SIAM J. APPL. MATH.
Vol. 65, No. 4, pp. 1388–1406

MODELING OF SEISMIC DATA IN THE DOWNWARD CONTINUATION APPROACH*

CHRISTIAAN C. STOLK† AND MAARTEN V. DE HOOP‡

Abstract. Seismic data are commonly modeled by a high-frequency single scattering approximation. This amounts to a linearization in the medium coefficient about a smooth background. The discontinuities are contained in the medium perturbation. The high-frequency part of the wavefield in the background medium is described by a geometrical optics representation. It can also be described by a one-way wave equation. Based on this we derive a downward continuation operator for seismic data. This operator solves a pseudodifferential evolution equation in depth, the so-called double-square-root equation. We consider the modeling operator based on this equation. If the rays in the background that are associated with the reflections due to the perturbation are nowhere horizontal, the singular part of the data is described by the solution to an inhomogeneous double-square-root equation.

Key words. seismic modeling, microlocal analysis, double-square-root equation

AMS subject classifications. 86A15, 35R30

DOI. 10.1137/S0036139904439545

1. Introduction. In reflection seismology one places point sources and point receivers on the earth's surface. The source generates acoustic waves in the subsurface, which are reflected where the medium properties vary discontinuously. In seismic imaging, one tries to reconstruct the properties of the subsurface from the reflected waves that are observed. There are various approaches to seismic imaging, each based on a different mathematical model for seismic reflection data with underlying assumptions. In general, seismic scattering and inverse scattering have been formulated in the form of a linearized inverse problem for the medium coefficient in the acoustic wave equation. The linearization is around a smoothly varying background, called the velocity model, which is a priori also unknown.

In this paper and a companion paper [24] we study a method of seismic imaging introduced by Clayton [6] and Claerbout [5]. The key concept in this method is the construction of data of fictitious experiments carried out in the subsurface, at increasing depths, from data observed at the earth's surface. These so-called downward continued data are then used for imaging the medium contrast as well as for a reflection tomographic procedure to estimate the smoothly varying background (known as migration velocity analysis). The downward continuation approach to seismic imaging has received much attention in the geophysical research literature, and it is currently widely used in practice in various approximations [3, 19, 16].

The downward continuation of data is derived from the factorization of the wave equation into two one-way wave equations. This factorization is closely connected to the notion of wave splitting [28]. One-way wave equations, in various approximations,

*Received by the editors January 7, 2004; accepted for publication (in revised form) October 8, 2004; published electronically April 26, 2005. This work was supported in part by the Mathematical Sciences Research Institute through National Science Foundation grant DMS-9810361.
http://www.siam.org/journals/siap/65-4/43954.html

†Department of Applied Mathematics, University of Twente, Drienerlolaan 5, 7522 NB Enschede, The Netherlands (c.c.stolk@ewi.utwente.nl).
‡Center for Wave Phenomena, Colorado School of Mines, Golden, CO 80401 (mdehoop@Mines.EDU).

have been extensively used in applications other than seismics: for integrated optics (see, e.g., [12]) and for underwater acoustics (see, e.g., [25, 7]).

There are basically two categories of seismic imaging methods. One category is associated with the evolution of waves and data in time; the other is associated with the evolution in depth (or another principal spatial direction). The first category contains approaches known under the collective names of Kirchhoff migration [4] or generalized Radon transform inversion, and reverse-time migration [21]; the second category comprises the downward continuation approach. There are great computational advantages of the downward continuation approach to seismic imaging over the Kirchhoff approaches. There are fundamental, theoretical advantages as well, in particular with a view to the problem of estimating the smoothly varying background. These are analyzed in a separate paper [24]. For the Kirchhoff approach to seismic imaging there is a solid mathematical theory, which treats seismic imaging as an inverse problem and shows that singularities can be reconstructed [2, 20]. For the downward continuation approach much research has gone into the development of numerical one-way wave equations, but little is known from an analysis point of view. For a constant coefficient background, the downward continuation method was cast into an inverse problem in [1]. For the case of variable coefficients, which of course is the case of interest in practice, there has been no such theory.

The purpose of this paper is to develop a mathematical theory for modeling seismic reflection data in the downward continuation approach. As was done in the analysis of Kirchhoff methods, we make use of techniques and concepts from microlocal analysis, such as wave front set, denoted by $\mathrm{WF}(\cdot)$, and Fourier integral operators; see, e.g., [10] for background information on these concepts. We introduce the main concepts and operators involved in the method. We then study the double-square-root modeling operator. This modeling operator and its properties will be the point of departure for the development of an inverse scattering theory [24].

In our notation we will distinguish the vertical coordinate $z \in \mathbb{R}$ from the horizontal coordinates $x \in \mathbb{R}^{n-1}$ and write $(z, x) \in \mathbb{R}^n$. In these coordinates the scalar acoustic wave equation with wave speed function $c_0(z, x)$ is given by

$$Pu = f, \qquad P = c_0(z,x)^{-2}\partial_t^2 - \partial_z^2 - \sum_{j=1}^{n-1} \partial_{x_j}^2, \tag{1.1}$$

where $u = u(z, x, t)$ is the acoustic pressure. The equation is considered for t in a time interval $]0, T[$, together with an initial condition $u(\cdot, \cdot, 0) = 0$. The solution to (1.1) can be written as

$$u(z,x,t) = \int_0^t \int G(z,x,t-t_0,z_0,x_0) f(z_0,x_0,t_0)\, \mathrm{d}z_0 \mathrm{d}x_0 \mathrm{d}t_0, \tag{1.2}$$

where G is the Green's function of (1.1). The source f can be a distribution.

To model the scattering of waves, we adopt the linearized scattering or Born approximation. The linearization is in the wavespeed, around a smooth (C^∞) background c_0; for the full wavespeed function we write $c = c_0 + \delta c$. The perturbation δc may contain singularities. The perturbation in G at the acquisition surface $z = 0$ is given by (see, e.g., [2])

$$\begin{aligned}\delta G(0,r,t,0,s) = \int_{\mathbb{R}_+ \times \mathbb{R}^{n-1}} \int_0^t & G(0,r,t-t_0,z_0,x_0)\, 2c_0^{-3}(z_0,x_0)\delta c(z_0,x_0) \\ & \times \partial_{t_0}^2 G(z_0,x_0,t_0,0,s)\, \mathrm{d}t_0 \mathrm{d}z_0 \mathrm{d}x_0,\end{aligned} \tag{1.3}$$

where both $s, r \in \mathbb{R}^{n-1}$. We assume that the acquisition manifold Y, which contains the set of values of (s, r, t) used in the acquisition, is a bounded open subset of $\mathbb{R}^{2n-2} \times \mathbb{R}_+$. The modeled data are then a function of $(s, r, t) \in Y$ given by (1.3). We define the Born modeling map F through (1.3) as the map from δc to δG evaluated at $z = 0$. Since Y is bounded and the waves propagate with finite speed we may assume that δc is supported in a bounded open subset X of $\mathbb{R}_+ \times \mathbb{R}^{n-1}$. Furthermore, we assume that $\overline{X} \cap \{z = 0\} = \emptyset$. Naturally, (1.3) is, in general, not a complete model for raw data measured in seismic experiments. It models data that are the input for imaging and inversion and have undergone some processing.

We summarize some results in the literature about the modeling map, F. The solution operator (1.2) is such that singularities in the solution propagate along bicharacteristics. Denote by $p(z, x, \zeta, \xi, \tau) = -c(z, x)^{-2}\tau^2 + \zeta^2 + \|\xi\|^2$ the principal symbol of P. Propagating singularities are in the characteristic set, given by the points $(z, x, t, \zeta, \xi, \tau) \in T^*\mathbb{R}^{n+1}$ with

$$p(z, x, \zeta, \xi, \tau) = -c(z, x)^{-2}\tau^2 + \zeta^2 + \|\xi\|^2 = 0. \tag{1.4}$$

The bicharacteristics are the solution curves of a Hamilton system with Hamiltonian given by p,

$$\frac{\mathrm{d}(z, x, t)}{\mathrm{d}\lambda} = \frac{\partial p}{\partial(\zeta, \xi, \tau)}, \quad \frac{\mathrm{d}(\zeta, \xi, \tau)}{\mathrm{d}\lambda} = -\frac{\partial p}{\partial(z, x, t)}. \tag{1.5}$$

Assuming that $\tau \neq 0$, the time t is strictly increasing or decreasing with λ and can be used as parameter for the solution curve. To parameterize points on the solution curves, we use the initial position (z_0, x_0), the take-off direction $\alpha \in S^{n-1}$, the frequency τ, which together define the initial cotangent vector $(\zeta_0, \xi_0) = -\tau c(z_0, x_0)^{-1}\alpha$, and the time t (instead of λ). Points on the solution curves will be denoted by

$$\begin{aligned} \eta(t, z_0, x_0, \alpha, \tau) = (&\eta_z(t, z_0, x_0, \alpha, \tau), \eta_x(t, z_0, x_0, \alpha, \tau), t, \\ &\eta_\zeta(t, z_0, x_0, \alpha, \tau), \eta_\xi(t, z_0, x_0, \alpha, \tau), \tau). \end{aligned} \tag{1.6}$$

The variable τ is invariant along the Hamilton flow. We take $t = 0$ as the initial value for t (note that (1.5) are time translation invariant).

To ensure that δG defines a continuous map from $\mathcal{E}'(X)$ to $\mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n \times]0, T[)$ and that the restriction of δG to Y is a Fourier integral operator we make the following assumption on c_0.

Assumption 1. *There are no rays from $(0, s)$ to $(0, r)$ with travel time t such that $(s, r, t) \in Y$. For all ray pairs connecting $(0, r)$ via some $(z, x) \in X$ to $(0, s)$ with total time t such that $(s, r, t) \in Y$, the rays intersect the plane $z = 0$ transversally at r and s.*

We also assume that rays from such a point $(z, x) \in X$ intersect the surface $z = 0$ only once, because all reflections must come from the region $z > 0$ (the subsurface). The first part of the assumption excludes direct rays, or a pair of incident and reflected rays with scattering angle π. The second part of the assumption excludes rays grazing the plane $z = 0$. Concerning the second part, strictly only caustics grazing the plane $z = 0$ have to be excluded. In practice the wave speed near the surface is much lower than in the interior of the earth, and waves from the interior arrive under small angles with the vertical. So from a geophysical point of view one is only interested in incoming rays that intersect the measurement surface transversally. We have the following theorem. (See [10] for a general reference on Fourier integral operators.)

THEOREM 1.1 (see [20, 17]). *With Assumption* 1 *the map* F *is a Fourier integral operator* $\mathcal{E}'(X) \to \mathcal{D}'(Y)$ *of order* $(n-1)/4$ *with canonical relation*

(1.7)
$$\begin{aligned}&\big\{(\eta_x(t_\mathrm{s},z,x,\beta,\tau),\eta_x(t_\mathrm{r},z,x,\alpha,\tau),t_\mathrm{s}+t_\mathrm{r},\eta_\xi(t_\mathrm{s},z,x,\beta,\tau),\eta_\xi(t_\mathrm{r},z,x,\alpha,\tau),\tau;z,x,\zeta,\xi)\,|\\ &\quad t_\mathrm{s},t_\mathrm{r}>0,\ \eta_z(t_\mathrm{s},z,x,\beta,\tau)=\eta_z(t_\mathrm{r},z,x,\alpha,\tau)=0,\ (\zeta,\xi)=-\tau c_0(z,x)^{-1}(\alpha+\beta),\\ &\qquad (z,x,\alpha,\beta,\tau)\in \text{ subset of } X\times(S^{n-1})^2\times\mathbb{R}\backslash 0\big\}\subset T^*\mathbb{R}^{2n-1}_{(s,r,t)}\times T^*\mathbb{R}^n_{(z,x)}.\end{aligned}$$

In this paper, we express F in terms of a depth-continuation operator, and we study the properties of this operator. The main contributions of this paper are the following:

(i) We define an *upward continuation operator* $H(z, z_0)$ using the solution operators to one-way wave equations. Its adjoint will be the downward continuation operator. Intuitively this operator maps data from a fictitious experiment carried out at depth z_0 to data from an experiment carried out at depth z, $z < z_0$. Subject to Assumption 2 in the main text—stating, essentially, that the rays in the background that are associated with the reflections are nowhere tangent to horizontal—we prove that the data $F\delta c$ are given by $\int_0^\infty(\dots)H(0,z)(\dots)g(z,\cdot,\cdot,\cdot)\,\mathrm{d}z$, where the dots are pseudodifferential factors specified below and $g = g(z,s,r,t)$ is given by mapping $c_0^{-3}\delta c$ to a function $E_2E_1(c_0^{-3}\delta c)$ of (z,s,r,t) using the maps

(1.8)
$$\begin{aligned}&E_1:\mathcal{D}'(\mathbb{R}^n)\to\mathcal{D}'(\mathbb{R}^{2n-1})\ :\ (c_0^{-3}\delta c)(z,x)\mapsto h(z,\bar{x},x)=\delta(x-\bar{x})(c_0^{-3}\delta c)(z,\tfrac{\bar{x}+x}{2}),\\ &E_2:\mathcal{D}'(\mathbb{R}^{2n-1})\to\mathcal{D}'(\mathbb{R}^{2n})\ :\ h(z,\bar{x},x)\mapsto\delta(t)h(z,\bar{x},x)\end{aligned}$$

(Theorem 5.1).

(ii) We show that the operator $H(z, z_0)$ solves the initial value problem for a first-order pseudodifferential evolution equation in depth, known as the *double-square-root (DSR) equation*. The data can be identified with the solution to an inhomogeneous DSR equation, with inhomogeneous term g (section 3). The computation of the map from g to data and the computation of its adjoint can be done by marching in depth using the DSR equation. This is the basis of DSR modeling and imaging methods in geophysics.

(iii) The modeling operator can be written as the composition of a Fourier integral operator representing *depth-to-time conversion*, with a locally invertible canonical relation (Theorem 4.2) and the operator E_1.

It should be mentioned that our Assumption 2 can be quite restrictive. However, the limited aperture of seismic acquisition yields a natural cutoff so that, in general, a large part of the *observed* data can be modeled with the approach presented in this paper.

In general, the downward continuation approach results in a more complete computation of the wave propagation and diffraction in the modeling of seismic reflection data than the one based on the geometrical optics approximation underlying the Kirchhoff approach. Fast algorithms have been designed to solve the DSR equation; as compared with numerical algorithms solving the full wave equation, the advantage of using the DSR equation becomes significant in space dimension 3 (and higher).

The outline of the paper is as follows. In section 2 we discuss one-way acoustic equations. In section 3 we use these to define the upward/downward continuation operator H, and we describe some of its properties. Section 4 contains our result on depth-to-time conversion. In section 5 we show that the data can be modeled using the downward continuation method. The last section is about the relation between our assumption and the Bolker condition that occurs in the inversion.

2. Directional decomposition, single-square-root equations. Singularities of solutions to the wave equation, which propagate with velocity with nonzero vertical (z) component, are described by a first-order pseudodifferential evolution equation in z. This follows from a well-known factorization argument; see, e.g., [26]. In [22] the approximation of solutions to the wave equation by solutions to an evolution equation in z is discussed. Such an equation is called a one-way wave equation or single-square-root (SSR) equation. We summarize the structure and properties of this one-way wave equation that we need for the upward/downward continuation approach to seismic data processing.

To determine whether the velocity vector at some point of a ray (cf. (1.5)) is close to horizontal, we use the angle with the vertical, defined to be in $[0, \pi/2]$ and given by $\tan(\theta) = \frac{\|\xi\|}{|\zeta|}$. We recall that the propagating singularities are microlocally in the characteristic set given by (1.4). Given a point (z, x, ξ, τ) with $\|\xi\| < c(z, x)^{-1}|\tau|$, there are two solutions ζ to (1.4), given by $\zeta = \pm b$, where $b = b(z, x, \xi, \tau)$ is defined by

$$b(z, x, \xi, \tau) = -\tau\sqrt{c(z, x)^{-2} - \tau^{-2}\xi^2}. \tag{2.1}$$

The sign is chosen such that $\zeta = \pm b$ corresponds to propagation with $\pm\frac{\mathrm{d}z}{\mathrm{d}t} > 0$. There is also an angle associated with (z, x, ξ, τ) given by the solution $\theta \in [0, \pi/2]$ of the equation

$$\sin(\theta) = c(z, x)\|\tau^{-1}\xi\|. \tag{2.2}$$

When this angle is smaller than $\pi/2$ along a ray segment, then the vertical velocity $\frac{\mathrm{d}z}{\mathrm{d}t}$ does not change sign, and the ray segment can be parameterized by z. The maximal z-interval such that $\arcsin(c(z, x)\|\tau^{-1}\xi\|) < \theta$ for given θ along the bicharacteristic determined by the initial values $(z, x, \pm b, \xi, \tau)$ will be denoted by

$$]z_{\min,\pm}, z_{\max,\pm}[=]z_{\min,\pm}(z, x, \xi, \tau, \theta), z_{\max,\pm}(z, x, \xi, \tau, \theta)[; \tag{2.3}$$

see also Figure 1. Furthermore, we define a set

$$I_\theta = \{(z, x, t, \zeta, \xi, \tau) \mid \arcsin(c(z, x)\|\tau^{-1}\xi\|) < \theta, |\zeta| < C|\tau|\}, \tag{2.4}$$

where C is some constant that is everywhere larger than $c(z, x)^{-1}$.

The SSR equation. To obtain a one-way wave equation, the wave equation is written as the following first-order system in z:

$$\frac{\partial}{\partial z}\begin{pmatrix} u \\ \frac{\partial u}{\partial z} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -A(z, x, D_x, D_t) & 0 \end{pmatrix}\begin{pmatrix} u \\ \frac{\partial u}{\partial z} \end{pmatrix} + \begin{pmatrix} 0 \\ f \end{pmatrix}, \tag{2.5}$$

where $D_x = -\mathrm{i}\frac{\partial}{\partial x}, D_z = -\mathrm{i}\frac{\partial}{\partial z}$, and $A(z, x, D_x, D_t) = c_0(z, x)^{-2}D_t^2 - D_x^2$. Then the system is transformed by using a family of matrix pseudodifferential operators $\mathsf{Q}(z) = \mathsf{Q}(z, x, D_x, D_t)$ with

$$\begin{pmatrix} u_+ \\ u_- \end{pmatrix} = \mathsf{Q}(z)\begin{pmatrix} u \\ \frac{\partial u}{\partial z} \end{pmatrix}, \qquad \begin{pmatrix} f_+ \\ f_- \end{pmatrix} = \mathsf{Q}(z)\begin{pmatrix} 0 \\ f \end{pmatrix}. \tag{2.6}$$

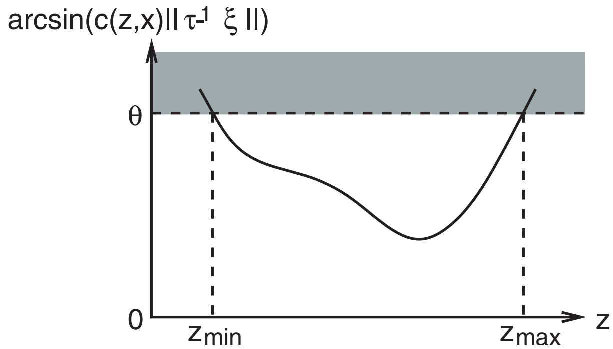

FIG. 1. *Definition of* $z_{\mathrm{min},\pm}$ *and* $z_{\mathrm{max},\pm}$*, which give the maximal interval where* $\arcsin(c(z,x)\|\tau^{-1}\xi\|)$ *is in the interval* $[0,\theta]$*. Here,* (z,x,ξ,τ) *lies on a bicharacteristic.*

The functions (u_+,u_-) satisfy a pseudodifferential system of equations. Let $\theta_2<\pi/2$ be a given angle. (In the next subsection we need another angle, θ_1, with $0<\theta_1<\theta_2<\pi/2$, hence the subscript 2.) With suitably chosen Q it is shown in [22] that the system that results from applying the transformation (2.6) to (2.5) is diagonal on I_{θ_2}. It then follows that (2.5) is equivalent to two equations of the form

$$\left(\frac{\partial}{\partial z}-\mathrm{i}B_\pm(z,x,D_x,D_t)\right)u_\pm=f_\pm, \tag{2.7}$$

microlocally on I_{θ_2}. These are called the one-way wave or SSR equations. The principal part of $B_\pm$ is equal to $\pm b$, while its subprincipal part depends on the normalization of $\mathsf{Q}(z)$. We choose the normalization such that $B_\pm$ are self-adjoint and Q satisfies

$$\begin{aligned}\mathsf{Q}(z,x,\xi,\tau)&=\tfrac12\begin{pmatrix}a^{1/4} & -\mathrm{i}\,\mathrm{sgn}(\tau)a^{-1/4}\\ a^{1/4} & \mathrm{i}\,\mathrm{sgn}(\tau)a^{-1/4}\end{pmatrix}+\text{order}\begin{pmatrix}-\frac12 & -\frac32\\ -\frac12 & -\frac32\end{pmatrix},\\ \mathsf{Q}(z,x,\xi,\tau)^{-1}&=\begin{pmatrix}a^{-1/4} & a^{-1/4}\\ \mathrm{i}\,\mathrm{sgn}(\tau)a^{1/4} & -\mathrm{i}\,\mathrm{sgn}(\tau)a^{1/4}\end{pmatrix}+\text{order}\begin{pmatrix}-\frac32 & -\frac32\\ -\frac12 & -\frac12\end{pmatrix}\end{aligned} \tag{2.8}$$

with $a=a(z,x,\xi,\tau)=c_0(z,x)^{-2}\tau^2-\xi^2$.

It appears that only two components of $\mathsf{Q}(z)$ and $\mathsf{Q}(z)^{-1}$ are needed in the analysis. To clarify this, we first observe that multiplication by $\mathrm{i}\,\mathrm{sgn}(\tau)$ in the frequency domain corresponds to the application of a Hilbert transform with respect to the time variable, which we denote by $\mathcal{H}$. Next, we use the relation between $\mathsf{Q}(z)^*$ and $\mathsf{Q}(z)^{-1}$,

$$\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\mathsf{Q}(z)^{-1*}=2\mathsf{Q}(z)\begin{pmatrix}0 & -\mathcal{H}\\ \mathcal{H} & 0\end{pmatrix},$$

shown to hold microlocally in [22, (59)]. (This relation also appears in [8, (II.49)].) We let

$$Q_+=Q_+(z,x,D_x,D_t)=2\mathsf{Q}_{1,2}\mathcal{H},\ \ Q_-=Q_-(z,x,D_x,D_t)=-2\mathsf{Q}_{2,2}\mathcal{H},$$

where we choose a convenient normalization such that both $Q_\pm$ have principal symbol $a^{1/4}$. It follows with these definitions that

$$u=Q_+^*u_++Q_-^*u_-, \tag{2.9}$$

$$f_\pm=\mp\tfrac12\mathcal{H}Q_\pm f. \tag{2.10}$$

The above procedure does not prescribe the symbol of the operator B_- for $\arcsin(c(z,x)\|\tau^{-1}\xi\|) > \theta_2$. We will assume that B_- is a first-order family of pseudodifferential operators with real homogeneous principal symbol. This implies that the evolution problem (2.7) has well-defined solutions satisfying energy estimates.

Propagation of singularities and introduction of a microlocal cutoff. Here, we discuss how the wave field is approximated by solutions of (2.7). This approximation is valid microlocally on part of the cotangent bundle $T^*\mathbb{R}^{n+1}_{(z,x,t)}$. We consider the approximation of upward traveling waves using the equation for u_-, where we assume that there are only upward traveling singularities at depth z_0, hence $u_+(z_0,\cdot) \in C^\infty$. The treatment of downward traveling waves using the equation for u_+ is analogous.

Consider the initial value problem for $P_{0,-} \stackrel{\text{def}}{=} \partial_z - \mathrm{i}B_-$,

$$P_{0,-}u_- = 0, \quad z < z_0, \quad Q_-^* u_-(z_0,\cdot) = u(z_0,\cdot). \tag{2.11}$$

Let $J_-(z_0,\theta)$ be defined by

$$J_-(z_0,\theta) = \{(z,x,t,\zeta,\xi,\tau) \in I_\theta \,|\, \tau^{-1}\zeta > 0 \text{ and } z_{\max,-}(z,x,\xi,\tau,\theta) \geq z_0\}. \tag{2.12}$$

The solutions to (2.11) agree with the solutions to the original wave equation microlocally on the set $J_-(z_0,\theta_2)$ in the following way. Suppose that $\mathrm{WF}(u) \cap \{z = z_0, \tau^{-1}\zeta < 0\} = \emptyset$ (i.e., at depth z_0 all singularities are propagating in the $-$ direction), and let u_- be a solution to (2.11); then it follows from the propagation of regularity/propagation of singularities result that

$$u \equiv Q_-^* u_- \tag{2.13}$$

microlocally on the set $J_-(z_0,\theta_2)$ [22]. Here, we say that $u \equiv v$ microlocally on a set $\Gamma \subset T^*\mathbb{R}^n$ if $\mathrm{WF}(u-v) \cap \Gamma = \emptyset$.

The solutions to (2.11) have propagating singularities, also in the part of the phase space where $\arcsin(c(z,x)\|\tau^{-1}\xi\|) \geq \theta_2$, but there the singularities of the solution are in general incorrect in the sense that they do not correspond to solutions of the original wave equation. For such singularities we introduce a pseudodifferential cutoff. Let θ_1 be given with $0 < \theta_1 < \theta_2$. We assume we have a pseudodifferential cutoff $\psi_1 = \psi_1(z,z_0,x,D_x,D_t)$ with symbol satisfying

$$\psi_1(z,z_0,x,\xi,\tau) \sim 1 \text{ on } J_-(z_0,\theta_1), \tag{2.14}$$

$$\psi_1(z,z_0,x,\xi,\tau) \in S^\infty \text{ outside } J_-(z_0,\theta_2), \text{ if } z - z_0 > \delta > 0. \tag{2.15}$$

Then we have

$$\psi_1 u \equiv \psi_1 Q_-^* u_-. \tag{2.16}$$

We reformulate this result in terms of the solution operators, the propagators. By $G_{0,-}(z,z_0)$ we will denote the solution operator to the evolution problem (2.11), defined to map $u_-(z_0,\cdot)$ to $u_-(z,\cdot)$. We assume that the full one-way propagator is then given by

$$G_-(z,z_0) = \psi_1(z,z_0)G_{0,-}(z,z_0). \tag{2.17}$$

Here, we let $z < z_0$. This can also be written as a pseudodifferential cutoff applied prior to $G_{0,-}$. We denote this different cutoff also by ψ_1 but with the order of z, z_0 interchanged, so that

$$G_-(z, z_0) = G_{0,-}(z, z_0)\psi_1(z_0, z). \tag{2.18}$$

In this paper this is all we need to know about the pseudodifferential cutoff ψ_1. But it raises the question of an explicit recipe for computing ψ_1: Can it, for example, be computed with a modified evolution equation in depth? This is indeed the case. It was established in [22, 23] that such a pseudodifferential cutoff can be generated by adding a dissipative term to $P_{0,-}$. Instead of $P_{0,-}$ one considers the operator

$$P_- = \partial_z - \mathrm{i}B_\pm(z, x, D_x, D_t) - C(z, x, D_x, D_t) \tag{2.19}$$

with C a first-order pseudodifferential operator with homogeneous, nonnegative real principal symbol, satisfying certain conditions. The operator $\psi_1(z, z_0)$ is then a (z, z_0)-family of pseudodifferential operators with symbol in $S^0_{\rho,1-\rho}(\mathbb{R}^n \times \mathbb{R}^n)$, such that the derivatives $\frac{\partial^{j+k}\psi_1}{\partial z_0^j \partial_z^k}$ are in $S^{(j+k)(1-\rho)}_{\rho,1-\rho}(\mathbb{R}^n \times \mathbb{R}^n)$ for $z \neq z_0$, where ρ can be any number satisfying $\frac{1}{2} < \rho < 1$ (see [23]). For the theory of such operators, see, e.g., [27, 14].

Let the elements $(z, x, t, \zeta, \xi, \tau)$ of the wave front set of f be such that $\tau^{-1}\zeta > 0$ (corresponding to propagation direction $\frac{\partial z}{\partial t} < 0$). Consider u_- defined by

$$u_-(z, \cdot) = \int_z^\infty G_-(z, z_0)\left(\frac{1}{2}\mathcal{H}Q_-(z_0)\right) f(z_0, \cdot)\,\mathrm{d}z_0, \tag{2.20}$$

assuming also that $f = 0$ on a neighborhood of the plane given by z. We have that $Q_-^* u_-(z, \cdot) \equiv u(z, \cdot)$, where u is the solution to (1.1) with f replaced by $(\psi_1(z_0, z) - Q_-^{-1}[Q_-, \psi_1(z_0, z)])f$. Here the square brackets denote a commutator.

We use the notation $\gamma(z, z_0, x_0, t_0, \xi_0, \tau)$ for the bicharacteristic of $P_{0,-}$ parameterized by z. In components we write them as (note that they are time translation invariant)

$$\begin{aligned}\gamma(z, z_0, x_0, t_0, \xi_0, \tau) = (z, \gamma_x(z, z_0, x_0, \xi_0, \tau), \gamma_t(z, z_0, x_0, \xi_0, \tau) + t_0,\\ - b(z, \gamma_x, \gamma_\xi, \tau), \gamma_\xi(z, z_0, x_0, \xi_0, \tau), \tau).\end{aligned} \tag{2.21}$$

Properties of G_-. The operator $G_-(z, z_0)$ is a Fourier integral operator with canonical relation

$$\{(\gamma_x, t_0 + \gamma_t, \gamma_\xi, \tau; x_0, t_0, \xi_0, \tau)\} \subset T^*\mathbb{R}^n \times T^*\mathbb{R}^n, \tag{2.22}$$

where $\gamma_x = \gamma_x(z, z_0, x_0, \xi_0, \tau)$ and the same for γ_t, γ_ξ as in (2.21).

The operators $B_\pm$ are self-adjoint. It follows that $G_{0,-}(z, z_0)$ is unitary. But then

$$G_-(z, z_0)^* G_-(z, z_0) = \psi_1(z_0, z)^* \psi_1(z_0, z), \tag{2.23}$$

and $G_-(z, z_0)^* G_-(z, z_0)$ is one microlocally where $\psi_1(z_0, z)$ is one.

Numerical methods for one-way wave propagation are described, e.g., in [9] and [13] and in the references given in those papers.

3. Downward/upward continuation and the DSR equation. In this section we construct the data downward/upward continuation operator, and we establish some of its properties.

Data model. In preparation of the downward/upward continuation approach to seismic data modeling, we rewrite (1.3) in the form

$$
\begin{aligned}
\delta G(0,r,t,0,s) = \int_{\mathbb{R}^{n-1}\times\mathbb{R}_+} \int_{\mathbb{R}^{n-1}} \int_{-\infty}^{t} \int_{\mathbb{R}_+} & G(0,r,t-t_0,z,x) \\
& \times 2\partial_{t_0}^2 R(z,x,\bar{x},t_0-\bar{t}_0) \\
& \times G(z,\bar{x},\bar{t}_0,0,s)\,\mathrm{d}\bar{t}_0\,\mathrm{d}t_0\,\mathrm{d}\bar{x}\,\mathrm{d}x\mathrm{d}z,
\end{aligned} \tag{3.1}
$$

where

$$
R(z,x,\bar{x},t_0) = \delta(t_0)\delta(x-\bar{x})\left(\frac{\delta c}{c_0^3}\right)\left(z,\frac{\bar{x}+x}{2}\right) \tag{3.2}
$$

so that

$$
R = E_2 E_1 c_0^{-3}\delta c \tag{3.3}
$$

with the definitions in (1.8). Changing variables of integration, i.e., $t_0 \mapsto t_0' = t_0 - \bar{t}_0$, (3.1) can be written in the form of an integral operator acting on the distribution R,

$$
\begin{aligned}
\delta G(0,r,t,0,s) = \int_{\mathbb{R}_+} \Bigg\{ \int_{\mathbb{R}} \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}^{n-1}} \Bigg(\int_{\mathbb{R}_+} & G(0,r,t-t_0'-\bar{t}_0,z,x) \\
& \times G(z,\bar{x},\bar{t}_0,0,s)\,\mathrm{d}\bar{t}_0 \Bigg) \\
\times 2\partial_{t_0'}^2 R(z,x,\bar{x},t_0')\,\mathrm{d}\bar{x}\,\mathrm{d}x\,\mathrm{d}t_0' \Bigg\} \mathrm{d}z,
\end{aligned} \tag{3.4}
$$

in between the braces, the contributions of which are integrated over depth z.

Using the reciprocity relation of the time-convolution type for the Green's function, we arrive at the integral representation

$$
\begin{aligned}
\delta G(0,r,t,0,s) = \int_{\mathbb{R}_+} \Bigg\{ \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \Bigg(\int_0^{t-t_0} & G(0,r,t-t_0-\bar{t}_0,z,x) \\
& \times G(0,s,\bar{t}_0,z,\bar{x})\,\mathrm{d}\bar{t}_0 \Bigg) \\
\times 2\partial_{t_0}^2 R(z,x,\bar{x},t_0)\,\mathrm{d}\bar{x}\,\mathrm{d}x\,\mathrm{d}t_0 \Bigg\} \mathrm{d}z.
\end{aligned} \tag{3.5}
$$

Upon substituting (3.3) into this representation we obtain a mapping $\delta c(z,x) \to \delta G(0,r,t,0,s)$ as encountered in Theorem 1.1. The associated operator kernel appears to propagate singularities from two different scattering points, $\bar{x}$ and x, at each depth z, to the surface at $z = 0$.

To arrive at an upward continuation formulation of data modeling, the idea is to substitute in (3.5) for the Green's functions their upward propagating constituents. Thus we replace these Green's functions in accordance with (2.9), (2.10), using only the u_- constituent. So, for the Green's functions $G(z,x,t-t_0,z_0,x_0)$ in (3.5) we substitute the kernel of the operator

$$
\frac{1}{2}\mathcal{H}Q_-^*(z,x,D_x,D_t)G_-(z,z_0)Q_-(z_0,x_0,D_{x_0},D_{t_0}), \tag{3.6}
$$

viewed as a function of (z, x, t, z_0, x_0, t_0). Naturally, a time convolution of two one-way Green's functions appears. This is the motivation of the definition, by its kernel, of an operator $H(z, z_0), z < z_0$ on functions of (s, r, t),

$$(3.7)\qquad \begin{aligned}&(H(z, z_0))(s, r, t, s_0, r_0, t_0)\\ &\quad = \int_{\mathbb{R}} (G_-(z, z_0))(s, t - t_0 - \bar{t}_0, s_0)(G_-(z, z_0))(r, \bar{t}_0, r_0)\, \mathrm{d}\bar{t}_0.\end{aligned}$$

Here $(G_-(z, z_0))(r, \bar{t}_0, r_0, 0)$ denotes the distribution kernel of $G_-(z, z_0)$, and $(H(z, z_0))$ (s, r, t, s_0, r_0, t_0) denotes the distribution kernel of $H(z, z_0)$.

As an alternative formulation, we can write the operator $H(z, z_0)$ as the composition of two operators obtained by a tensor product. We recall that if ψ_1, ψ_2 are two operators with kernels $K_{\psi_1}(x, \bar{x})$, $K_{\psi_2}(y, \bar{y})$, then their tensor product $\psi_1 \otimes \psi_2$ has kernel given by the product $K_{\psi_1}(x, \bar{x}) K_{\psi_2}(y, \bar{y})$ and maps functions of $(\bar{x}, \bar{y})$ to functions of (x, y). We denote the identity operator acting on functions of s by Id_s and similarly for Id_r. If ψ is an operator acting in the (x, t) variables, then we will write ψ_s, ψ_r for the operator acting in the (s, t) variables or the (r, t) variables, respectively. Then we can also write (3.7) as

$$(3.8)\qquad H(z, z_0) = (\mathrm{Id}_s \otimes G_{-,r}(z, z_0)) \circ (G_{-,s}(z, z_0) \otimes \mathrm{Id}_r).$$

Since the tensor product of two operators is a well-defined operator, this shows that $H(z, z_0)$ is well defined. If ψ is an operator on functions of (x, t), then we will often simply write ψ_s instead of $\psi_s \otimes \mathrm{Id}_r$. The map $H(z, z_0)$, $z < z_0$, is the upward continuation operator.

If ψ_1 and ψ_2 are operators on functions of (x, t) and are time translation invariant, then $\psi_{1,s}$ and $\psi_{2,r}$ commute, which can be derived by writing out the distribution kernel of the compositions. The factors $G_{-,s}$ and $G_{-,r}$ can be written as compositions $\psi_{1,s} G_{0,-,s}$, $G_{0,-,s} \psi_{1,s}$ (and similarly for r) using (2.17), (2.18). It follows that the operator H can be written as a composition $\psi_2(z, z_0) H_0(z, z_0)$, where H_0 is given by (3.7) with G_- replaced by $G_{-,0}$ and $\psi_2(z, z_0) = \psi_{1,s}(z, z_0) \psi_{1,r}(z, z_0)$. The operator $\psi_2(z, z_0)$ is pseudodifferential with symbol

$$(3.9)\qquad \psi_2(z, z_0, s, r, \sigma, \rho, \tau) = \psi_1(z, z_0, s, \sigma, \tau) \psi_1(z, z_0, r, \rho, \tau).$$

We can also write $H(z, z_0) = H_0(z, z_0) \psi_2(z_0, z)$ with ψ_2 defined by (3.9) as well, but with z, z_0 interchanged.

Replacing both source and receiver Green's functions, the result is the replacement of the integral in the parentheses of (3.5) by $-\frac{1}{4} Q^*_{-,s}(0) Q^*_{-,r}(0) H(0, z) Q_{-,s}(z) Q_{-,r}(z)$, where we denote $Q_{-,s}(z) = Q_-(z, s, D_s, D_t)$, and similarly for $Q_{-,r}(z)$. Therefore, we define the DSR modeling operator as

$$(3.10)\qquad F_{\mathrm{D}} \delta c = Q^*_{-,s}(0) Q^*_{-,r}(0) \int_0^Z H(0, z) Q_{-,s}(z) Q_{-,r}(z) \tfrac{1}{2} D_t^2 (E_2 E_1 c_0^{-3} \delta c)(z, \cdot, \cdot, \cdot) \mathrm{d}z,$$

where Z is some large number such that $\mathrm{supp}(\delta c)$ is contained in $]0, Z[\times \mathbb{R}^{n-1}$.

In Theorem 5.1 we will show that, in general, F_{D} differs from F by a pseudodifferential cutoff and that under a certain assumption F_{D} models the singular part of the data. We first derive some important properties of H.

The DSR equation. It follows from differentiating expression (3.8) for H with respect to z, using the fact that $B_-(z,r,D_r,D_t)$ and $G_{-,s}(z,z_0)$ commute, that the operator $H_0(z,z_0)$ is a solution operator for the Cauchy initial value problem for the so-called DSR equation, given by

$$\left(\frac{\partial}{\partial z} - \mathrm{i}B_-(z,s,D_s,D_t) - \mathrm{i}B_-(z,r,D_r,D_t)\right) u = 0. \tag{3.11}$$

Using Duhamel's principle (cf. (1.2)), it follows that

$$u(z,s,r,t) = \int_z^Z (H(z,z_0)g(z_0,\cdot,\cdot,\cdot))(s,r,t)\,\mathrm{d}z_0 \tag{3.12}$$

solves the inhomogeneous DSR equation,

$$\begin{aligned}&\left(\frac{\partial}{\partial z} - \mathrm{i}B_-(z,s,D_s,D_t) - \mathrm{i}B_-(z,r,D_r,D_t) - C(z,s,D_s,D_t) - C(z,r,D_r,D_t)\right) u\\ &\quad = g(z,s,r,t), \quad 0 \le z < Z,\end{aligned} \tag{3.13}$$

with zero initial condition, $u(Z,s,r,t) = 0$. It follows from (3.10) that $F_{\mathrm{D}}\delta c$ is given by $Q^*_{-,s}(0)Q^*_{-,r}(0)$ acting on the solution u at $z=0$ of an inhomogeneous DSR equation with

$$g = Q_{-,s}(z)Q_{-,r}(z)\tfrac{1}{2}D_t^2 R \tag{3.14}$$

and Z such that δc is supported in $0 < \delta < z < Z$ as before.

The bicharacteristics associated with (3.13) are, in the notation of (2.21), given by

$$\begin{aligned}\Gamma(z,z_0;s_0,r_0,t_0,\sigma_0,\rho_0,\tau) = (&\gamma_x(z,z_0,s_0,\sigma_0,\tau),\gamma_x(z,z_0,r_0,\rho_0,\tau),t_0\\ &+\gamma_t(z,z_0,s_0,\sigma_0,\tau)+\gamma_t(z,z_0,r_0,\rho_0,\tau),\gamma_\xi(z,z_0,s_0,\sigma_0,\tau),\gamma_\xi(z,z_0,r_0,\rho_0,\tau),\tau).\end{aligned} \tag{3.15}$$

They are defined on the intersection of the maximal intervals associated with source ray coordinates (z,s,σ,τ) and receiver ray coordinates (z,r,ρ,τ); let θ be given as in the previous section. The intersection will be denoted by $]Z_{\min},Z_{\max}[=]Z_{\min}(z,s,r,\sigma,\rho,\tau,\theta),Z_{\max}(z,s,r,\sigma,\rho,\tau,\theta)[$, where we have

$$Z_{\min}(z,s,r,\sigma,\rho,\tau,\theta) = \max(z_{\min,-}(z,s,\sigma,\tau,\theta),z_{\min,-}(z,r,\rho,\tau,\theta)), \tag{3.16}$$

$$Z_{\max}(z,s,r,\sigma,\rho,\tau,\theta) = \min(z_{\max,-}(z,s,\sigma,\tau,\theta),z_{\max,-}(z,r,\rho,\tau,\theta)). \tag{3.17}$$

Let $g(z,s,r,t)$ be supported in the set $0 < \delta < z < Z$. As mentioned, the map $g \mapsto u$ given by (3.12) maps g to the solution of the inhomogeneous DSR equation (3.13) at $z=0$. Motivated by (3.10), we define an operator L by modifying (3.12) with pseudodifferential factors $Q_{-,s}, Q_{-,r}$ and setting $z=0$ as follows:

$$Lg = Q^*_{-,s}(0)Q^*_{-,r}(0)\int_0^Z H(0,z)Q_{-,s}(z)Q_{-,r}(z)g(z,\cdot,\cdot,\cdot)\,\mathrm{d}z. \tag{3.18}$$

Our next result states that H and L are Fourier integral operators and gives a representation of the kernel of H as an oscillatory integral. Consider the following set:

$$\begin{aligned}\{(&\Gamma(0,z,s,r,t,\sigma,\rho,\tau);z,s,r,t,-b(z,s,\sigma,\tau)-b(z,r,\rho,\tau),\sigma,\rho,\tau)\mid\\ &(s,r,t,\sigma,\rho,\tau)\in T^*\mathbb{R}^{2n-1}_{(s,r,t)}, 0 > Z_{\min}(z,s,r,t,\sigma,\rho,\tau,\theta_2)\}.\end{aligned} \tag{3.19}$$

As will be clear from the proof below, this set is a canonical relation. Let $y_0 = (s_0, r_0, t_0)$, $\eta_0 = (\sigma_0, \rho_0, \tau)$. A convenient choice of phase function for the canonical relation is described by Maslov and Fedoriuk [18]. They state that one can always use a subset of the cotangent vector components as phase variables. There is always a set of local coordinates for the canonical relation of the form

$$(z, y_{0I}, \eta_{0J}, s, r, t), \tag{3.20}$$

where $I \cup J$ is a partition of $\{1, \ldots, 2n-1\}$. It follows from Theorem 4.21 in Maslov and Fedoriuk [18] that there is a function $S = S(z, y_{0I}, \eta_{0J}, s, r, t)$, such that locally the canonical relation (3.19) is given by

$$y_{0J} = -\frac{\partial S}{\partial \eta_{0J}}, \qquad \zeta = \frac{\partial S}{\partial z}, \tag{3.21}$$

$$\eta_{0I} = \frac{\partial S}{\partial y_{0I}}, \qquad (\sigma, \rho, \tau) = -\frac{\partial S}{\partial (s, r, t)}. \tag{3.22}$$

Here we take into account the fact that we have a canonical relation, which introduces a minus sign for (σ, ρ, τ).

LEMMA 3.1. *$H(z, z_0)$ is a Fourier integral operator with canonical relation*

$$\{(\Gamma(z, z_0, s, r, t, \sigma, \rho, \tau); s, r, t, \sigma, \rho, \tau) \mid \\ (s, r, t, \sigma, \rho, \tau) \in T^*\mathbb{R}^{2n-1}_{(s,r,t)} \backslash 0,\ z_0 > Z_{\min}(z, s, r, t, \sigma, \rho, \tau, \theta_2)\}. \tag{3.23}$$

The operator L is a Fourier integral operator with canonical relation (3.19). *The kernel of $H(0, z)$ admits microlocally an oscillatory integral representation with phase variables η_{0J}, given by*

$$(H(0, z))(s_0, r_0, t_0, s, r, t) \\ = (2\pi)^{-(2n-1+|I|)/2} \int A(z, y_0, \eta_{0J}, s, r, t) \exp[\mathrm{i}(S(z, y_{0I}, \eta_{0J}, s, r, t) + \langle \eta_{0J}, y_{0J} \rangle)]\, \mathrm{d}\eta_{0J} \tag{3.24}$$

such that the principal part a of the amplitude A satisfies

$$|a(z, y_0, \eta_{0J}, s, r, t)| = \left| \frac{\partial(\sigma, \rho, \tau)}{\partial(y_{0I}, \eta_{0J})} \right|^{1/2} \tag{3.25}$$

with

$$(\sigma(z, y_{0I}, \eta_{0J}, s, r, t), \rho(z, y_{0I}, \eta_{0J}, s, r, t), \tau(z, y_{0I}, \eta_{0J}, s, r, t)) \\ = -\frac{\partial S}{\partial(s, r, t)}(z, y_{0I}, \eta_{0J}, s, r, t) \tag{3.26}$$

in accordance with (3.22).

Proof. The operators $G_{-,s}(z, z_0)$ and $G_{-,r}(z, z_0)$ are Fourier integral operators as noted at the end of section 2 (subject to the substitution of x by s or r, respectively). We consider $G_{-,s}(z, z_0)$. Locally there are Maslov phase functions for its canonical relation (cf. (2.22)), similar to the one described above, here with phase variables $(\tau, \sigma_{0J'})$, where $I' \cup J'$ is a partition of $\{1, \ldots, n-1\}$. Thus $G_{-,s}(z, z_0)$ is a locally finite sum $\sum_j G^{(j)}_{-,s}(z,z_0)$, where the kernels of $G^{(j)}_{-,s}(z,z_0)$ admit oscillatory integral representations of the form

$$(3.27)\quad (G^{(j)}_{-,s}(z,z_0))(s,t,s_0) = \int A'(s,s_0,\sigma_{0,J'},\tau)\exp[\mathrm{i}(S'(z,z_0,s,s_{0I'},\sigma_{0J'},\tau)-\langle\sigma_{0J'},s_{0J'}\rangle-\tau t)]\mathrm{d}\sigma_{0J'}\mathrm{d}\tau.$$

We denote the canonical relation of $G^{(j)}_{-,s}(z,z_0)$ by $\Lambda^{(j)}_s$ (cf. (2.22)). Similarly, we have $G_{-,r}(z,z_0) = \sum_k G^{(k)}_{-,r}(z,z_0)$ in which the kernels of $G^{(k)}_{-,r}(z,z_0)$ admit oscillatory integral representations of the above type with phase variables $(\tau,\rho_{0J''})$, amplitude A'', and phase function $S''(z,z_0,r,r_{0I''},\rho_{0J''},\tau)-\langle\rho_{0J''},r_{0J''}\rangle-\tau t$. We denote the canonical relation of $G^{(k)}_{-,r}(z,z_0)$ by $\Lambda^{(k)}_r$. But then the kernel of $H(z,z_0)$ is given by a sum $\sum_{j,k} H^{(j,k)}(z,z_0)$. Entering expressions of the type (3.27) for $G^{(j)}_{-,s}(z,z_0)$ and $G^{(k)}_{-,r}(z,z_0)$ into (3.7), and performing the $\bar{t}_0$ integration, we find the following expression for the kernel of $H^{(j,k)}(z,z_0)$:

$$(3.28)\quad (H^{(j,k)}(z,z_0))(s,r,t,s_0,r_0,t_0) = \int 2\pi A'(s,s_0,\sigma_{0,J'},\tau)A''(r,r_0,\rho_{0,J''},\tau) \times\exp[\mathrm{i}(S'(z,z_0,s,s_{0I'},\sigma_{0J'},\tau)-\langle\sigma_{0J'},s_{0J'}\rangle + S''(z,z_0,r,r_{0I''},\rho_{0J''},\tau)-\langle\rho_{0J''},r_{0J''}\rangle-\tau t)]\,\mathrm{d}\sigma_{0J'}\,\mathrm{d}\rho_{0J''}\,\mathrm{d}\tau.$$

It is not difficult to verify that $-\mathrm{i}$ times the argument in the exponent is a nondegenerate phase function. Because A' and A'' are symbols supported inside a region with $\|\sigma\| < C|\tau|$ and $\|\rho\| < C|\tau|$ it follows that $A'A''$ is a symbol and that (3.28) is a Fourier integral operator. From the phase function it follows that the canonical relation of $H^{(j,k)}(z,z_0)$ is given by the points

$$(s,r,t_0+t_1+t_2,\sigma,\rho,\tau;s_0,r_0,t_0,\sigma_0,\rho_0,\tau)$$

with

$$(s,t_1,\sigma,\tau;s_0,0,\sigma_0,\tau)\in\Lambda^{(j)}_s \text{ and } (r,t_2,\rho,\tau;r_0,0,\rho_0,\tau)\in\Lambda^{(k)}_r.$$

Taking the union over (j,k) results in (3.23).

Using (3.18) and the fact that H is given by a sum of terms of the form (3.28), it also follows that L is a Fourier integral operator with canonical relation (3.19), as usual for the solution operators of first-order hyperbolic equations.

The phase function $S(z,y_{0I},\eta_{0J},s,r,t)-\langle\eta_{0J},y_{0J}\rangle)$, with S as described in (3.21)–(3.22), describes locally the canonical relation of $H(0,z)$. Therefore the kernel of $H(0,z)$ has microlocally an oscillatory integral representation of the form

$$(3.29)\quad (H(0,z))(y_0,s,r,t) = (2\pi)^{-(2n-1+|I|)/2} \times\int A(z,y_0,\eta_{0J},s,r,t)\exp[\mathrm{i}(S(z,y_{0I},\eta_{0J},s,r,t)+\langle\eta_{0J},y_{0J}\rangle)]\,\mathrm{d}\eta_{0J}.$$

Then the adjoint $H(0,z)^*$ has amplitude $\overline{A(z,y_0,\eta_{0J},s,r,t)}$ and phase $-S(z,y_{0I},\eta_{0J},s,r,t)-\langle\eta_{0J},y_{0J}\rangle$. Hence, the kernel of the composition $H(0,z)^*H(0,z)$ has the oscillatory integral representation

$$(3.30)\quad (2\pi)^{-(2n-1)} \int \overline{A(z,y_0,\eta_{0J},s',r',t')} A(z,y_0,\eta_{0J},s,r,t) \times \exp(\mathrm{i}[-S(z,y_{0I},\eta_{0J},s',r',t') + S(z,y_{0I},\eta_{0J},s,r,t)])\, \mathrm{d}y_{0I}\mathrm{d}\eta_{0J}.$$

We expand the phase as a function of (s',r',t') in a Taylor series about (s,r,t) and identify the gradient

$$(3.31)\quad -\frac{\partial S}{\partial(s,r,t)}(z,y_{0I},\eta_{0J},s,r,t) = (\sigma(z,y_{0I},\eta_{0J},s,r,t), \rho(z,y_{0I},\eta_{0J},s,r,t), \tau(z,y_{0I},\eta_{0J},s,r,t)).$$

Applying a change of variables, $(y_{0I},\eta_{0J}) \mapsto (\sigma,\rho,\tau)$, the phase takes the form

$$(3.32)\qquad \langle(\sigma,\rho,\tau),(s'-s,r'-r,t'-t)\rangle.$$

In the text preceding (2.23) it was noted that $G_{0,-}(z,z_0)$ is unitary. It follows using (3.8) that $H_0(z,z_0)$ is also unitary. Therefore, the operator $H(0,z)^*H(0,z)$ must be a pseudodifferential operator (in (s,r,t)) with symbol 1 in the set of (s,r,t,σ,ρ,τ), where ψ_2 is equal to 1. We conclude that the principal part a of the amplitude A is given by

$$(3.33)\qquad |a(z,y_0,\eta_{0J},s,r,t)| = \left|\frac{\partial(\sigma,\rho,\tau)}{\partial(y_{0I},\eta_{0J})}\right|^{1/2}. \qquad \square$$

4. Depth-to-time conversion. For $h = h(z,s,r)$ we consider the mapping

$$(4.1)\qquad K \;:\; h \mapsto Q^*_{-,s}(0)Q^*_{-,r}(0)\int_0^Z H(0,z)Q_{-,s}(z)Q_{-,r}(z)(E_2h)(z,\cdot,\cdot,\cdot)\mathrm{d}z;$$

we have $K = LE_2$ (cf. (3.18)). The DSR modeling operator (cf. (3.10)) is then given by

$$(4.2)\qquad F_{\mathrm{D}}\delta c = \frac{1}{2}D_t^2 K E_1 c_0^{-3}\delta c.$$

This factorization is exploited in seismic applications such as imaging.

First we make the following observation. We use the notation $\Theta = \Theta(z,s,r,\sigma,\rho,\tau)$ for the sum $-b(z,s,\sigma,\tau) - b(z,r,\rho,\tau)$ appearing in the canonical relation (3.19) of L,

$$(4.3)\qquad \Theta(z,s,r,\sigma,\rho,\tau) = -b(z,s,\sigma,\tau) - b(z,r,\rho,\tau).$$

Because of expression (2.1) the map $\tau \mapsto \zeta = \Theta$ is strictly monotone when Θ is real. Taking as domain only the τ where the two square roots are real, we find the following lemma. The inverse of this map will be denoted by Θ^{-1}.

LEMMA 4.1. *Suppose (z,s,r,σ,ρ) are given, let $c = \max(c(z,s)\|\sigma\|, c(z,r)\|\rho\|)$, and let $d = \sqrt{|\sigma^2\frac{c(z,s)^2}{c(z,r)^2} - \rho^2|}$ if $c(z,s)\|\sigma\| \geq c(z,r)\|\rho\|$ and $d = \sqrt{|\rho^2\frac{c(z,r)^2}{c(z,s)^2} - \sigma^2|}$ otherwise. The map $\tau \mapsto \Theta(z,s,r,\sigma,\rho,\tau)$ is a diffeomorphism $]-\infty,-c[\cup]c,\infty[\; \to\;]-\infty,-d[\cup]d,\infty[$.*

The maximal depth $Z_{\max}$ associated with $(z,s,r,\sigma,\rho,\tau,\theta)$ also has an associated maximal time, given by

$$(4.4)\qquad T_{\max}(z,s,r,\sigma,\rho,\tau,\theta) = -\Gamma_t(Z_{\max}(z,s,r,\sigma,\rho,\tau,\theta),s,r,\sigma,\rho,\tau).$$

We define a subset Ω_θ of $T^*\mathbb{R}^{2n-1}_{(s,r,t)}$, such that t is bounded by $T_{\max}$,

$$\Omega_\theta = \{(s,r,t,\sigma,\rho,\tau) \,|\, 0 < t < T_{\max}(0,s,r,\sigma,\rho,\tau,\theta)\}. \tag{4.5}$$

We have the following result about K.

THEOREM 4.2. *The operator K is microlocally a Fourier integral operator with canonical relation consisting of a set of points,*

$$\{(\Gamma(0,z,s,r,0,\sigma,\rho,\tau); z,s,r,\Theta(z,s,r,\sigma,\rho,\tau),\sigma,\rho) \,| \\ z,s,r,\sigma,\rho,\tau \in \mathbb{R}^{4n-2}, 0 < Z_{\min}(z,s,r,\sigma,\rho,\tau,\theta_2)\}. \tag{4.6}$$

This canonical relation is the graph of an invertible map Σ:

$$\{(z,s,r,\zeta,\sigma,\rho) \,|\, 0\langle z,0\rangle Z_{\min}(z,s,r,\sigma,\rho,\Theta^{-1}(z,s,r,\zeta,\sigma,\rho),\theta_2)\} \to \Omega_{\theta_2}. \tag{4.7}$$

The map K converts depth to time, which is indeed the way seismologists often look at modeling.

Proof. The operator K is the composition of (3.18) and E_2. The first is a Fourier integral operator with canonical relation given by (3.19). The operator E_2 is a Fourier integral operator with canonical relation given by

$$\{(z,s,r,0,\zeta,\sigma,\rho,\tau; z,s,r,\zeta,\sigma,\rho) \,|\, (z,s,r,\zeta,\sigma,\rho) \in T^*\mathbb{R}^{2n-1}_{(z,s,r)}\backslash 0, \tau \in \mathbb{R}\backslash\{0\}\}. \tag{4.8}$$

In general, the composition of two canonical relations $\Lambda_1 \subset T^*(X \times Y)\backslash 0$, $\Lambda_2 \subset T^*(Y \times Z)\backslash 0$, X, Y, Z open subsets of $\mathbb{R}^{n_X}, \mathbb{R}^{n_Y}$, respectively, $\mathbb{R}^{n_Z}$, is said to be transversal if

$$\Lambda_1 \times \Lambda_2 \text{ intersects } T^*X \times (\operatorname{diag} T^*Y) \times T^*Z \text{ transversally.}$$

In the particular case of the canonical relations of L and E_2, their composition is transversal if at the solution τ of

$$-b_s - b_r = \zeta \tag{4.9}$$

we have $\frac{\mathrm{d}\Theta}{\mathrm{d}\tau} \neq 0$; see, e.g., Theorem 2.4.1 in [10]. Because by the previous lemma this is the case, it then follows that the composition LE_2, hence K, is a Fourier integral operator. The composition of the canonical relations is equal to (4.6).

The canonical relation of K is parameterized by (z,s,r,σ,ρ,τ) in a subset of $\mathbb{R}^{4n-2}$. To show that it is invertible we must show that the projections of (4.6) on the two sets given in (4.7) are both diffeomorphisms. By the previous lemma this is clear for the projection on the right-hand side of (4.7). For the projection on the left-hand side of (4.7) it follows from Lemma 25.3.6 of [15] and the fact that the right projection has maximal rank that the linearization of this projection is invertible. Thus it remains to be shown that the equation

$$(s_0,r_0,t_0,\sigma_0,\rho_0,\tau_0) = \Gamma(0,z,s,r,0,\sigma,\rho,\tau) \tag{4.10}$$

determines a unique point (z,s,r,σ,ρ,τ) when $(s_0,r_0,t_0,\sigma_0,\rho_0,\tau_0)$ is in Ω_{θ_2}, the right-hand side of (4.7). The point $(s_0,r_0,t_0,\sigma_0,\rho_0,\tau_0)$ determines a DSR bicharacteristic $\Gamma(z,0,s_0,r_0,t_0,\sigma_0,\rho_0,\tau_0)$. The t component will be denoted here by $\Gamma_t = t_0 - \gamma_t(z,0,s_0,\sigma_0,\tau) - \gamma_t(z,0,r_0,\rho_0,\tau)$. We have a solution to (4.10) if and only if

$$\Gamma(z,0,s_0,r_0,t_0,\sigma_0,\rho_0,\tau_0) - (s,r,0,\sigma,\rho,\tau) = 0. \tag{4.11}$$

In particular, we have that $\Gamma_t(z, 0, s_0, r_0, t_0, \sigma_0, \rho_0, \tau_0) = 0$. Because Γ_t depends strictly monotonically on z, this equation uniquely determines z. The other equations uniquely determine s, r, σ, ρ, τ. If $z < Z_{\max}(0, s_0, r_0, \sigma_0, \rho_0, \tau)$, then there is a DSR bicharacteristic connecting $(s_0, r_0, t_0, \sigma_0, \rho_0, \tau)$ at depth 0 with $(s, r, t, \sigma, \rho, \tau) = \Gamma(z, 0, s, r, t_0, \sigma, \rho, \tau)$ at depth z; hence it follows that then $0 > Z_{\min}(z, s, r, \sigma, \rho, \tau)$, and vice versa. So using definition (4.4) this point $(z, s, r, \sigma, \rho, \tau)$ is such that $0 > Z_{\min}(z, s, r, \sigma, \rho, \tau, \theta_2)$ precisely when $t < T_{\max}(0, s_0, r_0, \sigma_0, \rho_0, \tau_0, \theta_2)$. This completes the proof of the theorem. ◻

5. Modeling in the single-scattering approximation. The replacement of the wave equation Green's function by a pair of one-way Green's functions leads to a cutoff in the modeling of the scattered wave field. To describe when all the singularities of the data are modeled by the DSR method, we need the following assumption. We use some angle θ, $0 < \theta < \pi/2$, with the vertical as introduced in section 2.

ASSUMPTION 2 (DSR assumption). *If $(z, x) \in X$ and $\alpha, \beta \in S^{n-1}$, $t_{\mathrm{s}}, t_{\mathrm{r}} > 0$ depending on (z, x, α, β) are such that $\eta_z(t_{\mathrm{s}}, z, x, \beta, \tau) = \eta_z(t_{\mathrm{r}}, z, x, \alpha, \tau) = 0$, then*

$$c(z,x)^{-1}\frac{\partial \eta_z}{\partial t}(t, z, x, \beta, \tau) < -\cos(\theta), t \in [0, t_{\mathrm{s}}], \tag{5.1}$$

$$c(z,x)^{-1}\frac{\partial \eta_z}{\partial t}(t, z, x, \alpha, \tau) < -\cos(\theta), t \in [0, t_{\mathrm{r}}]. \tag{5.2}$$

It is clear that this assumption is stronger than Assumption 1. In general the set of rays violating this assumption is not small, but it can contain an open subset of the canonical relation (1.7), depending on the properties of the background medium. This limits the applicability of the method discussed here, which however is still useful in many cases, as discussed in the introduction.

In the following theorem we give the DSR modeling formula, and we give the result in terms of a cutoff acting on $F\delta c$. The symbol $\psi_2(0, z, s, r, 0, \sigma, \rho, \tau)$ can be pulled back to a symbol that is a function of $(s, r, t, \sigma, \rho, \tau)$ by the inverse of the map Σ given by (4.6), (4.7).

THEOREM 5.1. *If Assumption 2 is satisfied with $\theta = \theta_1$, then $F_{\mathrm{D}}\delta c \equiv F\delta c$. There is a pseudodifferential operator $\psi_{\mathrm{D}} = \psi_{\mathrm{D}}(s, r, t, D_s, D_r, D_t)$ with principal symbol given by the pull back mentioned just above of ψ_2, that is, 1 on Ω_{θ_1}, and is in $S^{-\infty}$ outside Ω_{θ_2}, such that*

$$F_{\mathrm{D}}\delta c \equiv \psi_{\mathrm{D}} F\delta c. \tag{5.3}$$

Proof. We reconsider the modeling operator F of Theorem 1.1 and use its description by (3.5). In this proof, we denote by G_s the map from a function $f(z, s, t)$ to $(G_s f)(z, s, t) = \int_{\mathbb{R}\times\mathbb{R}^{n-1}\times\mathbb{R}} G(z, s, t - t_0, z_0, s_0) f(z_0, s_0, t_0)\, \mathrm{d}z_0 \mathrm{d}s_0 \mathrm{d}t_0$; cf. (1.2). Motivated by (3.5) and the introduction of H in (3.8), we consider the operator $M = (\mathrm{Id}_r \otimes G_s) \circ (G_r \otimes \mathrm{Id}_s)$, which maps functions of (z_s, z_r, s, r, t) to functions of (z_s, z_r, s, r, t). In our application, we consider M as a map of functions in $z > \delta$ to functions on a small neighborhood of $z_s = 0, z_r = 0$. By an argument similar to the first part of the proof of Lemma 3.1, it follows that M is a Fourier integral operator, with canonical relation consisting of a set of points

$$\{(\eta_{z,s}, \eta_{z,r}, \eta_{x,s}, \eta_{x,r}, t + t_s + t_r, \eta_{\zeta,s}, \eta_{\zeta,r}, \eta_{\xi,s}, \eta_{\xi,r}, \tau; z_s, z_r, s, r, t, \zeta_s, \zeta_r, \rho, \sigma, \tau)\}, \tag{5.4}$$

where $(\zeta_s, \sigma) = -\tau c(z_s, s)^{-1}\beta$, $(\zeta_r, \rho) = -\tau c(z_r, r)^{-1}\alpha$, $\eta_{z,s} = \eta_z(t_s, z_s, s, \beta, \tau)$, and similar for the other components, and for the r-components, cf. (1.6); $\alpha, \beta \in S^{n-1}$ as in Theorem 1.1.

Denote by $R_4(z)$ the restrictions to $z_s = z$ and $z_r = z$ of functions $f(z_s, z_r, s, r, t)$ and by $E_4(z)$ the map that maps a function $f(s, r, t)$ to $(E_4(z)f)(z_s, z_r, s, r, t) = \delta(z_s - z)\delta(z_r - z)f(s, r, t)$. It follows, from writing out the distribution kernel of M, and using the remark below (2.20), that for distributions in (z_s, z_r, s, r, t) with singularities with $\tau^{-1}\zeta_s > 0$ and $\tau^{-1}\zeta_r > 0$, we have

$$R_4(0)ME_4(z)\psi_2'(z, 0) = -\frac{1}{4}Q_{-,s}^*(0)Q_{-,r}^*(0)H(0, z)Q_{-,s}(z)Q_{-,r}(z), \tag{5.5}$$

modulo a regularizing operator, where $\psi_2' = \psi_2 - Q_{-,s}^{-1}Q_{-,r}^{-1}[Q_{-,s}Q_{-,r}, \psi_2]$. Since for F the rays come from one side of the surface $z = 0$, we can apply this to (3.5). Denote by E_5 the map that maps a function $f(z, s, r, t)$ to $(E_5(z)f)(z_s, z_r, s, r, t) = \delta(z_s - z_r)f(\frac{z_s+z_r}{2}, s, r, t)$. It follows that we have

$$R_4(0)ME_5\psi_2'(z, 0)E_2E_1(c_0^{-3}\delta c) \equiv \tfrac{1}{4}KE_1(c_0^{-3}\delta c), \tag{5.6}$$

modulo a regularizing operator.

We can find an operator $\psi'(z, s, r, D_z, D_s, D_r)$ such that the principal symbol $\psi_2' - \psi'$ is zero on the set $\zeta = -b(z, s, \sigma, \tau) - b(z, r, \rho, \tau)$. Namely, first set (for the principal symbol) $\psi'(z, s, r, \zeta, \sigma, \rho) = \psi_2'(z, 0, s, r, \sigma, \rho, \Theta^{-1})$. Then the map $ME_5(\psi_2' - \psi')$ is a Fourier integral operator with highest-order amplitude equal to zero. With lower-order terms in ψ' we find that we can replace ψ_2' in (5.6) by an operator $\psi' = \psi'(z, s, r, D_z, D_s, D_r)$. The operator ψ' commutes with E_2. Hence, if $h = h(z, s, r)$, we have that

$$R_4(0)ME_5E_2\psi'h = Kh, \tag{5.7}$$

modulo a smoothing operator. Because of equality (5.7), $R_4(0)ME_5E_2$ is an invertible Fourier integral operator with canonical relation given by (4.6), microlocally on a neighborhood of the set where ψ' is not in $S^{-\infty}$. Now define microlocally on a neighborhood of the set where ψ' is not in $S^{-\infty}$,

$$\psi_{\mathrm{D}} = R_4(0)ME_5E_2\psi'(R_4(0)ME_5E_2)^{-1}. \tag{5.8}$$

By Egorov's theorem this is a pseudodifferential operator with symbol as in the theorem and we have

$$\psi_{\mathrm{D}}R_4(0)ME_5E_2 = K, \tag{5.9}$$

modulo a smoothing operator. It follows that (5.3) is satisfied. ∎

6. The Bolker condition. It follows from (4.2) and from Theorem 4.2 that the canonical relation of F_{D} in (3.10) satisfies Guillemin's [11] Bolker condition: The projection of the canonical relation (1.7) on $T^*Y\backslash 0$ is an embedding.

Indeed, Assumption 2 is stronger than this condition, as can be seen from the arguments in the proof of Theorem 4.2. This fact will be important for the inverse scattering based on modeling data by F_D.

Acknowledgments. We thank Margaret Cheney for her interest in the work presented in this paper and for encouraging us to write up the results in the present form. We also thank the members of the Mathematical Sciences Research Institute, and in particular Gunther Uhlmann, for providing a very stimulating environment during the Inverse Problems program in Fall 2001.

REFERENCES

[1] A. J. Berkhout and C. P. A. Wapenaar, *A unified approach to acoustical reflection imaging. II: The inverse problem*, J. Acoust. Soc. Amer., 93 (1993), pp. 2017–2023.
[2] G. Beylkin, *Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform*, J. Math. Phys., 26 (1985), pp. 99–108.
[3] B. Biondo and G. Palacharla, *3-D prestack migration of common-azimuth data*, Geophysics, 61 (1996), pp. 1822–1832.
[4] N. Bleistein, J. K. Cohen, and J. W. Stockwell, *Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion*, Springer-Verlag, New York, 2001.
[5] J. F. Claerbout, *Imaging the Earth's Interior*, Blackwell Scientific Publications, Oxford, UK, 1985.
[6] R. W. Clayton, *Common Midpoint Migration*, Technical report SEP-14, Stanford University, Stanford, CA, 1978.
[7] M. D. Collins, and R. B. Evans, *A two-way parabolic equation for acoustic backscattering in the ocean*, J. Acoust. Soc. Amer., 91 (1992), pp. 1357–1368.
[8] M. V. de Hoop, *Generalization of the Bremmer coupling series*, J. Math. Phys., 37 (1996), pp. 3246-3282.
[9] M. V. de Hoop, J. H. Le Rousseau, and R.-S. Wu, *Generalization of the phase-screen approximation for the scattering of acoustic waves*, Wave Motion, 31 (2000), pp. 43–70.
[10] J. J. Duistermaat, *Fourier Integral Operators*, Birkhäuser Boston, Boston, 1996.
[11] V. Guillemin, *On some results of Gel′fand in integral geometry*, in Pseudodifferential Operators and Applications, AMS, Providence, RI, 1985, pp. 149–155.
[12] G. R. Hadley, *Multistep method for wide-angle beam propagation*, Opt. Lett., 17 (1992), pp. 1743–1745.
[13] L. Halpern and L. N. Trefethen, *Wide-angle one-way wave equations*, J. Acoust. Soc. Amer., 84 (1988), pp. 1397–1404.
[14] L. Hörmander, *The Analysis of Linear Partial Differential Operators*, vol. 3, Springer-Verlag, Berlin, 1985.
[15] L. Hörmander, *The Analysis of Linear Partial Differential Operators*, vol. 4, Springer-Verlag, Berlin, 1985.
[16] S. Jin, C. C. Mosher, and R.-S. Wu, *Offset-domain pseudoscreen prestack depth migration*, Geophysics, 67 (2002), pp. 1895–1902.
[17] A. P. E. ten Kroode, D. J. Smit, and A. R. Verdel, *A microlocal analysis of migration*, Wave Motion, 28 (1998), pp. 149–172.
[18] V. P. Maslov and M. V. Fedoriuk, *Semi-Classical Approximation in Quantum Mechanics*, D. Reidel, Boston, 1981.
[19] A. M. Popovici, *Prestack migration by split-step DSR*, Geophysics, 61 (1996), pp. 1412–1416.
[20] Rakesh, *A linearized inverse problem for the wave equation*, Comm. Partial Differential Equations, 13 (1988), pp. 573–601.
[21] P. S. Schultz and J. W. C. Sherwood, *Depth migration before stack*, Geophysics, 45 (1980), pp. 376–393.
[22] C. C. Stolk, *A pseudodifferential equation with damping for one-way wave propagation in inhomogeneous acoustic media*, Wave Motion, 40 (2004), pp. 111–121.
[23] C. C. Stolk, *Parametrix for a hyperbolic initial value problem with dissipation in some region*, Asymptot. Anal., to appear.
[24] C. C. Stolk, and M. V. de Hoop, *Seismic Inverse Scattering in the Downward Continuation Approach*, 2004. preprint.
[25] F. D. Tappert, *The parabolic approximation method*, in Wave Propagation and Underwater Acoustics, Lecture Notes in Phys. 70, J. B. Keller and J. S. Papadakis, eds., Springer-Verlag, New York, 1977, pp. 224–287.

[26] M. E. Taylor, *Reflection of singularities of solutions to systems of differential equations*, Comm. Pure Appl. Math., 28 (1975), pp. 457–478.
[27] M. E. Taylor, *Pseudodifferential Operators*, Princeton University Press, Princeton, NJ, 1981.
[28] V. H. Weston, *Factorization of the wave equation in higher dimensions*, J. Math. Phys., 28 (1987), pp. 1061–1068.