

Inverse multiple scattering in the downward continuation approach: Preface

Alison E. Malcolm and Maarten V. De Hoop

Moses (1956) constructed a series to represent the quantum scattering potential in terms of measured reflection coefficients. Razavy (1975) extended this work to scalar wave scattering. These two papers show that it is formally possible to represent both the data as a series in the medium contrast (forward series) and the medium contrast as a series in the data (inverse series). In seismology the second series has been exploited by several authors to represent the medium contrast, or reflectors, as a series in the data. Taking only the first term of this series leads to the Born approximation. The analysis of the forward and inverse series is truly incomplete. For example, in general the series are not identifiable as Neumann series. Only for very simple obstacle scattering problems have properties of the forward series been proven. (There is a much better developed theory for the Schrödinger equation.) As a starting point for our approach we use the generalized Bremmer series as a forward series. De Hoop (1996) proved that this series is identifiable as a Neumann series. The first term of the generalized Bremmer series models seismic reflection data in the downward-continuation approach [with the double-square-root (DSR) equation]. Here, we develop a theory for inverse multiple scattering using the third term of a hybrid series derived from the generalized Bremmer series and the Lippmann-Schwinger series.

Fokkema and Van den Berg (1993) developed a rigorous theory for the suppression of surface-related-multiples. Their approach admits a series expansion in terms of the reflection data that is close to the inverse scattering series approach of Weglein (1997). Weglein (1997; 2003) has also used these ideas in an attempt to attenuate internal multiples in 1D along with a restricted class of 2D velocity models. Ten Kroode (2002) describes in more detail the theory underlying Weglein's approach and highlights some its fundamental limitations. One of the goals in our approach is to remove some of the limitations explained by Ten Kroode. For the sake of simplicity, we also focus on contributions from singly and triply scattered waves,

though the extension to higher order is straightforward.

Weglein and ten Kroode use the Lippmann-Schwinger series to model triply scattered data. In that approach, ten Kroode makes the assumption that each scattering results in a change in direction between the incoming and outgoing waves. In other words, he enforces that the first scatterer is strictly below the second scatterer which is strictly above the third scatterer. To incorporate this assumption naturally in the formulation of the problem, we construct a hybrid series between the Lippmann-Schwinger series and the Bremmer series. The Bremmer series has the advantage that it directly splits the wavefield into its up- and down-going constituents. In addition, the convergence properties on the Bremmer series (de Hoop, 1996) justify the truncation of the series after only a few terms. The Lippmann-Schwinger series alone is not even asymptotic (though for surface-related multiples it can be identified as a convergent series).

In his work, ten Kroode (2002) has to make two essential simplifying assumptions. The first is that there are no caustics in the wavefield. In our formulation we find it necessary to exclude only turning rays; this is necessary because the downward-continuation approach removes horizontally traveling waves. The second assumption is his so-called *traveltime monotonicity assumption*. This assumption requires, essentially, that the traveltime increase monotonically with depth. In other words, if primary energy arrives later in the data it must have come from a deeper reflector. As ten Kroode points out in his paper it is not difficult to violate this assumption.

In the absence of caustics and if the traveltime monotonicity assumption is valid, the work presented in the following paper reduces easily to the results of ten Kroode and thus Weglein. Our implementation is truly different, however, as we use the DSR equation explicitly. An advantage of using the DSR approach is that there are no artifacts in the image domain (de Hoop *et al.*, 2003b). When there are caustics and the traveltime monotonicity

assumption is not satisfied, such as in subsalt imaging for example, the theory presented here becomes important for successful suppression of multiples. In the absence of the traveltime monotonicity assumption, multiple attenuation requires knowledge of the velocity model to the depth of the shallowest scatter. Since it is not in general known at which depth the shallowest scatter occurs, this means that the full velocity model is required to model internal multiples.

Another problem in attenuating multiples is that multiples and primaries have different subsurface illumination properties. This means that in the removal of multiples, adaptive subtraction techniques must be employed. We analyze this issue in the paper, but do not present an algorithm for adaptive subtraction. We are considering techniques to improve adaptive subtraction methods to enhance the success of the subtraction process.

Another way in which our technique differs from other multiple attenuation methods is that we propose subtracting the multiples in the image domain rather than the data domain. It is easier to account for illumination issues in common image gathers (CIGs) than it is on the data, because correctly imaged primary events are flat in CIGs, making it easier to differentiate between primaries and multiples. Also, since we have found internal multiple attenuation to be dependent on the velocity model, attenuating the multiples in the image domain may be more efficient computationally because the modeling need not be carried to the surface.

Our final goal with the research presented in the following paper is to arrive at a scheme whose terms can be properly estimated as multilinear operators acting on the data. This paper is just a first step toward that goal.

Inverse multiple scattering in the downward continuation approach

Alison E. Malcolm and Maarten V. De Hoop

Center for Wave Phenomena, Colorado School of Mines

ABSTRACT

Imaging with seismic data is typically done in the single-scattering approximation. We move beyond this assumption to allow for triply scattered waves in the imaging process. We develop a scattering series that is a hybrid of the Lippmann-Schwinger scattering series and the Bremmer coupling series. From the third term of this hybrid series an approximation of data scattered three times is constructed. From the inverse hybrid series it is also evident that subtracting an image constructed from triply scattered data from an image constructed assuming singly scattered data results in an image that is third-order in the data. This is in contrast to the standard first-order approximation. We discuss both a standard inverse scattering operator and the wave equation angle transform as possible imaging techniques.

Key words: internal multiples, Bremmer series, Lippmann-Schwinger series

1 INTRODUCTION

In a seismic experiment, the source generates both compressional and shear waves that travel through the subsurface reflecting at discontinuities in the medium properties. These waves are then recorded at the surface. A collection of such experiments, recorded as a function of source position, receiver position, and time make up seismic data. It is generally assumed in seismic imaging that the recorded signals have reflected only once between the source and receiver. The goal of this paper is to move beyond this assumption to allow for waves which have scattered three times to be included in seismic imaging. Here, we consider only compressional waves and back scattering; i.e., sources and receivers are on the same surface. Although we discuss only the case of triple scattering, the extension to any finite odd-order scattering is straightforward. Constructing even-order scattered waves is fundamentally different, for a reflection experiment, because transmission in addition to reflection must be considered.

We use a series derived from the Lippmann-

Schwinger-Dyson equation; this framework gives us both a forward and an inverse series representation. We also make use of the generalized Bremmer series (de Hoop, 1996), which is a forward series that splits the wavefield into its up- and down-going constituents. From these two series, we develop a hybrid between the Lippmann-Schwinger and Bremmer series. The hybrid series uses the directional decomposition of the Bremmer series along with the Lippmann-Schwinger medium decomposition into a known, smooth reference and unknown, singular perturbation. This allows us to trace waves through their up and down scatters while still preserving the contrast source formulation of the Lippmann-Schwinger construction. From the third term of the forward series, we model the triply scattered data. The third term of the inverse series describes the third-order (in the data) contribution to the image. We use the volume scattering framework as in Bremmer rather than the surface scattering model of Lippmann-Schwinger. We make this choice because de Hoop (2004) has shown that the downward continuation approach

followed here does not allow the surface-scattering-based wave-equation Kirchhoff approximation.

In this paper, we discuss first-order internal multiples, as distinguished from surface-related multiples. We define a reflection as a scattering in which the wave changes direction, and transmission as a scattering in which the wave does not change direction. From this, we define a primary event as a wave that contains only one reflection, along with any number of transmissions. This is distinguished from a singly scattered event, which is an event with one reflection or transmission. Extending this, a first-order internal multiple is an event with three reflections, all below the Earth's surface, and no transmissions. A surface-related multiple is one in which at least one reflection comes from the Earth's surface. Fokkema and van den Berg (1993) developed a rigorous theory for the suppression of surface multiples, derived from the reciprocity theorem in integral form. If assumptions allowing the construction of data at zero-offset, such as those given by de Hoop *et al.* (2003a) are satisfied, then the theory given by Fokkema and van den Berg (1993) solves the theoretical surface multiple attenuation problem. There is as yet no complete theory for internal multiples; a goal of this paper is to begin developing such a theory. In this paper, we consider the reflection seismology application of this theory. There are other applications, however, a few of which we mention in the following two paragraphs.

In earthquake seismology, surface-related multiples are estimated as part of the receiver function analysis. Park and Levin (2000) give a method of estimating surface-related multiples in the transverse and horizontal components of earthquake data using the multiples seen on the vertical component. Although they consider only surface-related multiples they also see signal-generated noise, which they speculate could be partly generated by internal multiples (Park & Levin, 2001). Surface-related multiples in the mantle were observed by Revenaugh and Jordan (1987); they use these multiples to estimate the attenuation quality factor, Q of the upper mantle. In reflection seismology, Verschuur and Berkhout (1997) demonstrate the abilities of their de-multiple theory (Berkhout & Verschuur, 1997). In this theory, they use a Neumann series expansion technique to predict surface-related multiples, which can then be subtracted from the original data set. They also discuss the extension of their technique to the internal multiples case, with a layer stripping approach.

All these examples involve surface-related multiples. Here, we are concerned with internal multiples, which have also been exploited in other appli-

cations. In earthquake seismology, Burdick and Orcutt (1979) investigate the truncation of the generalized ray sum, which accounts for increasing orders of internal multiples. They compare this generalized ray sum with the reflectivity method for modeling waveforms and mention earth models in which the inclusion of internal multiples becomes important. Revenaugh and Jordan (1987) also observe internal multiples. From these wave-paths they estimate the attenuation of the lower mantle. Revenaugh and Jordan (1989) also used both internal and surface-related multiples to investigate layering in the mantle. They used waves which have scattered multiple times between the core-mantle boundary and the surface, including at least one scatter from discontinuities in the mantle. This scatter from discontinuities in the mantle can be either a surface-related or internal multiple. They then perform a linear inversion, similar to a Kirchhoff migration, on the scattered waveforms to estimate the depth of the discontinuity from which the wave scattered. In addition, for synthetic aperture radar (SAR) data Cheney and Borden (2002) derive a theory to relate the singular structure (wavefront set) of the object to the singular structure of the data. They do this for three cases: single scattering, double scattering with two discrete scatterers, and for wave-guiding within the object, in which they find that their results are similar to the case of three discrete scatterers. The main goal in their work is to identify objects via their wavefront sets without constructing an image.

Historically, in reflection seismology, two distinct methods have been used to attenuate multiples to obtain an approximation to singly scattered data. The first predicts the triply scattered data and then subtracts it from the data set. The second uses differences in event shape in a particular domain (e.g., the time-space domain or the frequency-wavenumber domain) to attenuate multiples through filtering. We first discuss the wave-theoretic (prediction) approach, and then summarize the signal-processing (filtering) approach. The work discussed here falls into the first category.

In the wave-theoretic approach, Kennett (1974; 1979b) used the Thomson-Haskell (Kennett, 1983) method in horizontally layered media to model synthetic seismograms containing both surface and internal multiples. He also applied his theory to suppress surface-related multiples in plane-layered elastic media (Kennett, 1979a). In a horizontally layered medium, Aminzadeh used the Bremmer series to model the seismic wavefield (Aminzadeh & Mendel, 1980) and construct filters to remove surface-related multiples (Aminzadeh & Mendel, 1981). Fokkema *et al.* (1994) ex-

tended their work on surface-related multiples to suppress internal multiples generated at a single layer in the subsurface, given the medium parameters above the particular layer. Weglein and others (Weglein *et al.*, 1997; Weglein *et al.*, 2003) have used the Lippmann-Schwinger series to model and process seismic data, including the suppression of both surface-related and internal multiples. In ten Kroode (2002) the theory behind this approach is given in both one and two dimensions. He extends the work of Weglein *et al.* to a two-dimensional subsurface model satisfying two conditions: ten Kroode's traveltime monotonicity assumption, and the condition that the wavefield does not contain caustics. The work of Berkhout and Verschuur (1997) extends, via layer stripping, from the surface-related multiple case to that of internal multiples. Jakubowicz (1998) proposes a method for modeling internal multiples by correlating one primary reflection with the convolution of two other primary reflections. When the two assumptions of ten Kroode are satisfied, our method can be reduced to this construction with the addition of a time-windowing procedure; this is discussed further in remark 7.1. Kelamis *et al.* (2002) use an approach similar to Jakubowicz, in which the multiples are constructed from a combination of different data sets, both at the surface and in the subsurface. Van Borselen (2002), provides another extension of the surface-related multiple-attenuation algorithm proposed by Fokkema and van den Berg (1993) to the case of internal multiples. His method removes internal multiples by identifying either the shallowest layer involved in the multiple generation or a pseudo-boundary through which the multiples are assumed to have passed. The surface-related multiple attenuation method is then applied at the identified layer. In any method that predicts the internal multiples and subtracts them, an adaptive subtraction technique such as that suggested by Guitton (2004) must be used to compensate for illumination effects in the predicted multiples.

Many signal-processing methods are also available for suppressing multiples. The work of Buttkus (1979), which uses the different apparent stacking velocities for multiples and primaries, is an example of such filtering in the data domain. Sava and Guitton (2004) show how such techniques can be applied in the image domain. Symes (1999) provides a method of velocity analysis that also suppresses multiples through a data perturbation applied to remove events in the data that do not fit the assumptions used in generating a velocity model. Another approach, described by Essenreiter *et al.* (2001), uses computer learning to predict

which events in the data are multiples given an initial set of identified primary and multiple events.

Although we are primarily concerned with the inverse problem, we discuss also the forward problem of modeling multiples. Since we are concerned here with the propagation of singularities, we construct only the most singular part of each term in the hybrid Lippmann-Schwinger-Bremmer series. From the third term of this series, we specifically construct distributions to model triply scattered data. These modeled data are used in the formulation of the inverse problem. The modeled data are needed in our approach since we construct the inverse problem through the prediction and subtraction of the triply scattered data in the image domain. The theory presented here does not account for multiple scattering from large numbers of fine layers. Instead, we assume a finite collection of scatterers with a separation large compared to the wavelength.

The work presented here is motivated by that of Razavy (1975) and Aminzadeh (1981). Razavy developed a method to estimate the wave velocity from reflection coefficients, with a series approach to the inverse problem. Aminzadeh used a Bremmer series approach to suppress surface-related multiples. In our hybrid series, we use the techniques from the generalized Bremmer series (de Hoop, 1996), although we remain in the framework of the Lippmann-Schwinger series as discussed by ten Kroode (2002). In this methodology we allow for lateral variation of the background velocity, as well as the presence of caustics. We require knowledge of the velocity model to the depth of the shallowest reflector involved in the triple scattering, and our approach to the suppression of multiples is applied in the image domain rather than the data domain. This is preferred as it allows for compensation of variations in illumination, which is different for triply scattered waves than for singly scattered waves. A by-product of the construction developed here is the prediction of internal multiples from subsets of the 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's functions. Following this we describe 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 two main results in (89). Following this, we describe the methodology for the inverse series arriving at our second result, (113). In the final section we illustrate through isochrons how the propagation of

singularities in the construction of triply scattered data is related to the propagation of singularities in the single scattering case.

2 DIRECTIONAL DECOMPOSITION

In the Bremmer series formulation, the wavefield is split into up- and down-going constituents. This is done by separating the vertical, z , derivative from the horizontal derivatives, writing the wave equation as a first-order system in z and then diagonalizing the resulting system. We begin with the scalar acoustic wave equation

$$\begin{aligned} \mathcal{P}u &= f, \\ \mathcal{P}(z, x, D_x, \partial_z, D_t) &= \\ &= -c(z, x)^{-2} D_t^2 + \sum_{j=1}^{n-1} D_{x_j}^2 - \partial_z^2, \end{aligned} \quad (1)$$

where $D_{x_j} \equiv -i\partial_{x_j}$, $D_t \equiv i\partial_t$ and $c(x, z)$ is the isotropic velocity function. (See table 1 for a complete list of symbols.) In order to separate the up- and down-going waves we consider the wave equation written as a system of first-order equations,

$$\partial_z \begin{pmatrix} u \\ \partial_z u \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -A(z, x, D_x, D_t) & 0 \end{pmatrix} \begin{pmatrix} u \\ \partial_z u \end{pmatrix} + \begin{pmatrix} 0 \\ -f \end{pmatrix}, \quad (2)$$

where $A(z, x, \xi, \tau) = c(z, x)^{-2} \tau^2 - \|\xi\|^2$ is the symbol of the transverse Helmholtz operator. In general, we use Greek letters for cotangent variables, which are dual variables of space variables (e.g. τ is radial frequency, the dual of time, t). The symbol $\|\cdot\|$ indicates the euclidean norm (length) of a vector. To simplify the notation in (2), we re-write it in matrix form

$$\partial_z D = -AD + M, \quad (3)$$

where

$$D = \begin{pmatrix} u \\ \partial_z u \end{pmatrix}, \quad A = \begin{pmatrix} 0 & -1 \\ A(z, x, D_x, D_t) & 0 \end{pmatrix} \quad (4)$$

and $M = \begin{pmatrix} 0 \\ -f \end{pmatrix}$.

We diagonalize the operator matrix A , which can be done microlocally*, away from the zeros of $A(z, x, \xi, \tau)$ modulo a smoothing operator. In this framework, there is a z -family of pseudodifferential

operator matrices $Q(z)$ such that microlocally we may write,

$$U = \begin{pmatrix} u_+ \\ u_- \end{pmatrix} = Q(z)D \quad X = \begin{pmatrix} f_+ \\ f_- \end{pmatrix} = Q(z)M \quad (5)$$

In this notation, u_{\pm} satisfy the one-way wave equations

$$(I\partial_z + Q(z)\partial_z Q^{-1}(z) + B)U = X, \quad (6)$$

where I is the identity matrix, introducing

$$B = Q(z)AQ^{-1}(z) = \begin{pmatrix} iB_+ & 0 \\ 0 & iB_- \end{pmatrix}, \quad (7)$$

where B_{\pm} has symbol

$$b_{\pm}(z, x, \xi, \tau) = \pm \sqrt{A(z, x, \xi, \tau)} = \pm b(z, x, \xi, \tau).$$

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 a generalization of eigenvectors and thus we are free to choose their normalization. We choose the vertical power flux normalization (de Hoop, 1996) so as to make B_{\pm} in (6) self-adjoint (the normalization changes the sub-principal part of the equation), we then have

$$\begin{aligned} Q &= \frac{1}{2} \begin{pmatrix} (Q_+^*)^{-1} & -iQ_+ \\ (Q_-^*)^{-1} & iQ_- \end{pmatrix} \\ Q^{-1} &= \begin{pmatrix} Q_+^* & Q_-^* \\ iQ_+^{-1} & -iQ_-^{-1} \end{pmatrix}, \end{aligned} \quad (8)$$

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

$$u = Q_+^* u_+ + Q_-^* u_-, \quad \text{and } f_{\pm} = \pm \frac{1}{2} i Q_{\pm} f. \quad (9)$$

The Q_{\pm} operators act, in the time variable, as time convolutions.

Noting that, in the flux normalization, the term $Q^{-1}\partial_z Q$ in (6) is of lower-order in the singularities, we omit it and introduce the propagators for the one-way wave equations (6) as

$$\overbrace{(I\partial_z + B)}^{P'} L' = I\delta, \quad L' = \begin{pmatrix} G_+^* & 0 \\ 0 & G_-^* \end{pmatrix}. \quad (10)$$

We use primes here because these operators do not yet have cut-offs applied to remove horizontal propagation. The cut-off procedure is explained in the following section. We can now write the solution of (6) as $U' = L' X$, using Duhamel's principle. In

*Something is true microlocally if it is true in a neighborhood both in space and direction. See Sjöstrand and Grigis (1994) for an introduction to microlocal analysis.

integral form this is

$$u'_+ = \int_{-\infty}^z G'_+(z, z_0) f_+(z_0) dz_0 \quad (11)$$

$$u'_- = \int_z^{\infty} G'_-(z, z_0) f_-(z_0) dz_0.$$

3 THE GREEN FUNCTION

In the previous section, we diagonalized the wave equation into two first-order equations. To do this, we implicitly assumed that the diagonal system is equivalent to the original equation. This is nearly the case, but the choice of a preferential direction alters the ability of the system to propagate singularities in directions orthogonal to the preferred direction. Here, we have chosen the vertical direction as the preferential direction and are thus unable to propagate singularities horizontally. To ensure that the diagonal system does not propagate singularities incorrectly, we need to attenuate the horizontally propagating singularities. The details of the method are given by Stolk and de Hoop (2004a); we give only a brief description here.

In order to identify horizontal propagation, we define the phase angle $\theta = \arcsin(c(z, x) \|\tau^{-1} \xi\|)$, where (ζ, ξ) is the cotangent vector associated with (z, x) and $c(z, x)$ is the velocity. Note that if the angle θ , is less than $\pi/2$ on a ray segment, the vertical velocity $\frac{dz}{dt}$ does not change sign, allowing the parameterization of the ray segment by z . Thus, for any ray segment and any given angle $\theta < \pi/2$, we can define a maximal interval,

$$(z_{\min, \pm}(z, x, \xi, \tau, \theta), z_{\max, \pm}(z, x, \xi, \tau, \theta)), \quad (12)$$

for which the propagation away from a particular point (z, x, ξ, τ) can be parameterized by z . In Figure 1, the interval $(z_{\min, -}, z_{\max, -})$ is illustrated; it is the maximal interval containing the point (z, x) such that the angle of the ray with the vertical, θ , does not exceed a given value; in this case that value is θ_2 .

The angle θ can be given physical meaning by looking at the ray picture, as is done in Figure 1. The wave equation is solved in all of phase space, however, which means that any restrictions we put on the wave operator must be done in all of phase space. To this end we introduce the set

$$I_\theta = \{(z, x, t, \zeta, \xi, \tau) \mid \arcsin(c(z, x) \|\tau^{-1} \xi\|) < \theta, |\zeta| < C|\tau|\}, \quad (13)$$

illustrated in Figure 2, where C is some constant everywhere larger than $c(z, x)^{-1}$. Finally, we con-

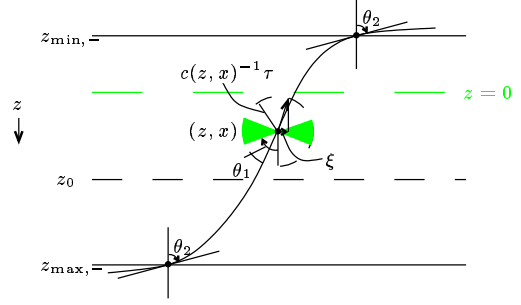


Figure 1. Removing horizontal propagations. The symbol of the cut-off operator ψ is one up to an angle of θ_1 and then tapers smoothly to zero at the angle θ_2 . This removes all propagation at angles larger than θ_2 .

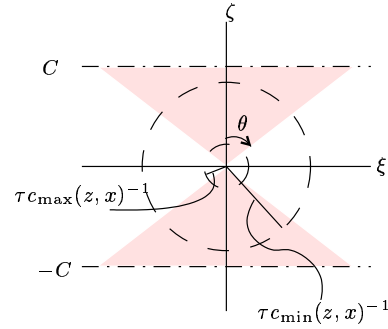


Figure 2. Illustration of I_θ . The shaded region represents the ray directions in the set. The minimum velocity in the region is c_{\min} and the maximum is c_{\max} .

struct the sets

$$J_-(z_0, \theta) = \{(z, x, t, \zeta, \xi, \tau) \in I_\theta \mid \tau^{-1} \zeta < 0 \text{ and } z_{\max, -}(z, x, \xi, \tau, \theta) \geq z_0\}, \quad (14)$$

and

$$J_+(z_0, \theta) = \{(z, x, t, \zeta, \xi, \tau) \in I_\theta \mid \tau^{-1} \zeta > 0 \text{ and } z_{\max, +}(z, x, \xi, \tau, \theta) \geq z_0\}, \quad (15)$$

which exclude all parts of phase space in which singularities propagate horizontally. Figure 1 illustrates the set $J_-(z_0, \theta_2)$, considering the shaded region as excluded from the set. The condition $z_{\max, -}(z, x, \xi, \tau, \theta) \geq z_0$, along with the implicit requirement that $z_{\min, -} < 0$ ensure that the two points between which one propagates the wavefield are within the allowed propagation interval $(z_{\min, -}, z_{\max, -})$.

The sets J_\pm describe the portions of phase space that must be excluded to remove horizontally propagating singularities. To actually remove singularities from these regions, we define a pseudodifferential cutoff $\psi_{-,1} = \psi_{-,1}(z, z_0, x, D_x, D_t)$

with symbol satisfying

$$\psi_{-,1}(z, x, \xi, \tau) \sim 1 \text{ on } J_-(z_0, \theta_1), \quad (16)$$

$$\psi_{-,1}(z, x, \xi, \tau) \in S^\infty \text{ outside } J_-(z_0, \theta_2), \text{ if } z - z_0 > \delta. \quad (17)$$

The angle θ_1 can be freely chosen within the interval $(0, \theta_2)$. Singularities propagating at an angle less than θ_1 are unaffected by the cutoff; at angles greater than θ_2 , the operator is infinitely smooth. We can then write

$$\psi_{-,1}u \equiv \psi_{-,1}Q_-u_-, \quad (18)$$

which suppresses the singularities outside of J_- . (We may do the equivalent for the $+$ operators.) We may now rewrite the operators defined above, with the singularities outside of J_- (or J_+) suppressed. For the wave operator, we add a dissipative term to P' , giving

$$P = \partial_z + B + C, \quad (19)$$

where C is a diagonal operator matrix in which each entry is first-order homogeneous with non-negative, real principal symbol. It was shown by Stolk and de Hoop (2004a) and references therein that the solution operator L to P is

$$L = \begin{pmatrix} G_+ & 0 \\ 0 & G_- \end{pmatrix} = \begin{pmatrix} \psi_{+,1}G'_+ & 0 \\ 0 & \psi_{-,1}G'_- \end{pmatrix}. \quad (20)$$

From this point onward we will assume that the above procedure has been followed and thus singularities do not propagate horizontally.

4 SCATTERING: CONTRAST SOURCE

The Bremmer formulation assumes a degree of smoothness in the velocity model. In the contrast formulation, the velocity is split into a smooth background and the singularities (reflectors) in the model are contained in a separate term. This form leads to the Lippmann-Schwinger equation. We use a hybrid of the two; the Lippmann-Schwinger contrast source and the Bremmer decomposition into up- and down-going waves. We begin with the wave equation in the smooth background and the true medium respectively

$$(\partial_z + A_0)D_0 = M, \quad (\partial_z + A)D = M, \quad (21)$$

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 in the smooth background from that in the true medium gives the contrast formulation

$$(\partial_z + A_0)\delta D = -\delta A D, \quad (22)$$

where $D = \delta D + D_0$ and $A = \delta A + A_0$. We have (cf. (4))

$$\delta A = \begin{pmatrix} 0 & 0 \\ \delta A & 0 \end{pmatrix} \quad (23)$$

where

$$\delta A = -2c_0^{-3}\delta c D_t^2 = -a D_t^2. \quad (24)$$

We insert the Bremmer formulation into that above by diagonalizing the A_0 -matrix. We apply the (smooth background) diagonalizing Q operator matrices to transform the system in (22). Using the diagonalization procedure of section 2 equation (7) in particular, we find

$$(\partial_z + B_0)\delta U = -Q(z)\partial_z Q^{-1}(z)\delta U - Q(z)\delta A Q^{-1}(z)U, \quad (25)$$

recalling, from section 2, the definition of U

$$U = QD \quad (26)$$

while,

$$U_0 = QD_0 \quad (27)$$

$$\delta U = Q\delta D. \quad (28)$$

The Q operator matrix is common in all the transformations. Note that δA will not, in general, be diagonalized by Q as the Q operators diagonalize A in the background velocity model only.

Since we have used the flux normalization, the $-Q(z)\partial_z Q^{-1}(z)\delta U$ term is of lower-order as before; hence we omit this contribution so that

$$(\partial_z + B_0)\delta U = -Q(z)\delta A Q^{-1}(z)U, \quad (29)$$

where $Q(z)\delta A Q^{-1}(z)$ is given explicitly as

$$Q(z)\delta A Q^{-1}(z) = \frac{i}{2} \begin{pmatrix} Q_+(z) a Q_+^*(z) & Q_+(z) a Q_-^*(z) \\ -Q_-(z) a Q_+^*(z) & -Q_-(z) a Q_-^*(z) \end{pmatrix} D_t^2. \quad (30)$$

In (29), we can make the analogy with (6) where δU plays the role of U and $-Q(z)\delta A Q^{-1}(z)U$ that of X , the contrast source.

We make the comparison between the elements of V and the reflection and transmission coefficients of de Hoop (1996) viz,

$$V = Q(z)a D_t^2 Q^{-1}(z) = \frac{i}{2} \begin{pmatrix} Q_+ a Q_+^* & Q_+ a Q_-^* \\ -Q_- a Q_+^* & -Q_- a Q_-^* \end{pmatrix} D_t^2 = \begin{pmatrix} S_{++} & S_{-+} \\ S_{+-} & S_{--} \end{pmatrix} D_t^2. \quad (31)$$

Here, S_{++} and S_{--} are interpreted as transmission operators since they govern scatterings between singularities traveling in the same direction.

In contrast, S_{+-} and S_{-+} are interpreted as reflection operators because they govern scatterings that result in a change of direction; from up-going to down-going and down-going to up-going respectively. (The the off-diagonal entries in the V operator matrix are true adjoints of one another whereas the diagonal entries are adjoints only to principal parts.)

To simplify the notation, we define

$$P_0 = \partial_z + B_0, \quad (32)$$

its forward parametrix[†],

$$L_0 = \begin{pmatrix} G_+ & 0 \\ 0 & G_- \end{pmatrix} \quad (33)$$

recalling

$$V = Q(z) \delta A Q^{-1}(z). \quad (34)$$

In this notation, (29) reduces to

$$P_0 \delta U = -VU, \quad (35)$$

or

$$\delta U = -L_0(VU). \quad (36)$$

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

$$\delta U = -L_0(VU_0) - L_0(V\delta U), \quad (37)$$

or equivalently,

$$(I + L_0 V) \delta U = -L_0(VU_0). \quad (38)$$

In analogy with (24), we introduce \widehat{V} , the matrix of $S_{\pm\pm}$ coefficients, viz,

$$V(z, x, D_t) = \widehat{V}(z, x)D_t^2, \quad (39)$$

which gives

$$(I + D_t^2 L_0 \widehat{V}) \delta U = -D_t^2 L_0(\widehat{V}U_0), \quad (40)$$

where $\widehat{V}\delta U$ and $\widehat{V}U_0$ are products of distributions. This is the resolvent equation in our hybrid Lippmann-Schwinger-Bremmer formulation. (See (Yoshida, 1995) for details on resolvent equations.)

5 SCATTERING SERIES

5.1 Forward series

Having recognized (40) as a resolvent equation, we set up the recursion

$$\delta U = \sum_{m \in \mathbb{N}} (-1)^m \delta U_m(\widehat{V}), \quad (41)$$

[†]The parametrix of an operator is an asymptotic approximation to its inverse.

where

$$\delta U_1(\widehat{V}) = D_t^2 L_0(\widehat{V}U_0), \quad \text{and} \quad \delta U_m(\widehat{V}) = D_t^2 L_0(\widehat{V} \delta U_{m-1}(\widehat{V})). \quad (42)$$

This recursion is motivated by the Neumann series. The expressions in (41) and (42) 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 arriving at all depth levels. We therefore define a restriction operator, R , which restricts the operator L_0 to the acquisition surface, $z = 0$. To avoid difficulties with the free boundary at the surface we will assume that the medium contrast, contained in V , has its support away from $z = 0$. In addition, (42) is in the diagonal system; to return to the 'true' coordinate system in which we make observations we must also apply the Q^{-1} operator to the expression for δU . We thus rewrite (41)

$$RQ^{-1}\delta U = \sum_{m \in \mathbb{N}} (-1)^m RQ^{-1}\delta U_m(\widehat{V})$$

where $RQ^{-1}\delta U_1 = D_t^2 M_0(\widehat{V}U_0)$, and

$$RQ^{-1}\delta U = -D_t^2 M_0(\widehat{V}(U_0 + \sum_{m \in \mathbb{N}} (-1)^m \delta U_m(\widehat{V}))), \quad (43)$$

upon introducing the operator $M_0 = RQ^{-1}L_0$. The composition RL_0 is transversal provided there are no grazing rays (Stolk & de Hoop, 2000), which are already excluded by the ψ_1 cut-off. The composition with Q^{-1} also does not change the properties of the composite operator provided we satisfy the assumption in the generalized Bremmer series (de Hoop, 1996). Note that the recursion in (43) gives an expression for the data at the surface in terms of the unrestricted data from the previous term; the restriction is applied after the recursion is constructed.

For the leading-order, single scattering term we introduce the shorthand notation

$$d_1 = -RQ^{-1}\delta U_1(\widehat{V}) = F(\widehat{V}) = -D_t^2 M_0(\widehat{V}U_0). \quad (44)$$

Note that d_1 describes only the singly scattered part of the data. The operator F is the usual Born modeling operator. We can construct a left inverse to F , by $N^{-1}F^*$, where $N = F^*F$ is the normal operator, which is elliptic pseudodifferential under the DSR assumption of Stolk and de Hoop (2004a), and F^* is the adjoint of the modeling operator, i.e., the imaging/migration operator. This gives the estimate

$$\widehat{V}_1 = N^{-1}F^*(d_1), \quad (45)$$

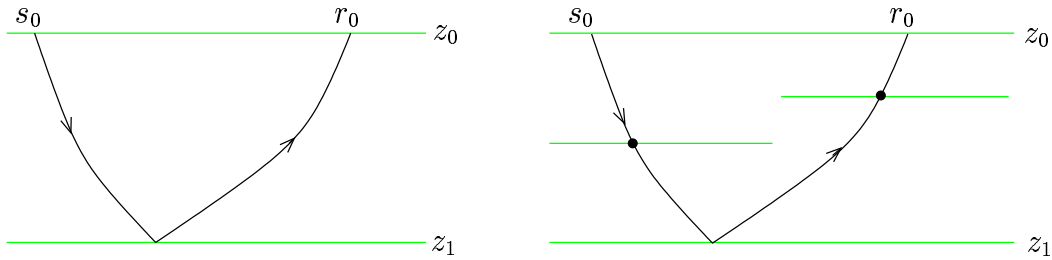


Figure 3. Single scattering versus primary reflection. The black dots indicate transmission events.

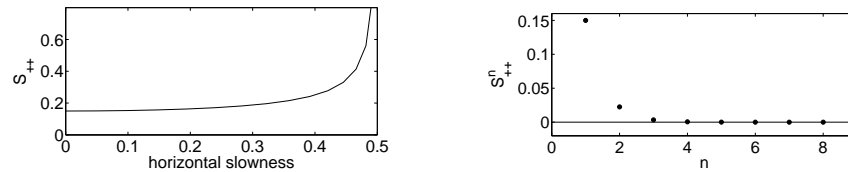


Figure 4. Left panel: transmission coefficient. Right panel: vertical transmission coefficient raised to the power n .

of \widehat{V} , via the single scattering inverse.

To compare the Bremmer series formulation ((VII.1)-(VII.22) of de Hoop (1996)) to the recursion in (41) we first make the following identifications. From (VII.1) and (VII.12) we note that W_0 of (de Hoop, 1996) is equivalent to δU_1 . From this formulation we now see that $-D_t^2 L_0 \widehat{V}$ corresponds with K of equation (VII.15) in (de Hoop, 1996) and (41) corresponds to equation (VII.22). Thus we have shown that there is a correspondence between the Bremmer series and the series representation to the solution of the Lippmann-Schwinger equation, with the exception that the Bremmer series contains an $m = 0$ term associated with waves that propagate directly downward (without ever reflecting) away from the source. We use this hybrid formulation to derive operators that model both ‘singly’ and ‘triplly’ scattered waves.

By a singly scattered wave, we mean a wave that has been reflected or transmitted once, such as that shown on the left in Figure 3. By a primary reflection we mean any ‘ray-path’ (more accurately wave-path since we use full-wave solutions rather than ray theory) that has reflected only once but may have gone through several transmission, or scatterings where the direction of the wave does not change. This type of contribution is depicted in the right panel of Figure 3. Primary reflections have the same traveltime as singly scattered waves but the amplitude will be smaller because of the scaling by additional factors of S_{--} or S_{++} . Figure 4 illustrates just how much smaller this contribution would be. The same distinction can be made be-

tween first-order internal multiples and triply scattered waves. The diagram on the right of Figure 3 is a triply scattered event, generated by the third term of the series in (41). The third-order contributions that we take into account are those for which each scattering event is a reflection, i.e., after the scattering the singularities propagate in the opposite direction to that in which they were propagating before the scattering. We refer to these contributions as first-order internal multiples.

5.2 Inverse scattering series using all the data

The forward scattering series (43) models the data, given a representation of the medium in terms of a smooth background and a contrast. To address the inverse problem, we return to (40) and write it as a resolvent equation in terms of the data, d , rather than the medium contrast, \widehat{V} ,

$$D_t^2 L_0(\widehat{V}(U_0 - (-\delta U))) = -\delta U, \tag{46}$$

or, returning to observables via the RQ^{-1} operator,

$$D_t^2 M_0(\widehat{V}(U_0 - (-\delta U))) = -d, \tag{47}$$

where we notice that $d = RQ^{-1}\delta U$. We then assume that the medium can be represented in terms of a series of linear operators acting on the data viz,

$$\widehat{V} = \sum_{m \in \mathbb{N}} \widehat{V}_m(\delta U), \text{ or, } \widehat{V} = \sum_{m \in \mathbb{N}} \widehat{V}_m(d), \tag{48}$$

where m indicates the order of \widehat{V}_m in the data, δU , given in (41), or d in observables. This series rep-

resentation is suggested for quantum mechanical problems by Moses (1956), where the analogue of (48) is his equation (3.12). It is also suggested by Razavy (1975) for wave problems, in which the analogue of (48) is his equation (33).

Substituting (48) into (41) yields a recursion for \widehat{V}_m in terms of δU

$$\delta U = -D_t^2 L_0(\widehat{V}_1 U_0) \quad (49)$$

$$0 = -D_t^2 L_0(\widehat{V}_2 U_0) + D_t^4 L_0(\widehat{V}_1 L_0(\widehat{V}_1 U_0)) \quad (50)$$

$$\begin{aligned} 0 &= -D_t^2 L_0(\widehat{V}_3 U_0) + D_t^4 L_0(\widehat{V}_2 L_0(\widehat{V}_1 U_0)) \\ &\quad + D_t^4 L_0(\widehat{V}_1 L_0(\widehat{V}_2(U_0))) \\ &\quad - D_t^6 L_0(\widehat{V}_1 L_0(\widehat{V}_1 L_0(\widehat{V}_1 U_0))), \end{aligned} \quad (51)$$

etc.

Substituting (48) into (43) yields a recursion for V_m in terms of the true data d

$$d = -D_t^2 M_0(\widehat{V}_1 U_0) \quad (52)$$

$$0 = -D_t^2 M_0(\widehat{V}_2 U_0) + D_t^4 M_0(\widehat{V}_1 L_0(\widehat{V}_1 U_0)) \quad (53)$$

$$\begin{aligned} 0 &= -D_t^2 M_0(\widehat{V}_3 U_0) + D_t^4 M_0(\widehat{V}_2 L_0(\widehat{V}_1 U_0)) \\ &\quad + D_t^4 M_0(\widehat{V}_1 L_0(\widehat{V}_2(U_0))) \\ &\quad - D_t^6 M_0(\widehat{V}_1 L_0(\widehat{V}_1 L_0(\widehat{V}_1 U_0))), \end{aligned} \quad (54)$$

etc.

Equation (54) can be simplified using (50), since $D_t^2 M_0(\widehat{V}_2 U_0)$ and $D_t^4 M_0(\widehat{V}_1 L_0(\widehat{V}_1 U_0))$ are identical distributions and $D_t^2 M_0 \widehat{V}_1$ is a linear operator, to read

$$D_t^2 M_0(\widehat{V}_3 U_0) = D_t^4 M_0(\widehat{V}_2 L_0(\widehat{V}_1 U_0)). \quad (55)$$

Solving these recursions for \widehat{V} gives a solution for the medium contrast. The first term in the series, given in (52), models singly scattered data. The third term, in (54), models first-order internal multiples as well as other primary events. (The second term, given in (53), models, among other things, primary events.) We have the following relation between the $\widehat{V}_j = \widehat{V}_j(d)$

$$\begin{aligned} D_t^2 M_0(\widehat{V}_2 U_0) &= D_t^4 M_0(\widehat{V}_1 L_0(\widehat{V}_1 U_0)) \\ &= -D_t^2 M_0(\widehat{V}_1 \delta U), \end{aligned} \quad (56)$$

$$\begin{aligned} D_t^2 M_0(\widehat{V}_3 U_0) &= D_t^4 M_0(\widehat{V}_2 L_0(\widehat{V}_1 U_0)) \\ &= -D_t^2 M_0(\widehat{V}_2 \delta U) \end{aligned} \quad (57)$$

$$D_t^2 M_0(\widehat{V}_j U_0) = -D_t^2 M_0(\widehat{V}_{j-1} \delta U). \quad (58)$$

The general term in the recursion (58) follows because higher order terms are built from lower-order terms through the application of additional operators of the form $M_0 \widehat{V}_i$ to the $(j-i)^{\text{th}}$ -order terms to form terms of order j . Thus the sum of terms of order j will contain sub-series of the form

$$D_t^2 M_0 \widehat{V}_1 (\text{sum of terms of order } j-1 \text{ from (49-51)}),$$

$$D_t^2 M_0 \widehat{V}_2 (\text{sum of terms of order } j-2 \text{ from (49-51)}),$$

etc. For $j \geq 2$ the sub-series in brackets sum to zero because of the zero on the left-hand side of (50). Thus, the only terms of order j remaining in the series are of the form

$$D_t^2 M_0 \widehat{V}_j U_0$$

and

$$D_t^4 M_0 \widehat{V}_{j-1} L_0 \widehat{V}_1,$$

from which the general term (58) follows. Note the similar structure between (58) and (42); (58) constructs the medium contrast in terms of the data, while (42) constructs the data in terms of the medium contrast.

From these relations, we deduce that the only inverse operator we need is $N^{-1} F^*$. Using this we can write the \widehat{V} -series as

$$\widehat{V} = N^{-1} F^* d - N^{-1} F^* \left(\sum_{m \in \mathbb{N}} M_0(\widehat{V}_m \delta U) \right). \quad (59)$$

We identify the series in the second term as a correction to the first term, which is the single scattering inverse. Additionally, $N^{-1} F^*$ can be replaced with A_{WE} , the wave equation angle transform (Stolk & de Hoop, 2004b), which computes an image $\Psi \widehat{V}(z, x, p)$, where $p = \tau^{-1} \xi$ is the horizontal slowness. An accurate reconstruction of $\Psi \widehat{V}(z, x, p)$ will be independent of p ; this property can be exploited to improve the knowledge of the smooth background medium. The pseudodifferential operator Ψ corrects for illumination effects in the reconstruction of \widehat{V} ; we cannot reconstruct a portion of the image, $\widehat{V}(z_0, x_0, p_0)$, for example, if the data do not scatter from the point (z_0, x_0) with slowness p_0 .

Solving (56) for \widehat{V}_2 gives

$$\widehat{V}_2 = N^{-1} F^* (D_t^4 M_0(\widehat{V}_1 L_0(\widehat{V}_1 U_0))) = D_t^2 \widehat{V}_1 L_0 \widehat{V}_1. \quad (60)$$

This allows us to simplify (55) to

$$D_t^2 M_0(V_3 U_0) = D_t^6 M_0(V_1 L_0(V_1 L_0(V_1 U_0))). \quad (61)$$

In the above we have nowhere assumed the absence of caustics in the wavefield. In fact, thus far we have made only the DSR assumption (Stolk & de Hoop, 2004a) excluding horizontal propagation.

6 MODELING

6.1 Single scattering to double square-root

The first term in the forward scattering series given in (41) can be used to construct data in the

Born approximation in accordance with (Stolk & de Hoop, 2004aequation 3.10). We formulate the solution only for the upward propagating portion of δU_1 , which we denote by δu_- because we assume that only waves arriving from below the surface are recorded. We first determine the form of the down-going portion 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 f_+ in (9) and that for u_+ in (11) we find that

$$u_{+,0}(z_1, x_1, t_1, z_0, s_0, t_{s_0}) = \frac{i}{2} \int_{-\infty}^{z_1} d\tilde{z}_0 \int_{\mathbb{R}_+} d\tilde{s}_0 \int d\tilde{t}_{s_0} G_+(z_1, x_1, t_1 - \tilde{t}_{s_0}, \tilde{z}_0, \tilde{s}_0) Q_{+,\tilde{s}_0}(\tilde{z}_0) f(\tilde{z}_0, \tilde{s}_0, \tilde{t}_{s_0}, z_0, s_0, t_{s_0}), \quad (62)$$

where we will adopt the convention that an integral without bounds is assumed to be an integration over \mathbb{R}^2 . In general, s represents a source parameter, r represents a receiver parameter, t is a time parameter and z is depth, regardless of subscripts and superscripts; a complete table of symbols is given in table 8. The operator G_+ propagates between the levels z_0 and z_1 , with its action being in the lateral variables s_0 , and t_{s_0} ; we will also use the notation $G_+(z_1, z_0)$ for the operator 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 G_{\pm} as the difference of elapsed time and initial source time as the wave equation is time translation invariant. Expression (62) is valid for $z_1 > z_0$. We note that z_0, s_0 , and t_{s_0} are parameters assumed to be known. Here $Q_{-,s}(z)$ is short for $Q_-(z, s, D_s, D_t)$, suppressing the time dependence as well as the separate integrations introduced by this pseudodifferential operator. If we assume a point source, we have $f(\tilde{z}_0, \tilde{s}_0, \tilde{t}_{s_0}, z_0, s_0, t_{s_0}) = \delta(\tilde{s}_0 - s_0)\delta(\tilde{z}_0 - z_0)\delta(\tilde{t}_{s_0} - t_{s_0})$.

Next, we compute c_- , the up-going term in the contrast source defined by,

$$\begin{pmatrix} c_+ \\ c_- \end{pmatrix} = VU_0 = V \begin{pmatrix} u_{+,0} \\ u_{-,0} \end{pmatrix}. \quad (63)$$

Using the expression for V in (31), and recalling that $u_{-,0} = 0$ since the source generates no up-going waves, we obtain an expression for c_- , viz,

$$c_-(z_1, x_1, t_1) = -\frac{i}{2} D_{t_1}^2 Q_{-,x_1}(z_1) a Q_{+,x_1}^*(z_1) u_{+,0}(z_1, x_1, t_1, z_0, s_0, t_{s_0}). \quad (64)$$

Substituting the expression above for f_- in (11) gives

$$\delta u_{-,1}(z_0, r_0, t_{r_0}, z_0, s_0, t_{s_0}) = -\frac{i}{2} D_{t_{r_0}}^2 \int_z d z_1 \int d x_1 \int_{\mathbb{R}_+} d t_1 G_-(z_0, r_0, t_{r_0} - t_1, z_1, x_1) Q_{-,x_1}(z_1) a(z_1) Q_{+,x_1}^*(z_1) u_{+,0}(z_1, x_1, t_1, z_0, s_0, t_{s_0}) \quad (65)$$

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 (41). To return to observables, we apply the RQ^{-1} operator to obtain

$$d_1(s_0, r_0, t_{r_0} - t_{s_0}) = \int d\tilde{s}_0 \int_{\mathbb{R}_+} d\tilde{t}_{s_0} \left\{ \frac{1}{4} D_{t_{r_0}}^2 \int_0^\infty d z_1 \int d x_1 \int_{\mathbb{R}_+} d t_1 Q_{-,r_0}^*(0) G_-(0, r_0, t_{r_0} - t_1, z_1, x_1) \underbrace{Q_{-,x_1}(z_1) a(z_1) Q_{+,x_1}^*(z_1)}_{\widehat{S}_{+-}} G_+(z_1, x_1, t_1 - \tilde{t}_{s_0}, 0, \tilde{s}_0) Q_{+,\tilde{s}_0}(0) \right\} f(0, \tilde{s}_0, \tilde{t}_{s_0}, 0, s_0, t_{s_0}). \quad (66)$$

This gives the Born modeled data distribution in terms of the G_{\pm} , the solutions of the single square-root equation. This is the first term in the series in (43); the expression within the braces, is equivalent to the operator $F(\widehat{V})$. Since a is compactly supported in z_1 , the integral over z_1 is actually over a compact set.

In the downward continuation approach to seismic modeling, source and receiver wavefronts are propagated simultaneously upward from the scattering point to the surface. To mimic this, we apply reciprocity to the expression for F given in (66) so that we may write it in terms of G_- only. It is possible to do this only because we have returned to observables. This is because reciprocity holds for the full G operators and not for the G_{\pm} individually. Applying reciprocity gives

$$(F\widehat{V})(s_0, r_0, t_{r_0} - \tilde{t}_{s_0}) = \frac{1}{4} D_{t_{r_0}}^2 \int_0^\infty d z_1 \int d x_1 \int_{\mathbb{R}_+} d t_1 Q_{-,r_0}^*(0) G_-(0, r_0, t_{r_0} - t_1, z_1, x_1) Q_{-,x_1}(z_1) a(z_1) Q_{-,\tilde{s}_0}^*(0) G_-(0, \tilde{s}_0, t_1 - \tilde{t}_{s_0}, z_1, x_1) Q_{-,x_1}(z_1). \quad (67)$$

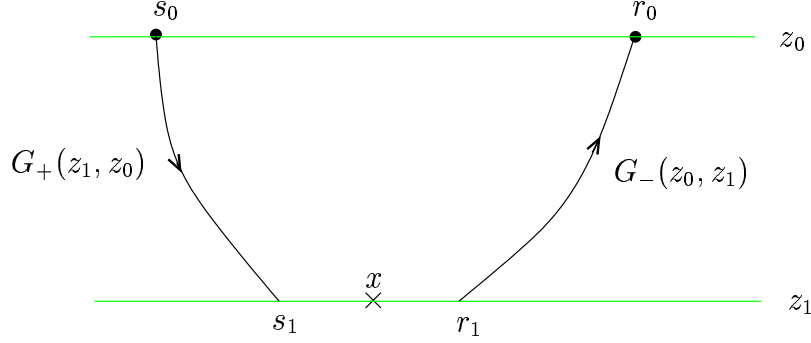


Figure 5. Notation and conventions.

To write (67) in terms of the Green function for the double-square-root equation, there must be integrations in (x_1, t_1) for each of the Green functions. To introduce these integrations we introduce two extension operators

$$E_1 : u(z, x) \mapsto \delta(r - s)u(z, \frac{r+s}{2}), \quad (68)$$

$$E_2 : u(z, r, s) \mapsto \delta(t)u(z, r, s). \quad (69)$$

These operators map the medium contrast, a , into data. They also apply the adjoint imaging condition, namely assuring that the two rays meet at the scattering point at time 0. The δ -functions allow the splitting of the single integration joining the two G_- operators into two integrals. This allows the two Green functions to both act on these data as we expect intuitively; each Green function continues the data from the subsurface to the surface. This condition is illustrated in Figure 5. With these operators, we re-write (67)

$$(F\hat{V})(s_0, r_0, t_{r_0} - \tilde{t}_{s_0}) = \frac{1}{4}D_{t_{r_0}}^2 \int_0^\infty dz_1 \int ds_1 \int dr_1 \int_{\mathbb{R}_+} dt_0 \int_{\mathbb{R}_+} dt_1 Q_{-,r_0}^*(0)G_-(0, r_0, t_{r_0} - t_1 - t_0, z_1, r_1)Q_{-,r_1}(z_1) \\ Q_{-,s_0}^*(0)G_-(0, \tilde{s}_0, \underline{t_1 - \tilde{t}_{s_0}}, z_1, s_1, 0)Q_{-,s_1}(z_1)(E_1 E_2 a)(z_1, s_1, r_1, t_0). \quad (70)$$

We note that the two one-way Green's functions are connected through time convolution. To make this explicit we change integration variables from t_1 to $t' = t_1 - \tilde{t}_{s_0}$ (the underlined expression in (70)), giving

$$(F\hat{V})(\tilde{s}_0, r_0, t_{r_0} - \tilde{t}_{s_0}) = \frac{1}{4}D_t^2 \int_0^\infty dz_1 \int ds_1 \int dr_1 \int_{\mathbb{R}_+} dt_0 \int_{\mathbb{R}_+} dt' Q_{-,r_0}^*(0)G_-(0, r_0, t_{r_0} - \tilde{t}_{s_0} - t' - t_0, z_1, r_1)Q_{-,r_1}(z_1) \\ Q_{-,s_0}^*(0)G_-(0, \tilde{s}_0, t', z_1, s_1)Q_{-,s_1}(0)(E_1 E_2 a)(z_1, s_1, r_1, t_0). \quad (71)$$

To write (71) in terms of the Green function for the double square-root operator, we first re-arrange terms giving

$$(F\hat{V})(\tilde{s}_0, r_0, t_{r_0} - \tilde{t}_{s_0}) = \frac{1}{4}D_t^2 Q_{-,r_0}(0)Q_{-,s_0}^*(0) \int_0^\infty dz_1 \int ds_1 \int dr_1 \int_{\mathbb{R}_+} dt_0 \\ \left\{ \int_{\mathbb{R}_+} dt' G_-(0, r_0, t_{r_0} - \tilde{t}_{s_0} - t' - t_0, z_1, r_1)G_-(0, \tilde{s}_0, t', z_1, s_1) \right\} \\ Q_{-,r_1}(z_1)Q_{-,s_1}(z_1)(E_1 E_2 a)(z_1, s_1, r_1, t_0). \quad (72)$$

We then introduce the operator $H(z_0, z_1)$, for $z_1 > z_0$, as

$$(H(z_0, z_1))(s_0, r_0, t - t_0, s_1, r_1) = \int_{\mathbb{R}_+} (G_-(z_0, z_1))(r_0, t - t' - t_0, r_1)(G_-(z, z_0))(s_0, t', s_1) dt'. \quad (73)$$

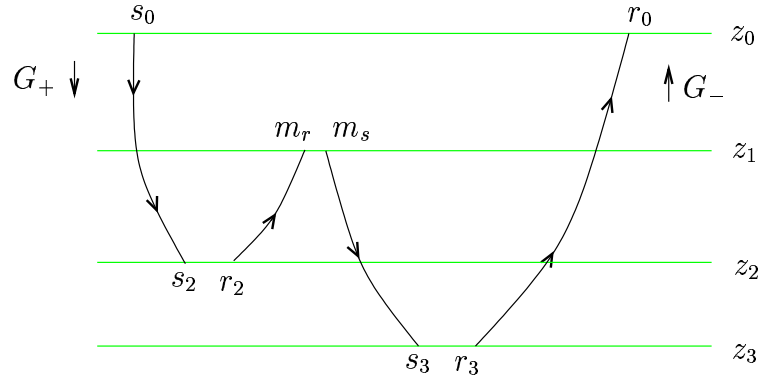


Figure 6. Triple scattering notations and conventions. This illustration assumes that the E operators have not yet been applied to be clear which variable refers to which leg of the interactions. The time notations are illustrated in Figure 7.

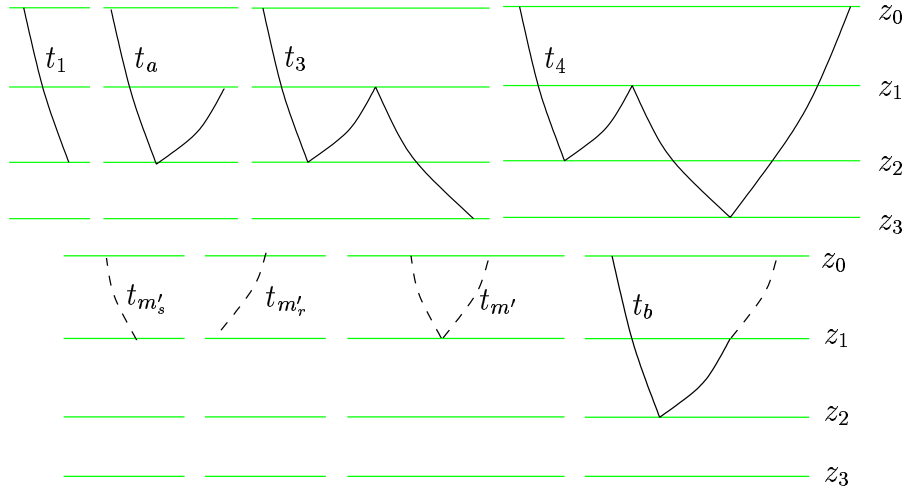


Figure 7. Triple scattering notations and conventions: time notations.

Using this operator we arrive at the final expression for d_1 , writing H in operator notation

$$d_1(s_0, r_0, t_{r_0} - t_{s_0}) = \int d\tilde{s}_0 \int_{\mathbb{R}_+} d\tilde{t}_{s_0} \left\{ \frac{1}{4} D_t^2 Q_{-,r_0}^*(0) Q_{-,s_0}^*(0) \int_0^\infty dz_1 H(0, z_1) \right. \\ \left. Q_{-,r_1}(z_1) Q_{-,s_1}(z_1) (E_1 E_2 a)(z_1, s_1, r_1, t_0) \right\} f(0, \tilde{s}_0, \tilde{t}_{s_0}, 0, s_0, t_{s_0}). \quad (74)$$

This expression is equivalent to (Stolk & de Hoop, 2004a equation 3.10). The two Green functions in this final modeling operator give it the form of a double-square-root equation (Claerbout, 1985).

6.2 triple scattering to double square-roots

In (74), we show 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 first-order internal multiples. To do this, we need a good idea of how the first-order internal multiples are modeled; we compute them using a method similar to that used to arrive at (74).

Following the diagram in Figure 6, we see that 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, since the second leg (from r_2 to m_r) does not reach the surface. Thus, we define

$$\delta u_{-,1}(z_1, m, t_a, 0, s_0, t_{s_0}) = \frac{1}{4} D_{t_a}^2 \int_{\mathbb{R}_+} d\tilde{s}_0 \int_{\mathbb{R}_+} d\tilde{t}_{s_0} \int_{z_1}^{\infty} dz_2 \int ds_2 \int dr_2 \int_{\mathbb{R}_+} dt_0 \int_{\mathbb{R}_+} dt' G_-(z_1, m, t_a - \tilde{t}_{s_0} - t' - t_0, z_2, r_2) \\ G_-(0, \tilde{s}_0, t', z_2, s_2) Q_{-,r_2}(z_2) Q_{-,s_2}(z_2) (E_1 E_2 a)(z_2, s_2, r_2, t_0) Q_{-,s_0}^*(0) f(0, \tilde{s}_0, \tilde{t}_{s_0}, 0, s_0, t_{s_0}), \quad (75)$$

where $t' = t_1 - \tilde{t}_{s_0}$ and t_a is the total time traveled along the ray (see Figure 7). In (75), we have applied the R operator to restrict the source to $z_0 = 0$, but have not returned to observables as the second leg, $G_-(z_1, m, t_a - \tilde{t}_{s_0} - t' - t_0, z_2, r_2)$, does not reach the surface. The field, δu_1 acts as the source of waves for propagation from m_s to s_3 , through the contrast source formulation used in the single-scattering case. This gives,

$$\delta u_{+,2}(z_3, x_3, t_3, 0, s_0, t_{s_0}) = \frac{i}{2} D_{t_3}^2 \int_{z_0}^{z_3} dz_1 \int_{\mathbb{R}_+} dm \int dt_a \\ G_+(z_3, x_3, t_3 - t_a, z_1, m) Q_{+,m}(z_1) a(z_1) Q_{-,m}^*(z_1) \delta u_{-,1}(z_1, m, t_a, 0, s_0, t_{s_0}), \quad (76)$$

which acts as a contrast source for the final ray, connecting r_3 with r_0 ,

$$d_3(s_0, r_0, t_4 - t_{s_0}) = -\frac{i}{2} D_{t_4}^2 Q_{-,r_0}^*(0) \int_0^{\infty} dz_3 \int dx_3 \int_{\mathbb{R}_+} dt_4 \\ G_-(0, r_0, t_4 - t_3, z_3, x_3) Q_{-,x_3}(z_3) a(z_3) Q_{+,x_3}^*(z_3) \delta u_{+,2}(z_3, x_3, t_3, 0, s_0, t_{s_0}), \quad (77)$$

where we have returned to observables through the operator RQ^{-1} , introduced in (43). For the above construction to be valid, it is necessary that the three scattering positions are sufficiently far apart. This ensures that the wavefront sets of the G_{\pm} distribution kernels do not align, making their product well-defined (Friedlander & Joshi, 1998 proposition 11.2.3).

We can now apply reciprocity to the G_+ of (76). We do this by substituting the expression for $\delta u_{+,2}$ in (76) into (77), which allows us to use the Q_+ operators from both expressions. This gives

$$d_3(s_0, r_0, t_4 - t_{s_0}) = -\frac{1}{16} D_{t_4}^4 \int_0^{\infty} dz_3 \int ds_3 \int dr_3 \int_{\mathbb{R}_+} dt_3 \int_0^{z_3} dz_1 \int dm_s \int dm_r \int_{\mathbb{R}_+} dt_a \\ Q_{-,r_0}^*(0) G_-(0, r_0, t_4 - t_3, z_3, r_3) Q_{-,r_3}(z_3) \boxed{Q_{-,m_s}^*(z_1) G_-(z_1, m_s, t_3 - t_a, z_3, s_3) Q_{-,s_3}(z_3)} \\ (E_1 a)(z_3, s_3, r_3) Q_{-,m_r}^*(z_1) (E_1 a)(z_1, m_s, m_r) \delta u_{-,1}(z_1, m_r, t_a, 0, s_0, t_{s_0}); \quad (78)$$

we have also introduced the extension operator E_1 , to split each of the m and x_3 integrations into two. The portion of the expression that has changed through reciprocity is in the box.

The insertion of the relation $G(z_a, x_a, t, z_b, x_b) = Q_{-,x_a}(z_a) G_-(z_a, x_a, t, z_b, x_b) Q_{-,z_b}(z_b)$ into (78), where G is the Green's function for the double square-root equation, shows the correspondence of (78) with expression (8) in (ten Kroode, 2002).

Noting that t_3 is independent of t_a , we can interchange their order of integration, which allows us to change integration variables from t_3 to $t'_3 = t_3 - t_a$ and introduce the E_2 operator at the third scatter. This results in

$$d_3(s_0, r_0, t_4 - t_{s_0}) = -\frac{1}{16} D_{t_4}^4 \int_0^{\infty} dz_3 \int ds_3 \int dr_3 \int_{\mathbb{R}_+} dt_{30} \int_{\mathbb{R}_+} dt_a \int_0^{z_3} dz_1 \int dm_s \int dm_r \int_{\mathbb{R}_+} dt'_3 \\ Q_{-,r_0}^*(0) G_-(0, r_0, t_4 - t_a - t'_3 - t_{30}, z_3, r_3) Q_{-,r_3}(z_3) Q_{-,m_s}^*(z_1) G_-(z_1, m_s, t'_3, z_3, s_3) Q_{-,s_3}(z_3) \\ Q_{-,m_r}^*(z_1) (E_1 E_2 a)(z_3, s_3, r_3, t_{30}) (E_1 a)(z_1, m_s, m_r) \delta u_{-,1}(z_1, m_r, t_a, 0, s_0, t_{s_0}), \quad (79)$$

which is a modeling operator for triply scattered waves. We have not yet introduced the E_2 operator at the m_s, m_r scattering point as it is not clear in which G the time t_{m_0} should be placed. It will become necessary in what follows. In ten Kroode (2002), equation (8) is the Lippmann-Schwinger equivalent of (79); ten Kroode's equation is valid for wave propagation in a medium of up to 2 dimensions without caustics.

Thus far, we have constructed an operator (79) that is tri-linear in the medium contrast a . From a practical viewpoint, however, it is advantageous to construct an operator that acts on the data, as this is known, rather than the medium contrast. To this end, we reformulate (79) so that all the G operators

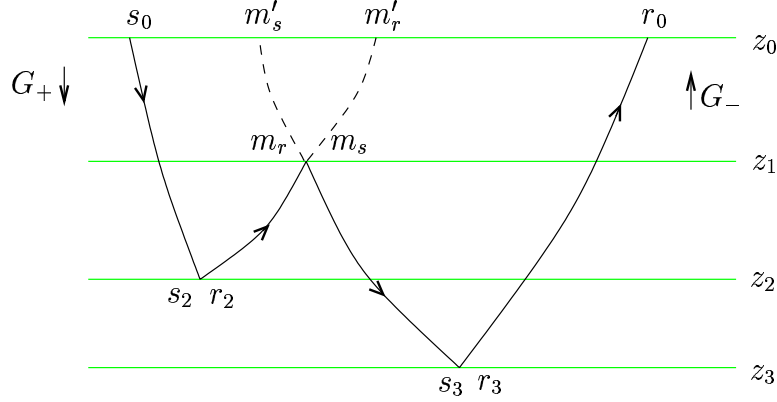


Figure 8. Triple scattering notations and conventions. This illustration assumes that the E operators have already been applied.

communicate with the surface, because data are collected at the surface $z_0 = 0$. This is similar to the layer stripping approach proposed by Fokkema (1994) to extend the work of Berkhout and Verschuur for surface multiples (Berkhout & Verschuur, 1997; Verschuur & Berkhout, 1997) to the internal multiple case. In order to write (79) in terms of the data we first extend the two legs meeting at (z_1, m_s, m_r) to the surface (see Figure 8). To do this, we make use of the semi-group property: $G_-(z, z')G_-(z', z'') = G_-(z, z'')$. The extension to the surface must be done twice, once for the path from s_0 through z_2 to m_r and once for the path from m_s through z_3 to r_0 . Following this, we perform some simplifications to enable us to write (79) in terms of the H operators.

We first return to (75), rewriting it as

$$\begin{aligned} \delta u_{-,1}(z_1, m_r, t_a, 0, s_0, t_{s_0}) &= \frac{1}{4}D_t^2 \int d\tilde{s}_0 \int_{\mathbb{R}_+} d\tilde{t}_{s_0} \int_{z_1}^{\infty} dz_2 \int ds_2 \int dr_2 \int_{\mathbb{R}_+} dt_0 \int_{\mathbb{R}_+} dt' \int dm'_r \int_{\mathbb{R}_+} dt_{m'_r} \\ &G_-^*(z_1, m_r, t_{m'_r}, 0, m'_r) G_-(0, m'_r, t_a + t_{m'_r} - t_{\tilde{s}_0} - t' - t_0, z_2, r_2) \\ &Q_{-,r_2}(z_2) Q_{-,s_2}(z_2) (E_1 E_2 a)(z_2, s_2, r_2, t_0) Q_{-,s_0}^*(0) f(0, \tilde{s}_0, \tilde{t}_{s_0}, 0, s_0, t_{s_0}), \end{aligned} \quad (80)$$

where $t_a + t_{m'_r}$ is the time required to travel from the source at \tilde{s}_0 to the pseudo-receiver at m'_r , as illustrated in Figure 7. Since t' is independent of the other variables we bring this integration to the inside so that the two G_- operators may be replaced by the H operator giving,

$$\begin{aligned} \delta u_{-,1}(z_1, m_r, t_a, 0, s_0, t_{s_0}) &= \frac{1}{4}D_t^2 \int d\tilde{s}_0 \int_{\mathbb{R}_+} d\tilde{t}_{s_0} \int_{z_1}^{\infty} dz_2 \int dm'_r \int_{\mathbb{R}_+} dt_{m'_r} \\ &G_-^*(z_1, m_r, t_{m'_r}, 0, m'_r) \int_{\mathbb{R}_+} ds_2 \int_{\mathbb{R}_+} dr_2 \int_{\mathbb{R}_+} dt_0 H(0, s_0, m'_r, t_a + t_{m'_r} - t_0 - t_{\tilde{s}_0}, z_2, s_2, r_2) \\ &Q_{-,r_2}(z_2) Q_{-,s_2}(z_2) (E_1 E_2 a)(z_2, s_2, r_2, t_0) Q_{-,s_0}^*(0) f(0, \tilde{s}_0, \tilde{t}_{s_0}, 0, s_0, t_{s_0}). \end{aligned} \quad (81)$$

Next, we apply the same procedure to (79)

$$\begin{aligned} d_3(s_0, r_0, t_4 - t_{s_0}) &= -\frac{1}{4}D_t^4 \int_0^{\infty} dz_3 \int ds_3 \int_{\mathbb{R}_+} dr_3 \int_{\mathbb{R}_+} dt_{30} \int_{\mathbb{R}_+} dt'_3 \int_0^{z_3} dz_1 \int dm_s \int dm_r \int_{\mathbb{R}_+} dt_a \\ &Q_{-,r_0}^*(0) G_-(0, r_0, t_4 - t_a - t_{\tilde{s}_0} - t'_3 - t_{30}, z_3, r_3) Q_{-,r_3}(z_3) \\ &Q_{-,m_s}^*(z_1) \int_{\mathbb{R}_+} dm'_s \int_{\mathbb{R}_+} dt_{m'_s} G_-(z_1, m_s, t_{m'_s}, 0, m'_s) G_-(0, m'_s, t'_3 + t_{m'_s}, z_3, s_3) \\ &Q_{-,s_3}(z_3) Q_{-,m_r}^*(z_1) (E_1 E_2 a)(z_3, s_3, r_3, t_{30}) (E_1 a)(z_1, m_s, m_r) \delta u_{-,1}(z_1, m_r, t_a, 0, s_0, t_{s_0}), \end{aligned} \quad (82)$$

where $t_{m'_s}$ is defined by analogy to $t_{m'_r}$ (see Figure 7). Since G_-^* does not depend on the same variables as the propagator preceding it, we interchange their order. Now, we change variables from t'_3 to $t''_3 = t'_3 + t_{m'_s}$,

interchanging the t'_3 and $t_{m'_s}$ integrations giving

$$\begin{aligned}
 d_3(s_0, r_0, t_4 - t_{s_0}) &= -\frac{1}{4} D_t^4 \int_0^\infty dz_3 \int ds_3 \int dr_3 \int_{\mathbb{R}_+} dt_{30} \int_0^{z_3} dz_1 \int dm_s \int dm_r \int_{\mathbb{R}_+} dt_a \int_{\mathbb{R}_+} dm'_s \int_{\mathbb{R}_+} dt_{m'_s} \int_{\mathbb{R}_+} dt''_3 \\
 Q_{-,r_0}^*(0) G_-^*(z_1, m_s, t_{m'_s}, 0, m'_s) G_-(0, r_0, t_4 - t_a - t''_3 + t_{m'_s} - t_{30}, z_3, r_3) Q_{-,r_3}(z_3) Q_{-,m_s}^*(z_1) G_-(0, m'_s, t''_3, z_3, s_3) \\
 Q_{-,s_3}(z_3) Q_{-,m_r}^*(z_1) (E_1 E_2 a)(z_3, s_3, r_3, t_{30}) (E_1 a)(z_1, m_s, m_r) \delta u_{-,1}(z_1, m_r, t_a, 0, s_0, t_{s_0}). \quad (83)
 \end{aligned}$$

We now substitute the H operator for the two G_- operators, interchanging the order of integration to obtain

$$\begin{aligned}
 d_3(s_0, r_0, t_4 - t_{s_0}) &= -\frac{1}{4} D_t^4 Q_{-,r_0}^*(0) \int_0^\infty dz_3 \int_0^{z_3} dz_1 \int dm_s \int dm_r \int_{\mathbb{R}_+} dt_a \int_{\mathbb{R}_+} dm'_s \int_{\mathbb{R}_+} dt_{m'_s} \\
 Q_{-,m_s}^*(z_1) Q_{-,m_r}^*(z_1) G_-^*(z_1, m_s, 0, 0, m'_s, t_{m'_s}) \\
 \int ds_3 \int dr_3 \int_{\mathbb{R}_+} dt_{30} H(0, m'_s, r_0, t_4 - t_a + t_{m'_s} - t_{30}, z_3, s_3, r_3) Q_{-,s_3}(z_3) Q_{-,r_3}(z_3) \\
 (E_1 a)(z_1, m_s, m_r) (E_1 E_2 a)(z_3, s_3, r_3, t_{30}) \delta u_{-,1}(z_1, m_r, t_a - t_{s_0}, 0, s_0). \quad (84)
 \end{aligned}$$

In (84), the G_-^* term in $\delta u_{-,1}$ does not depend on any of the variables in the operators preceding it. Thus, we may interchange operators to combine the two G_-^* terms. We do this, as well as changing the order of integration to move the t_a integral inside the $t_{m'_s}$ one and also introduce E_2 , giving

$$\begin{aligned}
 d_3(s_0, r_0, t_4 - t_{s_0}) &= \frac{1}{16} D_t^6 Q_{-,r_0}^*(0) \int d\tilde{s}_0 \int_{\mathbb{R}_+} d\tilde{t}_{s_0} \int_0^\infty dz_3 \int_0^{z_3} dz_1 \int_{z_1}^\infty dz_2 \\
 \int dm_s \int dm_r \int_{\mathbb{R}_+} dt_{m_0} \int_{\mathbb{R}_+} dm'_s \int_{\mathbb{R}_+} dm'_r \int_{\mathbb{R}_+} dt_{m'_s} \int_{\mathbb{R}_+} dt_{m'_r} (E_1 E_2 a)(z_1, m_s, m_r, t_{m_0}) \\
 Q_{-,m_s}^*(z_1) Q_{-,m_r}^*(z_1) G_-^*(z_1, m_s, t_{m'_s} - t_{m_0}, 0, m'_s) G_-^*(z_1, m_r, t_{m'_r}, 0, m'_r) \left\{ \int_{\mathbb{R}_+} dt_a \right. \\
 \int ds_3 \int dr_3 \int_{\mathbb{R}_+} dt_{30} H(0, m'_s, r_0, t_4 - t_a + t_{m'_s} - t_{30}, z_3, s_3, r_3) Q_{-,s_3}(z_3) Q_{-,r_3}(z_3) \\
 (E_1 E_2 a)(z_3, s_3, r_3, t_{30}) \int ds_2 \int dr_2 \int_{\mathbb{R}_+} dt_0 H(0, \tilde{s}_0, m'_r, t_a + t_{m'_r} - \tilde{t}_{s_0} - t_0, z_2, s_2, r_2) \\
 \left. Q_{-,r_2}(z_2) Q_{-,s_2}(z_2) (E_1 E_2 a)(z_2, s_2, r_2, t_0) \right\} Q_{-,s_0}^*(0) f(0, \tilde{s}_0, \tilde{t}_{s_0}, 0, s_0, t_{s_0}). \quad (85)
 \end{aligned}$$

In (85), the expression in braces is a convolution in time. Because of this, we may shift time variables between the two H operators. To do this we change time variables from t_a to $t_b = t_a + t_{m'_r} - \tilde{t}_{s_0}$. We then introduce the distribution w ,

$$\begin{aligned}
 w(\tilde{s}_0, m'_r, t, m'_s, r_0; z_2, z_3) &= \int_{\mathbb{R}_+} dt_b \int dr_3 \int ds_3 \int_{\mathbb{R}_+} dt_{30} H(0, m'_s, r_0, t - t_b - t_{30}, z_3, s_3, r_3) \\
 Q_{-,s_3}(z_3) Q_{-,r_3}(z_3) (E_1 E_2 a)(z_3, s_3, r_3, t_{30}) \int ds_2 \int dr_2 \int_{\mathbb{R}_+} dt_0 H(0, \tilde{s}_0, m'_r, t_b - t_0, z_2, s_2, r_2) \\
 Q_{-,r_2}(z_2) Q_{-,s_2}(z_2) (E_1 E_2 a)(z_2, s_2, r_2, t_0), \quad (86)
 \end{aligned}$$

where we may allow the lower bound on the t_b integral to extend to 0, rather than $t_{m'_r}$ because $t_b > t_{m'_r}$ by definition. To overlay the operator w with the expression in braces in (85) we need only make the identification $t = t_4 + t_{m'_r} + t_{m'_s} - \tilde{t}_{s_0}$. By changing the lower bound of the z_3 integral in (85) to z_1 and the upper bound of the z_1 integral to ∞ we may define

$$W(z_1; \tilde{s}_0, m'_r, t, m'_s, r_0) = \int_{z_1}^\infty dz_3 \int_{z_1}^\infty dz_2 w(\tilde{s}_0, m'_r, t, m'_s, r_0; z_2, z_3). \quad (87)$$

In the definition of w , we now see the emergence of a new time variable $t_{m'} = t_{m'_r} + t_{m'_s}$ in the expression

for t . To introduce this variable, we return to (85) and change variables from $t_{m'_s}$ to $t_{m'}$.

$$d_3(s_0, r_0, t_4 - t_{s_0}) = \frac{1}{16} D_t^6 Q_{-,r_0}^*(0) \int d\tilde{s}_0 \int_{\mathbb{R}_+} d\tilde{t}_{s_0} \int_0^\infty dz_1 \int dm_s \int dm_r \int_{\mathbb{R}_+} dt_{m_0} \int dm'_s \int dm'_r \int_{\mathbb{R}_+} dt_{m'} \int_0^{t_{m'}} dt_{m'_r} \\ (E_1 E_2 a)(z_1, m_s, m_r, t_{m_0}) Q_{-,m_s}^*(z_1) Q_{-,m_r}^*(z_1) G_-^*(z_1, m_s, t_{m'} - t_{m'_r} - t_{m_0}, 0, m'_s) G_-^*(z_1, m_r, t_{m'_r}, 0, m'_r) \\ W(z_1; \tilde{s}_0, m'_r, m'_s, r_0, t_4 + t_{m'} - \tilde{t}_{s_0}) Q_{-,s_0}^*(0) f(0, \tilde{s}_0, \tilde{t}_{s_0}, 0, s_0, t_{s_0}). \quad (88)$$

The two G_-^* operators in (88) along with the integration in $t_{m'_r}$ are nearly in the form of the H operator.

We can extend the integration in $t_{m'_r}$ to ∞ as $t_{m'_r} > t_{m'}$ results in a negative time in the first G_-^* , which is then 0 by the causality of the Green's function. This allows us to introduce the H operator,

$$d_3(s_0, r_0, t_4 - t_{s_0}) = \frac{1}{16} D_t^6 Q_{-,r_0}^*(0) \int d\tilde{s}_0 \int_{\mathbb{R}_+} d\tilde{t}_{s_0} \int_0^\infty dz_1 \left\{ \int dm_s \int dm_r \int_{\mathbb{R}_+} dt_{m_0} (E_1 E_2 a)(z_1, m_s, m_r, t_{m_0}) \right. \\ \left. \int dm'_s \int dm'_r \int_{\mathbb{R}_+} dt_{m'} Q_{-,m_s}^*(z_1) Q_{-,m_r}^*(z_1) H^*(z_1, m_s, m_r, t_{m'} - t_{m_0}, 0, m'_s, m'_r) \right. \\ \left. W(z_1; \tilde{s}_0, m'_r, t_4 + t_{m'} - \tilde{t}_{s_0}, m'_s, r_0) \right\} Q_{-,s_0}^*(0) f(0, \tilde{s}_0, \tilde{t}_{s_0}, 0, s_0, t_{s_0}). \quad (89)$$

In (89), we recognize the expression in braces as an inner product in the (m_s, m_r, t_{m_0}) variables.

7 INVERSE SCATTERING

7.1 single scattering – summary

Stolk and de Hoop (2004b) give a method construct an image of the subsurface from singly scattered data. The construction involves the depth-to-time conversion operator, K , defined as

$$K : a \mapsto \int_0^\infty H(0, z) (E_2 a)(z, t_0) dz. \quad (90)$$

Stolk and de Hoop (2004a; 2004b) denote this operator as \bar{K} ; K is an invertible Fourier integral operator. Upon substitution of $f(0, \tilde{s}_0, \tilde{t}_{s_0}, 0, s_0, t_{s_0}) = \delta(\tilde{s}_0 - s_0) \delta(\tilde{t}_{s_0} - t_{s_0})$ in (74), we see that the singly scattered data can be written in terms of this K operator, the advantage of which is that K is invertible. Performing this substitution gives

$$d_1(s_0, r_0, t) = \frac{1}{4} D_t^2 Q_{-,s_0}^*(0) Q_{-,r_0}^*(0) K Q_{-,r_1}(z_1) Q_{-,s_1}(z_1) (E_1 a)(z_1, s_1, r_1). \quad (91)$$

Seismic data are reflected from reflectors much deeper than the depth of sources and receivers. Thus, we are in the far-field of the source, justifying the point source assumption.

To construct an image of the subsurface, we will take the adjoint of the K and E_1 operators. Thus, we introduce R_2 and R_1

$$\begin{aligned} g(z, s, r, t) &\mapsto (R_2 g)(z, s, r) = g(z, s, r, 0) \\ h(z, s, r) &\mapsto (R_1 h)(z, x) = h(z, x, x) \end{aligned} \quad (92)$$

as the transpose of E_2 and E_1 , and note that $R_2 H^*(0, z) = K^*(0, z)$. We introduce the pseudodifferential operator, $\Xi = K^* K$, the ‘normal operator’, which corrects for the amplitude errors made in using the adjoint rather than the inverse in the imaging. Its symbol is given in (Stolk & de Hoop, 2004b Lemma 2.1) and is (recalling that Ξ here is $\bar{\Xi}$ in (Stolk & de Hoop, 2004b))

$$\left(\frac{1}{c_0(z, s)^2 b(z, s, \sigma, \tau)} + \frac{1}{c_0(z, r)^2 b(z, r, \rho, \tau)} \right) \Big|_{\tau = \Theta^{-1}(z, s, r, \zeta, \sigma, \rho)}, \quad (93)$$

where c_0 is the background velocity, (ζ, σ, ρ) are the cotangent variables associated with (z, s, r) , and Θ is the frequency to vertical wavenumber map $\tau \mapsto \zeta = \Theta(z, s, r, \zeta, \sigma, \rho, \tau)$ (Stolk & de Hoop, 2004a Lemma 4.1). From this we find that the inverse of the modeling operator can be written as

$$\Phi(z, x, D_z, D_x) a(z) = \underbrace{R_1 Q_{-,s}(z)^{-1} Q_{-,r}(z)^{-1} \Xi^{-1} R_2 H^*(0, z) Q_{-,s_0}^*(0)^{-1} Q_{-,r_0}^*(0)^{-1} D_t^{-2}}_J d_1 \left(\begin{smallmatrix} s_0 & r_0 \\ \cdot & \cdot \\ \cdot & \cdot \end{smallmatrix}, t \right). \quad (94)$$

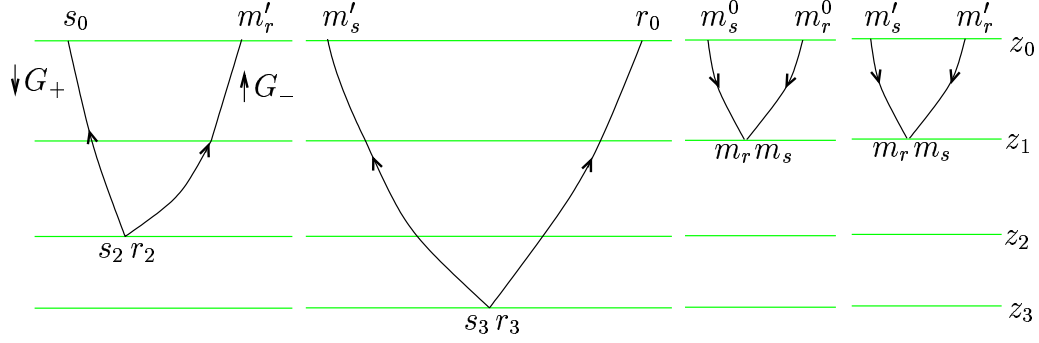


Figure 9. Illustration of the four operators used to construct the final modeling operator. Notice that R and H act at different surface points.

The operator $\Phi(z, x, D_z, D_x)$, which is also a pseudodifferential operator, corrects for illumination effects. Its symbol is 1 where there is illumination and zero elsewhere. In fact, the analysis requires that we replace the data, d_1 , with $\psi_Y d_1$, where ψ_Y is a cut-off function which is zero near the boundary of the acquisition surface, Y . Further details of these cut-off operators can be found in (Stolk & de Hoop, 2004b).

In the single scattering approximation, the data in (74) are used as an approximation of the data d , in (52). To reconstruct V_1 we must apply the above inverse to these model data. This gives an estimate for the a distribution,

$$\langle a_1 \rangle(z, x) = \Phi(z, x, D_z, D_x)a, \quad (95)$$

which is, as expected, simply an amplitude change, or ‘non-stationary’ filter applied to the original medium contrast. We use the $\langle \cdot \rangle$ notation to indicate that this is an estimate of a rather than its true value, the subscript 1 indicates that this estimate is obtained through the single scattering approximation. From this estimate of a , we obtain an estimate of the matrix V_1 using (31).

7.2 triple scattering

In the framework of the series expansion in (59), we require only the single scattering inverse to estimate the third order contribution to V : we write the first and third terms of (59) as

$$\langle a_1 \rangle + \langle a_3 \rangle = N^{-1}F^*d - N^{-1}F^*(D_t^4 M_0(\widehat{V}_1 L_0(\widehat{V}_1 \delta U))). \quad (96)$$

We will use the operator, J , on the right hand side in (94) in place of $N^{-1}F^*$ to estimate this inverse. The inverse series, (59) is a sum of terms of different order in the data, with the first term being first-order in the data and the third term third-order in the data. In the modeling formula, (89), the data are written as a trilinear operator on the medium contrast, a , as in the forward series (43). To estimate a_3 we need to write $D_t^4 M_0(\widehat{V}_1 L_0(\widehat{V}_1 \delta U))$, for which an expression was constructed in section 6.2, as a trilinear operator acting on the data. We then apply the single scattering inverse to this expression to estimate a_3 , from which we can compute \widehat{V}_3 .

To write d_3 in terms of the data, we insert the instantaneous point source into (89) as in the single scattering case. This gives

$$d_3(s_0, r_0, t_4 - t_{s_0}) = \frac{1}{16} D_t^6 Q_{-,r_0}^*(0) Q_{-,s_0}^*(0) \int_0^\infty dz_1 \int dm_s \int dm_r \int_{\mathbb{R}_+} dt_{m_0} (E_1 E_2 a)(z_1, m_s, m_r, t_{m_0}) \int dm'_s \int dm'_r \int_{\mathbb{R}_+} dt_{m'} Q_{-,m_s}^*(z_1) Q_{-,m_r}^*(z_1) H^*(z_1, m_s, m_r, t_{m'} - t_{m_0}, 0, m'_s, m'_r) W(z_1; s_0, m'_r, t_4 + t_{m'} - t_{s_0}, m'_s, r_0). \quad (97)$$

There are three equivalent ways in which one could proceed to write (89) in terms of the data only. First, one could replace each of the a terms with J applied to the data, giving a trilinear operator acting on the data. Second, one can re-arrange terms to introduce d directly in place of the a term at both z_2 and z_3 , and use J to write the a term at z_1 in terms of the data. Third, one can rearrange terms to write all three a directly in terms of d . We will go through the second method in detail.

We make use of the relationship

$$W(z_1; s_0, m'_r, t_4 + t_{m'} - t_{s_0}, m'_s, r_0) = 16D_t^2 Q_{-,s_0}^*(0)^{-1} Q_{-,r_0}^*(0)^{-1} Q_{-,m'_s}^*(0)^{-1} Q_{-,m'_r}^*(0)^{-1} \int_{\mathbb{R}_+} dt_b \mathbf{d}_1(z_1; m'_s, r_0, t_4 + t_{m'} - t_b - t_{s_0}) \mathbf{d}_1(z_1; s_0, m'_r, t_b), \quad (98)$$

between the W operator and the singly scattered data (given in (74)) to write d_3 in terms of two data sets and the medium contrast at one point. In (98), we have introduced

$$\mathbf{d}_1(z_1; s, r, t - t_s) = \frac{1}{4} D_t^2 Q_{-,r}^*(0) Q_{-,s}^*(0) \int_{z_1}^{\infty} dz H(0, z) Q_{-,r}(z) Q_{-,s}(z) (E_1 E_2 a)(z, s, r, t_0), \quad (99)$$

as the subset of the singly scattered data with reflection below the level z_1 . Assuming the traveltime monotonicity assumption as done by ten Kroode (2002), would allow the restriction in z_1 to be translated to a restriction on the time $t - t_s$. If this assumption is not satisfied one could generate $\mathbf{d}_1 = d_1 - \mathbf{D}(a)$, where

$$(\mathbf{D}(a))(z_1, s, r, t) = \frac{1}{4} D_t^2 Q_{-,r}^*(0) Q_{-,s}^*(0) \int_0^{z_1} dz H(0, z) Q_{-,r}(z) Q_{-,s}(z) (E_1 E_2 \langle a \rangle)(z, s, r, t_0), \quad (100)$$

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

By inserting the expression for W in (98) into the modeling formula, (89) we arrive at an expression for d_3 in terms of the singly scattered data as well as the medium contrast at z_1 ,

$$d_3(s_0, r_0, t_4 - t_{s_0}) = D_t^2 \int_0^{\infty} dz_1 \int dm_s \int dm_r \int_{\mathbb{R}_+} dt_{m_0} (E_1 E_2 a)(z_1, m_s, m_r, t_{m_0}) \int_{\mathbb{R}_+} dt_{m'} Q_{-,m_s}^*(z_1) Q_{-,m_r}^*(z_1) H^*(z_1, m_s, m_r, t_{m'} - t_{m_0}, 0, m'_s, m'_r) Q_{-,m'_s}^*(0)^{-1} Q_{-,m'_r}^*(0)^{-1} \int_{\mathbb{R}_+} dt_b \mathbf{d}_1(z_1; m'_s, r_0, t_4 + t_{m'} - t_b - t_{s_0}) \mathbf{d}_1(z_1; s_0, m'_r, t_b). \quad (101)$$

To write (101) explicitly in terms of the data only we require an expression for the medium contrast, $(E_1 E_2 a)(z_1, m_s, m_r, t_{m_0})$ in terms of the data. We use the estimate, $\langle a_1 \rangle$, given in (95).

Thus far we have assumed that we have singly scattered data. In reality, we record data, d , which contain all orders of events. Comparing equations (44) and (52) we observe that the recorded data, d , are a first-order approximation of the singly scattered data, d_1 . We can then define $\mathbf{d} = \mathbf{d}_1 - \mathbf{D}\langle a_1 \rangle$. The quantity $\mathbf{D}\langle a_1 \rangle$ is computed by composing a migration, to compute $\langle a_1 \rangle$, with a demigration. It is necessary to go through this procedure because the 'data' \mathbf{D} must scatter off of reflectors at *depths* above z_1 ; the only way to know if data satisfy this assumption is to model them from such depths.

Performing the above two steps, we obtain an approximation, $\langle d_3 \rangle$, for d_3 ,

$$\langle d_3 \rangle(s_0, r_0, t_4 - t_{s_0}) = D_t^2 \int_0^{\infty} dz_1 \int dm_s \int dm_r \int_{\mathbb{R}_+} dt_{m_0} E_2 Q_{-,m_s}(z_1)^{-1} Q_{-,m_r}(z_1)^{-1} \left\{ \Xi^{-1} R_{2,t'} \int dm_s^0 \int dm_r^0 \int_{\mathbb{R}_+} dt_{m_0} H^*(z_1, m_s, m_r, t_{m_0} - t', 0, m_s^0, m_r^0) Q_{-,m_s^0}^*(0)^{-1} Q_{-,m_r^0}^*(0)^{-1} D_t^{-2} d(m_s^0, m_r^0, t_{m_0}) \right\} \int_{\mathbb{R}_+} dt_{m'} Q_{-,m_s}^*(z_1) Q_{-,m_r}^*(z_1) H^*(z_1, m_s, m_r, t_{m'} - t_{m_0}, 0, m'_s, m'_r) Q_{-,m'_s}^*(0)^{-1} Q_{-,m'_r}^*(0)^{-1} \int_{\mathbb{R}_+} dt_b \mathbf{d}(z_1; m'_s, r_0, t_4 + t_{m'} - t_b - t_{s_0}) \mathbf{d}(z_1; s_0, m'_r, t_b). \quad (102)$$

The above formulation depends on the smooth background velocity model only from the surface, $z = 0$, to the first scattering point at z_1 .

The form of (102) is similar to that of an inner product in the variables (m_s, m_r, t_{m_0}) . We will make

use of this fact and the notion of adjoint to rewrite (102) as

$$\begin{aligned}
 \langle d_3 \rangle (s_0, r_0, t_4 - t_{s_0}) &= \int dm'_s \int dm'_r \int_{\mathbb{R}_+} dt_{m'} \int_0^\infty dz_1 \int dm_s \int dm_r \int_{\mathbb{R}_+} dt_{m^0} H(0, m'_s, m'_r, t_{m'} - t_{m^0}, z_1, m_s, m_r) E_2 \\
 &Q_{-,m_s}(z_1) Q_{-,m_r}(z_1) Q_{-,m_s}(z_1)^{-1} Q_{-,m_r}(z_1)^{-1} \left\{ \Xi^{-1} R_{2,t'} \int dm_s^0 \int dm_r^0 \int_{\mathbb{R}_+} dt_{m^0} \right. \\
 &H^*(z_1, m_s, m_r, t_{m^0} - t', 0, m_s^0, m_r^0) Q_{-,m_s^0}^*(0)^{-1} Q_{-,m_r^0}^*(0)^{-1} d(m_s^0, m_r^0, t_{m^0}) \left. \right\} \\
 &Q_{-,m_s^0}^*(0)^{-1} Q_{-,m_r^0}^*(0)^{-1} \int_{\mathbb{R}_+} dt_b \mathbf{d}(z_1; m'_s, r_0, t_4 + t_{m'} - t_b - t_{s_0}) \mathbf{d}(z_1; s_0, m'_r, t_b), \quad (103)
 \end{aligned}$$

which simplifies to

$$\begin{aligned}
 \langle d_3 \rangle (s_0, r_0, t_4 - t_{s_0}) &= \int dm'_s \int dm'_r \int_{\mathbb{R}_+} dt_{m'} \int_0^\infty dz_1 \int dm_s \int dm_r \int_{\mathbb{R}_+} dt_{m^0} H(0, m'_s, m'_r, t_{m'} - t_{m^0}, z_1, m_s, m_r) E_2 \Xi^{-1} R_{2,t'} \\
 &\int dm_s^0 \int dm_r^0 \int_{\mathbb{R}_+} dt_{m^0} H^*(z_1, m_s, m_r, t_{m^0} - t', 0, m_s^0, m_r^0) Q_{-,m_s^0}^*(0)^{-1} Q_{-,m_r^0}^*(0)^{-1} d(m_s^0, m_r^0, t_{m^0}) \\
 &Q_{-,m_s^0}^*(0)^{-1} Q_{-,m_r^0}^*(0)^{-1} \int_{\mathbb{R}_+} dt_b \mathbf{d}(z_1; m'_s, r_0, t_4 + t_{m'} - t_b - t_{s_0}) \mathbf{d}(z_1; s_0, m'_r, t_b). \quad (104)
 \end{aligned}$$

We define

$$\begin{aligned}
 d'_1(z_1, s, r, t) &= D_t^2 Q_{-,s}^*(0) Q_{-,r}^*(0) \int ds_1 \int dr_1 \int_{\mathbb{R}_+} dt_0 \\
 &H(0, s, r, t - t_0, z_1, s_1, r_1) Q_{-,s_1}(z_1) Q_{-,r_1}(z_1) (E_1 E_2 a)(z_1, s_1, r_1, t_0); \quad (105)
 \end{aligned}$$

we use the notation d' because it is close to the derivative with respect to z_1 of d . The quantity d' is not one that can be measured, because one cannot tell directly from which depth the data come. To compute d' , the expression in (94) must be substituted for a to write it in terms of what can be measured, d . Because of this, we can write (102) in terms of this d' 'data' set

$$\begin{aligned}
 d_3(s_0, r_0, t_4 - t_{s_0}) &= \int_0^\infty dz_1 \int dm'_s \int dm'_r \int_{\mathbb{R}_+} dt_{m'} Q_{-,m'_s}^*(0)^{-1} Q_{-,m'_r}^*(0)^{-1} d'_1(z_1, m'_s, m'_r, t_{m'}) \\
 &Q_{-,m'_s}^*(0)^{-1} Q_{-,m'_r}^*(0)^{-1} \int_{\mathbb{R}_+} dt_b \mathbf{d}_1(z_1; m'_s, r_0, t_4 + t_{m'} - t_b - t_{s_0}) \mathbf{d}_1(z_1; s_0, m'_r, t_b). \quad (106)
 \end{aligned}$$

We have nearly succeeded in writing $D_t^4 M_0(\widehat{V}_1 L_0(\widehat{V}_1 \delta U))$ in terms of the data d . We find, however that we cannot write (106) in terms of the actual data because of the z_1 dependence of each of the three 'data' sets.

Remark 7.1. If we ignore the z_1 dependence of \mathbf{d} in (104), the composition

$$\int_0^\infty dz_1 \int dm_s \int dm_r \int_{\mathbb{R}_+} dt_{m^0} H(0, m'_s, m'_r, t_{m'} - t_{m^0}, z_1, m_s, m_r) E_2 \Xi^{-1} R_2 H^*(z_1, m_s, m_r, t_{m^0}, 0, m_s^0, m_r^0),$$

gives microlocally,

$$\delta(m'_s - m_s^0) \delta(m'_r - m_r^0) \delta(t_{m'} - t_{m^0}).$$

Then, (104) can be written as

$$\begin{aligned}
 \langle d_3 \rangle (s_0, r_0, t_4 - t_{s_0}) &\sim \int dm'_s \int dm'_r \int_{\mathbb{R}_+} dt_{m'} Q_{-,m'_s}^*(0)^{-1} Q_{-,m'_r}^*(0)^{-1} d(m'_s, m'_r, t_{m'}) \\
 &Q_{-,m'_s}^*(0)^{-1} Q_{-,m'_r}^*(0)^{-1} \int_{\mathbb{R}_+} dt_b d(m'_s, r_0, t_4 + t_{m'} - t_b - t_{s_0}) d(s_0, m'_r, t_b), \quad (107)
 \end{aligned}$$

with the approximation $d \approx d_1$.

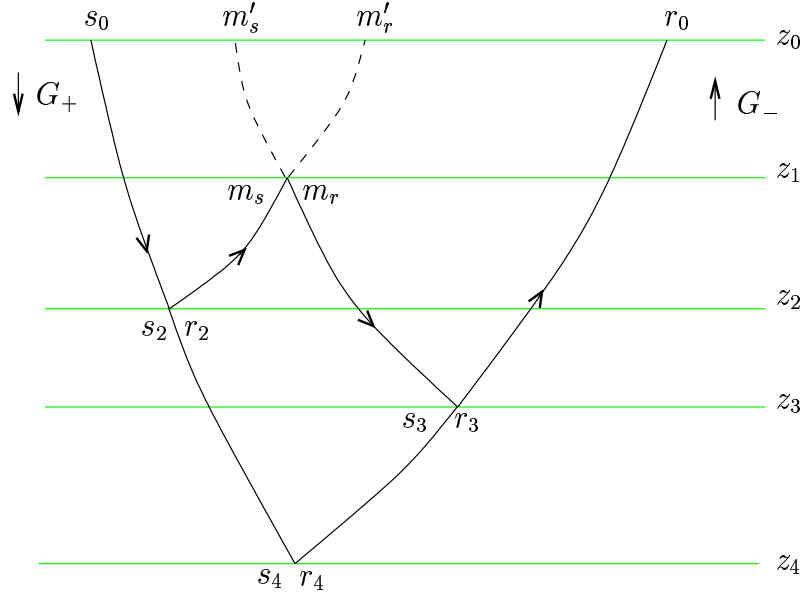


Figure 10. Applying the single scattering inverse to triply scattered data results in the mis-positioning of the imaging point on z_4 rather than z_2 , z_1 and z_3 .

Remark 7.2. The approach presented here is a generalization of the Weglein-ten Kroode approach (Weglein *et al.*, 1997; Weglein *et al.*, 2003; ten Kroode, 2002). To demonstrate this, we specifically compare (104) in this paper with (120) of (ten Kroode, 2002). To do this it is first necessary to establish a correspondence between our notation and ten Kroode's notation. The easiest way to do this is to compare Figure 8 of this paper with Figure 4 of (ten Kroode, 2002). We then identify the t_1 variable of ten Kroode with the t_b variable here, the t_2 variable of ten Kroode with $t_{m'}$ and the t_3 variable with $t_4 + t_{m'} - t_b$. Then we note that $t_1 - t_2 + t_3$, which would be the time argument of d_3^{LM} in (117) of ten Kroode, is equal to t_4 here. This establishes the correspondence between the time dependence of the final result, (120) in ten Kroode, with (104) here.

To make the correspondence between the pseudo-data \mathbf{d} here and the integration bounds on (117) of ten Kroode we observe that Z'_2 of ten Kroode is a time parameterization of the scattering depth denoted here by z_1 . Thus, as is done in ten Kroode, under the travel-time monotonicity assumption[†], we can replace the restrictions on the depth of the scattering points in the definition of \mathbf{d} with the restriction $t_b > t_{m'}$ on the t_b integration. Using this we can replace \mathbf{d} with d in (104), allowing us to perform the composition as in remark 7.1, which brings it into correspondence with (120) of ten Kroode.

Ten Kroode performs stationary phase analysis in three sets of variables, corresponding 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 semi-group property used to extend the modeled data from the scattering point at z_1 to the surface.

We do not expect the amplitudes of (104) here and (120) of (ten Kroode, 2002) to be the same since here we use the wavefield continuation approach and ten Kroode uses the Kirchhoff common shot approach.

To this point, we have constructed an expression for $M_0(\widehat{V}_1 L_0(\widehat{V}_1 \delta U))$. Computing the inverse requires one more step however; the application of the single-scattering inverse, J . To do this we introduce a new depth level, z_4 , which is the point at which the triply scattered data will be incorrectly imaged if single

[†]Ten Kroode's travel-time monotonicity assumption states that the traveltimes for a ray leaving a position (z, x) in direction α arrives later than a ray leaving position (z', x) in direction α whenever $z > z'$. He assumes this to hold for all x and α and gives an example of the violation of this assumption.

scattering is assumed (see Figure 10). The application of J to $\langle d_3 \rangle$ gives

$$\langle a_3 \rangle(z_4, x_4) = -R_1 Q_{-,s_4}(z_4)^{-1} Q_{-,r_4}(z_4)^{-1} \Xi^{-1} H^*(0, s_0, r_0, t_4 - t_{s_0}, z_4, s_4, r_4) \\ Q_{-,s_0}^*(0)^{-1} Q_{-,r_0}^*(0)^{-1} D_t^{-2} \langle d_3 \rangle \begin{pmatrix} s_0 & r_0 & t_4 - t_{s_0} \\ \cdot & \cdot & \cdot \end{pmatrix}. \quad (108)$$

As can be expected, this expression is not close to the true medium contrast at (z_4, x_4) since J is the single scattering inverse and thus will not correctly treat multiply scattered waves. Returning to (96) we see that the term \widehat{V}_3 has the opposite sign to \widehat{V}_1 . Equation (108) is meant to remove the erroneous contribution to $\langle a \rangle$ made by imaging all the data as though they had scattered only once as is done in the construction of $\langle a_1 \rangle$.

7.3 Inverse series in angle gathers

The J operator discussed above computes an image of the subsurface given seismic data at the surface. There is more information available in seismic data than it exploits, however. For example, studying amplitude variations with scattering angle (the angle between the incoming and outgoing rays at the scattering point) an image, which is averaged over angle, is not sufficient. In addition, the imaging operator depends on the background velocity model, which is not in general known. One way of estimating this model is to look at the data as a function of angle at the scattering point, called an angle gather. With the correct velocity model, these angle gathers are independent of angle.

To construct these angle gathers, we use the wave-equation angle transform, \tilde{A}_{WE} , as described by Stolk and de Hoop (2004b). Here, in place of \tilde{A}_{WE} we use the notation A_{WE} , where

$$(A_{\text{WE}}d)(z, x, p) = j^{-1} R_3 \Xi^{-1} Q_{-,s}(z)^{-1} Q_{-,r}(z)^{-1} H(0, z)^* Q_{-,s_0}^*(0)^{-1} Q_{-,r_0}^*(0)^{-1} D_t^{-2} d, \quad (109)$$

with p the slowness vector. We define j , as the pseudodifferential operator compensating for amplitude effects. Its symbol is derived in Proposition 3.2 of (Stolk & de Hoop, 2004b). We also introduce R_3 , through

$$R_3 : g(z, s, r, t) \mapsto (R_3 g)(z, x, p) = \int_{\mathbb{R}^{n-1}} g(z, x - \frac{h}{2}, x + \frac{h}{2}, ph) \chi(z, x, h) dh, \quad (110)$$

where $h = \frac{s-r}{2}$ and χ is a pseudodifferential cutoff with compact support containing the point $h = 0$. Note that R_3 is a restriction operator, its properties are described by Stolk and de Hoop (2004b); it takes the place of R_1 and R_2 used previously. Applying the A_{WE} operator to data results in a family of images, parameterized by p . In this construction, the data $(A_{\text{WE}}d)(z, x, p)$ are independent of the parameter p under a number of assumptions. One of these assumptions is that the data consist of only primary events. First-order internal multiples will violate these assumptions and thus will give contributions to $(A_{\text{WE}}d)(z, x, p)$ that do depend on the parameter p .

In the series solution for \widehat{V} , we can use the angle transform to construct a series of solution families, dependent on the parameter p . From the singly scattered term (74), we get the reconstruction for a_1

$$\langle\langle a_1 \rangle\rangle(z, x, p) = (A_{\text{WE}} d)(z, x, p), \quad (111)$$

where we use the notation $\langle\langle \cdot \rangle\rangle$ to denote an estimate obtained through the angle transform. We know that this can be written as

$$\Psi_{\text{WE}}(z, x, p, D_z, D_x, 0)(a) = A_{\text{WE}}d, \quad (112)$$

(Stolk & de Hoop, 2004b) where Ψ_{WE} is the composition of all of the pseudodifferential operators in (94), transformed via pullback into the angle coordinates.

The third-order contribution comes from $\langle d_3 \rangle$. Applying the wave-equation angle transform in place of the inverse from (94), we arrive at the following analogue of (108)

$$\langle\langle a_3 \rangle\rangle(z, x, p) = (A_{\text{WE}} \langle d_3 \rangle)(z, x, p) \quad (113)$$

Again, $\langle\langle a_3 \rangle\rangle(z, x, p)$ should not be interpreted as a true medium contrast, but rather as a correction term to $\langle\langle a_1 \rangle\rangle$. Recalling from (48) that \widehat{V}_1 and \widehat{V}_3 have opposite sign, our best estimate of a comes from

$$\langle\langle a \rangle\rangle = \langle\langle a_1 \rangle\rangle - \langle\langle a_3 \rangle\rangle. \quad (114)$$

Although it is possible to subtract the two angle gathers, $\langle\langle a_1 \rangle\rangle$ and $\langle\langle a_3 \rangle\rangle$, (114) does not necessarily give a good estimate of a . This is because, although (112) assures us that $\langle\langle a_1 \rangle\rangle$ is a pseudodifferential operator

applied to the true medium contrast, we have no such assurance for $\langle\langle a_3 \rangle\rangle$. To alleviate this problem, we define a new operator, $\Psi'_{\text{WE}}(z, x, p, D_x, D_z, 0)$, which accounts for amplitude and illumination effects in the estimate $\langle\langle a_3 \rangle\rangle$. We then difference the two estimates via

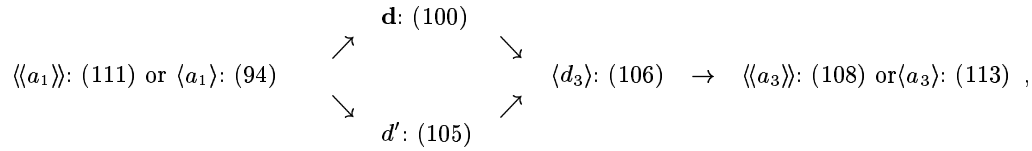
$$\langle\langle a \rangle\rangle = \langle\langle a_1 \rangle\rangle - \Psi'_{\text{WE}} \langle\langle a_3 \rangle\rangle, \quad (115)$$

where we anticipate accounting for Ψ' in an adaptive manner.

8 DISCUSSION

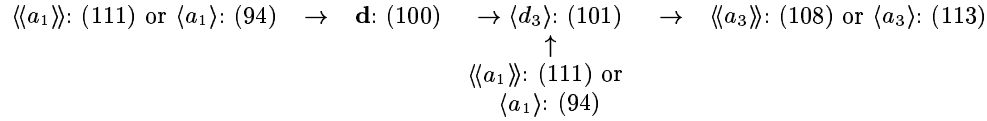
The following flowcharts illustrate two different possible methods of computing the final estimate $\langle a \rangle$ or $\langle\langle a \rangle\rangle$, from the observed data. Each step shows the quantity computed and the equation used to compute it; quantities at the tail of arrows are used to compute the quantity at the head of the arrow. (For example $\mathbf{d}:(100) \rightarrow \langle d_3 \rangle:(101)$ indicates that \mathbf{d} is estimated with (100) and is then used in (101) to estimate $\langle d_3 \rangle$.) In each step we substitute the recorded data d for the single scattered data d_1 ; this approximation is not stated explicitly in several of the referenced equations.

The first method,



computes both \mathbf{d} and d' from the estimate, $\langle a_1 \rangle$ or $\langle\langle a_1 \rangle\rangle$. It is this method that reduces, under the traveltime monotonicity condition, to a velocity independent method.

The second method,



requires the computation of \mathbf{d} only, along with an additional downward continuation of the two \mathbf{d} 'data' sets in (101). Although this method is not as directly tied to the data itself, it avoids computing d' at each depth level. The downward continuation of the \mathbf{d} data can be done using the relation

$$\begin{aligned}
 \frac{1}{4} D_t^2 H^*(z_1, 0) \int_0^\infty dz H(0, z) Q_{-,r}(z) Q_{-,s}(z) (E_1 E_2 a)(z, s, r, t_0) = \\
 \frac{1}{4} D_t^2 \int_{z_1}^\infty dz H(z_1, z) Q_{-,r}(z) Q_{-,s}(z) (E_1 E_2 a)(z, s, r, t_0), \quad (116)
 \end{aligned}$$

to allow the downward continuation to be done on the data d rather than on \mathbf{d} .

Once an estimate of $\langle d_3 \rangle$ is obtained, either of the following subtractions complete the procedure

$$\begin{array}{l}
 \langle\langle a \rangle\rangle = \langle\langle a_1 \rangle\rangle - \langle\langle a_3 \rangle\rangle \\
 \langle a \rangle = \langle a_1 \rangle - \langle a_3 \rangle
 \end{array}$$

The quantity $\langle a_1 \rangle$ or $\langle\langle a_1 \rangle\rangle$ is estimated as part of the usual seismic imaging procedure. The procedure discussed here requires an additional imaging step to compute $\langle a_3 \rangle$ or $\langle\langle a_3 \rangle\rangle$ from $\langle d_3 \rangle$. For this estimate, \mathbf{d} must be computed. This requires an upward continuation to the surface at each depth step. This procedure identifies portions of the data which scatter above the depth z_1 , so that they can be removed. If the velocity model is such that the traveltime monotonicity assumption of ten Kroode (2002) is valid the upward continuation may be replaced by a simple time-windowing procedure.

Figures 11(a) and (b) illustrate the construction of the third-order data, $\langle d_3 \rangle$, from three singly scattered data points. In the construction of $\langle\langle a_3 \rangle\rangle$, these data contribute to the single scattering isochron with the same source position, receiver position and total traveltime (dashed curve), rather than the three partial time single scattering isochrons (solid curves). The estimate $\langle\langle a_1 \rangle\rangle$ also puts a contribution from these rays onto the dashed isochron. The subtraction of $\langle\langle a_3 \rangle\rangle$ from $\langle\langle a_1 \rangle\rangle$, removes this incorrect contribution. In (c), a set of rays with identical source-receiver positions and slopes is shown. In our theory, these kinematically

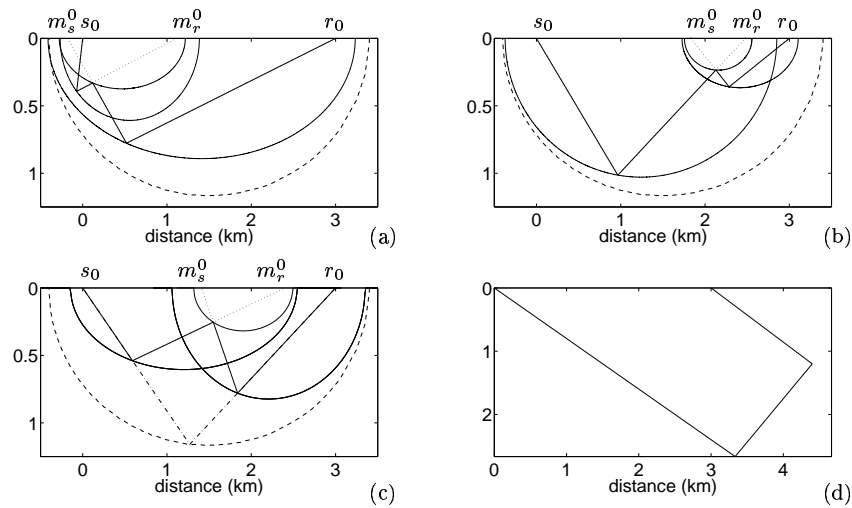


Figure 11. Two different situations contributing to the third-order term are shown in (a) and (b). The three solid curves are isochrons corresponding to the three singly scattered data points contributing to a third-order scattering event. The dashed curve is the corresponding single-scattering isochron (i.e. the same source and receiver positions and travel time as the triply scattered event assuming single scattering). The solid ray is the true travel path and the dashed rays illustrate the additional travel used to write the triply scattered data in terms of three single scatters. In (c) we show a triple scattering contribution difficult to distinguish from a singly scattered event (dash-dot ray), in data space, since the surface positions and initial slopes are the same. In (d), a contribution that violates our assumptions is shown; this is a doubly scattered event that would be recorded at the surface.

analogous contributions (Hron, 1972) will not pose a problem; the subtraction, $\langle\langle a_3 \rangle\rangle - \langle\langle a_1 \rangle\rangle$ will result in an amplitude correction. The event shown in (d), is not accounted for by our theory. The event shown is a doubly scattered event, and thus will contribute to a_2 , which is not estimated here. Events like this make an important contribution to seismic data, especially near salt. However, the contribution from the majority of doubly scattered events is lost to the interior of the Earth. Such contributions are therefore more important for transmission experiments than reflection experiments like those studied here.

In Figure 12, the triple scattering isochron is shown. This isochron represents all points in the subsurface which could contribute to a particular point in the triply scattered data. In (a), points from the first scattering point at (z_2, s_2, r_2) are shown; the points in (b) are from the third scatter at (z_3, s_3, r_3) . These two plots are mirror images of one another as expected from reciprocity. The points shown in (c), mark the position of the central scattering event at (z_1, m_s, m_r) . In (c), the points do not cluster on lines like those in (a) and (b). This scattered distribution of points fills the interior of the corresponding single scattering isochron (dashed line). Thus we observe that while singly scattered data at a single source, receiver, and time sample the subsurface along an ellipse, triply scattered data at the same source, receiver, and time sample the entire interior of the same ellipse.

The algorithm described above requires knowledge of the velocity model to the depth z_1 of the up-to-down scatter. This knowledge is necessary to compute \mathbf{d} , when the traveltime monotonicity assumption is not valid. In addition, an adaptive subtraction technique is required to compensate for differences in illumination between the singly and triply scattered data. In other words, a technique must be developed to subtract $\langle\langle a_3 \rangle\rangle$ from $\langle\langle a_1 \rangle\rangle$ in a way that is robust to differences in the subsurface sampling between the singly and triply scattered waves.

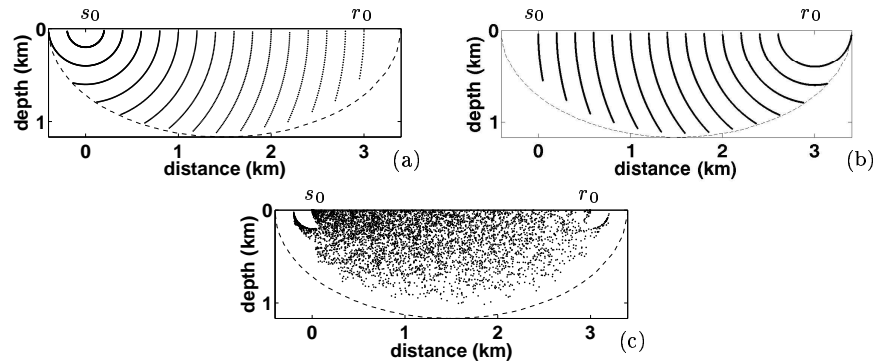


Figure 12. Triple scattering isochron. In (a), only the points from the (z_2, s_2, r_2) are shown. In (b), only the points from the (z_3, s_2, r_2) scattering point are shown. In (c), the (z_1, m_s, m_r) scattering points are shown, although only 1/50 of the number of points in the other plots are shown here. The dashed line is the same single-scattering isochron shown in Figure 11.

REFERENCES

- 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.
- Berkhout, A. J., & Verschuur, D. J. 1997. Estimation of multiple scattering by iterative inversion, Part I: Theoretical considerations. *Geophysics*, **62**(5), 1586–1595.
- Burdick, L. J., & Orcutt, J. A. 1979. A comparison of the generalized ray and reflectivity methods of waveform synthesis. *Geophys. J. Int.*, **58**, 261–278.
- Buttkus, B. 1979. Coherency weighting – an effective approach to the suppression of long leg multiples. *Geophys. Prosp.*, **27**, 29–39.
- Cheney, M., & Borden, B. 2002. Microlocal structure of inverse synthetic aperture radar data. *Inverse Problems*, **19**, 173–194.
- Claerbout, J. F. 1985. *Imaging the Earth's Interior*. Blackwell Sci. Publ.
- de Hoop, M. V. 1996. Generalization of the Bremmer coupling series. *J. Math. Phys.*, **37**, 3246–3282.
- de Hoop, M. V. 2004. The downward continuation approach to modeling and inverse scattering of seismic data in the Kirchhoff approximation. *in print, Contemp. Math.*
- 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.
- Essenreiter, R., Karrenbach, M., & Treitel, S. 2001. Identification and classification of multiple reflections with self-organizing maps. *Geoph. Prosp.*, **49**, 341–352.
- 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.
- Friedlander, F. G., & Joshi, M. 1998. *Introduction to the theory of distributions*. Cambridge university press.
- Guillon, A., & Verschuur, D. J. 2004. Adaptive subtraction of multiples using the L_1 -norm. *Geoph. Prosp.*, **52**, 27–38.
- Hron, F. 1972. Numerical methods of ray generation in multilayered media. *Methods in computational physics*, **12**, 1–34.
- Jakubowicz, H. 1998. Wave equation prediction and removal of interbed multiples. *Pages 1527–1530 of: Expanded Abstracts. Soc. Explor. Geophys.*
- 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.
- Moses, H. E. 1956. Calculation of scattering potential from reflection coefficients. *Phys. Rev.*, **102**, 559–567.
- Park, J., & Levin, V. 2000. Receiver functions from multiple-taper spectral correlation estimates. *Bull. Seis. Soc.*

- Am.*, 1507–1529.
- Park, J., & Levin, V. 2001. Receiver functions from regional P waves. *Geophys. J. Int.*, **147**, 1–11.
- Razavy, M. 1975. Determination of the wave velocity in an inhomogeneous medium from reflection data. *J. Acoust. Soc. Am.*, **58**, 956–963.
- Revenaugh, J., & Jordan, T. H. 1987. Observations of first-order mantle reverberations. *Bull. Seism. Soc. Am.*, **77**, 1704–1717.
- Revenaugh, J., & Jordan, T. H. 1989. A study of mantle layering beneath the western pacific. *Jour. Geoph. Res.*, **94**, 5787–5813.
- Sava, P., & Guitton, A. 2004. Multiple attenuation in the image space. *submitted, Geophysics*.
- Sjöstrand, J., & Grigis, A. 1994. *Microlocal Analysis for Differential Operators : An Introduction*. Cambridge: Cambridge University Press.
- Stolk, C. C., & de Hoop, M. V. 2000. Modeling and inversion of seismic data in anisotropic elastic media. *Comm. Pure Appl. Math.* submitted.
- Stolk, C. C., & de Hoop, M. V. 2004a. Modeling of seismic data in the downward continuation approach. *SIAM J. Appl. Math.* CWP468P.
- Stolk, C. C., & de Hoop, M. V. 2004b. Seismic inverse scattering in the downward continuation approach. *SIAM J. Appl. Math.* CWP469P.
- Symes, W., & Gockenbach, M. 1999. *Coherent noise suppression in velocity inversion*. Soc. of Expl. Geophys. Pages 1719–1722.
- ten Kroode, A. P. E. 2002. Prediction of internal multiples. *Wave Motion*, **35**, 315–338.
- van Borselen, R. 2002. Data-driven interbed multiple removal: Strategies and examples. *In: Expanded Abstracts*. Society of Exploration Geophysicists.
- Verschuur, D. J., & Berkhout, A. 1997. Estimation of multiple scattering by iterative inversion, Part II: Practical aspects and examples. *Geophysics*, **62**(5), 1596–1611.
- Weglein, A., Gasparotto, F. A., Carcalho, P. M., & Stolt, R. H. 1997. An inverse-scattering series method for attenuating multiples in seismic reflection data. *Geophysics*, **62**, 1975–1989.
- Weglein, A., Araújo, F. B., Carvalho, P. M., Stolt, R. H., Matson, K. H., Coates, R. T., Corrigan, D., Foster, D. J., Shaw, S. A., & Zhang, H. 2003. Inverse scattering series and seismic exploration. *Inverse Problems*, **19**, R27–R83.
- Yoshida, K. 1995. *Functional Analysis*. Berlin: Springer.

Table 1. table of symbols

symbol	in or near	meaning
\mathcal{P}	(1)	wave operator
u	(1)	particle displacement
f	(1)	source
$c(z, x)$	(1)	velocity at position x and depth z
$A(z, x, \xi, \tau)$	(2)	principal symbol of transverse Helmholtz operator
$D = \begin{pmatrix} u \\ \partial_z u \end{pmatrix}$	(4)	field
\mathbf{A}	(4)	second order system operator matrix
$M = \begin{pmatrix} 0 \\ -f \end{pmatrix}$	(4)	source vector
U	(5)	up- and down-going wave vector
u_+, u_-	(5)	down- and up-going displacements respectively
X	(5)	up- and down-going source vector
f_+, f_-	(5)	down- and up-going source respectively
\mathbf{Q}	(5),(8)	diagonalization matrix for the \mathbf{A} -matrix
\mathbf{B}	(6)	wave operator matrix in diagonal (first order) system
B_+, B_-	(7)	down- and up-going wave operators respectively
b_{\pm}	(7)	principle symbols of the one-way operators
Q_+, Q_-	(8)	entries in the \mathbf{Q} -matrix $Q_{\pm} = B_{\pm}^{-1/2}$
\cdot^*	(8)	operator adjoint
\cdot'		quantity without dissipative term
\mathbf{P}'	(10)	one-way wave operator (without dissipative term)
L'	(10)	parametrix of system of one-way equations (without dissipative term)
G'_+, G'_-	(10)	down- and up-going propagators respectively
$\theta = \arcsin(c(z, x) \tau^{-1}\xi)$	12	angle between the vertical and the ray
τ, ξ, ζ	12	cotangent variables associated with t, x, z , respectively
$(z_{\min, \pm}, z_{\max, \pm})$	12	maximal interval in which ray can be parameterized by z
θ_1, θ_2	Figure 1	specific values of θ for pseudodifferential cutoff
I_{θ}	(13), Figure 2	set of phase space
C	(13)	constant everywhere larger than $c(z, x)^{-1}$
$J_{\pm}(z_0, \theta)$	(15)	set of phase space excluding horizontal propagation
$\psi_{-,1}$	(16)	pseudodifferential cutoff

symbol	in or near	meaning
C	(19)	dissipative term
L, G_{\pm}, P	(19)	same as above with dissipative term excluding horizontal propagation
\cdot_0		operator in the background medium
$\delta \cdot$		difference between operator in true and background media
$a = 2c_0^{-3} \delta c$	(24)	medium contrast
V	(34)	contrast source
$S_{\pm, \pm}$	(34)	reflection/transmission operators; elements of the V matrix
\tilde{V}	(38)	V without the D_t^2
R	(43)	restriction operator
$M_0 = RQ^{-1}L_0$	(43)	parametrix of one-way equation restricted to acquisition surface and returned to the original second order system
d_1	(44)	single-scattered data
F	(44)	single scattering modeling operator
V_m ($V = \delta P$)	(48)	contrast operator of order m in the data
N	(45)	normal operator
d	(48)	data
$\delta u_{\pm, j}$	(62)	wavefield after j scatters going in the \pm direction
t_{s_0}	(62)	source time
$\tilde{\cdot}$	(62)	variables of integration over source parameters
z_j	(62)	j^{th} datum (see Figure 6)
s_j, r_j	(62)	j^{th} source/receiver (see Figure 6)
x_j	(62)	lateral position at j^{th} datum, generally $x_j = s_j = r_j$
c_{\pm}	(64)	components of contrast source
t_{r_0}	(65)	total single scattering traveltime
t_1	(65)	start time of single-scattering receiver ray
E_1, E_2	(69)	wave coupling operators
t_0	(70)	zero time at contrast for single scattering, at z_2 in triple scattering
t'	(71)	convolution variable
H	(73)	Green's function convolution operator
t_a	(75) Figure 7	time from s_0 through s_2, r_2 to m_r
m_s, m_r	(76)	lateral position of upward scatter (see Figure 6)
t_3	(76) Figure 7	t_a plus time from m_r, m_s to s_3, r_3
t_4	(77) Figure 7	total time on triple-scattered ray
t'_3	(79)	convolutional variable, equal to time from m_s, m_r to s_3, r_3
t_{30}	(79)	zero time at contrast at z_3
t_{m_0}	(79),(88)	zero time at contrast at z_1
$t_{m'_r}$	(80) Figure 7	time from m'_r to m_r
$t_{m'_s}$	(82) Figure 7	time from m'_s to m_s
t'_3	(83)	convolutional variable, time from m'_s to s_3, r_3
$t_b = t_a + t_{m'_r}$	(85) Figure 7	time from s_0 through s_2, r_2 to m'_r
w	(86)	convolution of two 'data' sets
W	(87)	integral of w with respect to z_2 and z_3
$t_{m'} = t_{m'_s} + t_{m'_r}$	(88) Figure 7	time from m'_s through m_s, m_r to m'_r
K	(90)	downward continuation forward modelling operator
$R_{1,2}$	(91)	restriction operator, inverse of $E_{1,2}$
$\tilde{\Xi}, \psi_Y$ and $\tilde{\Phi}(z, x, D_z, D_x)$	(94)	pseudodifferential operators to correct for amplitude affects
$\langle \cdot \rangle$	(95)	estimate of quantity, for a using traditional inverse
$\mathbf{d}_1(z_1; s, r, t)$	(99)	single-scattered data with scatters only below the z_1 level
$(\mathbf{D}(a))(z_1, s, r, t)$	(99)	data with scatters only above the z_1 level
$R_{2, t'}$	(102)	R_2 operator acting in the t' variable
m_s^0, m_r^0	(102)	surface points for imaging at z_1
t_{m_0}	(102)	time from m_s^0 through m_s, m_r to m_r^0
d'_1	(104)	single scattered data without integration over z_1
A_{WE}	(109)	wave equation angle transform
J	(110)	jacobian
R_3	(110)	operator to map (z, s, r, t) to (z, x, p)
h	(110)	half-offset
χ	(110)	pseudodifferential cutoff about $h = 0$
$\langle \langle \cdot \rangle \rangle$	(111)	estimate of medium contrast using angle transform
Ψ_{WE}	(112)	composition of pseudodifferential operators
p	(112)	slowness vector

