) Q^*_{-,m_r}(z_1) \int \mathrm{d}m'_s \int \mathrm{d}m'_r \int_{\mathbb{R}} \mathrm{d}t_{m'} \int_0^{t'_m} \mathrm{d}t_{m'_s} \\ G^*_-(z_1, m_s, t_{m'_s}, 0, m'_s) G^*_-(z_1, m_r, t_{m'} - t_{m'_s} - t_{m_0}, 0, m'_r) \\ Q^*_{-,m'_r}(0)^{-1} Q^*_{-,m'_s}(0)^{-1} W(z_1; s_0, m'_r, t_4 + t_{m'}, m'_s, r_0)\,. \quad \text{(J.10)}$$

The two G^*_- operators in (J.10) along with the integration in $t_{m'_s}$ are nearly in the form of the H operator.

The integration in $t_{m'_s}$ is extended to ∞ as $t_{m'_s} > t_{m'}$ results in a negative time in the second G^*_- making it 0 by the anti-causality of G^*_- (Remark 4.4.1). This allows us to introduce the H operator, which gives the result.

Appendix K

Comparison with the Weglein/ten Kroode approach[1]

If no caustics form in the background medium, and the traveltime monotonicity of ten Kroode is satisfied, our results can be brought into correspondance with those of Weglein *et al.* (1997), and ten Kroode (2002). To facilitate this comparison, we will write (4.84) in terms of the data only.

We begin by recalling from the discussion following Theorem 4.7.1, that the integration in (m_r, m_s, t_m) is an inner product in these variables. We then identify $Q^*_{-,m_s}(z_1)$ $Q^*_{-,m_r}(z_1)(H(z_1,0))^*$ as an operator acting on $Q^*_{-,m'_r}(0)^{-1}Q^*_{-,m'_s}(0)^{-1}$ $W(z_1; s_0, m'_r, t_4 + t'_m, m'_s, r_0)$; this makes up the second entry in the inner product. The first entry in this inner product is $(E_2E_1a)(z_1, m_s, m_r, t_{m_0})$. An equivalent form of (4.84) is then

$$d_3(s_0,r_0,t_4) = D_t^2 \int_0^\infty \mathrm{d}z_1 \left(\int \mathrm{d}m'_s \int \mathrm{d}m'_r \int_{\mathbb{R}} \mathrm{d}t_{m'} \left\{ \int \mathrm{d}m_s \int \mathrm{d}m_r \int_{\mathbb{R}} \mathrm{d}t_{m_0} \right.\right.$$
$$H(0,m'_s,m'_r,t_{m'}-t_{m_0},z_1,m_s,m_r)Q_{-,m_r}(z_1)Q_{-,m_s}(z_1)$$
$$\left.\left.(E_2E_1a)(z_1,m_s,m_r,t_{m_0})\right\} Q^*_{-,m'_r}(0)^{-1}Q^*_{-,m'_s}(0)^{-1}W(z_1;s_0,m'_r,t_4+t'_m,m'_s,r_0)\right), \quad \text{(K.1)}$$

where $H(0,m'_s,m'_r,t_{m'}-t_{m_0},z_1,m_s,m_r)Q_{-,m_r}(z_1)Q_{-,m_s}(z_1)$ now acts on (E_2E_1a) and the inner product is in the (m'_s,m'_r,t'_m) variables. We define (for the expression in braces in (K.1)

$$\bar{d}_1(z_1,s,r,t) = -D_t^2 Q^*_{-,s}(0)Q^*_{-,r}(0)\int \mathrm{d}s_1 \int \mathrm{d}r_1 \int_{\mathbb{R}} \mathrm{d}t_0$$
$$H(0,s,r,t-t_0,z_1,s_1,r_1)Q_{-,s_1}(z_1)Q_{-,r_1}(z_1)(E_2E_1a)(z_1,s_1,r_1,t_0)\,; \quad \text{(K.2)}$$

The quantity $\bar{d}_1$ is not one that can be measured directly from the data. To compute $\bar{d}_1$, the expression in (4.92) must be substituted for a to write it in terms of what can be measured, d.

[1]This appendix has been accepted, along with Chapter 4 and Appendix J, to Inverse Problems as: Malcolm, A. E. and de Hoop, M. V. A method for inverse scattering based on the generalized Bremmer coupling series.

to the position of each of the scattering points. From this he finds that the ray from (in the notation used here) r_2 to m'_r (s_3 to m'_s) must follow the same path as that from r_2 to m_r (s_3 to m_s). In the formulation described here this condition is automatically applied through the relation (4.23) used to extend the modeled data from the scattering point at z_1 to the surface.