INSTITUTE OF PHYSICS PUBLISHING INVERSE PROBLEMS

Inverse Problems **21** (2005) 1137–1167 doi:10.1088/0266-5611/21/3/021

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

Alison E Malcolm and Maarten V de Hoop

Center for Wave Phenomena, Colorado School of Mines, Golden, CO 80401, USA

Received 26 July 2004, in final form 21 February 2005
Published 13 May 2005
Online at stacks.iop.org/IP/21/1137

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

(Some figures in this article are in colour only in the electronic version)

1. Introduction

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

Fokkema and van den Berg [14] developed a rigorous theory for the suppression of surface-related multiples. A surface-related multiple is a wave that has been reflected at least three times, with at least one reflection at the surface. Their analysis is derived from the reciprocity theorem in integral form and results in a Neumann series representation to predict surface-related multiples. If assumptions allowing the construction of data at zero offset, such as those given by de Hoop *et al* [12] are satisfied, then, in theory, Fokkema and van den Berg's theory solves the surface-related multiple attenuation problem. This paper provides a theory for the suppression of leading-order internal multiples, which are waves that have been reflected three times with no reflections from the Earth's surface.

The work presented here is motivated by the series solutions to inverse scattering problems developed by Moses [28], Prosser [29] and Razavy [30], as well as the Bremmer series approach to multiple attenuation discussed by Aminzadeh [2]. Moses constructed a series to represent the quantum scattering potential in terms of measured reflection coefficients. Prosser discusses this methodology from the algorithm construction viewpoint and touches on convergence issues. Razavy extends this work to recovering the velocity from the reflection coefficient via the scalar wave equation. These three papers use the Lippmann–Schwinger series [26], which is a pair of series for the forward and inverse scattering problem. This series representation has been used in exploration seismology by Weglein *et al* [42, 41]. The Bremmer series was introduced by Bremmer [6] to solve the wave equation in a horizontally layered medium. The convergence of this one-dimensional series is discussed by Atkinson [3] and Gray [17]. Aminzadeh used the Bremmer series to model the seismic wavefield [1] and construct filters to attenuate surface-related multiples [2], both in horizontally layered media. The Bremmer series was extended to two-dimensional problems by Corones [10]; the convergence of this series is discussed by McMaken [27]. De Hoop [11] introduces a generalization of the Bremmer series to multi-dimensional laterally varying media. This generalization is a Neumann series for forward scattering, which motivates its use here.

From these two series, we develop a hybrid series that uses the directional decomposition (into up- and down-going constituents) of the Bremmer series along with the Lippmann–Schwinger medium decomposition into a known, smooth reference velocity model and unknown, singular perturbation or contrast. Using this hybrid series allows us to trace waves through their up and down scatters while still preserving the contrast-source formulation of the Lippmann–Schwinger construction.

We develop an explicit scheme for modelling and imaging with the triply scattered wave constituent that can be extended to higher-order scattering. This triple scattering scheme is naturally integrated in the downward continuation approach to inverse scattering in the Born approximation. This scheme requires knowledge of the velocity model only to the depth of the shallowest reflector involved in the triple scattering.

In reflection seismology, two distinct methods have been used to attenuate multiples to obtain an approximation of singly scattered data. The first predicts the triply scattered data and then subtracts them from the data set. The second filters out multiples, using filters designed to exploit the differences in moveout (change in arrival time with source–receiver separation) between primaries and multiples. The work discussed in this paper falls into the first category.

In the prediction approach, Kennett [22, 24] used the Thomson–Haskell [25] method in horizontally layered media to model synthetic seismograms containing both surface and internal multiples. In [23], he uses this theory to suppress surface-related multiples in plane-layered elastic media. There are several extensions of the surface-related multiple attenuation theory of Fokkema and van den Berg [14] to internal multiples [15, 4, 40, 39]. In these methods, a particular layer is identified as the multiple generator (i.e. the layer where the second reflection occurs) and the surface-related multiple attenuation is adapted to be applied at that layer. Dragoset and Jeričević give a practical algorithm for attenuating surface-related multiples in [13]; an algorithm such as that discussed by Dragoset and Jeričević could be used for internal multiples in any of the mentioned extensions. Weglein and others [42] have used the Lippmann–Schwinger series to model and process seismic data, including the suppression of both surface-related and internal multiples, without knowledge of the velocity model. In ten Kroode [38] the mathematical theory behind that approach is given in both one and two dimensions. He shows that internal multiples can be estimated without knowledge of the velocity model if the velocity model satisfies two conditions: ten Kroode's travel-time monotonicity assumption (this condition is described in appendix B), and the condition that the wavefield contains no caustics. When the two assumptions of ten Kroode are satisfied, our method can be rewritten in a form consistent with the method of Weglein *et al* [42]; this is discussed further in appendix B. Jakubowicz [20] proposes a method for modelling internal multiples by correlating one primary reflection with the convolution of two other primary reflections; his approach implicitly uses the Bremmer series and is similar to the work presented here under ten Kroode's travel-time monotonicity assumption. Kelamis *et al* [21] use an approach similar to that of Jakubowicz, in which the multiples are constructed from a combination of different data sets, both at the surface and in the subsurface. In any method that predicts internal multiples and subtracts them, an adaptive subtraction technique such as that suggested by Guitton [18] must be used.

Aside from reflection seismology, there are other applications in which multiply scattered waves are important. In earthquake seismology, Burdick and Orcutt [7] investigate the truncation of the generalized ray sum, from which they find earth models in which the inclusion of internal multiples becomes important. In [31], Revenaugh and Jordan observe both internal and surface-related multiples and use them to estimate the attenuation quality factor, Q, of the mantle. In [32, 33], the same authors use multiples to investigate layering in the mantle. Bostock *et al* [5] use incident teleseismic P-waves scattered from a free surface and then subsurface structure before being recorded in an inversion scheme in which the teleseismic P-wave coda is used to invert for subsurface structure. For synthetic aperture radar (SAR) data, Cheney and Borden [8] derive a theory to relate the singular structure (wavefront set) of the object to the singular structure of the multiply scattered data.

In the next section we describe the techniques of the directional decomposition used in the Bremmer series. In the third section, we describe some of the details of the construction of one-way Green functions. This is followed by a description of the contrast-source method used for the Lippmann–Schwinger series. In the fifth section, we construct the hybrid series. In the sixth section we use the hybrid series to model data, giving the first of our three main results in (83). The proof of this result is given in appendix A. Following this, we summarize a method of constructing an inverse to the modelling operator. We then describe, through a series of results in section 8, a method for estimating artefacts in the image caused by leading-order internal multiples. Appendix B shows the correspondence between the theory described here and that of ten Kroode [38] and Weglein [42] under certain assumptions.

2. Directional decomposition

In the Bremmer series formulation of scattering, the wavefield is split into up- and down-going constituents. This is done by separating the vertical, z, derivative from the horizontal, x, derivatives, and then writing the wave equation as a first-order system of partial differential equations in z. This system is then diagonalized completing the separation into up- and down-going constituents. We begin with the scalar acoustic wave equation

$$\left[-c(z,x)^{-2}D_t^2 + \sum_{j=1}^{n-1} D_{x_j}^2 - \partial_z^2\right] u = f, \tag{1}$$

where $x_1, \ldots, x_{n-1}$ denote the horizontal coordinates and $D_{x_j} \equiv -\mathrm{i}\partial_{x_j}$, $D_t \equiv -\mathrm{i}\partial_t$; $c(z, x)$ is the isotropic velocity function and f is a source density of injection rate. These equations do not account for attenuation in the medium. We write the wave equation as

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

where A is the transverse 'Helmholtz' operator, with symbol[1] $A(z, x, \xi, \tau) = c(z, x)^{-2}\tau^2 - \|\xi\|^2$. In general, we use Greek letters for cotangent variables, dual to the space/time variables (ξ is the horizontal wave number, dual to x, and τ is radial frequency, the dual of time, t). This notation is consistent with [36, 37] as this work builds upon these papers. To correspond with the notation of exploration seismology, τ is typically denoted as ω, ξ as k_x and ζ as k_z. The notation $\|\cdot\|$ indicates the norm of a vector.

To simplify the notation in (2), we re-write it in matrix form

$$\partial_z D = \mathsf{A} D + M, \tag{3}$$

where

$$D = \begin{pmatrix} u \\ \partial_z u \end{pmatrix}, \qquad \mathsf{A} = \begin{pmatrix} 0 & 1 \\ -A(z, x, D_x, D_t) & 0 \end{pmatrix} \qquad \text{and} \qquad M = \begin{pmatrix} 0 \\ -f \end{pmatrix}. \tag{4}$$

We diagonalize the operator matrix A, which can be done microlocally[2], away from the zeros of $A(z, x, \xi, \tau)$. There is a z-family of pseudodifferential operator matrices $\mathsf{Q}(z)$ such that microlocally,

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

and

$$\mathsf{B} = \mathsf{Q}(z)\mathsf{A}\mathsf{Q}^{-1}(z) = \begin{pmatrix} \mathrm{i}B_+(z, x, D_x, D_t) & 0 \\ 0 & \mathrm{i}B_-(z, x, D_x, D_t) \end{pmatrix}, \tag{6}$$

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

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

$$(I\partial_z + \mathsf{Q}(z)\partial_z \mathsf{Q}^{-1}(z) - \mathsf{B})U = X, \tag{7}$$

where I is the identity matrix.

With the conventions used here, u_+ represents downward propagating waves and u_- represents upward propagating waves. (As is standard in geophysics, we have chosen the positive z-axis downward.) The columns of the Q operator matrix are an operator generalization of eigenvectors and we are free to choose their normalization in the operator sense. We choose the vertical power flux normalization of de Hoop [11] so as to make $B_\pm$ in (6) self-adjoint (the normalization changes the sub-principal part of the operator). In this normalization, the decomposition and composition operators are

$$\mathsf{Q} = \frac{1}{2}\begin{pmatrix} (Q_+^*)^{-1} & -\mathcal{H}Q_+ \\ (Q_-^*)^{-1} & \mathcal{H}Q_- \end{pmatrix}, \qquad \mathsf{Q}^{-1} = \begin{pmatrix} Q_+^* & Q_-^* \\ \mathcal{H}Q_+^{-1} & -\mathcal{H}Q_-^{-1} \end{pmatrix}, \tag{8}$$

[1] The symbol of the differential operator, $P(x, D_x)$, is defined as $P(x, \xi)$ in which the D_x has been simply replaced with ξ. The principal symbol is generally denoted with the same symbol in lower case, i.e., $p(x, \xi)$.

[2] A statement is true microlocally, basically, if it is true in a neighbourhood of a point in phase space. See [34] for an introduction to microlocal analysis.

where * denotes the operator adjoint, $\mathcal{H}$ is the Hilbert transform in time and the principal symbol of both $Q_\pm$ is given by $\left(\frac{\tau^2}{c(z,x)^2} - \|\xi\|^2\right)^{-1/4}$. The $Q_\pm$ operators act in the time variable as time convolutions. From expressions (5) and (6) we find that

$$u = Q_+^* u_+ + Q_-^* u_-, \qquad \text{and} \qquad f_\pm = \pm\tfrac{1}{2}\mathcal{H}Q_\pm f. \tag{9}$$

In the flux normalization, the term $\mathsf{Q}^{-1}\partial_z\mathsf{Q}$ in (7) is of lower order in the singularities (i.e. the operator is smoothing in comparison with other terms), thus we suppress it. If required, its contribution can be accounted for by including it in the B matrix. We introduce the propagators for the one-way wave equations (7) as

$$(I\partial_z - \mathsf{B})\,\mathsf{L} = I\delta, \qquad \mathsf{L} = \begin{pmatrix} G_+ & 0 \\ 0 & G_- \end{pmatrix}. \tag{10}$$

We will denote $I\partial_z + \mathsf{B}$ by P. We can now write the solution of (7) as $U = \mathsf{L}X$, using Duhamel's principle, L is the forward parametrix of P. In components, in integral form this is

$$u_+(z,\cdot) = \int_{-\infty}^{z} G_+(z,z_0)\,f_+(z_0,\cdot)\,\mathrm{d}z_0 \qquad u_-(z,\cdot) = \int_{z}^{\infty} G_-(z,z_0)\,f_-(z_0,\cdot)\,\mathrm{d}z_0. \tag{11}$$

To make a connection to ray theory, the propagation of singularities by the one-way wave equations (7) is governed by their principal symbols. These yield the Hamiltonians, $\zeta \mp b$, for the system describing the rays in phase space; the evolution parameter along the rays is taken to be depth, z. In the following section we use this analogy to subject $u_\pm$ and $G_\pm$ to cut-offs removing near horizontally propagating constituents of the wavefield.

3. The Green functions

In the previous section we diagonalized the wave equation into two first-order equations. In doing this, we implicitly assume that the diagonal system is equivalent to the original system. This is nearly the case, but the choice of a principal direction alters the ability of the system to propagate singularities in directions orthogonal to this preferred direction. Here, we have chosen the vertical direction as the principal direction. To ensure that the diagonal system does not propagate singularities incorrectly, singularities that propagate somewhere horizontally must be attenuated. The details of the method are given in [36]; we give only a brief description here to introduce the double-square-root (DSR) assumption used by Stolk and de Hoop. This assumption states that there are no wave constituents that propagate horizontally at any time. At the end of this section, we give a brief summary of the essential properties of the Green functions.

In order to identify horizontal propagation, we work in the high frequency limit, i.e. we develop these ideas via ray theory. Thus we define the phase angle

$$\theta = \arcsin(c(z,x)\|\tau^{-1}\xi\|), \tag{12}$$

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 $\frac{\mathrm{d}z}{\mathrm{d}t}$ does not change sign, allowing the parametrization of the ray segment by z. Thus, for any ray segment and any given angle $\theta < \pi/2$, we can define a maximal interval,

$$(z_{\min\pm}(z,x,\xi,\tau,\theta), z_{\max\pm}(z,x,\xi,\tau,\theta)), \tag{13}$$

for which the propagation away from a particular point (z,x,ξ,τ) can be parametrized by z. In figure 1, the interval $(z_{\min-}, z_{\max-})$ is illustrated; it is the maximal interval such that a bicharacteristic passing through the point (z,x), with direction (ζ,ξ), propagates in a direction such that the angle of the ray with the vertical, θ, does not exceed a given value; in the figure

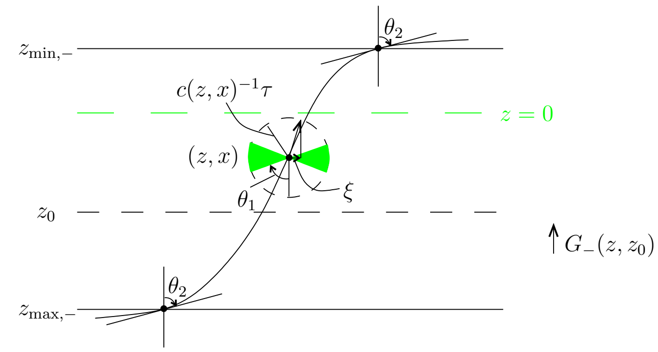

Figure 1. Removing horizontal propagations. The symbol of the cut-off operator ψ is one up to an angle θ_1 and then decays smoothly to zero at the angle θ_2. This removes all propagation at angles larger than θ_2, i.e., the region within the grey wedges.

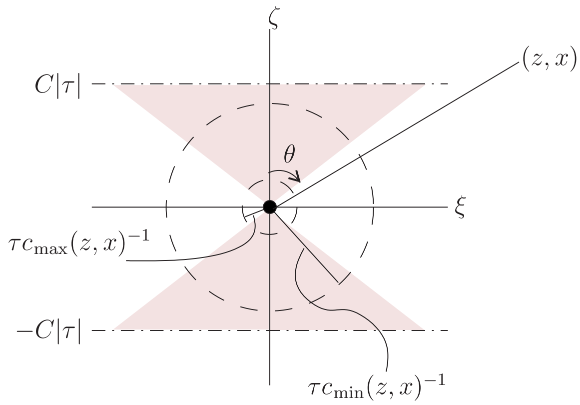

Figure 2. Illustration of I_θ. The shaded region represents the ray directions in the set. The minimum velocity in the region is $c_{\min}$ and the maximum is $c_{\max}$.

this value is θ_2. The angle θ can be given physical meaning by looking at the ray picture, in figure 1.

In phase space, we introduce the set

$$I_\theta = \{(z, x, t, \zeta, \xi, \tau) | \arcsin(c(z, x)\|\tau^{-1}\xi\|) < \theta, |\zeta| < C|\tau|\}, \tag{14}$$

illustrated in figure 2, where C is the maximum slowness. Finally, we construct the sets

$$J_-(z_0, \theta) = \{(z, x, t, \zeta, \xi, \tau) \in I_\theta | \tau^{-1}\zeta < 0 \text{ and } z_{\max-}(z, x, \xi, \tau, \theta) \geqslant z_0\}, \tag{15}$$

and

$$J_+(z_0, \theta) = \{(z, x, t, \zeta, \xi, \tau) \in I_\theta | \tau^{-1}\zeta > 0 \text{ and } z_{\min+}(z, x, \xi, \tau, \theta) \leqslant z_0\}. \tag{16}$$

Figure 1 illustrates the set $J_-(z_0, \theta_2)$, considering the shaded region as excluded from the set.

The sets $J_\pm$ encompass the regions of phase space that must be excluded in order to remove horizontally propagating singularities while analysing $G_\pm(z, z_0)$. To actually remove singularities from these regions, we define a pseudodifferential cutoff

$$\psi_- = \psi_-(z, z_0, x, D_x, D_t)$$

with symbol satisfying

$$\psi_-(z,x,\xi,\tau) \sim 1 \quad \text{on} \quad J_-(z_0,\theta_1), \tag{17}$$

$$\psi_-(z,x,\xi,\tau) \in S^{\infty} \quad \text{outside} \quad J_-(z_0,\theta_2), \quad \text{if } z - z_0 > \delta > 0; \tag{18}$$

here $0 < \theta_1 < \theta_2$. Singularities propagating at an angle less than θ_1 are unaffected by the cutoff; at angles greater than θ_2, the operator is smoothing. We then redefine u_- as

$$u_- \equiv \psi_- u'_-, \tag{19}$$

where u'_- is the wavefield u_- of the previous section. In u_- the singularities outside of J_- have been suppressed. There are equivalent expressions for the + constituents. We now rewrite the operators defined above with the singularities outside of J_- (or J_+) suppressed. It is shown in [36] and references therein that the solution operator L to the system of one-way equations P (cf [10]) is

$$\mathsf{L} = \begin{pmatrix} G_+ & 0 \\ 0 & G_- \end{pmatrix}, \tag{20}$$

redefining $G_\pm = \psi G'_\pm$ where $G'_\pm$ is the propagator described in the previous section. From this point onwards we will assume that the above procedure has been followed and will be reapplied if necessary.

The condition $z_{\max -}(z,x,\xi,\tau,\theta) \geqslant z_0$, in the definition of J_- combined with the implicit requirement that $z_{\min -} < 0$ ensures that the two points between which one propagates the wavefield are within the allowed propagation interval $(z_{\min -}, z_{\max -})$.

Remark 3.1. We denote the kernel of $G_-(z_0,z)$ as $(G_-(z_0,z))(x_0,t_0-t,x) = G_-(z_0,x_0,t_0-t,z,x)$. The adjoint propagator $(G_-(z_0,z)^*)(x,t-t_0,x_0) = G_-^*(z,x,t-t_0,z_0,x_0)$ follows from

$$\begin{aligned}
&\int \mathrm{d}s_0\,\mathrm{d}t_0 v(z_0,s_0,t_0)\left(\int \mathrm{d}s\,\mathrm{d}t G_-(z_0,s_0,t_0-t,z,s)u(z,s,t)\right) \\
&\quad = \int \mathrm{d}s\,\mathrm{d}t\left(\int \mathrm{d}s_0\,\mathrm{d}t_0 v(z_0,s_0,t_0)G_-(z_0,s_0,t_0-t,z,s)\right)u(z,s,t) \\
&\quad = \int \mathrm{d}s\,\mathrm{d}t\left(\int \mathrm{d}s_0\,\mathrm{d}t_0 (G_-(z_0,z)^*)(s,t-t_0,s_0)v(z_0,s_0,t_0)\right)u(z,s,t).
\end{aligned} \tag{21}$$

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

Remark 3.2. The $G_\pm$ propagators obey the reciprocity relation (of the time convolution type)

$$Q_+^*(z)G_+(z,z_0)Q_+(z_0) = -Q_-^*(z_0)G_-(z_0,z)Q_-(z). \tag{22}$$

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

Remark 3.3. We have

$$G_-(z,z')G_-(z',z'') = G_-(z,z''), \tag{23}$$

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

In the above, we have nowhere assumed the absence of caustics in the wavefield. This section has addressed the necessary assumption that rays are nowhere horizontal: the double-square-root assumption [36, assumption 2].

4. Scattering: contrast source formulation

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

$$(I\partial_z - \mathsf{A}_0)D_0 = M, \qquad (I\partial_z - \mathsf{A})D = M, \tag{24}$$

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

$$(I\partial_z - \mathsf{A}_0)\delta D = -\delta\mathsf{A}D, \tag{25}$$

where $D = D_0 + \delta D$ and $\mathsf{A} = \mathsf{A}_0 + \delta\mathsf{A}$. The right-hand side of (25) is the so-called contrast source. We have (cf (4))

$$\delta\mathsf{A} = \begin{pmatrix} 0 & 0 \\ \delta A & 0 \end{pmatrix} \qquad \text{where} \quad \delta A(z, x, D_t) = -2c_0^{-3}\delta c(z, x)D_t^2 = -a(z, x)D_t^2, \tag{26}$$

defining the contrast a. We insert the Bremmer formulation into that above by diagonalizing the A_0 matrix operator. We apply the (smooth background) diagonalizing Q operator matrices to transform the system in (25). Using the diagonalization procedure of section 2, equation (6) in particular, we find

$$(I\partial_z - \mathsf{B}_0)\delta U = -\mathsf{Q}(z)\partial_z\mathsf{Q}^{-1}(z)\delta U - \mathsf{Q}(z)\delta\mathsf{A}\mathsf{Q}^{-1}(z)U, \tag{27}$$

recalling, from section 2, the definition of U

$$U = \mathsf{Q}(z)D \tag{28}$$

while,

$$U_0 := \mathsf{Q}(z)D_0, \tag{29}$$

$$\delta U := \mathsf{Q}(z)\delta D. \tag{30}$$

The Q operator matrix is common in all the transformations. Note that $\delta\mathsf{A}$ will not, in general, be diagonalized by Q as the Q operators diagonalize A_0 in the background velocity model only.

Since we have used the flux normalization, the $-\mathsf{Q}(z)\partial_z\mathsf{Q}^{-1}(z)\delta U$ term is of lower order as before (discussion above (10)). This term could be absorbed in $\delta\mathsf{A}$ by $\delta\mathsf{A} := \delta\mathsf{A} + I\partial_z$. We omit this contribution so that

$$(I\partial_z - \mathsf{B}_0)\delta U = -\mathsf{Q}(z)\delta\mathsf{A}\mathsf{Q}^{-1}(z)U, \tag{31}$$

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

$$V = \mathsf{Q}(z)\delta\mathsf{A}\mathsf{Q}^{-1}(z) = \frac{1}{2}\mathcal{H}\begin{pmatrix} Q_+(z)aQ_+^*(z) & Q_+(z)aQ_-^*(z) \\ -Q_-(z)aQ_+^*(z) & -Q_-(z)aQ_-^*(z) \end{pmatrix} D_t^2. \tag{32}$$

In (31), we make the analogy with (7) where δU plays the role of U and $-\mathsf{Q}(z)\delta\mathsf{A}\mathsf{Q}^{-1}(z)U$ that of X, which is now the contrast source.

We make the comparison between the elements of V and the reflection and transmission operators of [11], namely,

$$V = \begin{pmatrix} S_{++} & S_{-+} \\ S_{+-} & S_{--} \end{pmatrix} D_t^2. \tag{33}$$

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

To simplify the notation, we define

$$P_0 = I\partial_z - \mathsf{B}_0, \tag{34}$$

and its forward parametrix,

$$\mathsf{L}_0 = \begin{pmatrix} G_+ & 0 \\ 0 & G_- \end{pmatrix}. \tag{35}$$

In this notation, (31) reduces to

$$P_0 \delta U = -VU, \tag{36}$$

or

$$\delta U = -\mathsf{L}_0(VU). \tag{37}$$

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

$$\delta U = -\mathsf{L}_0(VU_0) - \mathsf{L}_0(V\delta U), \tag{38}$$

or equivalently,

$$(I + \mathsf{L}_0 V)\delta U = -\mathsf{L}_0(VU_0). \tag{39}$$

As was done in (26), we take out the time derivative from V. Thus we introduce $\widehat{V}$, the matrix of $S_{\pm\pm}$ operators (cf (33)), namely,

$$V(z, x, D_t) = \widehat{V}(z, x) D_t^2, \tag{40}$$

which results in

$$\left(I + D_t^2 \mathsf{L}_0 \widehat{V}\right) \delta U = -D_t^2 \mathsf{L}_0(\widehat{V} U_0), \tag{41}$$

where $\widehat{V}\delta U$ and $\widehat{V}U_0$ are products of distributions (subject to the condition that their wavefront is favourably oriented [16, proposition 11.2.3], [19, theorem 8.2.10]). This is the resolvent equation in our hybrid Lippmann–Schwinger–Bremmer formulation for scattered waves. (See [43] for details on resolvent equations.)

5. Scattering series

5.1. Forward scattering series

In this section, we describe the construction of the forward scattering series for δU in terms of $\widehat{V}$, based on the discussion of the previous section. We arrive at expressions (43), (46) and (47) below, through which data are modelled.

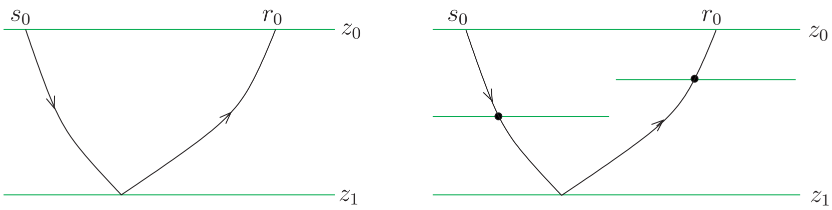

Figure 3. Single scattering (left panel) versus primary reflection (right panel). The black dots indicate transmission scatterings.

Here, we define a singly scattered wave as a wave that has been reflected or transmitted once, such as that shown on the left in figure 3. The term primary reflection is associated with any 'ray-path' (more accurately wave-path since we use wave solutions rather than ray theory) that has reflected only once but may have gone through several transmissions, 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 travel time as singly scattered waves but will have different amplitudes because of the transmissions. The same distinction can be made between leading-order internal multiples and triply scattered waves. The diagram on the right of figure 3 is a triply scattered event. 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 contributions such as these, where none of the three scattering events occurs at the acquisition surface, as leading-order internal multiples. The goal of this section is to develop a method for modelling such scattered wave constituents in the data.

Having identified (41) as a resolvent equation, we set up the recursion

$$\delta U = \sum_{m=1}^{M} (-1)^m \delta U_m(\widehat{V}), \tag{42}$$

where

$$\delta U_1(\widehat{V}) = D_t^2 \mathsf{L}_0(\widehat{V} U_0), \qquad \text{and} \qquad \delta U_m(\widehat{V}) = D_t^2 \mathsf{L}_0(\widehat{V} \delta U_{m-1}(\widehat{V})), \quad m = 2, 3, \ldots. \tag{43}$$

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

To compare the Bremmer series formulation ((VII.1)–(VII.22) of [11]) to the recursion in (42) we first make the following identifications. From (VII.1) and (VII.12) we note that W_0 of [11] corresponds to δU_1. From this formulation we note that $-D_t^2 \mathsf{L}_0 \widehat{V}$ corresponds to K of equation (VII.15) in [11] and (42) corresponds to equation (VII.22). The major difference between this hybrid series and the Bremmer series is in the coupling of the different components. In the Bremmer series the reflection and transmission operators come from derivatives of the medium contrast whereas in the hybrid series they come from differences between the reference and true model. We use this hybrid formulation to derive operators that model both 'singly' and 'triply' scattered waves.

The expressions in (42) and (43) are not quite in the form of observables, however; data are acquired only at the Earth surface, but the L_0 operator models data at all depths. We therefore define a restriction operator, R, which restricts a distribution to the acquisition surface, $z = 0$. This operator does not account for the free-surface boundary condition, thus we assume a continuation of the medium, with no reflectors, above the acquisition surface. In this way we have excluded incoming waves from above the acquisition surface. We assume that there are no reflectors at or near this surface, i.e., we assume that the support of the medium contrast, a, does not contain source or receiver points. The composition $R\mathsf{L}_0$ is well defined provided there are no grazing rays [35], which have been excluded already by the ψ cut-off from section 3. Observable quantities are obtained by applying Q^{-1} to δU, as in (30). We thus rewrite (42) as

$$R\mathsf{Q}^{-1}\delta U(\widehat{V}) = \sum_{m=1}^{M}(-1)^m R\mathsf{Q}^{-1}\delta U_m(\widehat{V}), \tag{44}$$

with

$$R\mathsf{Q}^{-1}\delta U_1(\widehat{V}) = D_t^2 R\mathsf{Q}^{-1}\mathsf{L}_0(\widehat{V}U_0), \tag{45}$$

introducing the operator

$$\mathsf{M}_0 = R\mathsf{Q}^{-1}\mathsf{L}_0. \tag{46}$$

We then can write

$$R\mathsf{Q}^{-1}\delta U(\widehat{V}) = -D_t^2\mathsf{M}_0\left(\widehat{V}\left(U_0 + \sum_{m=1}^{M}(-1)^m\delta U_m(\widehat{V})\right)\right), \tag{47}$$

for modelling the surface reflection data, in which δU_m is defined in (43). The first term on the right-hand side of (47) is the Born approximation. Using the notation introduced in (4) we have, from the leading-order term, an expression for the singly scattered data[3]

$$\delta D = \begin{pmatrix} d \\ \partial_z d \end{pmatrix} = -R\mathsf{Q}^{-1}\delta U_1(\widehat{V}) = -D_t^2\mathsf{M}_0(\widehat{V}U_0). \tag{48}$$

From this term the singly scattered data are modelled in section 6.1. In section 6.2 the second term of (47) is used to model internal multiples by examining the $m = 2$ term of the summation. Note that the recursion in (47) gives an expression for the data at the surface in terms of the unrestricted field from the previous step; the restriction is applied as a last step after the recursion is completed.

5.2. *Inverse scattering series using all the data*

The forward scattering series (47) models the data, given a representation of the medium as the sum of a smooth background and singular contrast. The inverse series estimates the medium contrast from the data. In this section we derive this inverse series, arriving at a recursion for the medium contrast in (61).

To motivate the inverse series, we return to (41) and write it as an equation for $\widehat{V}$ in terms of δU

$$D_t^2\mathsf{L}_0(\widehat{V}(U_0 - (-\delta U))) = -\delta U, \tag{49}$$

or, returning to observables via the $R\mathsf{Q}^{-1}$ operator,

$$D_t^2\mathsf{M}_0(\widehat{V}(U_0 - (-\delta U))) = -\delta D. \tag{50}$$

We then set up an inverse series, by assuming that the medium contrast can be represented in terms of a series of operators,

$$\widehat{V} = \sum_{m=1}^{M}\widehat{V}_m \tag{51}$$

[3] We continue to use the notation D in δD even though we have now applied the restriction operator, R.

where m indicates the 'order' of $\widehat{V}_m$ in the data. This series representation is suggested for quantum mechanical problems by Moses [28], where the analogue of (51) is his equation (3.12). It is also suggested by Razavy [30] for wave problems, in which the analogue of (51) is his equation (33). Perhaps the closest analogue to what is done here is given by Prosser [29], equations (7) and (8). It is this theory, for the Lippmann–Schwinger series, that is used extensively by Weglein *et al* [42].

Substituting (51) into (42) yields a recursion for $\widehat{V}_m$ in terms of δU,

$$\delta U = -D_t^2 \mathsf{L}_0(\widehat{V}_1 U_0) \tag{52}$$

$$0 = -D_t^2 \mathsf{L}_0(\widehat{V}_2 U_0) + D_t^4 \mathsf{L}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_1 U_0)) \tag{53}$$

$$0 = -D_t^2 \mathsf{L}_0(\widehat{V}_3 U_0) + D_t^4 \mathsf{L}_0(\widehat{V}_2 \mathsf{L}_0(\widehat{V}_1 U_0)) + D_t^4 \mathsf{L}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_2(U_0))) \\ - D_t^6 \mathsf{L}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_1 U_0))) \tag{54}$$

$$\vdots$$

These equations are assumed to hold anywhere in the interior of the scattering region. Restricting δU to the surface and transforming it into observables by applying $R\mathsf{Q}^{-1}$ to (52)–(54) yields a recursion for $\widehat{V}_m$ in terms of the data d,

$$\delta D = -D_t^2 \mathsf{M}_0(\widehat{V}_1 U_0) \tag{55}$$

$$0 = -D_t^2 \mathsf{M}_0(\widehat{V}_2 U_0) + D_t^4 \mathsf{M}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_1 U_0)) \tag{56}$$

$$0 = -D_t^2 \mathsf{M}_0(\widehat{V}_3 U_0) + D_t^4 \mathsf{M}_0(\widehat{V}_2 \mathsf{L}_0(\widehat{V}_1 U_0)) + D_t^4 \mathsf{M}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_2(U_0))) \\ - D_t^6 \mathsf{M}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_1 U_0))) \tag{57}$$

$$\vdots$$

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

The first term in the series, given in (55), models singly scattered data. The third term, in (57), models leading-order internal multiples as well as other primary events such as that shown on the right in figure 3. (The second term, given in (56), models events which have scattered twice, including primary events with one transmission and one reflection.)

Equation (57) can be simplified using (53). This is done by noting that the distributions $D_t^2 \mathsf{L}_0(\widehat{V}_2 U_0)$ from the second term of (57) and $D_t^4 \mathsf{L}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_1 U_0))$ from the third term are identical by (53) and $D_t^2 \mathsf{M}_0 \widehat{V}_1(\cdot)$, which acts on these distributions (again in the second and third terms) is a linear operator. With this simplification we have, for (57)

$$D_t^2 \mathsf{M}_0(\widehat{V}_3 U_0) = D_t^4 \mathsf{M}_0(\widehat{V}_2 \mathsf{L}_0(\widehat{V}_1 U_0)). \tag{58}$$

The general recursion follows from the fact that higher-order terms are built from lower-order terms through the application of $D_t^2 \mathsf{M}_0 \widehat{V}_i$ to $(j - i)$th-order terms to form terms of order j. For example, terms of order 4 are formed by subtracting $D_t^2 \mathsf{M}_0 \widehat{V}_1$ applied to (54), $D_t^2 \mathsf{M}_0 \widehat{V}_2$ applied to (53) and $D_t^2 \mathsf{M}_0 \widehat{V}_3$ applied to the right-hand side of (52), from $D_t^8 \mathsf{M}_0 \widehat{V}_4 U_0$. In general, terms of order j will contain sub-series of the form

$$D_t^2 \mathsf{M}_0 \widehat{V}_1 \text{ (sum of terms of order } j-1 \text{ from (52)–(54))}, \tag{59}$$

$$D_t^2 \mathsf{M}_0 \widehat{V}_2 \text{ (sum of terms of order } j-2 \text{ from (52)–(54))}, \tag{60}$$

etc. For $j \geqslant 2$ the sub-series in parentheses sum to zero because of the zero on the left-hand side of (53).

We obtain the final form of the recursion,

$$D_t^2 \mathsf{M}_0(\widehat{V}_j U_0) = D_t^4 \mathsf{M}_0(\widehat{V}_{j-1} \mathsf{L}_0(\widehat{V}_1 U_0)), \qquad j \geqslant 2, \tag{61}$$

while

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

Solving these recursions for $\widehat{V}_j$ gives, in principle, a solution for the medium contrast, $\widehat{V}$, in terms of the data, d, as in (51). Note the similarity in structure between (61) and (43); (61) constructs the medium contrast in terms of the data, while (43) constructs the data in terms of the medium contrast.

Remark 5.1. Using (61) along with the expression for δU in (52), we can write the $\widehat{V}$-series as

$$-D_t^2 \mathsf{M}_0(\widehat{V} U_0) = \delta D - \left(\sum_{m=1}^{M} D_t^2 \mathsf{M}_0(\widehat{V}_m \delta U) \right). \tag{62}$$

This expresses higher order terms in the series as a correction to the data, d. In what follows, we examine the correction obtained from the sum in (62); we specifically examine the $m = 2$ term in the series.

Remark 5.2. We verify the compatibility of (47) and (51) for $M = 2$. To this end we insert,

$$\widehat{V} \approx \widehat{V}_1 + \widehat{V}_2 + \widehat{V}_3, \tag{63}$$

into the truncated forward series

$$\delta D \approx -D_t^2 \mathsf{M}_0(\widehat{V} U_0) + D_t^4 \mathsf{M}_0(\widehat{V} \mathsf{L}_0(\widehat{V} U_0)) - D_t^6 \mathsf{M}_0(\widehat{V} \mathsf{L}_0(\widehat{V} \mathsf{L}_0(\widehat{V} U_0))). \tag{64}$$

Terms of first, second and third 'order' in the resulting sum cancel. The fourth 'order' term in this truncated sum is

$$\begin{aligned} &D_t^4 \mathsf{M}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_3 U_0)) + D_t^4 \mathsf{M}_0(\widehat{V}_2 \mathsf{L}_0(\widehat{V}_2 U_0)) + D_t^4 \mathsf{M}_0(\widehat{V}_3 \mathsf{L}_0(\widehat{V}_1 U_0)) \\ &\quad - D_t^6 \mathsf{M}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_2 U_0))) - D_t^6 \mathsf{M}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_2 \mathsf{L}_0(\widehat{V}_1 U_0))) \\ &\quad - D_t^6 \mathsf{M}_0(\widehat{V}_2 \mathsf{L}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_1 U_0))), \end{aligned} \tag{65}$$

which vanishes by (61). This implies that the error contains fifth 'order' to ninth 'order' terms.

6. Modelling multiply scattered data

This section illustrates the modelling of data based on the series discussed in the previous section. We consider two cases: modelling primaries in the single scattering approximation and modelling internal multiples from the third term of the series. Section 6.1 derives a method of modelling the primaries (singly scattered data), d_1, from the first term in (47). Section 6.2 derives a representation of internal multiples, d_3 using the $m = 2$ term of the sum in (47). In both these sections, we track the wavefield from the source, through the scattering(s) to the receiver. The results of this section are the expression for modelling singly scattered data given in (75) and that for modelling triply scattered data in (80).

6.1. Single scattering

The first term in the forward scattering series given in (47) is used to construct data in the Born approximation in accordance with equation (3.10) of [36]. We give here an alternate derivation of this equation, resulting in our equation (75). We formulate the solution only for the upward propagating constituent of δU_1, which we denote by $\delta u_{-,1}$. We first determine the form of the down-going constituent of U_0, denoted by $u_{+,0}$, which is the down-going wave excited at the surface and arriving at the scattering point. With the expression for the source f_+ in (9) and that for u_+ in (11) we find that

$$u_{+,0}(z_1, x_1, t_1, z_0, s_0) = \frac{1}{2}\int_{-\infty}^{z_1} \mathrm{d}\tilde{z}_0 \int \mathrm{d}\tilde{s}_0 \int_{\mathbb{R}} \mathrm{d}\tilde{t}_{s_0} G_+(z_1, x_1, t_1 - \tilde{t}_{s_0}, \tilde{z}_0, \tilde{s}_0) \times \mathcal{H} Q_{+,\tilde{s}_0}(\tilde{z}_0) f(\tilde{z}_0, \tilde{s}_0, \tilde{t}_{s_0}, z_0, s_0), \tag{66}$$

where we will adopt the convention that an integral without limits is assumed to be an integration over $\mathbb{R}^{n-1}$. In general, s represents a source position, r represents a receiver position, t is a time variable and z is depth, regardless of subscripts and superscripts. The notation $Q_{-,s}(z)$ is short for $Q_-(z, s, D_s, D_t)$. The t integrations are limited implicitly by the causality of the Green function. The operator G_+ in (66) propagates between the levels z_0 and z_1, with its action being in the lateral variables $\tilde{s}_0$, and $\tilde{t}_{s_0}$; we will also use the notation $G_+(z_1, z_0)$ for the propagator G_+ when the lateral positions at which it acts are unambiguous. We adopt the standard kernel notation that the input variables to an operator are written to the right of the output variables. We are justified in writing the time dependence of $G_\pm$ as the difference of elapsed time and initial source time as the wave equation is time translation invariant. Expression (66) is valid for $z_1 > z_0$. The parameters z_0, s_0 are assumed to be known.

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

$$\begin{pmatrix} c_+ \\ c_- \end{pmatrix} = V U_0 = V \begin{pmatrix} u_{+,0} \\ u_{-,0} \end{pmatrix}. \tag{67}$$

Using the expression for V in (33), and recalling that $u_{-,0} = 0$ for depths deeper than the source depth, we obtain an expression for c_-,

$$c_-(z_1, x_1, t_1) = -\tfrac{1}{2}\mathcal{H} D_{t_1}^2 Q_{-,x_1}(z_1) a(z_1, x_1) Q^*_{+,x_1}(z_1) u_{+,0}(z_1, x_1, t_1, z_0, s_0). \tag{68}$$

Substituting c_- from (68) for f_- in (11) gives

$$\delta u_{-,1}\big(z_0, r_0, t_{r_0}, z_0, s_0\big) = -\frac{1}{2}\mathcal{H} D_{t_{r_0}}^2 \int_{z_0}^{\infty} \mathrm{d}z_1 \int \mathrm{d}x_1 \int_{\mathbb{R}} \mathrm{d}t_1\, G_-\big(z_0, r_0, t_{r_0} - t_1, z_1, x_1\big) \times \underbrace{Q_{-,x_1}(z_1) a(z_1, x_1) Q^*_{+,x_1}(z_1)}_{S_{+-}} u_{+,0}(z_1, x_1, t_1, z_0, s_0) \tag{69}$$

in the diagonal system without the restriction to the Earth's surface, $z_0 = \tilde{z}_0 = 0$. This is the first term in the series in (42)–(43). Because a is compactly supported in z_1, the integral over z_1 is actually over a compact set. As in the previous section, we assume that the medium contrast, a, has its support away from $z = 0$. To obtain modelled data, we apply the $R\mathsf{Q}^{-1}$ operator as in (47),

$$d_1\big(s_0, r_0, t_{r_0}\big) = \int \mathrm{d}\tilde{s}_0 \int_{\mathbb{R}} \mathrm{d}\tilde{t}_{s_0} \frac{1}{4} D_{t_{r_0}}^2 \int_0^{\infty} \mathrm{d}z_1 \int \mathrm{d}x_1 \int_{\mathbb{R}} \mathrm{d}t_1\, Q^*_{-,r_0}(0) \\ G_-\big(0, r_0, t_{r_0} - t_1, z_1, x_1\big) Q_{-,x_1}(z_1) a(z_1, x_1) Q^*_{+,x_1}(z_1) \\ G_+\big(z_1, x_1, t_1 - \tilde{t}_{s_0}, 0, \tilde{s}_0\big) Q_{+,\tilde{s}_0}(0) f\big(0, \tilde{s}_0, \tilde{t}_{s_0}, 0, s_0\big), \tag{70}$$

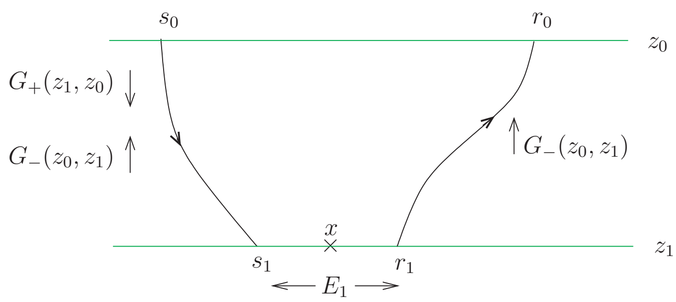

Figure 4. Notation for single scattering modelling.

yielding the Born modelled data in terms of the $G_\pm$, the solutions of the single square-root equation. This is the first entry in the δD vector, in the series in (47).

We apply reciprocity (22) to (70) to write d_1 in terms of G_- only, giving

$$d_1\big(s_0, r_0, t_{r_0}\big) = -\int \mathrm{d}\tilde{s}_0 \int_{\mathbb{R}} \mathrm{d}\tilde{t}_{s_0} f\big(0, \tilde{s}_0, \tilde{t}_{s_0}, 0, s_0\big) \frac{1}{4} D_{t_{r_0}}^2 \int_0^\infty \mathrm{d}z_1 \int \mathrm{d}x_1 \int_{\mathbb{R}} \mathrm{d}t_1 \, Q_{-,r_0}^*(0)$$
$$G_-\big(0, r_0, t_{r_0} - t_1, z_1, x_1\big) Q_{-,x_1}(z_1) Q_{-,\tilde{s}_0}^*(0)$$
$$G_-\big(0, \tilde{s}_0, t_1 - \tilde{t}_{s_0}, z_1, x_1\big) Q_{-,x_1}(z_1) a(z_1, x_1). \tag{71}$$

To write (71) in terms of a single Green function for the source and receiver together, there must be integrations in (x_1, t_1) for each of the Green functions. To introduce these integrations we introduce two extension operators,

$$E_1 : a(z, x) \mapsto \delta(r - s) a\left(z, \frac{r + s}{2}\right), \qquad E_2 : b(z, r, s) \mapsto \delta(t) b(z, r, s), \tag{72}$$

through their action on the functions a and b. These operators extend the medium contrast, $a(z, x)$, into fictitious data (now a function of (z, s, r, t)) in the subsurface as illustrated in figure 4. With these operators, we re-write (71), now assuming a point source in both space and time. This gives,

$$d_1\big(s_0, r_0, t_{r_0}\big) = -\frac{1}{4} D_{t_{r_0}}^2 \int_0^\infty \mathrm{d}z_1 \int \mathrm{d}s_1 \int \mathrm{d}r_1 \int_{\mathbb{R}} \mathrm{d}t_0 \int_{\mathbb{R}} \mathrm{d}t_1 \, Q_{-,r_0}^*(0)$$
$$G_-\big(0, r_0, t_{r_0} - t_1 - t_0, z_1, r_1\big) Q_{-,r_1}(z_1) \, Q_{-,s_0}^*(0)$$
$$G_-(0, s_0, t_1, z_1, s_1, 0) Q_{-,s_1}(z_1) (E_2 E_1 a)(z_1, s_1, r_1, t_0). \tag{73}$$

We note that the two one-way Green functions are connected through time convolution.

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

$$(H(z_0, z_1))(s_0, r_0, t - t_0, s_1, r_1)$$
$$= \int_{\mathbb{R}} (G_-(z_0, z_1))(r_0, t - t' - t_0, r_1)(G_-(z_0, z_1))(s_0, t', s_1) \, \mathrm{d}t', \tag{74}$$

the kernel of the propagator $H(z_0, z_1)$ associated with the so-called double-square-root equation [9], which propagates data from the depth z_1 to the depth z_0. Substituting this expression for the two Green functions in (73) gives equation (3.10) of [36, theorem 5.1],

$$d_1\big(s_0, r_0, t_{r_0}\big) = -\frac{1}{4} D_{t_{r_0}}^2 \, Q_{-,r_0}^*(0) Q_{-,s_0}^*(0) \int_0^\infty \mathrm{d}z_1 (H(0, z_1) Q_{-,r_1}(z_1)$$
$$Q_{-,s_1}(z_1)(E_2 E_1 a))\big(s_0, r_0, t_{r_0}\big). \tag{75}$$

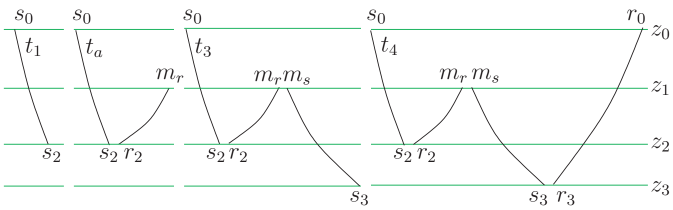

Figure 5. Triple scattering notation and convention. This illustration assumes that the E_2 and E_1 operators have been applied to be clear which variable refers to which leg of the interactions.

6.2. Leading-order internal multiple scattering

In (75), we showed how singly scattered data can be constructed given the medium perturbation. Our ultimate goal is to construct the medium contrast given data containing both primaries and leading-order internal multiples. In this section we establish a relation between the modelling of primaries and internal multiples.

Following the diagram in figure 5, the first scattering of the internal multiple, from s_0 through s_2, r_2 to m_r is nearly identical to the single scattering case. We cannot use the H operator however, because the second leg (from r_2 to m_r) does not reach the surface, $z = 0$. Thus,

$$\delta u_{-,1}(z_1, m, t_a, 0, s_0) = -\frac{1}{4} D_{t_a}^2 \int \mathrm{d}\tilde{s}_0 \int_{\mathbb{R}} \mathrm{d}\tilde{t}_{s_0} f\big(0, \tilde{s}_0, \tilde{t}_{s_0}, 0, s_0\big) Q^*_{-,\tilde{s}_0}(0)$$
$$\int_{z_1}^{\infty} \mathrm{d}z_2 \int \mathrm{d}s_2 \int \mathrm{d}r_2 \int_{\mathbb{R}} \mathrm{d}t_0 \int_{\mathbb{R}} \mathrm{d}t' G_-\big(z_1, m, t_a - \tilde{t}_{s_0} - t' - t_0, z_2, r_2\big)$$
$$\times G_-(0, \tilde{s}_0, t', z_2, s_2) Q_{-,r_2}(z_2) Q_{-,s_2}(z_2)(E_2 E_1 a)(z_2, s_2, r_2, t_0), \tag{76}$$

where $t' = t_1 - \tilde{t}_{s_0}$ and t_a is the running time variable along the ray (see figure 5). We assume that the three scattering points for multiple scattering are sufficiently far apart. We assume that the singular support of a consists of a countable set of hypersurfaces. This prevents an undefined multiplication of distributions from occurring (see [16, proposition 11.2.3] and [19, theorem 8.2.10]). In (76), we have not returned to observables as the second leg, $G_-\big(z_1, m, t_a - \tilde{t}_{s_0} - t' - t_0, z_2, r_2\big)$, does not reach the surface ($z_1 > 0$). The field, $\delta u_{-,1}$, acts as the source of waves propagating from m to s_3, through the contrast source formulation used in the single-scattering case. (The contrast source was explicitly defined in section 4 equation (25).) This gives,

$$\delta u_{+,2}(z_3, x_3, t_3, 0, s_0) = \frac{1}{2} \mathcal{H} D_{t_3}^2 \int_0^{z_3} \mathrm{d}z_1 \int \mathrm{d}m \int_{\mathbb{R}} \mathrm{d}t_a G_+(z_3, x_3, t_3 - t_a, z_1, m)$$
$$\times Q_{+,m}(z_1) a(z_1, m) Q^*_{-,m}(z_1) \delta u_{-,1}(z_1, m, t_a, 0, s_0), \tag{77}$$

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

$$d_3(s_0, r_0, t_4) = -\frac{1}{2} \mathcal{H} D_{t_4}^2 Q^*_{-,r_0}(0) \int_0^{\infty} \mathrm{d}z_3 \int \mathrm{d}x_3 \int_{\mathbb{R}} \mathrm{d}t_3 G_-(0, r_0, t_4 - t_3, z_3, x_3)$$
$$\times Q_{-,x_3}(z_3) a(z_3, x_3) Q^*_{+,x_3}(z_3) \delta u_{+,2}(z_3, x_3, t_3, 0, s_0), \tag{78}$$

where we have returned to observables through the operator $R\mathsf{Q}^{-1}$, introduced in (47). For the above construction to be valid, (z_1, x_1), (z_2, x_2) and (z_3, x_3) cannot be arbitrarily close to one another.

We now apply reciprocity (22) to the G_+ occurring in the expression for $\delta u_{+,2}$ in (77). We do this by substituting the expression for $\delta u_{+,2}$ in (77) into (78) to use the Q_+ operators from both expressions combined and introduce the extension operators E_1, E_2. This gives

$$\begin{aligned} d_3(s_0, r_0, t_4) = -\frac{1}{4} D_{t_4}^4 \int_0^\infty \mathrm{d}z_3 \int \mathrm{d}s_3 \int \mathrm{d}r_3 \int_{\mathbb{R}} \mathrm{d}t_3 \int_0^{z_3} \mathrm{d}z_1 \int \mathrm{d}m_s \int \mathrm{d}m_r \int_{\mathbb{R}} \mathrm{d}t_a \, Q_{-,r_0}^*(0) \\ G_-(0, r_0, t_4 - t_3, z_3, r_3) Q_{-,r_3}(z_3) \, Q_{-,m_s}^*(z_1) \\ G_-(z_1, m_s, t_3 - t_a, z_3, s_3) Q_{-,s_3}(z_3) \\ (E_1 a)(z_3, s_3, r_3)(E_1 a)(z_1, m_s, m_r) Q_{-,m_r}^*(z_1) \delta u_{-,1}(z_1, m_r, t_a, 0, s_0); \end{aligned} \tag{79}$$

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

Associating the propagator $Q_{-,x_a}(z_a) G_-(z_a, z_b) Q_{-,z_b}(z_b)$ in (79) with the function $G(z_a, x_a, t, z_b, x_b)$ in equation (8) of [38] along with the substitution of the expression for $\delta u_{-,1}$ in (77) shows the correspondence of (80) with expression (8) in [38]. (Note that $V(x)$ in [38] is $a(z, x)$ here.)

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

$$\begin{aligned} d_3(s_0, r_0, t_4) = -\frac{1}{4} D_{t_4}^4 \int_0^\infty \mathrm{d}z_3 \int \mathrm{d}s_3 \int \mathrm{d}r_3 \int_{\mathbb{R}} \mathrm{d}t_{30} \int_{\mathbb{R}} \mathrm{d}t_a \int_0^{z_3} \mathrm{d}z_1 \int \mathrm{d}m_s \int \mathrm{d}m_r \\ \int_{\mathbb{R}} \mathrm{d}t_3' \, Q_{-,r_0}^*(0) G_-(0, r_0, t_4 - t_a - t_3' - t_{30}, z_3, r_3) Q_{-,r_3}(z_3) Q_{-,m_s}^*(z_1) \\ \times G_-(z_1, m_s, t_3', z_3, s_3) Q_{-,s_3}(z_3)(E_2 E_1 a)(z_3, s_3, r_3, t_{30}) \\ \times (E_1 a)(z_1, m_s, m_r) Q_{-,m_r}^*(z_1) \delta u_{-,1}(z_1, m_r, t_a, 0, s_0), \end{aligned} \tag{80}$$

which is a modelling operator for triply scattered waves. We need not introduce E_2 at the m_s, m_r scattering point here, but it will be required later. Equations (80) and (76) are expressed entirely in terms of up-going propagators (G_-); they comprise the $m = 2$ term of the forward series, given in the summation in (47).

The recursion in equation (61) demonstrates that it is possible to express the triply scattered data, d_3, in terms of the singly scattered data d_1. The first step to writing d_3 in terms of the singly scattered data is to reformulate (80) so that propagation is always to the acquisition surface. This idea is motivated by the layer stripping approach proposed by Fokkema [15] to extend the work of Berkhout and Verschuur for surface multiples [4, 40] to the internal multiple case.

Theorem 6.1. *Let the data be modelled by (47) for $M = 2$. Let*

$$\begin{aligned} \mathbf{d}_1(z_1; s_0, r_0, t) = -\frac{1}{4} D_t^2 Q_{-,r_0}^*(0) Q_{-,s_0}^*(0) \int_{z_1}^\infty \mathrm{d}z \\ (H(0, z) Q_{-,r}(z) Q_{-,s}(z)(E_2 E_1 a))(s_0, r_0, t) \end{aligned} \tag{81}$$

represent the single scattered data constituent observed at the surface, but scattered below the depth z_1. Define the convolution

$$W(z_1; s_0, m_r', t, m_s', r_0) = \int_{\mathbb{R}} \mathrm{d}t_b \, \mathbf{d}_1(z_1; m_s', r_0, t - t_b) \mathbf{d}_1(z_1; s_0, m_r', t_b), \tag{82}$$

and let d_3 denote the triply scattered field corresponding to the $m = 2$ term in (47). Then,

$$\begin{aligned} d_3(s_0, r_0, t_4) = D_{t_4}^2 \int_0^\infty \mathrm{d}z_1 \int \mathrm{d}m_s \int \mathrm{d}m_r \int_{\mathbb{R}} \mathrm{d}t_{m_0} (E_2 E_1 a)\big(z_1, m_s, m_r, t_{m_0}\big) Q_{-,m_s}^*(z_1) \\ \times Q_{-,m_r}^*(z_1) H(0, z_1)^* Q_{-,m_r'}^*(0)^{-1} Q_{-,m_s'}^*(0)^{-1} W\big(z_1; s_0, \overset{m_r'}{\cdot}, t_4 + \overset{t_{m'}}{\cdot}, \overset{m_s'}{\cdot}, r_0\big). \end{aligned} \tag{83}$$

The proof is given in appendix A. The fictitious data $\mathbf{d}_1$ are the singly scattered data constituent predicting reflections below the level z_1. Because W is estimated at $z = 0$ but depends explicitly on z_1, the depth level that generates multiples, we separate z_1 from the other variables with a semi-colon.

Assuming the travel-time monotonicity assumption as done in [38], would allow the restriction in z_1 to be translated to a restriction on the time $t - t_s$, allowing $\mathbf{d}_1$ to be computed from d_1 by windowing in time. Expression (83) can be viewed as an inner product in the $\left(m_s, m_r, t_{m_0}\right)$ variables.

In appendix B, we write d_3 entirely in terms of the data, completing the correspondence with (61), and compare our approach to that of Weglein [42] and ten Kroode [38].

7. Inverse scattering method

Rather than following the approach of attenuating multiples in the data, we estimate and attenuate artefacts in the image caused by leading-order internal multiples. This requires an estimate of the multiples in the image rather than in the data as we have done thus far. To this end, we now discuss an inverse scattering theory. From the inverse series, constructed in section 5.2, we note that only a single-scattering inverse is required, because for each term in the series we estimate $\widehat{V}_j$ from $\mathsf{M}_0(\widehat{V}_j U_0)$ based on the recursion in (61). We therefore only need to determine the inverse of the linear mapping $\widehat{V} \mapsto -D_t^2 \mathsf{M}_0(\widehat{V} U_0)$.

A left inverse to the Born modelling operator, the inverse scattering operator, can be constructed under the double-square-root assumption (see section 3 and [36]). Stolk and de Hoop [37] give a method for inverse scattering from singly scattered data; here we give a brief summary. The construction involves the depth-to-time conversion operator, $\bar{K}$, defined as

$$\bar{K} : a \mapsto -\int_0^\infty H(0, z)(E_2 a)(z, \cdot, \cdot, \cdot)(s, r, t)\, \mathrm{d}z. \tag{84}$$

Stolk and de Hoop show that this operator is an invertible Fourier integral operator. Upon substitution of a point source in (75), we obtain

$$d_1 = \tfrac{1}{4} D_t^2 Q_{-,s}^*(0) Q_{-,r}^*(0) \bar{K} J(E_1 a). \tag{85}$$

The operator J (denoted V by Stolk and de Hoop [37]), has the symbol

$$\begin{aligned} J(z, s, r, \zeta, \sigma, \rho) = {} & |\tau|^{-1}(c_0(z, s)^{-2} - \tau^{-2}\|\sigma\|^2)^{-1/4} \\ & \times (c_0(z, r)^{-2} - \tau^{-2}\|\rho\|^2)^{-1/4}|_{\tau=\Theta^{-1}(z,s,r,\zeta,\sigma,\rho)}. \end{aligned} \tag{86}$$

This operator is related to the $Q_{-,s}(z) Q_{-,r}(z)$ appearing in (75); the difference is that J is applied before the E_2 extension operator whereas $Q_{-,s}(z) Q_{-,r}(z)$ is applied after. The map Θ is defined by (cf (6))

$$\Theta(z, s, r, \sigma, \rho, \tau) = -b(z, s, \sigma, \tau) - b(z, r, \rho, \tau). \tag{87}$$

Stolk and de Hoop [36, lemma 4.1] show that $\tau \mapsto \zeta = \Theta(z, s, r, \sigma, \rho, \tau)$ is a diffeomorphism. The mapping from frequency to vertical wavenumber described by this map is required for J to be applied before the E_2 extension operator.

After defining the adjoint operator in space (restriction to $s = r$) by $R_1 = E_1^*$, the adjoint operator in time (restriction to $t = 0$) by $R_2 = E_2^*$, and the normal operator $\bar{\Xi} = \bar{K}^* \bar{K}$ we have

$$\bar{\Phi}(z, x, D_z, D_x) a = R_1 J^{-1} \bar{\Xi}^{-1} \bar{K}^* Q_{-,s}^*(0)^{-1} Q_{-,r}^*(0)^{-1} D_t^{-2} d_1, \tag{88}$$

where $\bar{\Phi}$ is shown in [37, theorem 2.2, remark 2.4] to be a pseudodifferential operator.

The operator $\bar{\Phi}$ influences only the amplitudes of the image; its principal symbol is calculated by Stolk and de Hoop [37, lemma 2.1, theorem 2.2, remark 2.4].

We have essentially determined an inverse of the linear mapping $\widehat{V} \mapsto -D_t^2 \mathsf{M}_0(\widehat{V} U_0)$. From (61) we then have an estimate of the single scattering inverse

$$\langle a_1 \rangle = \bar{\Phi} a_1 = R_1 J^{-1} \bar{\Xi}^{-1} \bar{K}^* Q_{-,s}^*(0)^{-1} Q_{-,r}^*(0)^{-1} D_t^{-2} d. \tag{89}$$

In (89) we have used the single scattering approximation, in which the data in (55) are used as an approximation of the data in (48). 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 in the single scattering approximation. From this estimate of a, we obtain an estimate of the operator matrix V_1 using (32), with

$$\langle \widehat{V}_1 \rangle = \frac{1}{2} \mathcal{H} \begin{pmatrix} Q_+ \langle a_1 \rangle Q_+^* & Q_+ \langle a_1 \rangle Q_-^* \\ -Q_- \langle a_1 \rangle Q_+^* & -Q_- \langle a_1 \rangle Q_-^* \end{pmatrix}. \tag{90}$$

8. The downward continuation approach to inverse scattering for internal multiples

The construction of $\mathbf{d}_1$ with (81), at the surface, requires both an estimate of a and the modelling of the wavefield from this estimate. If the d_3 data set could be computed at the depth z_1 rather than at the surface $z = 0$ this modelling can be avoided. In this section, we give three results that form the framework of an algorithm to estimate artefacts caused by internal multiples in imaging. We assume that the DSR assumption (see below remark 3.3) holds throughout this section.

Lemma 8.1. *We define*

$$\tilde{d}_1(z, s, r, t) = -\frac{1}{4} D_t^2 \int_z^\infty \mathrm{d}z' \big(H(z, z') Q_{-,r'}(z') Q_{-,s'}(z') (E_2 E_1 a)\big(z', \overset{s'}{\cdot}, \overset{r'}{\cdot}, \overset{t'}{\cdot}\big)\big)(s, r, t). \tag{91}$$

For $t > 0$

$$(H(0, z)^* Q_{-,s}^*(0)^{-1} Q_{-,r}^*(0)^{-1} d_1)(s, r, t) = \tilde{d}_1(z, s, r, t), \tag{92}$$

where d_1 *is modelled by (75).*

This lemma is illustrated in figure 8.

We first define $\bar{a} = (1 - \chi) a$ where χ is the characteristic function of $(0, z)$. With this definition we write $\tilde{d}_1$ as

$$\tilde{d}_1(z, s, r, t) = -\frac{1}{4} D_t^2 \int_0^\infty \mathrm{d}z' \big(H(z, z') Q_{-,r'}(z') Q_{-,s'}(z') (E_2 E_1 \bar{a})\big(z', \overset{s'}{\cdot}, \overset{r'}{\cdot}, \overset{t'}{\cdot}\big)\big)(s, r, t). \tag{93}$$

We then examine

$$\begin{aligned} Q_{-,s}^*(0) Q_{-,r}^*(0) H(0, z) \tilde{d}_1 - d_1 &= -\frac{1}{4} D_t^2 Q_{-,s}^*(0) Q_{-,r}^*(0) \int_0^\infty \big(H(0, z') Q_{-,r'}(z') \\ &\quad \times Q_{-,s'}(z') (E_2 E_1 \chi a)\big(z', \overset{s'}{\cdot}, \overset{r'}{\cdot}, \overset{t'}{\cdot}\big)\big)(s, r, t) \\ &= -\frac{1}{4} D_t^2 Q_{-,s}^*(0) Q_{-,r}^*(0) \int_0^z \big(H(0, z') Q_{-,r'}(z') \\ &\quad \times Q_{-,s'}(z') (E_2 E_1 a)\big(z', \overset{s'}{\cdot}, \overset{r'}{\cdot}, \overset{t'}{\cdot}\big)\big)(s, r, t). \end{aligned} \tag{94}$$

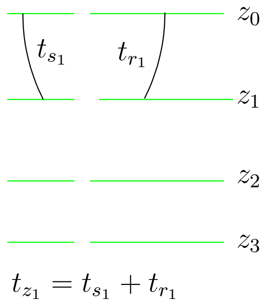

Figure 6. Time notation used to estimate d_3 at z_1.

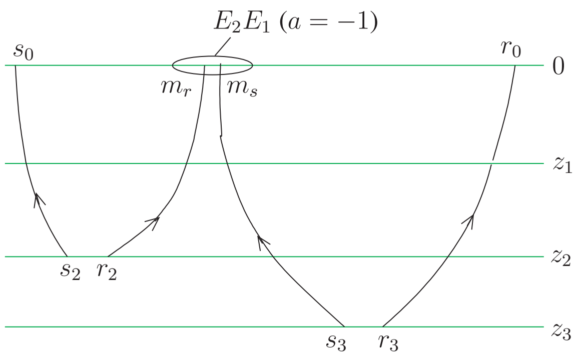

Figure 7. Illustration of the surface-related multiple elimination case (SRME).

Applying $H(0,z)^* Q^*_{-,s}(0)^{-1} Q^*_{-,r}(0)^{-1}$ to both sides of (94) gives

$$\tilde{d}_1 - H(0,z)^* Q^*_{-,s}(0)^{-1} Q^*_{-,r}(0)^{-1} d_1 = -\frac{1}{4} D_t^2 H(0,z)^* \int_0^z \big(H(0,z') Q_{-,r'}(z') Q_{-,s'}(z') \\ \times (E_2E_1a)\big(z', \overset{s'}{\cdot}, \overset{r'}{\cdot}, \overset{t'}{\cdot}\big)\big)(s,r,t). \tag{95}$$

We follow the propagation of singularities of $H(0,z)^* H(0,z')$, subject to $0 < z' \leqslant z$ and the DSR assumption within the integral on the right-hand side of (95). The operator $H(0,z)^*$ backpropagates the singularities generated by $H(0,z')$ at the surface along exactly the same bicharacteristics. Microlocally, $\tilde{d}_1 = H(0,z) Q^*_{-,s}(0)^{-1} Q^*_{-,r}(0)^{-1} d_1$ for $t > 0$ only.

Equation (92) describes a method of estimating (for single scattering) the data that would have been recorded had the experiment been performed at depth z from the data recorded at the surface; this is downward continuation.

We now define the convolution of the $\tilde{d}_1$ data sets, that have been restricted to $t > 0$, at the depth z, of the second scattering point for leading-order internal multiples

$$\tilde{d}_3(z,s,r,t) = D_t^2 \int \mathrm{d}s' \int \mathrm{d}r' Q^*_{-,s'}(z)(E_1a)(z,s',r') Q^*_{-,r'}(z) \\ \times \psi(t)\tilde{d}_1\big(z,s',r,\overset{t}{\cdot}\big) \overset{(t)}{*} \psi(t)\tilde{d}_1\big(z,s,r',\overset{t}{\cdot}\big). \tag{96}$$

We use the notation $\psi(t)\tilde{d}_1$ to indicate the $t > 0$ restriction. The operator E_1 contains $\delta(s' - r')$, thus the integral in (96) is over all possible source–receiver pairs with $s' = r'$.

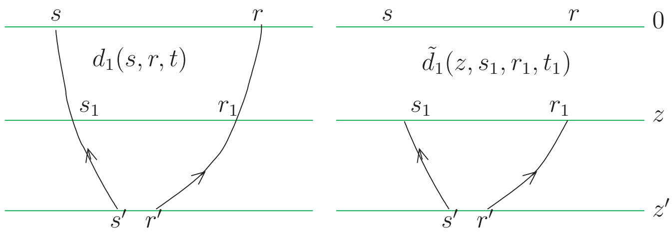

Figure 8. Illustration of lemma 8.1; the construction of $\tilde{d}_1$ from d_1.

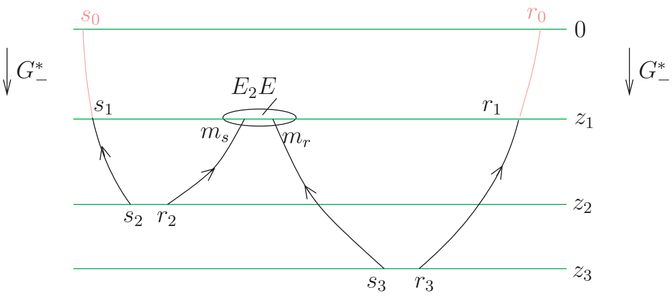

Figure 9. The d_3 data set at the depth z_1. The ellipse illustrates the application of the E_2E_1 operators to join the two data sets at m_s, m_r, t_{m_0}. This diagram illustrates the downward continuation of d_3 to form $\tilde{d}_3$ at depth z_1 as in theorem 8.3. The grey paths extending from z_1 to the surface illustrate the modelling of d_3 from $\tilde{d}_3$ with $\tilde{d}_3$ acting as a contrast source or the estimation of $\tilde{d}_3$ from d_3.

Remark 8.2. If we replace $D_t^2 a$ in (97) with -1 and the second scattering point is at the surface $z = 0$ then (96) becomes,

$$\tilde{d}_3^S(0,s,r,t) = -Q^*_{-,s}(0)Q^*_{-,r}(0)\int \mathrm{d}s' \int \mathrm{d}r'\, Q^*_{-,s'}(0)Q^*_{-,r'}(0)\delta(s'-r') \times \tilde{d}_1\big(0,s',r,\overset{t}{\cdot}\big) \overset{(t)}{*} \tilde{d}_1\big(0,s,r',\overset{t}{\cdot}\big), \tag{97}$$

returning to observables via $Q^*_{-,s}(0)Q^*_{-,r}(0)$. Noting that $Q^*_{-,s}(0)Q^*_{-,r}(0)\tilde{d}_1(0,s,r,t) = d_1(s,r,t)$ gives

$$d_3^S(s,r,t) = -\int \mathrm{d}s' \int \mathrm{d}r'\, d_1\big(s',r,\overset{t}{\cdot}\big) \overset{(t)}{*} d_1\big(s,r',\overset{t}{\cdot}\big)\delta(s'-r') \tag{98}$$

relating our method to the surface-related multiple elimination (SRME) procedure of Fokkema and van den Berg [14, chapter 12]. This is illustrated in figure 7.

The following theorem describes the relation between the internal multiple estimated at the surface through (83) given in theorem 6.1 and the estimate of $\tilde{d}_3$ defined in (96).

Theorem 8.3. *Let the data be modelled by the forward scattering series (47) for $M = 2$. Then there is the following correspondence between the leading-order internal multiple modelled at the surface and $\tilde{d}_3$*

$$d_3(s_0,r_0,t_0) = Q^*_{-,r_0}(0)Q^*_{-,s_0}(0)\int_0^\infty \mathrm{d}z_1 \big(H(0,z_1)\tilde{d}_3\big(z_1,\overset{s}{\cdot},\overset{r}{\cdot},\overset{t}{\cdot}\big)\big)(s_0,r_0,t_0). \tag{99}$$

The theorem is illustrated in figure 9.

We begin by returning to (76),

$$\delta u_{-,1}(z_1, m, t_a, 0, s_0) = -\frac{1}{4} D_{t_a}^2 Q_{-,s_0}^*(0) \int_{z_1}^{\infty} \mathrm{d}z_2 \int \mathrm{d}s_2 \int \mathrm{d}r_2 \int_{\mathbb{R}} \mathrm{d}t_0 \int_{\mathbb{R}} \mathrm{d}t' \\ G_-(z_1, m, t_a - t' - t_0, z_2, r_2) \int \mathrm{d}s_1 \int_{\mathbb{R}} \mathrm{d}t_{s_1} G_-\big(0, s_0, t_{s_1}, z_1, s_1\big) \\ \times G_-\big(z_1, s_1, t' - t_{s_1}, z_2, s_2\big) Q_{-,r_2}(z_2) Q_{-,s_2}(z_2)(E_2 E_1 a)(z_2, s_2, r_2, t_0), \quad (100)$$

assuming a point source and using relation (23); the time notation is illustrated in figure 6. We then change the order of integration in preparation for substituting H,

$$\delta u_{-,1}(z_1, m, t_a, 0, s_0) = -\frac{1}{4} D_{t_a}^2 Q_{-,s_0}^*(0) \int \mathrm{d}s_1 \int_{\mathbb{R}} \mathrm{d}t_{s_1} G_-\big(0, s_0, t_{s_1}, z_1, s_1\big) \\ \times \left\{ \int_{z_1}^{\infty} \mathrm{d}z_2 \int \mathrm{d}s_2 \int \mathrm{d}r_2 \int_{\mathbb{R}} \mathrm{d}t_0 \int_{\mathbb{R}} \mathrm{d}t' G_-(z_1, m, t_a - t' - t_0, z_2, r_2) \right. \\ \left. \times G_-\big(z_1, s_1, t' - t_{s_1}, z_2, s_2\big) Q_{-,r_2}(z_2) Q_{-,s_2}(z_2)(E_2 E_1 a)(z_2, s_2, r_2, t_0) \right\}. \\ (101)$$

Substituting H (cf (74)), for the two $G_-(z_1, x_2)$ propagators leads to the simplification

$$\delta u_{-,1}(z_1, m, t_a, 0, s_0) = Q_{-,s_0}^*(0) \int \mathrm{d}s_1 \int_{\mathbb{R}} \mathrm{d}t_{s_1} G_-\big(0, s_0, t_{s_1}, z_1, s_1\big) \\ \times \tilde{d}_1\big(z_1, s_1, m, t_a - t_{s_1}\big), \quad (102)$$

where we have substituted $\tilde{d}_1$ (given in lemma 8.1) for the expression in braces in (101). The same sequence of steps applied to (80) gives

$$d_3(s_0, r_0, t_4) = \frac{1}{4} D_{t_4}^2 \int_0^{\infty} \mathrm{d}z_1 \int \mathrm{d}r_1 \int \mathrm{d}t_{r_1} \; Q_{-,r_0}^*(0) G_-\big(0, r_0, t_{r_1}, z_1, r_1\big) \\ \times \int \mathrm{d}m_s \int \mathrm{d}m_r \; Q_{-,m_s}^*(z_1) \int_{\mathbb{R}} \mathrm{d}t_a \, \tilde{d}_1\big(z_1, m_s, r_1, t_4 - t_a - t_{r_1}\big) \\ \times (E_1 a)(z_1, m_s, m_r) Q_{-,m_r}^*(z_1) \delta u_{-,1}(z_1, m_r, t_a, 0, s_0), \quad (103)$$

where we have also interchanged the order of integration. Substituting the expression for $\delta u_{-,1}$ from (102) into (103) and re-ordering the Q operators and the G_- propagators results in

$$d_3(s_0, r_0, t_4) = D_{t_4}^2 Q_{-,r_0}^*(0) Q_{-,s_0}^*(0) \int_0^{\infty} \mathrm{d}z_1 \int \mathrm{d}m_s \int \mathrm{d}m_r \; Q_{-,m_s}^*(z_1)(E_1 a)(z_1, m_s, m_r) \\ \times Q_{-,m_r}^*(z_1) \int_{\mathbb{R}} \mathrm{d}t_a G_-(0, z_1) \tilde{d}_1\big(z_1, m_s, \overset{r_1}{\cdot}, t_4 - t_a - \overset{t_{r_1}}{\cdot}\big) \\ \times G_-(0, z_1) \tilde{d}_1\big(z_1, \overset{s_1}{\cdot}, m_r, t_a - \overset{t_{s_1}}{\cdot}\big). \quad (104)$$

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

Equation (99) is equivalent to (75) with $\tilde{d}_3$ taking the place of the contrast source.

Theorem 8.4. *Assume the inverse scattering series (62) for $M = 2$. If we replace d_1 in (92) in lemma 8.1 by d and a in equation (96) for $\tilde{d}_3$ by a_1 then*

$$\langle a_3(z, x) \rangle = \big(R_1 J^{-1} \Xi^{-1} R_2 D_t^{-2} \tilde{d}_3\big)(z, x). \quad (105)$$

Recall from the recursion in (61) that

$$D_t^2 \mathsf{M}_0(\widehat{V}_3 U_0) = D_t^6 \mathsf{M}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_1 \mathsf{L}_0(\widehat{V}_1 U_0))). \quad (106)$$

Theorem 6.1 shows that d_3 is third order in $\widehat{V}_1$ and thus third order in d. We then estimate $\widehat{V}_3$ directly from d_3 using (89).

$$\begin{aligned}\langle a_3\rangle &= R_1 J^{-1}\bar{\Xi}^{-1} K^* D_t^2 d_3 \\ &= R_1 J^{-1}\bar{\Xi}^{-1} R_2 H(0,z)^* Q^*_{-,r}(0)^{-1} Q^*_{-,s}(0)^{-1} D_t^{-2} d_3.\end{aligned} \tag{107}$$

The argument in the proof of lemma 8.1 can be repeated for the expression for d_3, recalling that $\tilde{d}_3$ is defined for $t > 0$, in (99) giving

$$\tilde{d}_3(z,s,r,t) = (H(0,z)^* Q^*_{-,s}(0)^{-1} Q^*_{-,r}(0)^{-1} d_3)(z,s,r,t), \tag{108}$$

for $t > 0$. We then have

$$\langle a_3(z,x)\rangle = \left(R_1 J^{-1}\bar{\Xi}^{-1} R_2 D_t^{-2}\tilde{d}_3\right)(z,x). \tag{109}$$

An estimate of $\widehat{V}_3$ is obtained from $\langle a_3\rangle$ by

$$\langle\widehat{V}_3\rangle = \frac{1}{2}\mathcal{H}\begin{pmatrix} Q_+\langle a_3\rangle Q_+^* & Q_+\langle a_3\rangle Q_-^* \\ -Q_-\langle a_3\rangle Q_+^* & -Q_-\langle a_3\rangle Q_-^*\end{pmatrix}, \tag{110}$$

so that the estimate of $\widehat{V}$ becomes

$$\widehat{V} \approx \widehat{V}_1 + \widehat{V}_3. \tag{111}$$

The estimate $\langle a_3\rangle$ corrects $\langle a_1\rangle$ by estimating and subtracting the erroneous contributions to $\langle a_1\rangle$ due to the single scattering assumption. Thus artefacts in the image, caused by internal multiples, are removed by subtracting an image of the multiples from an image of the full data set. Leading-order internal multiples and primaries have different illumination properties and therefore the estimated image artefacts will never be entirely accurate. We anticipate accounting for these illumination differences as well as errors in the estimate of d_3 via adaptive subtraction.

Remark 8.5. To estimate $\tilde{d}_3$ at depth z_1, knowledge of the velocity model is necessary only to the depth z_1; this knowledge is necessary to estimate $\tilde{d}_1$ at z_1. The same part of the velocity model is required to form an image at z_1, $\langle a_1\rangle$ or $\langle a_3\rangle$, from the data. To form a complete image of the subsurface a velocity model is necessary for all depths.

Remark 8.6. In this remark, we illustrate the estimation of $\langle a_3\rangle$ with an isochron construction. In figure 10 a contribution to $\langle a_3\rangle$ is shown. If the single scattering inverse is applied to the data d to estimate $\langle a_1\rangle$, the contributions from a particular source, receiver and time would be spread over the single scattering isochron (dashed curve). Although this is correct for a primary reflection, such as that shown with the dot-dash line, it is incorrect for a leading-order internal multiple, such as that shown with the solid rays. To correct these errors, $\langle a_3\rangle$ is estimated and subtracted, adaptively, from $\langle a_1\rangle$. The horizontal grey line in figure 10 shows the depth level z_1 at which d_3 is estimated. The first step in constructing d_3 is to remove the parts of the two data sets in the grey box. This also removes the part of the associated isochron in the grey box. (These isochrons are the solid curves in the figure.) Next the contribution spread over the remainder of the isochrons is combined through a time convolution, adding the contributions from the two single scattering isochrons. This constructs d_3 at the depth z_1. Applying the single scattering inverse to this data set spreads the contribution from this point along the single scattering isochron (dashed curve), giving $\langle a_3\rangle$. This contribution can then be subtracted from $\langle a_1\rangle$.

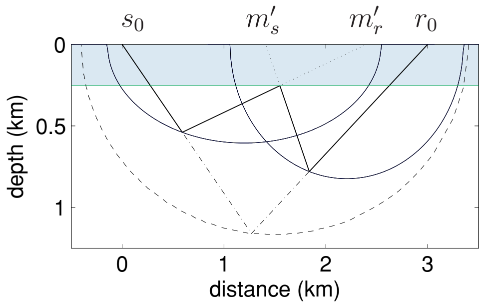

Figure 10. A contribution to $\langle\langle a_3\rangle\rangle$. The solid rays are the triply scattered rays. The dash-dot line is the singly scattered contribution with the same source and receiver positions as well as slopes. The dashed curve is the single scattering isochron, for the time t_4 corresponding to the amount of time required to travel along the triply scattered path. The shaded region extends to the depth level z_1 to which the entire wavefield is propagated before generating the image correction via $\langle\langle a_3\rangle\rangle$.

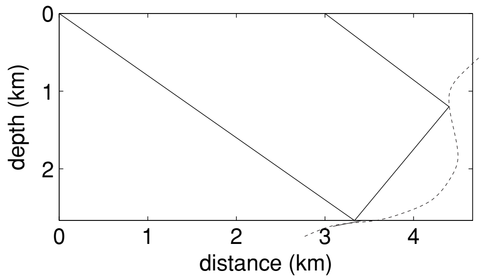

Figure 11. A contribution not accounted for by our theory is shown here; this is a doubly scattered event that would be recorded at the surface. The dashed line illustrates a surface that could generate such a scattering.

9. Discussion

We propose a method for attenuating artefacts in the image generated by leading-order internal multiples. We give two main results: a structure for modelling leading-order internal multiples in (83) and (96), and a system to estimate image artefacts due to leading-order internal multiples in (109). Our suggested algorithm is illustrated by the following flowchart

$$\begin{array}{ccccccc}
 & & d & & \overset{(e)}{} & & \\
 & & (a)\downarrow & & & & \\
a_1(z.,) & \overset{(d)}{\leftarrow} & \tilde{d}_1(z) & \overset{(b)}{\rightarrow} & \tilde{d}_3(z) & \overset{(d)}{\rightarrow} & a_3(z,.) \\
 & & (c)\downarrow & & \downarrow(c) & & \\
 & & \tilde{d}_1(z+\Delta z) & \overset{(+)}{\rightarrow} & \tilde{d}_3(z+\Delta z). & &
\end{array}$$

In (a) the data are downward continued to the depth z, through lemma 8.1. Following this, in (b) leading-order internal multiples are estimated via (96). In (c), both the data and the estimated multiple are propagated to the next depth, again through (92) in lemma 8.1. An image is formed, in (d), at this depth via (109). The image is also used to obtain an estimate of a used in the estimate of d_3 from (96). The theory discussed requires knowledge of the velocity model to the depth z_1 of the up-to-down scatter at which the image is formed. In addition, an adaptive subtraction technique is necessary to compensate for differences in illumination between the singly and triply scattered data. Throughout this paper we have assumed instantaneous point sources. When this assumption is not satisfied knowledge of the source wavelet is necessary because the source appears twice in the estimated first-order internal multiples and only once in the recorded first-order internal multiples. Under the travel-time monotonicity assumption, in the absence of caustics our theory is in correspondence with the velocity model independent theory of Weglein and ten Kroode. In figure 11, a contribution that is not accounted for by our theory is shown. The event is a doubly scattered event, and thus will contribute to a_2, which is not estimated here. Events like this may appear in 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.

Acknowledgments

We would like to thank John W Stockwell for his assistance in the preparation of this manuscript. We are also grateful to the reviewers for their suggestions. This work was supported by Total and the sponsors of the Consortium Project on Seismic Inverse Methods for Complex Structures at the Center for Wave Phenomena and by NSF grant EAR-0417910.

Appendix A. Proof of theorem 6.1

The proof rests on the semi-group property (23), discussed previously. The idea is to use this property to extend the two Green functions in (80) meeting at (z_1, m_s, m_r) to the surface (see figure 12). The resulting operators are then rearranged to pair the G_- operators to substitute the double-square-root Green function, H. We go through this procedure twice, once for $\delta u_{1,-}$ and once for the other elements of (80).

We start by applying the procedure outlined above to $\delta u_{1,-}$, beginning with the semi-group property applied to (76),

$$\begin{aligned}\delta u_{-,1}(z_1, m_r, t_a, 0, s_0) = -\frac{1}{4} D_{t_a}^2 Q_{-,s_0}^*(0) \int_{z_1}^{\infty} \mathrm{d}z_2 \int \mathrm{d}s_2 \int \mathrm{d}r_2 \int_{\mathbb{R}} \mathrm{d}t_0 \int_{\mathbb{R}} \mathrm{d}t' \int \mathrm{d}m_r' \int_{\mathbb{R}} \mathrm{d}t_{m_r'} \\ G_-^*(z_1, m_r, t_{m_r'}, 0, m_r') G_-(0, m_r', t_a + t_{m_r'} - t' - t_0, z_2, r_2) \\ \times G_-(0, s_0, t', z_2, s_2) Q_{-,r_2}(z_2) Q_{-,s_2}(z_2)(E_2 E_1 a)(z_2, s_2, r_2, t_0), \qquad \text{(A.1)}\end{aligned}$$

where $t_a + t_{m_r'}$ is the time required to travel from the source at s_0 to the pseudo-receiver at m_r', as illustrated in figure 13. We now begin to rearrange the terms in preparation for the H substitution.

We interchange the order of integration to

$$\begin{aligned}\delta u_{-,1}(z_1, m_r, t_a, 0, s_0) = -\frac{1}{4} D_{t_a}^2 Q_{-,s_0}^*(0) \int \mathrm{d}m_r' \int_{\mathbb{R}} \mathrm{d}t_{m_r'}\, G_-^*(z_1, m_r, t_{m_r'}, 0, m_r') \\ \times \int_{z_1}^{\infty} \mathrm{d}z_2 \int \mathrm{d}s_2 \int \mathrm{d}r_2 \int_{\mathbb{R}} \mathrm{d}t_0\, H(0, s_0, m_r', t_a + t_{m_r'} - t_0, z_2, s_2, r_2) \\ \times Q_{-,r_2}(z_2) Q_{-,s_2}(z_2)(E_2 E_1 a)(z_2, s_2, r_2, t_0). \qquad \text{(A.2)}\end{aligned}$$

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

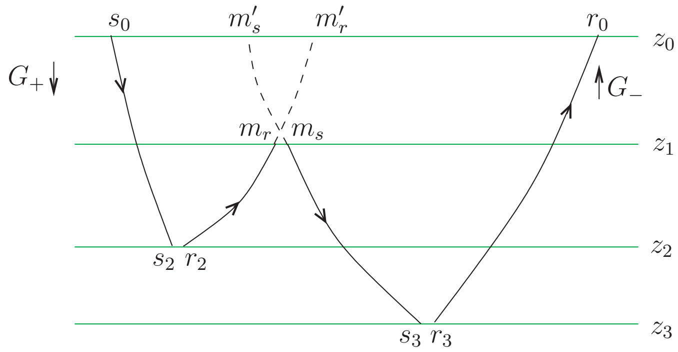

Figure 12. Triple scattering notation and conventions for the extensions via G_-^* operators to propagate to the surface.

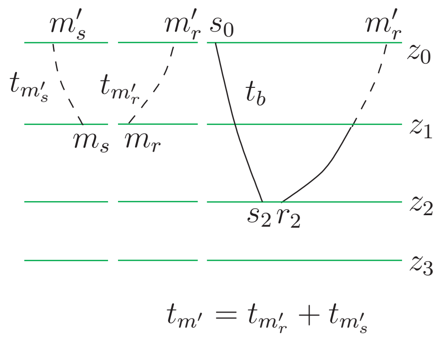

Figure 13. Time variables used in the continuation of the G_- operators to the surface.

Next, we apply the same procedure to the second Green function in (80),

$$\begin{aligned} d_3(s_0, r_0, t_4) = -\frac{1}{4} D_{t_4}^4 \int_0^\infty \mathrm{d}z_3 \int \mathrm{d}s_3 \int \mathrm{d}r_3 \int_{\mathbb{R}} \mathrm{d}t_{30} \int_{\mathbb{R}} \mathrm{d}t_a \int_0^{z_3} \mathrm{d}z_1 \int \mathrm{d}m_s \int \mathrm{d}m_r \int_{\mathbb{R}} \mathrm{d}t_3' \\ Q_{-,r_0}^*(0) G_-(0, r_0, t_4 - t_a - t_3' - t_{30}, z_3, r_3) Q_{-,r_3}(z_3) \\ \times Q_{-,m_s}^*(z_1) \int \mathrm{d}m_s' \int_{\mathbb{R}} \mathrm{d}t_{m_s'} G_-^*\big(z_1, m_s, t_{m_s'}, 0, m_s'\big) G_-(0, m_s', t_3' + t_{m_s'}, z_3, s_3) \\ \times Q_{-,s_3}(z_3)(E_2 E_1 a)(z_3, s_3, r_3, t_{30})(E_1 a)(z_1, m_s, m_r) \\ \times Q_{-,m_r}^*(z_1) \delta u_{-,1}(z_1, m_r, t_a, 0, s_0), \end{aligned} \tag{A.3}$$

where $t_{m_s'}$ is defined by analogy with $t_{m_r'}$ (see figure 13). We now begin to rearrange terms in (A.3) in preparation of the H substitution.

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

$$\begin{aligned} d_3(s_0, r_0, t_4) = -\frac{1}{4} D_{t_4}^4 \int_0^\infty \mathrm{d}z_3 \int \mathrm{d}s_3 \int \mathrm{d}r_3 \int_{\mathbb{R}} \mathrm{d}t_{30} \int_{\mathbb{R}} \mathrm{d}t_a \int_0^{z_3} \mathrm{d}z_1 \int \mathrm{d}m_s \int \mathrm{d}m_r \int_{\mathbb{R}} \mathrm{d}t_a \\ \int \mathrm{d}m_s' \int_{\mathbb{R}} \mathrm{d}t_{m_s'} \int_{\mathbb{R}} \mathrm{d}t_3''\, Q_{-,r_0}^*(0) Q_{-,m_s}^*(z_1) G_-^*\big(z_1, m_s, t_{m_s'}, 0, m_s'\big) \\ \times G_-\big(0, r_0, t_4 - t_a - t_3'' + t_{m_s'} - t_{30}, z_3, r_3\big) Q_{-,r_3}(z_3) G_-(0, m_s', t_3'', z_3, s_3) \end{aligned}$$

$$\times\, Q_{-,s_3}(z_3)(E_2E_1a)(z_3,s_3,r_3,t_{30})(E_1a)(z_1,m_s,m_r)$$
$$\times\, Q^*_{-,m_r}(z_1)\delta u_{-,1}(z_1,m_r,t_a,0,s_0). \tag{A.4}$$

We now substitute H from (74) for the time convolution of the two G_- kernels above, interchanging the order of integration, to obtain

$$\begin{aligned}d_3(s_0,r_0,t_4) = {}& -\frac{1}{4}D^4_{t_4}Q^*_{-,r_0}(0)\int_0^\infty \mathrm{d}z_3\int_0^{z_3}\mathrm{d}z_1\int \mathrm{d}m_s\int \mathrm{d}m_r\int_{\mathbb{R}}\mathrm{d}t_a\int \mathrm{d}m'_s\int_{\mathbb{R}}\mathrm{d}t_{m'_s}\\ & Q^*_{-,m_s}(z_1)G^*_-\big(z_1,m_s,t_{m'_s},0,m'_s\big)\int \mathrm{d}s_3\int \mathrm{d}r_3\int_{\mathbb{R}}\mathrm{d}t_{30}\,H\big(0,m'_s,r_0,t_4\\ & -t_a+t_{m'_s}-t_{30},z_3,s_3,r_3\big)Q_{-,s_3}(z_3)Q_{-,r_3}(z_3)(E_1a)(z_1,m_s,m_r)\\ & \times (E_2E_1a)(z_3,s_3,r_3,t_{30})Q^*_{-,m_r}(z_1)\delta u_{-,1}(z_1,m_r,t_a,0,s_0).\end{aligned} \tag{A.5}$$

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

To do this, we substitute (A.2) into (A.5). We then interchange operators to combine the two G^*_- terms, as well as changing the order of integration to move the t_a integral inside the $t_{m'_s}$ one and also introduce E_2. This results in

$$\begin{aligned}d_3(s_0,r_0,t_4) = {}& \frac{1}{16}D^6_{t_4}Q^*_{-,r_0}(0)Q^*_{-,s_0}(0)\int_0^\infty \mathrm{d}z_3\int_0^{z_3}\mathrm{d}z_1\int_{z_1}^\infty \mathrm{d}z_2\int \mathrm{d}m_s\int \mathrm{d}m_r\int_{\mathbb{R}}\mathrm{d}t_{m_0}\\ & \int \mathrm{d}m'_s\int \mathrm{d}m'_r\int_{\mathbb{R}}\mathrm{d}t_{m'_s}\int_{\mathbb{R}}\mathrm{d}t_{m'_r}\,Q^*_{-,m_s}(z_1)(E_2E_1a)\big(z_1,m_s,m_r,t_{m_0}\big)\\ & \times Q^*_{-,m_r}(z_1)G^*_-\big(z_1,m_s,t_{m'_s},0,m'_s\big)G^*_-\big(z_1,m_r,t_{m'_r}-t_{m_0},0,m'_r\big)\\ & \times \int_{\mathbb{R}}\mathrm{d}t_a\int \mathrm{d}s_3\int \mathrm{d}r_3\int_{\mathbb{R}}\mathrm{d}t_{30}\,H\big(0,m'_s,r_0,t_4-t_a+t_{m'_s}-t_{30},z_3,s_3,r_3\big)\\ & \times Q_{-,s_3}(z_3)Q_{-,r_3}(z_3)(E_2E_1a)(z_3,s_3,r_3,t_{30})\int \mathrm{d}s_2\int \mathrm{d}r_2\int_{\mathbb{R}}\mathrm{d}t_0\\ & H\big(0,s_0,m'_r,t_a+t_{m'_r}-t_{s_0}-t_0,z_2,s_2,r_2\big)Q_{-,r_2}(z_2)Q_{-,s_2}(z_2)\\ & \times (E_2E_1a)(z_2,s_2,r_2,t_0).\end{aligned} \tag{A.6}$$

Interchanging the z_1 and z_3 integrals gives

$$\begin{aligned}d_3(s_0,r_0,t_4) = {}& \frac{1}{4}D^6_{t_4}Q^*_{-,r_0}(0)\,Q^*_{-,s_0}(0)\\ & \times \int_0^\infty \mathrm{d}z_1\int \mathrm{d}m_s\int \mathrm{d}m_r\int_{\mathbb{R}}\mathrm{d}t_{m_0}\int \mathrm{d}m'_s\int \mathrm{d}m'_r\int_{\mathbb{R}}\mathrm{d}t_{m'_s}\int_{\mathbb{R}}\mathrm{d}t_{m'_r}\,Q^*_{-,m_s}(z_1)\\ & \times (E_2E_1a)\big(z_1,m_s,m_r,t_{m_0}\big)Q^*_{-,m_r}(z_1)G^*_-\big(z_1,m_s,t_{m'_s},0,m'_s\big)\\ & \times G^*_-\big(z_1,m_r,t_{m'_r}-t_{m_0},0,m'_r\big)\int_{\mathbb{R}}\mathrm{d}t_a\int_{z_1}^\infty \mathrm{d}z_3\int \mathrm{d}s_3\int \mathrm{d}r_3\int_{\mathbb{R}}\mathrm{d}t_{30}\\ & H\big(0,m'_s,r_0,t_4-t_a+t_{m'_s}-t_{30},z_3,s_3,r_3\big)Q_{-,s_3}(z_3)Q_{-,r_3}(z_3)\\ & \times (E_2E_1a)(z_3,s_3,r_3,t_{30})\int_{z_1}^\infty \mathrm{d}z_2\int \mathrm{d}s_2\int \mathrm{d}r_2\int_{\mathbb{R}}\mathrm{d}t_0\\ & H(0,s_0,m'_r,t_a+t_{m'_r}-t_0,z_2,s_2,r_2)Q_{-,r_2}(z_2)Q_{-,s_2}(z_2)\\ & \times (E_2E_1a)(z_2,s_2,r_2,t_0).\end{aligned} \tag{A.7}$$

Identifying the fictitious data set, $\mathbf{d}_1$, defined in (81) we simplify (A.7) to

$$d_3(s_0, r_0, t_4) = D_{t_4}^2 \int_0^\infty \mathrm{d}z_1 \int \mathrm{d}m_s \int \mathrm{d}m_r \int_{\mathbb{R}} \mathrm{d}t_{m_0} \int \mathrm{d}m_s' \int \mathrm{d}m_r' \int_{\mathbb{R}} \mathrm{d}t_{m_s'} \int_{\mathbb{R}} \mathrm{d}t_{m_r'}$$
$$Q_{-,m_s}^*(z_1)(E_2E_1a)\big(z_1, m_s, m_r, t_{m_0}\big)Q_{-,m_r}^*(z_1)G_-^*\big(z_1, m_s, t_{m_s'}, 0, m_s'\big)$$
$$\times G_-^*\big(z_1, m_r, t_{m_r'} - t_{m_0}, 0, m_r'\big)Q_{-,m_r'}^*(0)^{-1}Q_{-,m_s'}^*(0)^{-1}$$
$$\times \left\{\int_{\mathbb{R}} \mathrm{d}t_a \mathbf{d}_1\big(z_1; m_s', r_0, t_4 - t_a + t_{m_s'}\big)\mathbf{d}_1\big(z_1; s_0, m_r', t_a + t_{m_r'}\big)\right\}. \tag{A.8}$$

In (A.8), the expression in braces is a time convolution of two fictitious data sets. By shifting time variables between the two $\mathbf{d}_1$ fictitious data sets (the time convolution structure is time translation invariant) and changing time variables from t_a to $t_b = t_a + t_{m_r'}$ we arrive at a structure into which the distribution W defined in the theorem statement can be inserted. This W distribution is a new field constituent generated through the convolution of the two data sets on which the two Green functions in (A.8) act. To overlay the distribution W with the expression in braces in (A.8) we need only make the identification $t = t_4 + t_{m_r'} + t_{m_s'}$.

In the definition of W, we identify a new time variable $t_{m'} = t_{m_r'} + t_{m_s'}$ in the above expression for t. To introduce this variable we change variables from $t_{m_r'}$ to $t_{m'}$, substituting the expression for W from (82) into (A.8)

$$d_3(s_0, r_0, t_4) = D_{t_4}^2 \int_0^\infty \mathrm{d}z_1 \int \mathrm{d}m_s \int \mathrm{d}m_r \int_{\mathbb{R}} \mathrm{d}t_{m_0}\, Q_{-,m_s}^*(z_1)\,(E_2E_1a)\big(z_1, m_s, m_r, t_{m_0}\big)$$
$$\times Q_{-,m_r}^*(z_1) \int \mathrm{d}m_s' \int \mathrm{d}m_r' \int_{\mathbb{R}} \mathrm{d}t_{m'} \int_0^{t_m'} \mathrm{d}t_{m_s'}\, G_-^*\big(z_1, m_s, t_{m_s'}, 0, m_s'\big)$$
$$\times G_-^*\big(z_1, m_r, t_{m'} - t_{m_s'} - t_{m_0}, 0, m_r'\big)$$
$$\times Q_{-,m_r'}^*(0)^{-1}Q_{-,m_s'}^*(0)^{-1}W(z_1; s_0, m_r', t_4 + t_{m'}, m_s', r_0). \tag{A.9}$$

The two G_-^* kernels in (A.9) along with the integration in $t_{m_s'}$ are nearly in the form of the kernel of the H operator.

The integration in $t_{m_s'}$ is extended to ∞ as $t_{m_s'} > t_{m'}$ results in a negative time in the second G_-^* making it 0 by the anti-causality of G_-^* (remark 3.1). This allows us to introduce the H operator, which gives the result.

The $\mathbf{d}_1$ data constituents cannot be extracted directly from the data unless ten Kroode's travel-time monotonicity assumption is satisfied. If this assumption is not satisfied one could generate $\mathbf{d}_1$ as $d_1 - \mathbf{D}\langle a\rangle$, where

$$(\mathbf{D}\langle a\rangle)(z_1, s_0, r_0, t) = -\frac{1}{4}D_{t_4}^2 Q_{-,r}^*(0)Q_{-,s}^*(0)\int_0^{z_1} \mathrm{d}z \int \mathrm{d}s \int \mathrm{d}r \int_{\mathbb{R}} \mathrm{d}t_0$$
$$\times H(0, s_0, r_0, t - t_0, z, s, r)Q_{-,r}(z)Q_{-,s}(z)(E_2E_1\langle a\rangle)(z, s, r, t_0), \tag{A.10}$$

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

Appendix B. Comparison with the Weglein/ten Kroode approach

If no caustics form in the background medium, and the travel-time monotonicity of ten Kroode is satisfied, our results can be brought into correspondence with those of Weglein *et al* [42], and ten Kroode [38]. To facilitate this comparison, we will write (83) in terms of the data only.

We begin by recalling from the discussion following theorem 6.1, that the integration in (m_r, m_s, t_m) is an inner product in these variables. We then identify

$Q^*_{-,m_s}(z_1)Q^*_{-,m_r}(z_1)H(0,z_1)$ as an operator acting on $Q^*_{-,m_r'}(0)^{-1}Q^*_{-,m_s'}(0)^{-1}W(z_1;s_0,m_r', t_4+t_m', m_s', r_0)$; this makes up the second entry in the inner product. The first entry in this inner product is $(E_2E_1a)\big(z_1,m_s,m_r,t_{m_0}\big)$. An equivalent form of (83) is then

$$d_3(s_0,r_0,t_4) = D_{t_4}^2\int_0^\infty \mathrm{d}z_1\Bigg(\int \mathrm{d}m_s'\int \mathrm{d}m_r'\int_{\mathbb{R}}\mathrm{d}t_{m'}\Bigg\{\int \mathrm{d}m_s\int \mathrm{d}m_r\int_{\mathbb{R}}\mathrm{d}t_{m_0}\\ H\big(0,m_s',m_r',t_{m'}-t_{m_0},z_1,m_s,m_r\big)Q_{-,m_r}(z_1)Q_{-,m_s}(z_1)\\ \times(E_2E_1a)\big(z_1,m_s,m_r,t_{m_0}\big)\Bigg\}Q^*_{-,m_r'}(0)^{-1}Q^*_{-,m_s'}(0)^{-1}\\ \times W(z_1;s_0,m_r',t_4+t_m',m_s',r_0)\Bigg), \tag{B.1}$$

where $H(0,m_s',m_r',t_{m'}-t_{m_0},z_1,m_s,m_r)Q_{-,m_r}(z_1)Q_{-,m_s}(z_1)$ now acts on (E_2E_1a) and the inner product is in the (m_s',m_r',t_m') variables. We define (for the expression in braces in (B.1))

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

The quantity $\bar{d}_1$ is not one that can be measured directly from the data. To compute $\bar{d}_1$, the expression in (89) must be substituted for a to write it in terms of what can be measured, d.

Using the above definition and the expression for $\mathbf{d}_1$ in (81), we re-write (B.1) as

$$d_3(s_0,r_0,t_4) = -\int_0^\infty \mathrm{d}z_1\int \mathrm{d}m_s'\int \mathrm{d}m_r'\int_{\mathbb{R}}\mathrm{d}t_{m'}\,Q^*_{-,m_s'}(0)^{-1}Q^*_{-,m_r'}(0)^{-1}\bar{d}_1(z_1,m_s',m_r',t_{m'})\\ \times Q^*_{-,m_s'}(0)^{-1}Q^*_{-,m_r'}(0)^{-1}\int_{\mathbb{R}}\mathrm{d}t_b\mathbf{d}_1(z_1;m_s',r_0,t_4+t_{m'}-t_b)\mathbf{d}_1(z_1;s_0,m_r',t_b), \tag{B.3}$$

Although this expression is in terms of three quantities that are directly related to data, we find that we cannot write (B.3) in terms of the actual data because of the z_1 dependence of each of $\bar{d}_1$ and $\mathbf{d}_1$. It is this z_1 dependence that separates our approach from that of Weglein and ten Kroode. In the following remark we summarize how the comparison to their work is made in the absence of caustics, when the travel-time monotonicity assumption introduced by ten Kroode is satisfied. This travel-time monotonicity assumption states that the travel time for a ray leaving a position (z,x) in direction α arrives later than a ray leaving position (z',x') in direction α whenever $z>z'$. In his work, ten Kroode assumes this to hold for all x and α; of course this assumption can be violated.

If the travel-time monotonicity assumption is satisfied, we can replace the z_1 dependence of $\mathbf{d}$ in (B.3) with a time windowing procedure. In this case the z_1 integral in (B.3) can be combined with $\bar{d}_1$ resulting in

$$d_3(s_0,r_0,t_4) \approx -\int \mathrm{d}m_s'\int \mathrm{d}m_r'\int_{\mathbb{R}}\mathrm{d}t_{m'}\,Q^*_{-,m_s'}(0)^{-1}Q^*_{-,m_r'}(0)^{-1}d(m_s',m_r',t_{m'})\\ \times Q^*_{-,m_s'}(0)^{-1}Q^*_{-,m_r'}(0)^{-1}\int_{t_m'}^\infty \mathrm{d}t_b\,d(m_s',r_0,t_4+t_{m'}-t_b)d(s_0,m_r',t_b), \tag{B.4}$$

with the approximation $d\approx d_1$, substituting the definition of W. The time windowing is in the limits of integration.

Remark B.1. To show the correspondence of our method with that discussed in [42, 38], we specifically compare (B.4) in this paper with (120) of [38]. To do this it is first necessary to establish a correspondence between our notation and ten Kroode's notation. To do this we compare figure 12 of this paper with figure 4 of [38]. 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^{IM} 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 (B.4) here.

To make the correspondence between the pseudo-data **d** here and the integration bounds on (117) of ten Kroode we observe that Z_2' of ten Kroode is a time parametrization 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 **d** with the restriction $t_b > t_{m'}$ on the t_b integration. Using this we can replace **d** with d in (B.4), 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 relation (23) used to extend the modelled data from the scattering point at z_1 to the surface.

References

[1] Aminzadeh F and Mendel J M 1980 On the Bremmer series decomposition: equivalence between two different approaches *Geophys. Prospect.* **28** 71–84

[2] Aminzadeh F and Mendel J M 1981 Filter design for suppression of surface multiples in a non-normal incidence seismogram *Geophys. Prospect.* **29** 835–52

[3] Atkinson F V 1960 Wave propagation and the Bremmer series *J. Math. Anal. Appl.* **1** 225–76

[4] Berkhout A J and Verschuur D J 1997 Estimation of multiple scattering by iterative inversion, part I theoretical considerations *Geophysics* **62** 1586–95

[5] Bostock M G, Rondenay S and Shragge J 2001 Multiparameter two-dimensional inversion of scattered teleseismic body waves 1. Theory for oblique incidence *J. Geophys. Res.* **106** 30771–82

[6] Bremmer H 1951 The WKB approximation as the first term of a geometric-optical series *Commun. Pure Appl. Math.* **4** 105–15

[7] Burdick L J and Orcutt J A 1979 A comparison of the generalized ray and reflectivity methods of waveform synthesis *Geophys. J. Int.* **58** 261–78

[8] Cheney M and Borden B 2002 Microlocal structure of inverse synthetic aperture radar data *Inverse Problems* **19** 173–94

[9] Claerbout J F 1985 *Imaging the Earth's Interior* (Oxford: Blackwell)

[10] Corones J P 1975 Bremmer series that correct parabolic approximations *J. Math. Anal. Appl.* **50** 361–72

[11] de Hoop M V 1996 Generalization of the Bremmer coupling series *J. Math. Phys.* **37** 3246–82

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

[13] Dragoset W H and Jericević Z 1998 Some remarks on surface multiple attenuation *Geophysics* **63** 772–89

[14] Fokkema J T and van den Berg P M 1993 *Seismic Applications of Acoustic Reciprocity* (Amsterdam: Elsevier)

[15] Fokkema J T, van Borselen R G and van den Berg P M 1994 *Removal of Inhomogeneous Internal Multiples Proc. Eur. Assn. of Expl. Geophys. Annual Meeting H039*

[16] Friedlander F G and Joshi M 1998 *Introduction to the Theory of Distributions* (Cambridge: Cambridge University Press)

[17] Gray S H 1983 On the convergence of the time domain bremmer series *Wave Motion* **5** 249–55

[18] Guitton A and Verschuur D J 2004 Adaptive subtraction of multiples using the L_1-norm *Geophys. Prospect.* **52** 27–38

[19] Hörmander L 1983 *The Analysis of Linear Partial Differential Operators* vol 1 (Berlin: Springer)

[20] Jakubowicz H 1998 Wave equation prediction and removal of interbed multiples *Soc. Explor. Geophys.* 1527–30 Expanded abstracts

[21] Kelamis P, Erickson K, Burnstad R, Clark R and Verschuur D 2002 Data-driven internal multiple attenuation—applications and issues on land data *Soc. Expl. Geophys.* Expanded Abstracts 2035–8

[22] Kennett B L N 1974 Reflections, rays and reverberations *Bull. Seismol. Soc. Am.* **64** 1685–96
[23] Kennett B L N 1979 The suppression of surface multiples on seismic records *Geophys. Prospect.* **27** 584–600
[24] Kennett B L N 1979 Theoretical reflection seismograms for elastic media *Geophys. Prospect.* **27** 301–21
[25] Kennett B L N 1983 *Seismic Wave Propagation in Stratified Media* (Cambridge: Cambridge University Press)
[26] Lippmann B A 1956 Rearrangement collisions *Phys. Rev.* **102** 264–8
[27] McMaken H 1986 On the convergence of the Bremmer series for the Helmholtz equation in 2d *Wave Motion* **8** 277–83
[28] Moses H E 1956 Calculation of scattering potential from reflection coefficients *Phys. Rev.* **102** 559–67
[29] Prosser R T 1969 Formal solutions of inverse scattering problems *J. Math. Phys.* **10** 1819–22
[30] Razavy M 1975 Determination of the wave velocity in an inhomogeneous medium from reflection data *J. Acoust. Soc. Am.* **58** 956–63
[31] Revenaugh J and Jordan T H 1987 Observations of first-order mantle reverberations *Bull. Seismol. Soc. Am.* **77** 1704–17
[32] Revenaugh J and Jordan T H 1989 A study of mantle layering beneath the western pacific *J. Geophys. Res.* **94** 5787–813
[33] Revenaugh J and Jordan T H 1991 Mantle layering from ScS reverberations 1. Waveform inversion of zeroth-order reverberations *J. Geophys. Res.* **96** 749–19, 762
[34] Sjöstrand J and Grigis A 1994 *Microlocal Analysis for Differential Operators: An Introduction* (Cambridge: Cambridge University Press)
[35] Stolk C C and de Hoop M V 2002 Modeling and inversion of seismic data in anisotropic elastic media *Commun. Pure Appl. Math.* **55** 261–301
[36] Stolk C C and de Hoop M V 2005 Modeling of seismic data in the downward continuation approach *SIAM J. Appl. Math.* CWP468P at press
[37] Stolk C C and de Hoop M V 2005 Seismic inverse scattering in the downward continuation approach CWP469P submitted
[38] ten Kroode A P E 2002 Prediction of internal multiples *Wave Motion* **35** 315–38
[39] van Borselen R 2002 Data-driven interbed multiple removal: strategies and examples *Expanded Abstracts Society of Exploration Geophysicists*
[40] Verschuur D J and Berkhout A 1997 Estimation of multiple scattering by iterative inversion, part II. Practical aspects and examples *Geophysics* **62** 1596–611
[41] 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 and Zhang H 2003 Inverse scattering series and seismic exploration *Inverse Problems* **19** R27–83
[42] Weglein A, Gasparotto F A, Carvalho P M and Stolt R H 1997 An inverse-scattering series method for attenuating multiples in seismic reflection data *Geophysics* **62** 1975–89
[43] Yoshida K 1995 *Functional Analysis* (Berlin: Springer)