Efficient Computation of Extended Surface Sources

William W. Symes
Department of Computational and Applied Mathematics,
Rice University,Houston TX 77251-1892 USA,
email symes@rice.edu,
ORCID 0000-0001-6213-4272

ABSTRACT

Source extension is a reformulation of inverse problems in wave propagation, that at least in some cases leads to computationally tractable iterative solution methods. The core subproblem in all source extension methods is the solution of a linear inverse problem for a source (right hand side in a system of wave equations) through minimization of data error in the least squares sense with soft imposition of physical constraints on the source via an additive quadratic penalty. A variant of the time reversal method from photoacoustic tomography provides an approximate solution that can be used to precondition Krylov space iteration for rapid convergence to the solution of this subproblem. An acoustic 2D example for sources supported on a surface, with a soft contraint enforcing point support, illustrates the effectiveness of this preconditioner.

Keywords: inverse problems, wave propagation, time reversal, Krylov subspace methods, preconditioning

INTRODUCTION

Full Waveform Inversion (FWI) can be described in terms of

1. a linear wave operator $L[\mathbf{c}]$, depending on a vector of space-dependent coefficients $\mathbf{c}$ and acting on causal vector wavefields $\mathbf{u}$ vanishing in negative time:
$$\mathbf{u} \equiv 0, t \ll 0; \tag{1}$$
2. a trace sampling operator P acting on wavefields and producing data traces;
3. and a (vector) source function (of space and time) $\mathbf{f}$ representing energy input to the system.

The basic FWI problem is: given data d, find $\mathbf{c}$ so that
$$P\mathbf{u} \approx d \text{ and } L[\mathbf{c}]\mathbf{u} = \mathbf{f}. \tag{2}$$

In this formulation, the source function $\mathbf{f}$ may be given, or to be determined subject to some constraints.

A simple nonlinear least squares formulation is:

$$\text{choose } \mathbf{c} \text{ to minimize } \|PL[\mathbf{c}]^{-1}\mathbf{f} - d\|^2. \tag{3}$$

Practical optimization formulations typically augment the objective in 3 by additive penalties or other constraints.

As is well-known, local optimization methods are the only feasible approach given the dimensions of a typical instance of 2, and those have a tendency to stall due to "cycle-skipping". See for example Virieux and Operto (2009) and many references cited there. Source extension is one approach to avoiding this problem. It consists in imposing the wave equation as a soft as opposed to hard constraint, by allowing the source field $\mathbf{f}$ to have more degrees of freedom than is permitted by a faithful model model of the seismic experiment, and constraining these additional degrees of freedom by means of an additive quadratic penalty modifying the problem 3:

$$\text{choose } \mathbf{c}, \mathbf{f} \text{ to minimize } \|PL[\mathbf{c}]^{-1}\mathbf{f} - d\|^2 + \alpha^2\|A\mathbf{f}\|^2 \tag{4}$$

The operator A penalizes deviation from known (or assumed) characteristics of the source function - its null space consists of feasible (or "physical") source models.

Huang et al. (2019) present an overview of the literature on source extension methods, describing a variety of methods to add degrees of freedom to physical source model. The present paper concerns *surface source extension*: physical sources are presumed to be concentrated at points $\mathbf{x}_s$ in space, whereas their extended counterparts are permitted to spread energy over surfaces containing the physical source locations. A simple choice for the penalty operator A is then multiplication by the distance $|\mathbf{x} - \mathbf{x}_s|$ to the physical source location:

$$(A\mathbf{f})(\mathbf{x}, t) = |\mathbf{x} - \mathbf{x}_s|\mathbf{f}(\mathbf{x}, t) \tag{5}$$

I shall use this choice of penalty operator whenever a specific choice is necessary in the development of the theory below.

This paper presents a numerically efficient approach to solving the *source subproblem* of problem 4:

$$\text{given } \mathbf{c}, \text{ choose } \mathbf{f} \text{ to minimize } \|PL[\mathbf{c}]^{-1}\mathbf{f} - d\|^2 + \alpha\|A\mathbf{f}\|^2 \tag{6}$$

Solution of this subproblem is an essential component of *variable projection* algorithms for solution of the nonlinear inverse problem 4. Variable projection is not merely a convenient choice of algorithm for this purpose: it is in some sense essential, see for example Symes et al. (2020). It replaces the nonlinear least squares problem 4 with a *reduced* problem, to be solved iteratively. Each iteration involves solution of the subproblem 6. Therefore efficient solution of the subproblem is essential to efficient

solution of the nonlinear problem via variable projection.

The penalty operator A defined in 5 is linear, so the source subproblem is a linear least squares problem. Under some additional assumptions to be described below, I shall show how to construct an accurate approximate solution operator for problem 6. This approximate solution operator may be used to accelerate Krylov space methods for the solution of the surface source subproblem 6. Numerical examples suggest the effectiveness of this acceleration.

I will fully describe a preconditioner for a special case of the source subproblem 6, in which $\mathbf{u}$ is an acoustic field, $L[\mathbf{c}]$ is the wave operator of linear acoustodynamics, the spatial positions of traces extracted by P lie on a depth plane $z = z_r$, and the positions at which the extended source $\mathbf{f}$ is nonzero lie on another, parallel, depth plane $z = z_s$. This "crosswell" configuration simplifies the analysis underlying the construction of approximate solutions for the source subproblem 6. It is only one of many transmission configurations for which similar developments are possible. Perhaps the most important alternative example is the diving wave configuration, which plays a central role in contemporary FWI.

The preconditioner construction is very similar to the time-reversal method in photoacoustic tomography (Stefanov and Uhlmann, 2009). Preconditioning amounts to a change of norm in the domain and range spaces of the modeling operator. In this case, the modfied norms are weighted L^2, and the weight operators map pressure to corresponding surface source on the source and receiver planes. This pressure-to-source map is closely related to the "hyperbolic Dirichlet-to-Neumann" operator that plays a prominent role in photoacoustic tomography and other wave inverse problems (Rachele, 2000; Stefanov and Uhlmann, 2005). Hou and Symes (2016b) demonstrated a very similar preconditioner for Least Squares Migration, also for its subsurface offset extension (Hou and Symes, 2016a), motivated by ten Kroode (2012). These constructions also involve the Dirichlet-to-Neumann operator. This concept also turns up in hidden form in the work of Yu Zhang and collaborators on true amplitude migration (Zhang et al., 2014; Tang et al., 2013; Xu et al., 2012, 2011; Zhang and Sun, 2009).

The obvious computation of the pressure-to-source map - prescribe the pressure, solve the wave equation with this boundary condition, read off the equivalent source - suffers from intrinsic numerical inaccuracy. I suggest an alternative computationally feasible approach, via economical short-distance wave propagation. Since the map is symmetric only in an approximate, asymptotic sense, it must be symmetrized for use as a Krylov preconditioner. I describe a symmetrization procedure that requires no further wave computations beyond those necessary to compute the action of the operator itself.

The discussion in this paper is formal and incomplete, in the sense that some important mathematical underpinnings are only sketched. I will treat the modeling operator $PL[\mathbf{c}]^{-1}$ as if it mapped square integrable surface sources to square integrable sampled data. This is not true in full generality: while the surface source problem

has distribution solutions, they are not generally square integrable (finite acoustic field energy). Even if the solutions have finite energy, they do not in general have well-defined restrictions to lower-dimensional sets. In other words, the action of the sampling operator P on the space-time plane $z = z_r$ is not well-defined for arbitrary finite-energy acoustic fields. Thus the modeling operator envisioned above may not be well-defined.

This phenomenon is related to the ill-posedness of wave equations as evolution equations in spatial variables, an observation attributed to Hadamard (see Courant and Hilbert (1962), Chapter 6, section 17). Some constraint on the acoustic field, beyond finite energy, is mandatory in any precise mathematical formulation the inverse problems 4 and 6. In fact, the natural constraint in the "crosswell" geometry of this paper is that high-frequency energy travel *only* along rays crossing the surfaces $z = z_s, z = z_r$ transversally. That is, source functions on $z = z_s$ generate waves with energy traveling along rays leaving the surface at a non-zero angle, and energy arrives at the recording surface $z = z_r$ along rays making non-zero angle with it. I will call sources, sampled data, and acoustic fields with this property *downgoing* (even though the concept also encompasses *upcoming* propagation). Note that the downgoing property restricts the behaviour of acoustic fields near the source and receiver surfaces - what the fields do elsewhere is their own business.

Several works have explored the mathematics of the downgoing condition and its consequences in the context of the scalar second order wave equation, see for instance Payne (1975); Symes and Payne (1983); Lasiecka (1986); Lasiecka et al. (1986); Lasiecka and Trigianni (1989); Bao and Symes (1991). Elaboration of these mathematical details is beyond the scope of this paper, which aims instead to explore the algorithmic consequences of the mathematical structure implied by the downgoing condition.

The next section defines the modeling operator $PL[\mathbf{c}]$, its adjoint, and important specializations (pressure vs. normal velocity sources and data). The sections to follow define the source-to-pressure operator, construct an approximate inverse of the modeling operator by time reversal (as suggested by work in photoacoustic tomography), use the source-to-pressure operator to express the approximate inverse as the modeling operator adjoint in weighted norms (thus establishing that the modeling operator is *approximately unitary* in the sense of these norms), explain how to use this construction to precondition Conjugate Gradient iteration, and organize the preconditioning computation so as to involve only one extra and relatively inexpensive wave propagation calculation. The penultimate section displays simple 2D numerical examples of all of the key steps, culminating in a comparison of straight vs. preconditioned CG iteration. The paper ends with a brief discussion-and-conclusion section, reviewing what has been accomplished and listing a few of the many questions left open.

OPERATORS

For acoustic wave physics, the coefficient vector is $\mathbf{c} = (\kappa, \rho)^{\mathbf{T}}$, with components bulk modulus κ and density ρ, and the state vector $\mathbf{u} = (p, \mathbf{v})^T$ consists of pressure p (a scalar space-time field) and particle velocity $\mathbf{v}$ (a vector space-time field). The wave operator $L[\mathbf{c}]$ is:

$$L[\mathbf{c}]\mathbf{u} = \begin{pmatrix} \frac{1}{\kappa}\frac{\partial p}{\partial t} + \nabla \cdot \mathbf{v}, \\ \rho\frac{\partial \mathbf{v}}{\partial t} + \nabla p. \end{pmatrix} \tag{7}$$

That is,

$$L[\mathbf{c}] = \begin{pmatrix} \frac{1}{\kappa}\frac{\partial}{\partial t} & \nabla\cdot \\ \nabla & \rho\frac{\partial}{\partial t} \end{pmatrix} \tag{8}$$

$L[\mathbf{c}]$ has a well-defined inverse in the sense of distributions if it is restricted to either causal or anti-causal vector wavefields.

Most of what follows is valid for any space dimension $n > 0$. The coefficient vector $\mathbf{c} = (\kappa, \rho)$ is defined throughout space $\mathbf{R}^n$, the state vector $\mathbf{u}$ throughout space-time $\mathbf{R}^{n+1}$. Whenever convenient for mathematical manipulations, $n = 3$: for instance, I will write $\mathbf{x} = (x, y, z)^T$ for the spatial coordinate vector, and refer to the third (vertical) coordinate of particle velocity as v_z. Examples later in this paper will use $n = 2$ for computational convenience.

Since all of the operators in the discussion that follows depend on the coefficient vector $\mathbf{c}$, I will suppress it from the notation, for example, $L = L[\mathbf{c}]$.

The surface source extension replaces point sources on or near a surface in $\mathbf{R}^3$ with source functions confined to the same surface. The simplest example of this extended geometry specifies a plane $\{(x, y, z, t) : z = z_s\}$ at source depth z_s as the surface. For acoustic modeling, surface sources are combinations of constitutive law defects and loads normal to the surface, localized on $z = z_s$. That is, right-hand sides in the system $L\mathbf{u} = \mathbf{f}$ take the form $\mathbf{f}(\mathbf{x}, t) = (h_s(x, y, t)\delta(z - z_s), f_s(x, y, t)\mathbf{e_z}\delta(\mathbf{z} - \mathbf{z_s}))^{\mathbf{T}}$ for scalar defect h_s and normal force f_s ($\mathbf{e_z} = (\mathbf{0}, \mathbf{0}, \mathbf{1})$). With the choice L given in 8, the causal/anti-causal wave system $L\mathbf{u}^{\pm} = \mathbf{f}$ takes the form

$$\begin{aligned} \frac{1}{\kappa}\frac{\partial p^{\pm}}{\partial t} &= -\nabla \cdot \mathbf{v}^{\pm} + h_s\delta(z - z_s), \\ \rho\frac{\partial \mathbf{v}^{\pm}}{\partial t} &= -\nabla p^{\pm} + f_s\mathbf{e}\delta(z - z_s), \\ p^{\pm} &= 0 \text{ for } \pm t \ll 0, \\ \mathbf{v}^{\pm} &= 0 \text{ for } \pm t \ll 0. \end{aligned} \tag{9}$$

Remark: In system 9 and many similar systems to follow, I will use the shorthand

$$p^+ = 0 \text{ for } t \ll 0$$

to mean that p^+ is *causal*, that is,

$$\text{For some } T \in \mathbf{R}, p^+(\cdot, t) = 0 \text{ for all } t < T.$$

Similarly,

$$p^- = 0 \text{ for } t \gg 0$$

signifies that p^- is anti-causal.

Extended forward modeling consists in solving 9 and sampling the solution components at receiver locations. For simplicity, throughout this paper I will assume that the receivers are located on another spatial hyperplane $\{(x, y, z, t) : z = z_r\}$ at receiver depth $z_r > z_s$. The constructions to follow involve interchange of the roles of z_s and z_r (that is, locating sources on $z = z_r$ and receivers at $z = z_s$), so rather than the sampling operator P of the introduction, I will denote by P_s, P_r the sampling operators on $z = z_s$, $z = z_r$ respectively. In practice, sampling occurs at a discrete array of points (trace locations) on these surfaces, and over a zone of finite extent. In this theoretical discussion, I will neglect both finite sample rate and extent, and regard the data, for example $P_r p^+$, as continuously sampled and extending over the entire plane $z = z_r$.

As explained in the Introduction, the downgoing constraint on the square-integrable source functions h_s, f_s is essential both for finite energy solutions of the system 9 to exist, and for these solutions to have well-defined traces on the receiver surface $z = z_r$. This constraint will be assumed throughout, often tacitly. For downgoing solutions of system 9, the key components ($p^\pm$ and $v_z^\pm$) are continuous functions of z in the open slab $z_s < z < z_r$ with well-defined limits at the boundary planes, but may be discontinuous at the source plane $z = z_s$. Similarly, the roles of z_s and z_r will be interchanged in some of the constructions to come, and the corresponding solutions may be discontinuous at $z = z_r$. Accordingly, interpret P_s, P_r as the limit from right and left respectively: for $u = p^\pm$ or $v_z^\pm$,

$$\begin{aligned} P_s u(x, y, t) &= \lim_{z \to z_s^+} u(x, y, z, t), \\ P_r u(x, y, t) &= \lim_{z \to z_r^-} u(x, y, z, t). \end{aligned} \tag{10}$$

The causal/anti-causal vector modeling operators $\mathcal{S}^\pm_{z_s, z_r}$ are defined in terms of the solutions $(p^\pm, \mathbf{v}^\pm)$ of the systems 9 by

$$\mathcal{S}^\pm_{z_s, z_r}(h_s, f_s)^T = (P_r p^\pm, P_r v_z^\pm)^T, \tag{11}$$

The subscript signifies that sources are located on $z = z_s$, the receivers on $z = z_r$. It is necessary to include this information in the notation, as versions of $\mathcal{S}^\pm$ with sources and receivers in several locations will be needed in the discussion below.

Remark: To connect with the formulation presented in the introduction, note that for continuous u, $P_s u(x, y, t) = u(x, y, z_s, t)$, and therefore the adjoint of P_s (in the

sense of distributions) is $P_s^T h(x,y,z,t) = h(x,y,t)\delta(z-z_s)$. Write $\mathcal{P}_s = \text{diag}\ (P_s, P_s)$ and similarly for $\mathcal{P}_r$. Then

$$\mathcal{S}^+_{z_s,z_r} = \mathcal{P}_r L^{-1} (\mathcal{P}_s)^T,$$

in which L^{-1} is interpreted in the causal sense, and similarly for $\mathcal{S}^-$. Sources confined to $z = z_s$ are precisely those functions (distributions, really) output by $\mathcal{P}_s^T$, so the problem statements 4 and 6 can be rewritten in terms of $\mathcal{S}^+_{z_s,z_r}$, with P identified with $\mathcal{P}_r$.

$\mathcal{S}^\pm$ is not stably invertible: its columns are approximately linearly dependent, as will be verified below. The diagonal components of $\mathcal{S}^\pm$ thus carry essentially all of its information, and it is in terms of these that a sensible inverse problem is defined.

Denote by $\Pi_i, i = 0, 1$ the projection on the first, respectively second, component of a vector in $\mathbf{R}^2$. The forward modeling operator from pressure source to pressure trace is

$$S^\pm_{z_s,z_r} = \Pi_0 \mathcal{S}^\pm_{z_s,z_r} \Pi_0^T \tag{12}$$

and the forward modeling operator from velocity source (normal force) to velocity trace is

$$V^\pm_{z_s,z_r} = \Pi_1 \mathcal{S}^\pm_{z_s,z_r} \Pi_1^T \tag{13}$$

With these conventions, we can write the version of the source subproblem 6 studied in this paper as

$$\text{find } h_s \text{ to minimize } \|S^+_{z_s,z_r} h_s - d\|^2 + \alpha^2 \|A h_s\|^2. \tag{14}$$

It follows from the adjoint state method (see Appendix A for details) that

$$(\mathcal{S}^\pm_{z_s,z_r})^T = -\mathcal{S}^\mp_{z_r,z_s} \tag{15}$$

Define R to be the *time-reversal operator* on functions of space-time, $Rf(\mathbf{x},t) = f(\mathbf{x},-t)$, and $\mathcal{R}$ to be the *acoustic field time-reversal operator*

$$\mathcal{R} \begin{pmatrix} p \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} Rp \\ -R\mathbf{v} \end{pmatrix} \tag{16}$$

Then

$$\mathcal{R}\mathcal{S}^\mp = -\mathcal{S}^\pm_{z_r,z_s}\mathcal{R} \tag{17}$$

Since $R^2 = I$ and $\mathcal{R}^2 = I$, the identities 15 and 17 imply that

$$(\mathcal{S}^\pm_{z_s,z_r})^T = \mathcal{R}\mathcal{S}^\pm_{z_r,z_s}\mathcal{R} = -\mathcal{S}^\mp_{z_r,z_s}. \tag{18}$$

The relation 18 implies that

$$\begin{aligned}(S^{\pm}_{z_s,z_r})^T &= -S^{\mp}_{z_r,z_s}\\ &= RS^{\pm}_{z_r,z_s}R,\\ (V^{\pm}_{z_s,z_r})^T &= -V^{\mp}_{z_r,z_s}\\ &= RV^{\pm}_{z_r,z_s}R.\end{aligned} \tag{19}$$

PRESSURE-TO-SOURCE

Since the system 9 has a unique solution by standard theory (Lax, 2006), the source vector field (h_s, f_s) determines the acoustic field $(p^{\pm}, \mathbf{v}^{\pm})$ in space time, and in particular the limits from the right at $z = z_s$, $P_s p^{\pm}$ and $P_s v_z^{\pm}$. This relation is not invertible: it is not possible to prescribe both pressure and normal velocity on a surface such as $z = z_s$. So the columns of the matrix operator $\mathcal{S}^{\pm}_{z_s,z_r}$ must satisfy a linear relation. In this section I will explain this relation; it involves the *pressure-to-source* map. This operator also turns out to be the principal component of a preconditioning strategy for iterative solution of the optimization problem 6, so I will devote some effort to its proper definition. It is closely related to the Dirichlet-to-Neumann operator mentioned in the introduction.

While it is not possible to prescribe both pressure and velocity on $z = z_s$ in solutions of 9, it is possible to prescribe pressure only, for instance: if the function ϕ on the surface $z = z_s$ satisfies suitable conditions, for example the downgoing constraint mentioned earlier, a unique solution exists for the acoustic system in both half-spaces $\pm z > z_s$:

$$\begin{aligned}\frac{1}{\kappa}\frac{\partial p_{\pm}}{\partial t} &= -\nabla\cdot\mathbf{v}_{\pm},\\ \rho\frac{\partial \mathbf{v}_{\pm}}{\partial t} &= -\nabla p_{\pm},\\ p_{\pm} &= 0,\ \text{for } t \ll 0,\\ \mathbf{v}_{\pm} &= 0 \text{ for } t \ll 0,\\ \lim_{z\to z_s^{\pm}} p_{\pm}(x,y,t,z) &= \phi(x,y,t).\end{aligned} \tag{20}$$

Note that the subscript $\pm$ here refers to the sign of $z - z_s$, as opposed to the superscript $\pm$, which refers to the sign of t throughout this paper.

From the boundary condition (last equation in 20), one sees that the pressures $p_{\pm}$ in the two half-spaces have the same limit at the boundary $z = z_s$. Stick the two half-space solutions together to form an acoustic field $(p^+, \mathbf{v}^+)$ in all of space-time, that is,

$$p^+(x,y,z,t) = \begin{cases} p_+(x,y,z,t) \text{ if } z > 0,\\ p_-(x,y,z,t) \text{ if } z < 0,\end{cases} \tag{21}$$

and a similar definition for $\mathbf{v}^+$. Then p^+ is continuous across $z = z_s$, and the boundary

condition in system 20 may be written as $P_s p^+ = \phi$.

The same construction can be carried out in the anti-causal sense, with anti-causal half-space solutions glued together to form a full-space distribution solution $(p^-, \mathbf{v}^-)$, with the property that p^- is continuous across $z = z_s$ and $P_s p^- = \phi$.

The reader may object that the notation $(p^\pm, \mathbf{v}^\pm)$ is already in use, for the solution of 9. This objection is valid. However, *in the sense of distributions*, $(p^\pm, \mathbf{v}^\pm)$ as defined in display 21, is *exactly* the causal solution of 9 for the choice $h_s = -[v_z^\pm]|_{z=z_s}$, $f_s = 0$, as follows from a simple integration-by-parts calculation. So the notation is consistent!

The negative jump $-[v_z^\pm]|_{z=z_s}$ is thus a function of ϕ. Define the *pressure-to-source* operator $\Lambda_{z_s}^\pm$ by

$$\Lambda_{z_s}^\pm \phi = -[v_z^\pm]|_{z=z_s} \tag{22}$$

The conclusion: if $h_s = \Lambda_{z_s}^\pm \phi$ and $f_s = 0$ in the system 9, then $\phi = P_s p^\pm$.

Otherwise put, $S_{z_s,z_s}^\pm \Lambda_{z_s}^\pm \phi = \phi$, so $\Lambda_{z_s}^\pm$ is inverse to $S_{z_s,z_s}^\pm$. The relation 19 implies in turn that

$$(\Lambda_{z_s}^\pm)^T = -\Lambda_{z_s}^\mp \tag{23}$$

There is also a *velocity-to-source* operator. For the solution $(p^\pm, \mathbf{v}^\pm)$ of system 9 with $h_s = 0$, the normal component of velocity, $v_z^\pm$, is continuous across $z = z_s$, and the velocity source (vertical load) $f_s = -[p^\pm]_{z=z_s}$. I will not name the velocity-to-source operator, as it does not appear explicitly in the developments to follow. As will be seen, it is essentially the inverse of the pressure-to-source operator.

The quadratic form defined by $\Lambda_{z_s}^\pm$ has fundamental physical significance. Define the total acoustic energy $E^\pm(t)$ of the field $(p^\pm, \mathbf{v}^\pm)$, at time t by

$$E^\pm(t) = \frac{1}{2}\int d\mathbf{x} \left(\frac{(p^\pm)^2}{\kappa} + \rho |\mathbf{v}^\pm|^2 \right)(\mathbf{x}, t). \tag{24}$$

Then

$$\pm \lim_{\pm t \to \infty} E^\pm(t) = \langle P_s p^\pm, (\Lambda_{z_s}^\pm P_s p^\pm) \rangle_{L^2(z=z_s)}. \tag{25}$$

That is, the value of the quadratic form defined by $\Lambda_{z_s}^\pm$, evaluated at the pressure trace on $z = z_s$, gives the total energy transferred from the source to the acoustic field over time. Since E is itself a positive definite quadratic form in the acoustic field, it follows that $\pm\Lambda_{z_s}^\pm$ is positive semi-definite.

While $\Lambda_{z_s}^\pm$ is positive semi-definite, it is not symmetric. However, it is *approximately symmetric* in the high-frequency sense. This fact follows from a geometric optics analysis of the half-space solution. This leads to the identification of $\Lambda_{z_s}^\pm$ as a *pseudodifferential operator* of order zero on $z = z_s$, with principal symbol

$$\sigma_0(\Lambda_{z_s}^\pm) = \pm 2(\kappa(\mathbf{x})\rho(\mathbf{x}))^{1/2} \left(1 - \frac{\kappa(\mathbf{x})(\xi^2 + \eta^2)}{\rho(\mathbf{x})\omega^2} \right)^{-1/2}. \tag{26}$$

Here ξ, η, and ω are the dual Fourier variables to x, y, and t respectively. The downgoing assumptions means that for local planewave components of $P_s p$, the quantity inside the square root is positive. Thus $\Lambda^{\pm}_{z_s}$ has real principal symbol (in fact, the entire symbol is real) hence defines an asymptotically symmetric operator:

$$(\Lambda^{\pm}_{z_s})^T \approx \Lambda^{\pm}_{z_s}. \tag{27}$$

(For more on this, see Stefanov and Uhlmann (2005).) The analysis also reveals that the solution components not continuous at $z = z_s$ are odd there:

$$\lim_{z \to z_s^+} v_z^{\pm} \approx - \lim_{z \to z_s^-} v_z^{\pm} \tag{28}$$

for the solution of 9 with $f_s = 0$. Similarly,

$$\lim_{z \to z_s^+} p^{\pm} \approx - \lim_{z \to z_s^-} p^{\pm} \tag{29}$$

for the solution of 9 with $h_s = 0$. Here "$\approx$" means in the sense of high frequency asymptotics, that is, that the difference between the two sides is relatively smooth, hence small if the data is highly oscillatory. Therefore if $f_s = 0$ in system 9,

$$h_s = \Lambda^{\pm}_{z_s} P_s p^{\pm} = -[v_z^{\pm}]|_{z=z_s} \approx -2 P_s v_z^{\pm} \tag{30}$$

Similarly, if $h_s = 0$ in system 9, then

$$f_s = -[p^{\pm}]|_{z=z_s} \approx -2 P_s p^{\pm}. \tag{31}$$

Thus f_s determines approximately the boundary value of $p^{\pm}$, as a solution of the acoustic wave system in the half-space $z > z_s$. However, as repeated in equation 30, a solution with this boundary value is also the restriction to $z > z_s$ of a solution to 9 with $f_s = 0$ and $h_s = \Lambda^{\pm}_{z_s} P_s p^{\pm}$. Therefore if

$$h_s = -\frac{1}{2} \Lambda^{\pm}_{z_s} f_s, \tag{32}$$

then the pressure boundary value $P_s p^{\pm}$ is the same for the solutions of 9 for source vectors $(h_s, 0)$ and $(0, f_s)$. Since the pressure boundary values are the same, the solutions in $z > z_s$ are the same. In particular, since $z_r > z_s$ and $\mathcal{S}^{\pm}_{z_s, z_r}(h_s, f_s)^T = (P_r p^{\pm}, P_r v_z^{\pm})^T$, it follows that

$$\mathcal{S}^{\pm}_{z_s, z_r} \left(\frac{1}{2} \Lambda^{\pm}_{z_s} f_s, f_s \right)^T \approx 0. \tag{33}$$

Equation 33 states the relation between the columns of $\mathcal{S}^{\pm}_{z_s, z_r}$ mentioned in the introduction to this section.

TIME REVERSAL

Recall that the source vector (h_s, f_s) is assumed to produce a downgoing field $(p^+, \mathbf{v}^+)$, that is, emanates high-frequency energy only along rays that make an angle with the vertical bounded below by a common minimum angle. Such rays leave Ω within a common maximum time. Consequently (Appendix B), in the slab $z_s < z < z_r$, the field $(p^+, \mathbf{v}^+)$ approximates the solution of an anti-causal evolution equation. Choose $\chi(t)$ to be a smooth function that is $= 0$ for $t \gg 0$ and $= 1$ at times when near rays carrying high-frequency energy in $(p^+, \mathbf{v}^+)$ cross $z = z_r$. Define $(\tilde{p}^-, \tilde{\mathbf{v}}^-)$ to be the solution in the half-space $\Omega \times \mathbf{R}$ of

$$\begin{aligned}
\frac{1}{\kappa}\frac{\partial \tilde{p}^-}{\partial t} &= -\nabla \cdot \tilde{\mathbf{v}}^-, \\
\rho\frac{\partial \tilde{\mathbf{v}}^-}{\partial t} &= -\nabla \tilde{p}^-, \\
\tilde{p}^- &= 0, \text{ for } t \gg 0 \qquad (34)\\
\tilde{\mathbf{v}}^- &= 0 \text{ for } t \gg 0 \qquad (35)\\
P_r\tilde{p}^- &= \chi P_r p^+. \qquad (36)
\end{aligned}$$

That is, $\tilde{p}^-$ has the same boundary value on $z = z_r$ as p^+, except for low-frequency residue that is muted by χ. Therefore $p^+ \approx \tilde{p}^-, \mathbf{v}^+ \approx \tilde{\mathbf{v}}^-$ near $z = z_r$. Since the right-hand sides in the system 9 are singular only on $z = z_s$, and the high-frequency components of $(p^+, \mathbf{v}^+)$ are carried by downgoing rays, these differ negligibly from the the high-frequency components of $(\tilde{p}^-, \tilde{\mathbf{v}}^-)$ in the space-time slab $z_s < z < z_r$, and the approximation holds throughout this region. In particular $P_s v_z^+ \approx P_s \tilde{v}_z^-$. In view of the relation 30,

$$-2P_s\tilde{v}_z^- \approx h_s, \qquad (37)$$

so solution of the system 34 followed by restriction to $z = z_s$ and multiplication by -2 approximately inverts the map $S^+_{z_s,z_r} : h_s \mapsto P_r p^+$.

Next observe that in view of the relation 31, and the downgoing nature of the ray system carrying the high frequency energy in $(p^+, \mathbf{v}^+)$, the field $(\tilde{p}^-, \tilde{\mathbf{v}}^-)$ is actually the restriction to $z < z_r$ of the anti-causal solution of 9 with z_s replaced by z_r, zero constitutive defect, and vertical load given by the jump in pressure at $z = z_r$ - for this field, use the same notation. Continuity of vertical velocity $\tilde{v}_z^-$ at $z = z_r$ implies that the vertical load is

$$f_r = -[\tilde{p}^-]|_{z=z_r} = -(\lim_{z\to z_r^+} \tilde{p}^- - \lim_{z\to z_r^-} \tilde{p}^-)$$

$$\approx 2P_r\tilde{p}^- = 2P_r p^+$$

(from the definition 10, P_r is the limit from the left). Thus

$$P_s\tilde{v_z^-} \approx V^-_{z_r,z_s}(2P_r p^+) \approx 2V^-_{z_r,z_s}S^+_{z_r,z_s}h_s.$$

so

$$h_s \approx -2P_s v_z^+ \approx -2P_s \tilde{v}_z^- \approx -4V_{z_r,z_s}^- S_{z_r,z_s}^+ h_s$$

Combine this observation with 37 to obtain

$$-4V_{z_r,z_s}^- S_{z_s,z_r}^+ \approx I,$$

This relation combines with the identity 19 to yield the first main result of this section:

$$\begin{aligned}
(V_{z_s,z_r}^+)^T S_{z_s,z_r}^+ &\approx \frac{1}{4} I, \\
(S_{z_s,z_r}^+)^T V_{z_s,z_r}^+ &\approx \frac{1}{4} I, \\
V_{z_s,z_r}^+ (S_{z_s,z_r}^+)^T &\approx \frac{1}{4} I, \\
S_{z_s,z_r}^+ (V_{z_s,z_r}^+)^T &\approx \frac{1}{4} I.
\end{aligned} \tag{38}$$

. The second equation is simply the transpose of the first, and the last two follow by by an exactly analogous argument using time reversal and interchange of the roles of z_s andz_r.

The conclusion is significant enough to merit restating in English: provided that high-frequency energy in the various fields is carried along downgoing ray fields, the transpose of V^+ is an approximate inverse to S^+, modulo a factor of 4. To recover the pressure source h_s generating a pressure gather $P_r p$ at $z = z_r$, multiply the latter by -2, then apply the transpose of V_{z_s,z_r}^+ to this gather, reading out a vertical velocity field at $z = z_s$. Multiply again by -2 and you have a high-frequency approximation to h_s.

UNITARITY

The next chapter in this story recognizes the relations in display 38 as asserting the approximate unitarity of S_{z_s,z_r}^+.

The matrix identity 33 implies a relation between S, V, and Λ of some interest in itself. After minor re-arrangement, the second row of reads

$$-\frac{1}{2}\Pi_1 \mathcal{S}_{z_s,z_r}^{\pm} \Pi_0^T \Lambda_{z_s}^{\pm} \approx V_{z_s,z_r}^{\pm}. \tag{39}$$

In these relations, the projection on the left picks out the vertical velocity component of a downgoing wavefield at $z = z_r$: that is,

$$-\frac{1}{2}\Pi_1 \mathcal{S}_{z_s,z_r}^{\pm} \Pi_0^T \Lambda_{z_s}^{\pm} P_s p^+ = -\frac{1}{2} P_r v_z^+,$$

where $(p^+, \mathbf{v}^+)$ solve the system 9 with $f_s = 0$ and $h_s = \Lambda_{z_s}^{\pm} P_s p^+$. On the other hand,

from relation 30,

$$P_r v_z^+ = -\frac{1}{2}\Lambda_{z_r}^+ P_r p^+$$

where

$$\begin{aligned} P_r p^+ &= \Pi_0 \mathcal{S}_{z_s,z_r}^+ \Pi_0^T \Lambda_{z_s}^+ P_s p^+ \\ &= S_{z_s,z_r}^+ \Lambda_{z_s}^+ P_s p^+ \end{aligned}$$

Therefore combining the last two equations with 39, obtain

$$\frac{1}{4}\Lambda_{z_r}^+ S_{z_s,z_r}^+ \Lambda_{z_s}^+ = V_{z_s,z_r}^+. \tag{40}$$

This is the promised relation.

As shown in the last section, $4(V_{z_s,z_r}^+)^T$ is approximately inverse to S_{z_s,z_r}^+. Therefore, transposing both sides of equation 40 and using 38, obtain

$$4(V_{z_s,z_r}^+)^T S_{z_s,z_r}^+ = [(\Lambda_{z_s}^+)^T (S_{z_s,z_r}^+)^T (\Lambda_{z_r}^+)^T] S_{z_s,z_r}^+ \approx I. \tag{41}$$

The remarkable feature of the identity 41 is that it exhibits an approximate right inverse of S^+ as an adjoint with respect to a weighted inner product - or it would, if the operators (Λ^+) were symmetric positive definite. As noted earlier, these operators are only approximately symmetric, though they are positive semi-definite. That is not a great obstacle, however: symmetrizing them in the obvious way commits a negligible error, of the sort that this paper already neglects wholesale. That is,

$$[\frac{1}{2}((\Lambda_{z_s}^+)^T + \Lambda_{z_s}^+)(S_{z_s,z_r}^+)^T \frac{1}{2}((\Lambda_{z_r}^+)^T + \Lambda_{z_r}^+)] S_{z_s,z_r}^+ \approx I. \tag{42}$$

The symmetrized Λ operators are at least positive semi-definite, hence define (at least) semi-norms. Similar relations have been derived for other scattering operators, and have been used to accelerate iterative solutions of inverse scatering problems: Dafni and Symes (2018) review some of this literature.

ACCELERATED ITERATIVE INVERSION

For convenience, in this section write S in place of S_{z_s,z_r}^+. Also abbreviate the symmetrized Λ operators using notation suggesting weight operators in model and data spaces:

$$\begin{aligned} W_m^{-1} &= \frac{1}{2}((\Lambda_{z_s}^+)^T + \Lambda_{z_s}^+), \\ W_d &= \frac{1}{2}((\Lambda_{z_r}^+)^T + \Lambda_{z_r}^+). \end{aligned} \tag{43}$$

The identification of the symmetrized $\Lambda_{z_s}^+$ as the inverse of another operator W_m is formal, since the former operator is likely to have null (or nearly-null) vectors due to aperture-related amplitude loss. Since some version of W_m is essential in the formulation for effective preconditioning, I will derive a usable candidate to stand in for it below.

Adopting Hilbert norms defined by the operators W_m and W_d in its domain and range respectively, the adjoint of S is given by

$$S^\dagger = W_m^{-1} S^T W_d, \tag{44}$$

In this notation, the relation 42 takes the form

$$S^\dagger S \approx I. \tag{45}$$

That is to say, S is approximately unitary with respect to the weighted norms defined by W_m and W_d. Therefore a Krylov space method employing these norms will converge rapidly, at least for the well-determined components of the solution.

The most convenient arrangement the Conjugate Gradient (CG) algorithm taking advantage of the structure 44 is the *Preconditioned CG*. Allowing that the fit error will be measured by the data space norm, the least squares problem to be solved is not just $Sh \approx d$, but a regularized version:

$$\text{minimize}_h \|Sh - d\|_d^2 + \alpha^2 \|Ah\|_m^2 \tag{46}$$

Remark: recall that the modified data space norm $\|d\|_d^2 = \langle d, W_d d\rangle$ has physical meaning: for acoustics, it is proportional to the power transmitted to the fluid by the source.

The minimizer of the objective defined in equation 46 solves the normal equation

$$(S^\dagger S + \alpha^2 A^\dagger A)h = S^\dagger d \tag{47}$$

where the weighted adjoint $S^\dagger$ has already been defined in equation 44, and $A^\dagger$ is the adjoint of A in the weighted model space norm defined by W_m, namely

$$A^\dagger = W_m^{-1} A^T W_m. \tag{48}$$

Note that the normal operator appearing on the left-hand side of 47 is not an approximate identity, due to the presence of the regularization term: the spectrum increases in spread with increasing α, leading to slower convergence. Fortunately for the present setting, the operators W_m^{-1}, A, and W_m approximately commute (they are scalar *pseudodifferential*, once the difficulties with the definition of W_m, mentioned above, are taken care of). Scalar pseudodifferential operators approximately

commute, so $A^\dagger \approx A^T$. Therefore

$$S^\dagger S + \alpha^2 A^\dagger A \approx I + \alpha^2 A^T A \tag{49}$$

Recall that A is simply multiplication by the Euclidean distance to the physical source point $\mathbf{x}_s$: $Au(\mathbf{x}) = |\mathbf{x} - \mathbf{x}_s| u(\mathbf{x}), A^T A u(\mathbf{x}) = |\mathbf{x} - \mathbf{x}_s|^2 u(\mathbf{x})$. So the equation $(I + \alpha^2 A^T A)u = b$ is trivial to solve, and this is a key characteristic of a good preconditioner. However this observation must be combined with the weighted norm structure.

Rewrite the normal equation 47 as

$$W_m^{-1}(S^T W_d S + \alpha^2 A^T W_m A)h = W_m^{-1} S^T W_m d \tag{50}$$

Since W_m is self-adjoint and positive semidefinite, the common factor on both sides of 50 can be re-written as

$$Nh = (S^* S + \alpha^2 A^* A)h = S^* d \tag{51}$$

in which S^*, A^* are the adjoints with the original (Euclidean) inner product in the domains but the weighted inner product in data space:

$$S^* = S^T W_d, \tag{52}$$
$$A^* = A^T W_m. \tag{53}$$

Note the S^*S and A^*A are symmetric in the Euclidean sense, so equation 51 is a symmetric positive (semi-)definite linear system, just the sort of thing for which the The Preconditioned Conjugate Gradient ("PCG") algorithm was designed. PCG for solution of equation 51 with preconditioner M is usually written as Algorithm 1 (see for example Golub and van Loan (2012)):

Algorithm 1 Preconditioned Conjugate Gradient Algorithm, Standard Version

Choose $h_0 = 0$
$r_0 \leftarrow S^* d$
$p_0 \leftarrow M^{-1} r_0$
$g_0 \leftarrow p_0$
$q_0 \leftarrow N p_0$
$k \leftarrow 0$
repeat
  $\alpha_k \leftarrow \frac{\langle g_k, r_k \rangle}{\langle p_k, q_k \rangle}$
  $h_{k+1} \leftarrow h_k + \alpha_k p_k$
  $r_{k+1} \leftarrow r_k - \alpha_k q_k$
  $g_{k+1} \leftarrow M^{-1} r_{k+1}$
  $\beta_{k+1} \leftarrow \frac{\langle g_{k+1}, r_{k+1} \rangle}{\langle g_k, r_k \rangle}$
  $p_{k+1} \leftarrow g_{k+1} + \beta_{k+1} p_k$
  $q_{k+1} \leftarrow N p_{k+1}$
  $k \leftarrow k + 1$
until Error is sufficiently small, or max iteration count exceeded

The iteration converges rapidly if $M^{-1}N \approx I$. This is true if and only if the symmetrized operator $M^{-1/2} N M^{-1/2} \approx I$, which is in turn true if the eigenvalues of $M^{-1/2} N M^{-1/2}$ are close to 1 (actually works well is most of these eigenvalues are close to 1, and the rest are small - which is the case for the current problem).. Further, PCG is computationally effective is M is easy to invert.

From 49 and 50, it follows that

$$W_m^{-1}(S^T W_d S + \alpha^2 A^T W_m A) \approx I + \alpha^2 A^T A.$$

This observation suggests using $M = W_m(I + \alpha^2 A^T A)$. This choice is not symmetric, but since the operators on the right-hand side are scalar pseudodifferential hence commute, it is equivalent to use of

$$\begin{aligned} M &= (I + \alpha^2 A^T A)^{1/2} W_m (I + \alpha^2 A^T A)^{1/2}, \\ M^{-1} &= (I + \alpha^2 A^T A)^{-1/2} W_m^{-1} (I + \alpha^2 A^T A)^{-1/2}. \end{aligned} \tag{54}$$

With this choice, 49 implies that $M^{-1}N \approx I$, also M is symmetric. As already mentioned, powers of $I + \alpha^2 A^T A$ are trivial to compute, given the choice of A made here. We will examine fast algorithms for computing W_m^{-1} = the symmetrized pressure-to-source operator in the next section. Note that only M^{-1}, hence only W_m^{-1}, appears in Algorithm 1.

COMPUTING AND SYMMETRIZING Λ

Computations of $\Lambda^{\pm}_{z_s}$ and its transpose are clearly critical steps in an implementation of the PCG algorithm outlined in the preceding section. Direct computation of the pressure-to-source operator $\Lambda^{\pm}_{z_s}$, for instance by solving 9 and reading off $P_s v^{\pm}_z$, turns out to be numerically ill-behaved. The relation 33 provides and alternative approach, taking advantage of the accurate approximate inverse to $S^{+}_{z_s,z_r}$ constructed above. The first row of 33, slightly rearranged, is

$$\Pi_0 \mathcal{S}^{+}_{z_s,z_r} \Pi_1^T f_s \approx -\frac{1}{2} S^{+}_{z_s,z_r} \Lambda^{+}_{z_s} f_s. \tag{55}$$

The approximate inverse construction for $S^{+}+_{z_s,z_r}$ permits (approximate) solution of this equation for $\Lambda^{+}_{z_s} f_s$: apply $4(V^{+}_{z_s,z_r})^T$ to both sides of equation 55 and use the first equation in the list 38 to get

$$\Lambda^{+}_{z_s} \approx -8(V^{+}_{z_s,z_r})^T \Pi_0 \mathcal{S}^{+}_{z_s,z_r} \Pi_1^T. \tag{56}$$

This identity is the major result of this section: it shows how to compute that action of $\Lambda^{+}_{z_s}$ by propagating the input pressure trace, identified as a source for the velocity evolution, forward in time from z_s to z_r reading off the pressure trace on $z = z_r$, identifying it once more as a point load (source for velocity), propagating it backwards in time from z_r to z_s, and finally reading off the velocity trace, interpreted as a pressure evolution source on z_s.

The importance of this result lies in the failure of the obvious method for computing the action of $\Lambda^{\pm}_{z_s}$, namely to employ the pressure trace as a source in the velocity equation (f_s, in the notation used above) at $z = z_s$, and read off the velocity field also at $z = z_s$. This difficulty is related to the existence of tangentially propagating waves and the lack of continuity of the trace operator. The method implicit in equation 56 avoids this difficulty by propagating the fields a positive distance in z: assuming as always that the causal fields are downgoing, this step eliminates any tangentially propagating fields from consideration.

A deeper study of the pressure-to-source operator (or of the closely related Dirichlet-to-Neumann operator for the second order wave equation, see Stefanov and Uhlmann (2005)) shows that it is approximately dependent only on the model coefficients near the source surface ($z = z_s$ in this case). Since the homogeneous and lens models are identical near this surface, it is unsurprising that these figures are very close to the previous two. However an even more useful observation is that the calculations in the approximation 56 could just as well be carried out in a much smaller region around the source surface, and produce a result that is functionally identical in that it will serve as a source for the same acoustic fields globally, with small error. In effect, equation 56 involving propagation from source ($z = z_s$) to receiver ($z = z_r$) surfaces is altered by replacing z_r with a receiver datum $z_s + \Delta z$ considerably closers to z_s:

$$\Lambda^{+}_{z_s} \approx -8(V^{+}_{z_s,z_s+\Delta z})^T \Pi_0 \mathcal{S}^{+}_{z_s,z_s+\Delta z} \Pi_1^T. \tag{57}$$

Using a receiver datum closer to the source surface has two favorable consequences:

- The computational domain can be smaller than is necessary to simulate the target data, as it need only contain the source surface and the receiver datum implicit in equation 56. This shrinkage of the computational domain can lead to substantial improvements in computational efficiency.
- Since the receiver data may be chosen much closer to the source surface that is the case for the target data, the effective aperture active in the relation 56 can be much larger, producing an estimated source gather much less affected by aperture limitation.

As mentioned in the last section, computation of the transpose of Λ^+ (exact, not approximate in the high frequency sense) is critical to the successful construction of the preconditioner. The relation 56 does not provide a computation for this operator. However set

$$\tilde{\Lambda}^+_{z_s} = -8(V^+_{z_s,z_s+\Delta z})^T \Pi_0 \mathcal{S}^+_{z_s,z_s+\Delta z} \Pi_1^T. \tag{58}$$

Then 57 can be rewritten

$$\Lambda^+_{z_s} \approx \tilde{\Lambda}^+_{z_s}.$$

Of course, all of the examples so far show images of $\tilde{\Lambda}^+_{z_s}$.

Since successful preconditioning requires only approximate inversion, use of $\tilde{\Lambda}^+_{z_s}$ in place of $\Lambda^+_{z_s}$ will still yield a working preconditioner, and the former can be transposed to machine precision via the definition 58 and the adjoint state method (equations 18 19):

$$(\tilde{\Lambda}^+_{z_s})^T = -8\Pi_1(\mathcal{S}^+_{z_s,z_s+\Delta z})^T \Pi_0^T V^+_{z_s,z_s+\Delta z} \tag{59}$$

The model space weight operator W_m^{-1} introduced in the last section is replaced by its asymptotic approximation

$$\begin{aligned}
&\frac{1}{2}(\tilde{\Lambda}^+_{z_s} + (\tilde{\Lambda}^+_{z_s})^T) \\
&\approx -8\left((V^+_{z_s,z_s+\Delta z})^T \Pi_0 \mathcal{S}^+_{z_s,z_s+\Delta z} \Pi_1^T + \right. \\
&\qquad \left. \Pi_1(\mathcal{S}^+_{z_s,z_s+\Delta z})^T \Pi_0^T V^+_{z_s,z_s+\Delta z}\right) \\
&= -4(\Pi_1(\mathcal{S}^+_{z_s,z_s+\Delta z})^T(\Pi_0^T\Pi_1 + \Pi_1^T\Pi_0)\mathcal{S}^+_{z_s,z_s+\Delta z}\Pi_1^T) = \tilde{W}_m^{-1}.
\end{aligned} \tag{60}$$

with a similar definition for the replacement $\tilde{W}_d$ of W_d.

This identity shows that only one forward and one adjoint simulation are necessary to compute the action of $\tilde{W}_{m,d}$. The operator in the center of the expression on the right-hand side, $\Pi_0^T\Pi_1 + \Pi_1^T\Pi_0$, simply exchanges the components of the acoustic fields, passing the velocity field as a pressure source and the pressure field as a velocity source.

One more computation is required for the full implementation of the preconditioning strategy explained in the last section: W_m is required, not just W_m^{-1}. Note that W_m plays two roles in the second term in equation 50: it is the weight matrix for both the domain and range norms for A. It is perfectly OK for one of these to be replaced by an asymptotic approximation, so long as it is symmetric and computable (and at least semi-definite). The second row in equation 33 appears as 39 above: introducing (formally) the inverse of Λ^+,

$$-\frac{1}{2}\Pi_1\mathcal{S}^+_{z_s,z_r}\Pi_0^T \approx V^+_{z_s,z_r}(\Lambda^+_{z_s})^{-1} \tag{61}$$

whence from the second line in display 38

$$-\frac{1}{8}(S^+_{z_s,z_r})^T\Pi_1\mathcal{S}^+_{z_s,z_r}\Pi_0^T \approx (\Lambda^+_{z_s})^{-1} \tag{62}$$

and

$$-\frac{1}{8}\Pi_0(\mathcal{S}^+_{z_s,z_r})^T\Pi_1^T S^+_{z_s,z_r} \approx ((\Lambda^+_{z_s})^{-1})^T. \tag{63}$$

Using the definition 12 of $S^+_{z_s,z_r}$, the symmetrized Λ^{-1} is

$$\tilde{W}_m = -\frac{1}{16}\left(\Pi_0(\mathcal{S}^+_{z_s,z_r})^T(\Pi_1^T\Pi_0 + \Pi_0^T\Pi_1)\mathcal{S}^+_{z_s,z_r}\Pi_0^T\right) \approx \frac{1}{2}((\Lambda^+_{z_s})^{-1} + ((\Lambda^+_{z_s})^{-1})^T). \tag{64}$$

Comparison with the definition 60 shows that $\tilde{W}_m$ and $\tilde{W}_m^{-1}$ differ only in the initial and final projection factors (and overall scale), and in particular either can be computed for the cost of a forward/adjoint operator pair. Note that $\tilde{W}_m^{-1}$ is inverse to $\tilde{W}_m$ only in an approximate (asymptotic, aperture-limited) sense.

NUMERICAL EXAMPLES

This section illustrates the most important conclusions developed in the preceding sections by finite difference wavefield simulation.

Synthetic models and simulation

To illustrate the structure described in the preceding section, I introduce two 2D acoustic models, one spatially homogeneous, the other highly refractive. The first, homogenous model has $\kappa = 4$ GPa and $\rho = 1$ g/cm^3 throughout a rectangular domain of size 8 km (x) $\times$ 4 km (z). The second, refractive, model is a perturbation of the first by a low-velocity acoustic lens positioned in the center of the rectangle (Figure 1. To produce this structure, the density is chosen homogeneous as in the first model, while the bulk modulus decreases to from 4 GPa outside the lens to 1.6 GPA in its center, as shown in Figure 1.

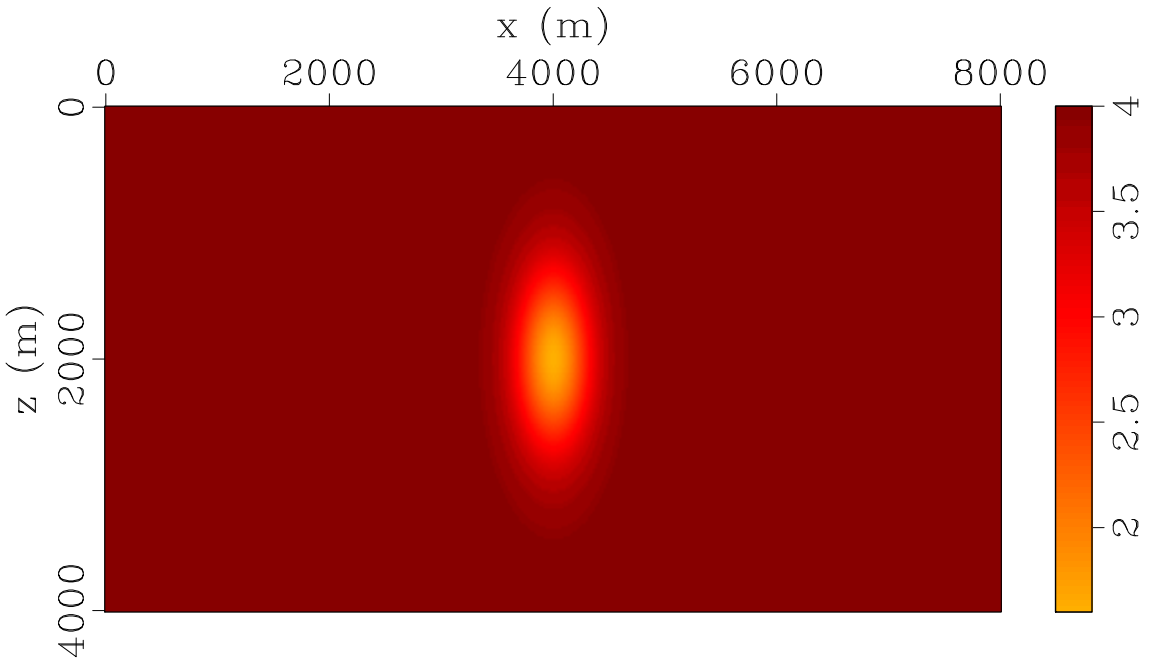

Figure 1: Bulk modulus, lens model. Color scale is in GPa. Positions of point source and receiver line indicated.

Discretization is conventional, with a rectangular grid and staggered finite difference scheme (Virieux, 1984) of order 2 in time and $2k$ in space; for most of the experiments reported below, $k = 4$. Absorbing boudary conditions of perfectly matched layer type are applied at all boundaries of the simulation rectangle (Hu et al., 2007). Sampling operators such as P_r are implemented via linear interpolation, and source insertion via adjoint linear interpolation (as noted above, in the continuum limit, sources are represented via adjoint sampling). Steps in x and z are the same. In the following examples, $\Delta x = 20$ m. This choice limits the temporal frequency of accurately computed fields to rougly 12 Hz.

Symes et al. (2011) gives a description of the code implementation, out-of-date in a few respects but overall accurate. The implementation uses the discrete adjoint state method and auto-generated code (Hascoët and Pascual, 2013), to assure that the computed adjoint operators are adjoint at the level of machine precision to the computed operators. The reverse-time storage issue is resolved through the optimal checkpointing technique (Griewank, 2000; Symes, 2007), again without loss of precision. This procedure results in computed adjoints for $S^+_{z_s,z_r}$ and other operators that pass usual test for adjoint accuracy, comparing inner products with pseudorandom input vectors, with errors well under machine precision.

The horizontal line of receivers sits at depth $z_r = 1000$ m, the (extended) sources at $z_s = 3000$ m. Source and receiver x ranges from 2000 to 6000 m. Note that we have reversed the order relation between z_s and z_r described in the text ($z_s < z_r$). This

difference is immaterial for the purpose of illustrating the mathematical structures developed in the preceding paragraphs.

Creating downgoing fields

The downgoing condition constrains high-frequency energy of localized plane wave components, hence could be enforced by dip¿ filtering. However, a simpler approach is to construct fields that must be entirely downgoing at the source and receiver surfaces by virtue of ray geometry.

Note that a point source on $z = z_s$ creates high frequency energy traveling on rays parallel and nearly parallel to $z = z_s$, so that won't do. However, placing a point source at a depth $z_d < z_s$ will work. Since the examples used here are homogeneous in $z < z_s$, and the sampling region for extended sources is a finite interval, all rays carrying high frequency energy cross the source surface $z = z_s$ at a postive angle, and the field and its traces are *a priori* downgoing. The same is obviously true at the receiver surface for the homogeneous model, but is also true for the lens model, as no rays are refracted horizontally *at the receiver surface*.

The choice of a point source at $z_d = 3500$ m, $x_d = 3500$ m, bandpass filter wavelet with corner frequencies $1, 2.5, 7.5, 12.5$ Hz, gives the causal pressure and velocity gathers at $z = z_s = 3000$ m shown in Figures 2 and 3. Since the mechanical parameters in the homogeneous and lens models are the same for $z < z_s$, and no rays return to this zone in either model, these data are asymptotically the same for both models, and I show only the homogenous medium results.

Equivalence of pressure, velocity sources

These gathers are the pressure and velocity traces $(P_s p^+, P_s v_z^+)$ on $z = z_s$ of a downgoing acoustic field in $z < z_s$, hence related by the operator $\Lambda_{z_s}^+$. Equations 33, 30 and 31 show that these differ by a factor of -2 from source functions f_s and h_s in the system 9, with $h_s = 0$ and $f_s = 0$ respectively, that generate the same acoustic field in $z < z_s$, and in particular the same receiver traces on $z = z_r$, at least asymptotically.

Figures 4, 5, 6 show the pressure gathers extracted at $z_r = 1000$ m for the point source at $z = z_d$ and for the two choices of extended source at $z = z_s$, on the same color scale. The obvious similarity between the fields generated by the two extended sources, predicted by equation 33, is confirmed by trace comparisons in figures 7,8. Other traces are equally similar.

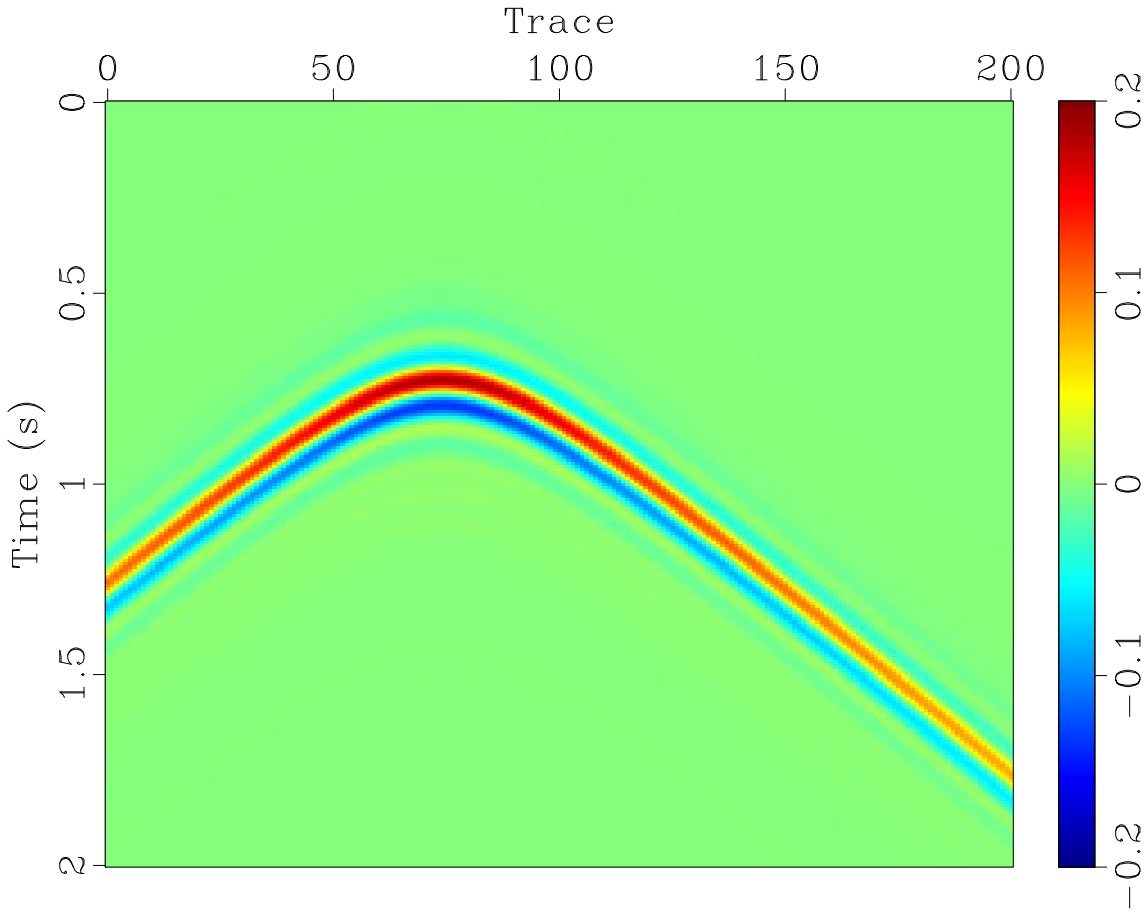

Figure 2: Trace $P_s p^+$ on $z = z_s = 3000$ m of pressure field from point source at $z_d = 3500$ m, $x_d = 3500$ m, bandpass filter source.

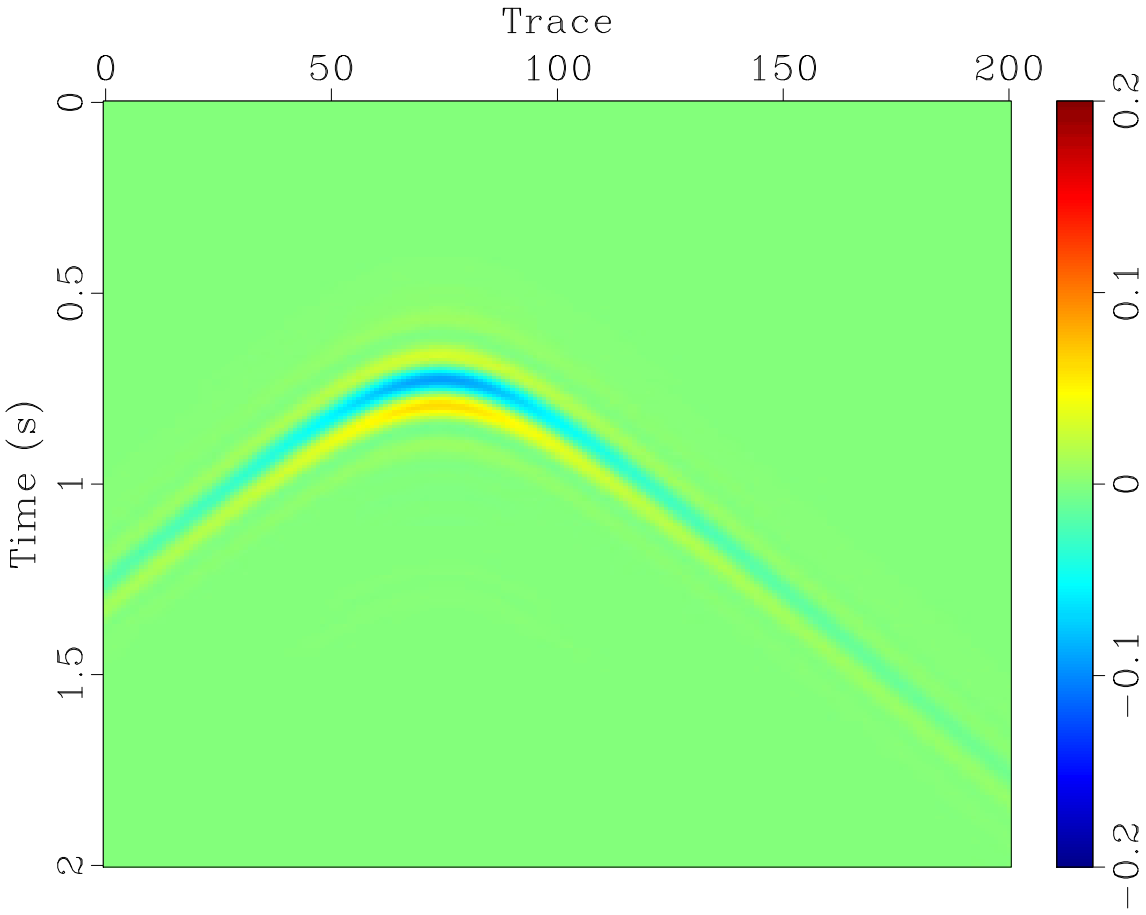

Figure 3: Trace $P_s v_z^+$ on $z = z_s = 3000$ m of vertical velocity field from point source at $z_d = 3500$ m, $x_d = 3500$ m, bandpass filter source.

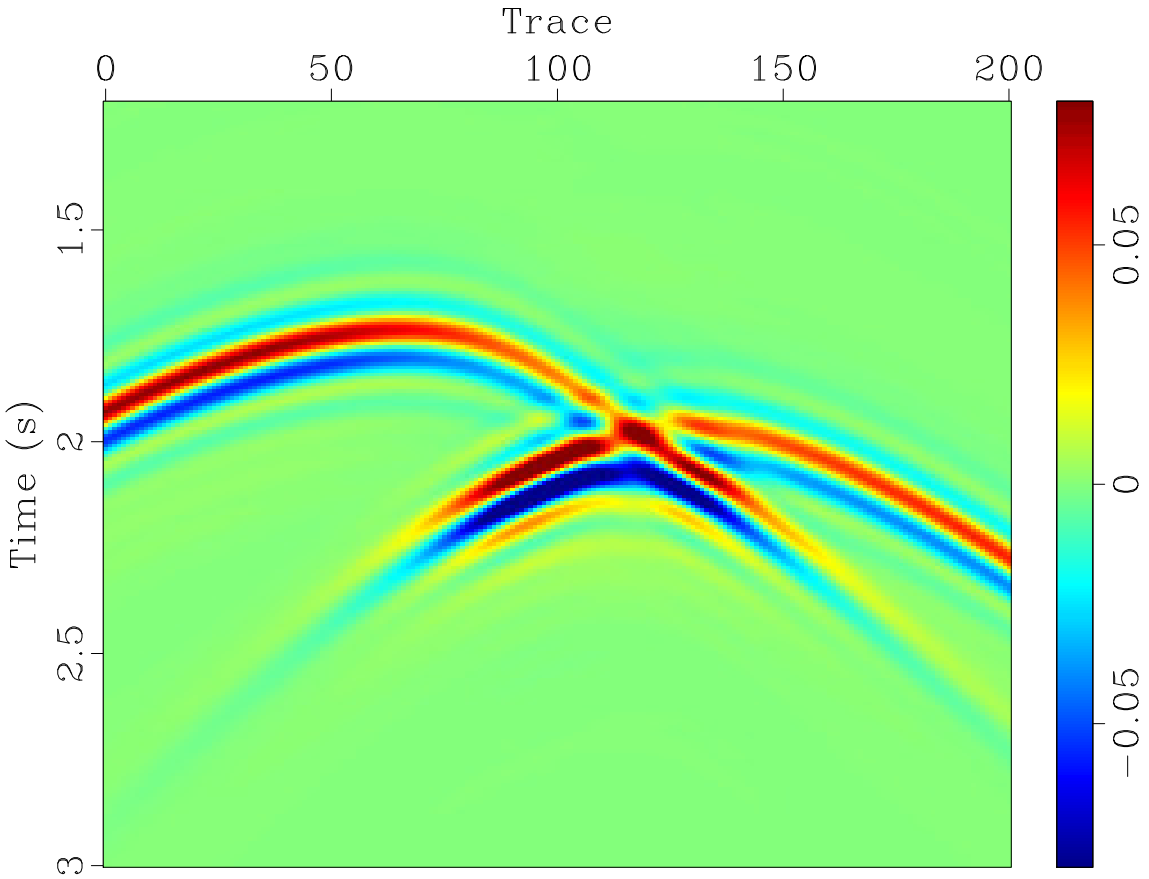

Figure 4: Pressure gather at receiver depth $z_r = 1000$ m from field generated by causal solution of acoustic system 9 in the lens model described in the text, with point pressure source (constitutive defect) at $z_d = 3500$ m, $x_d = 3500$ m.

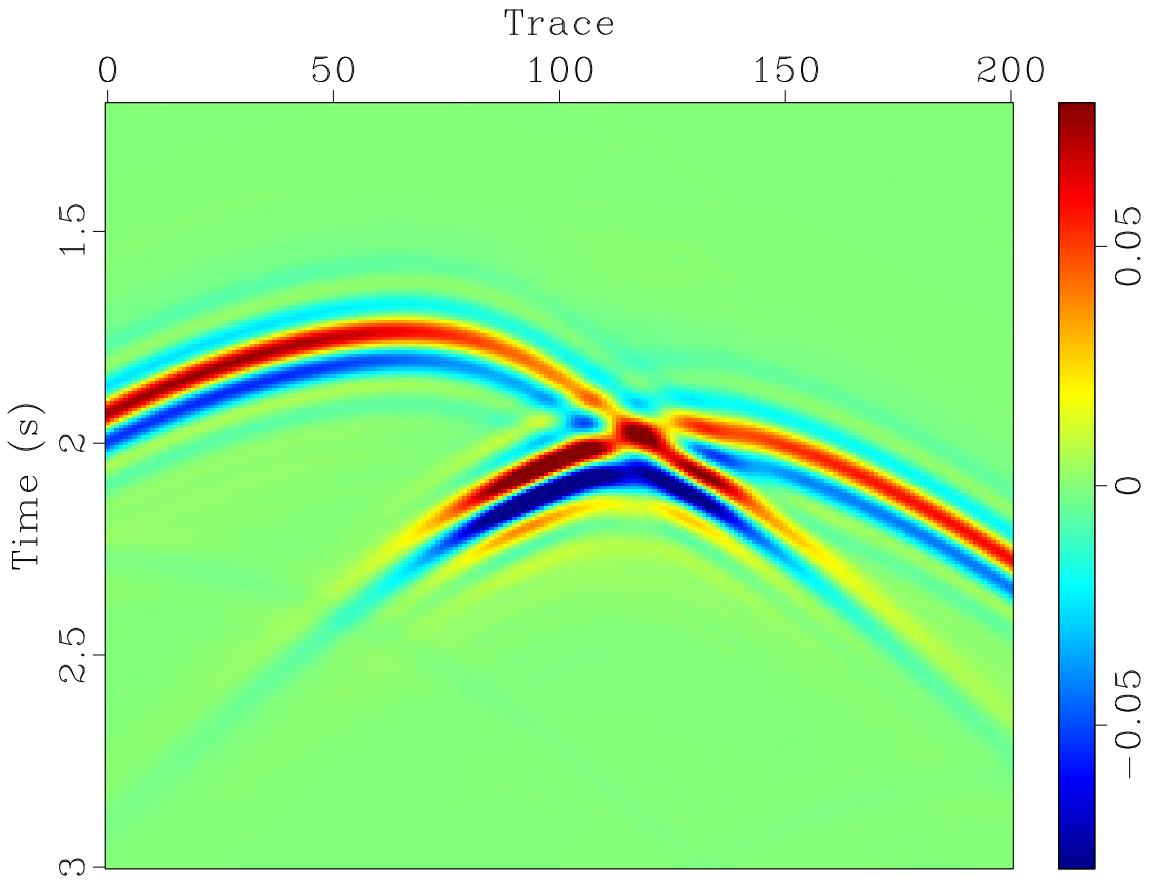

Figure 5: Pressure gather at receiver depth $z_r = 1000$ m from field generated by causal solution of acoustic system 9 in the lens model described in the text, with extended pressure source (constitutive defect) on $z = z_s = 3000$ m given bythe field depicted in Figure 3 scaled by -2 ($h_s = -2P_s v_z^+ = \Lambda_{z_s}^+ P_s p^+$) and zero velocity source (vertical load) ($f_s = 0$).

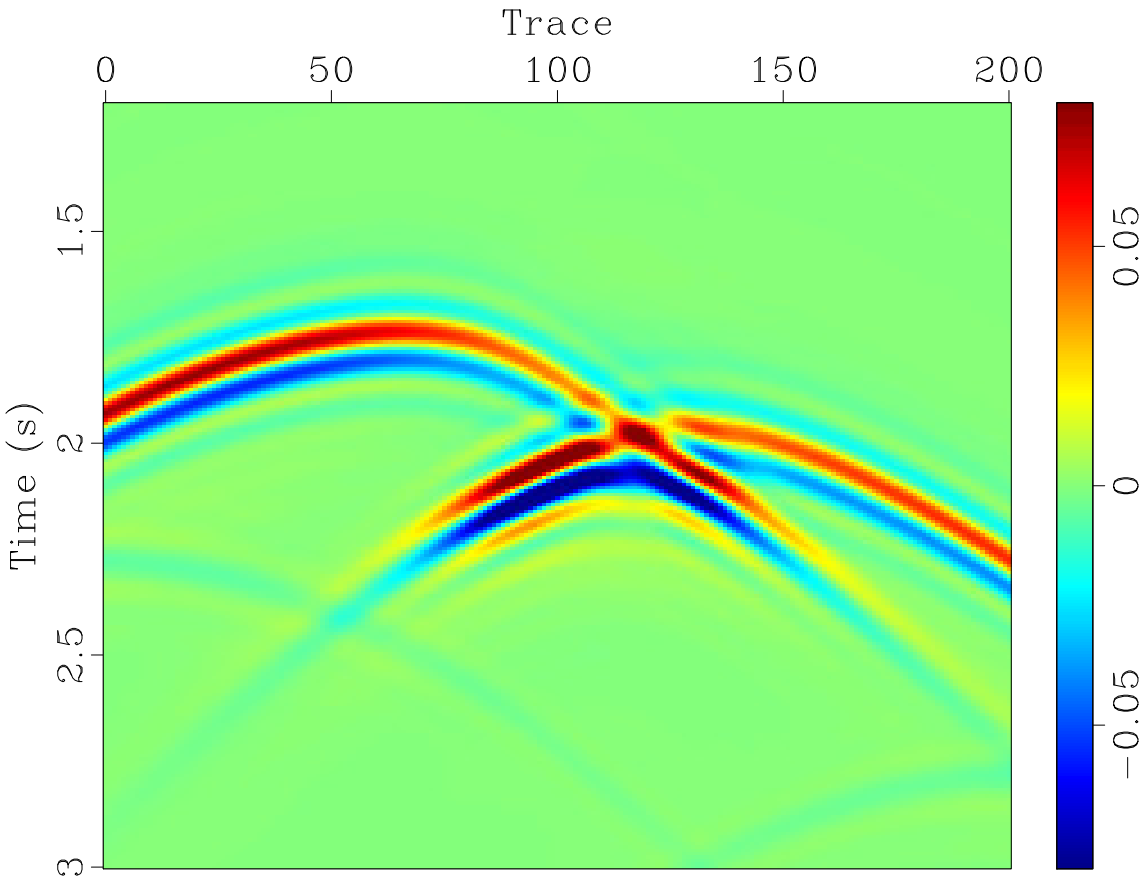

Figure 6: Pressure gather at receiver depth $z_r = 1000$ m from field generated by causal solution of acoustic system 9 in the lens model described in the text, with extended velocity source (vertical load) on $z = z_s = 3000$ m given by the field depicted in Figure 2 scaled by -2 ($f_s = -2P_s p^+$) and zero pressure source (constitutive defect) ($h_s = 0$).

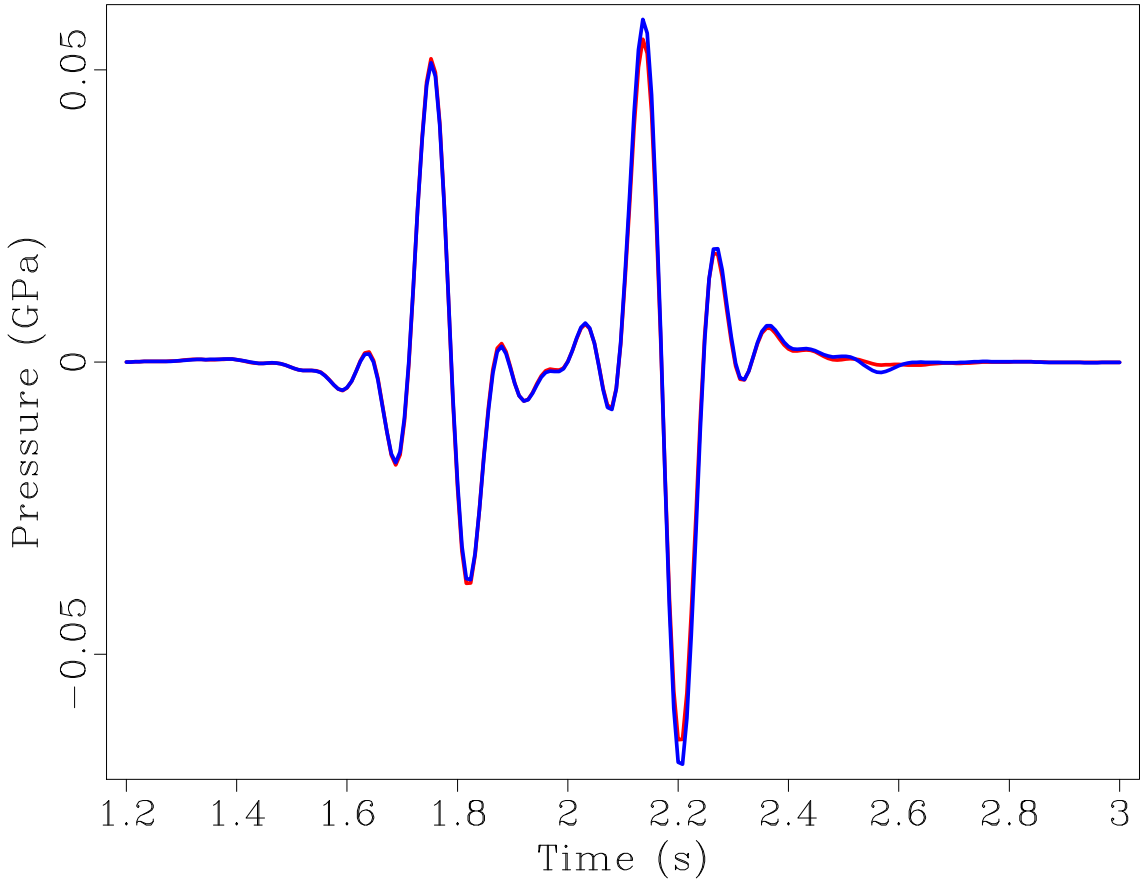

Figure 7: Overplot of traces 81 ($x = 3600$) from gathers shown in 4 (blue), 5 (red).

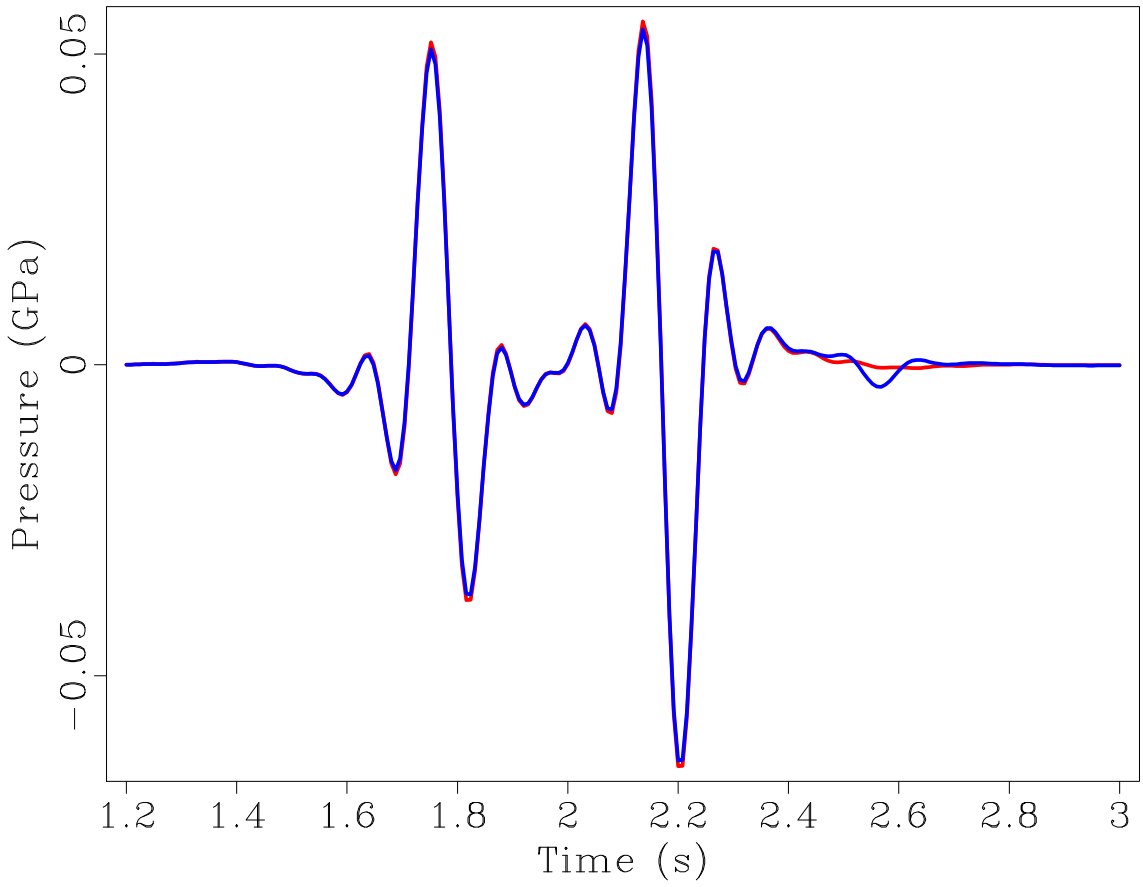

Figure 8: Overplot of traces 81 ($x = 3600$) from gathers shown in 4 (blue), 6 (red).

Inversion by time reversal

I have applied the approximate inversion procedure suggested in equation 38 to the pressure gather shown in Figure 4, generated by a point source at $z_d = 3500$ m, $x_s = 3500$ m, propagating in the lens model (Figure 1). I choose this example for two reasons. First, the success of the inversion demostrates the insensitivity of the time reversal method to ray multipathing (triplication), evident in the data (Figure 4). Second, I will invert this data in the homogeneous model, that is, construct sources that (approximately) reproduce the data using a different material model than the one in which it was produced. This capability is critically important in the application of the approximate inversion in nonlinear inversion, where the early iterations involve solution of the source estimation problem 6 at (possibly very) wrong material models $\mathbf{c}$. Successful extension methods maintain data fit throughout the course of the inversion.

As noted earlier, the acoustic field in this example is downgoing throughout the simulation range. It can be regarded as the result of either pressure or velocity source at $z = z_s$: the pressure source gather h_s (Figure 9) is -2 times the vertical velocity gather depicted in Figure 3, the velocity source gather f_s (Figure 10) is -2 times the pressure gather depicted in Figure 2.

Figure 11 shows the approximate inversion (via the first equation in display 38) of the pressure gather shown in Figure 4, inverted in the homogeneous model (rather than in the lens model used to generate the data). The result differs greatly from the pressure source shown in Figure 9, as it must since it results from inversion in the wrong material model. Some dip filter effect is unavoidable and is caused by the aperture limitation of the acquisition geometry: the steeper dips in the source gather (Figure 9) do not contribute to the data, nor to the inversion. Also, the limited receiver aperture causes truncation artifacts in the inversion. However, this result is an accurate inversion: re-simulation (application of $S^+_{z_s,z_r}$) *using the same (homogeneous) model as used in the inversion* results in accurate recovery (Figure 12) of the input pressure gather (Figure 4). The difference is shown on the same color scale in Figure 13.

Quasi-unitary property of the modeling operator

The identities 38 and 40 would together establish the approximately unitary property of $S^+_{z_s,z_r}$, if Λ were symmetric. Identity 38 was illustrated in the last subsection. Setting the symmetric issue aside for the moment, an illustration of the relation 40 proceeds as follows.

Relation 30 characterizes $\Lambda^+_{z_s}$ as connecting the pressure and velocity components of downgoing fields restricted to $z = z_s$. That is, the pressure source gather $h_s = \Lambda^+_{z_s} P_s p^+ = -2P_{z_s} v_z$, displayed in Figure 9, is the image of the pressure gather in Figure 2 under $\Lambda^+_{z_s}$. The pressure gather 5 is the image of this pressure source gather

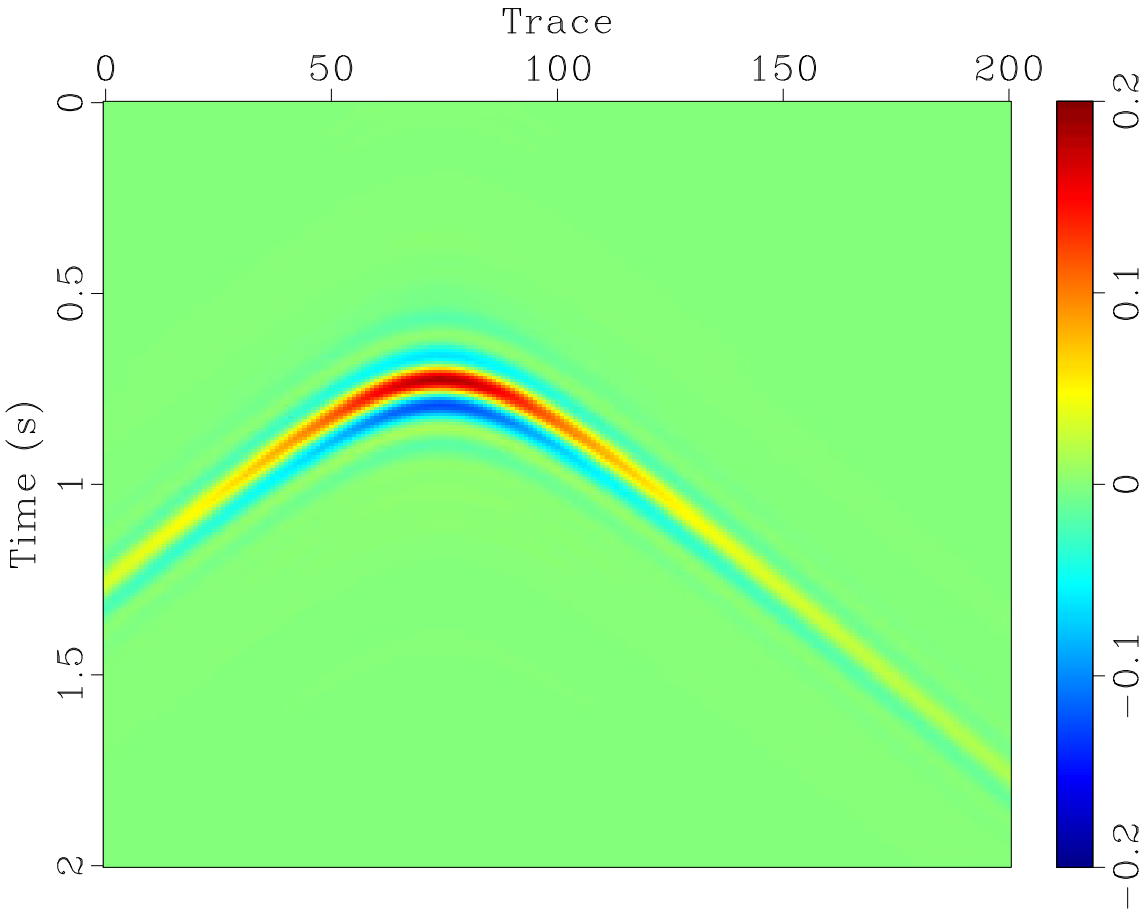

Figure 9: Pressure source gather = -2 × vertical velocity gather (Figure 3).

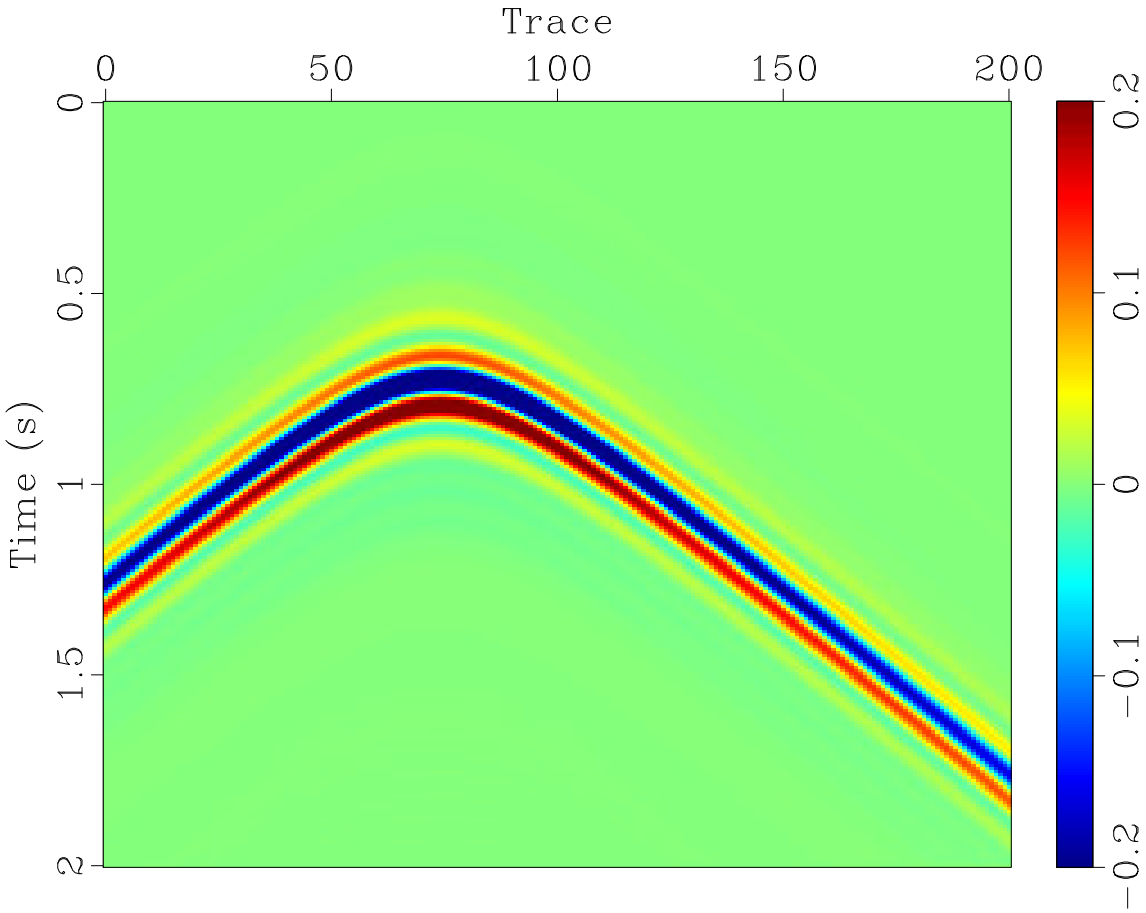

Figure 10: Velocity source gather = -2 × pressure gather (Figure 2).

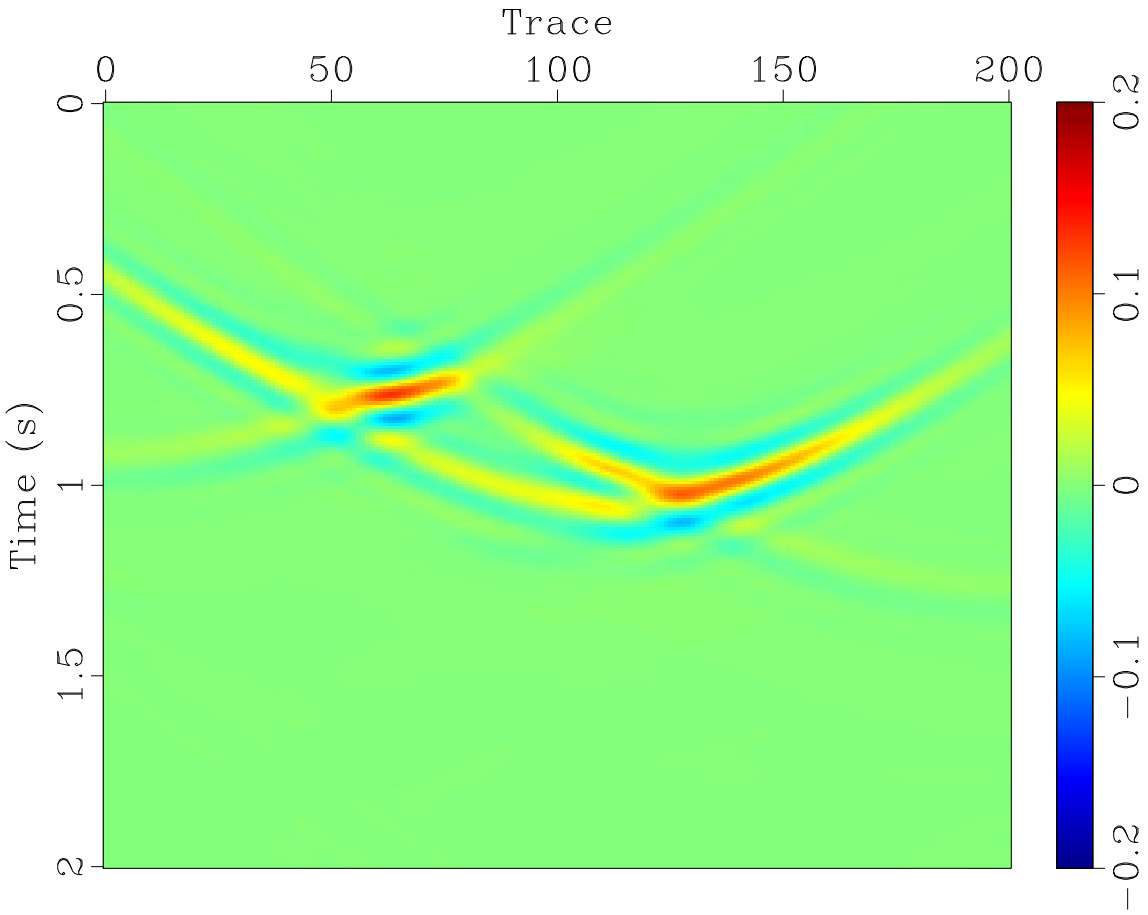

Figure 11: Approximate inversion via first equation in display 38. Inversion in homogenous model of the pressure gather in Figure 4, simulated with lens model. Scaled version of output v_z field obtained by applying transpose of $4V^+_{z_s,z_r}$. Quite different from pressure source gather (Figure 9) used to generate data - since inversion takes place in a different material model!

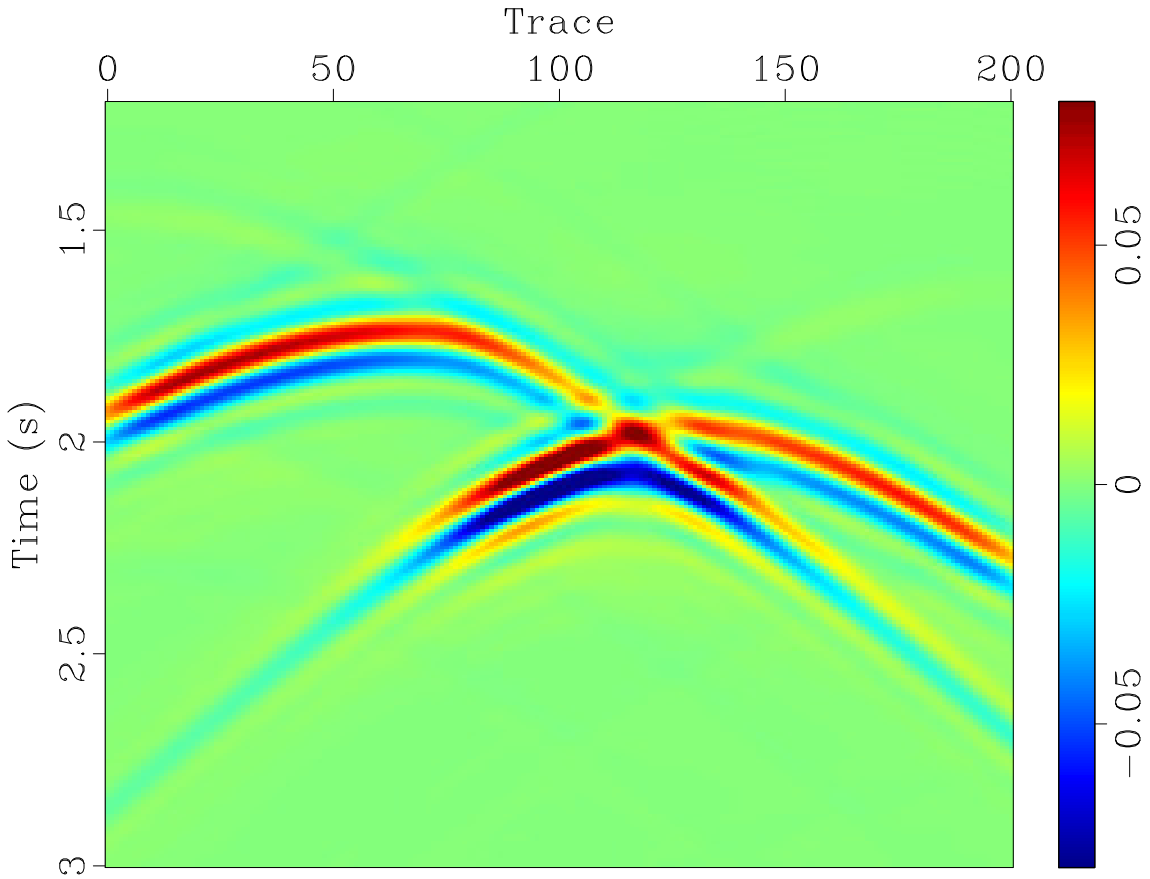

Figure 12: Re-simulated pressure gather produced from inverted source shown in Figure 11. Simulation in homogenous model used for inversion.

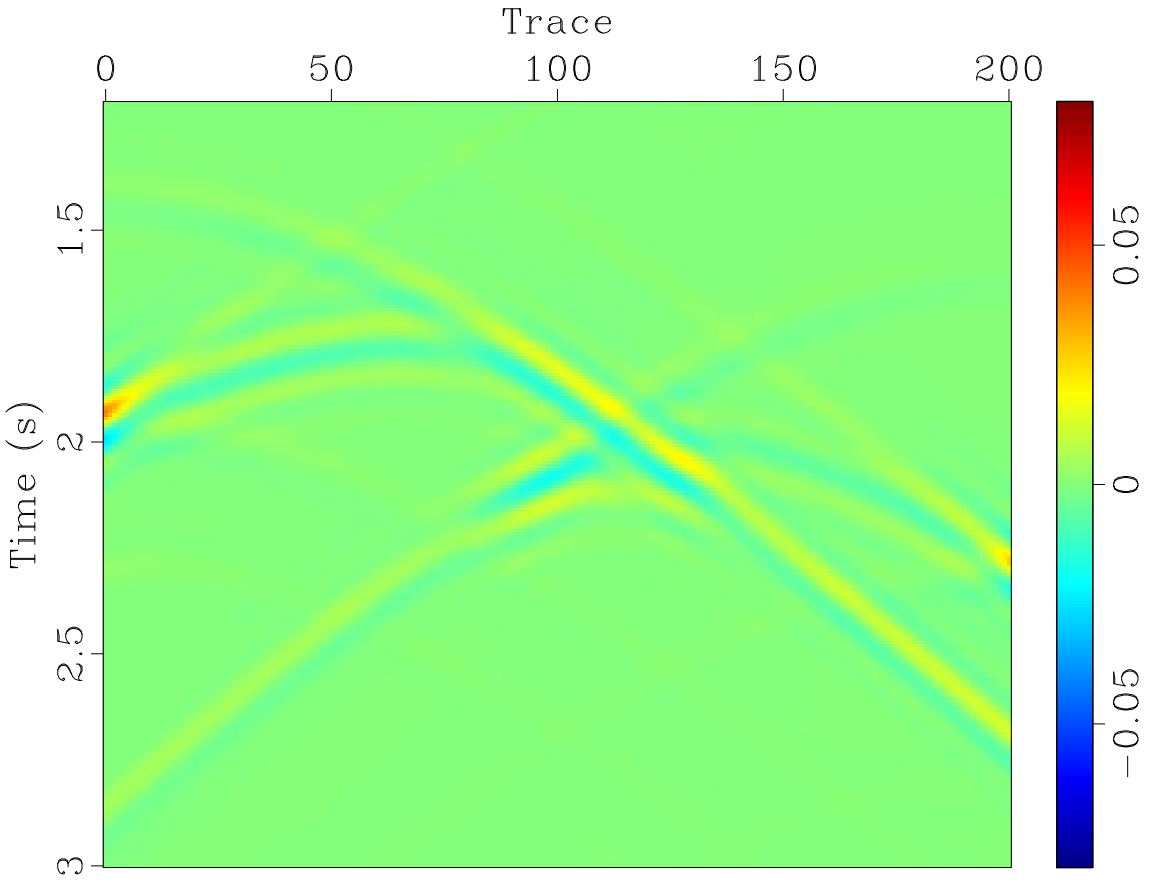

Figure 13: Difference between gathers displayed in Figures 4 and 12, plotted on same color scale.

under $S^+_{z_s,z_r}$ (using the lens model). The corresponding vertical velocity gather (Figure 14) is $-1/2$ times $\Lambda^+_{z_r} P_r p$. Therefore scaling the data in Figure 14 by -2 produces $\Lambda^+_{z_r} S^+_{z_s,z_r} \Lambda^+_{z_s} P_s p^+$. On the other hand, figure 15 shows the result of applying $V^+_{z_s,z_r}$ to $f_s = -2P_s p^+$. Therefore scaling the gather in Figure 15 by $-\frac{1}{2}$ produces $V^+_{z_s,z_r} P_z p^+$. Since the data in Figures 14 and 15 are essentially identical, the relation 40 holds for this example.

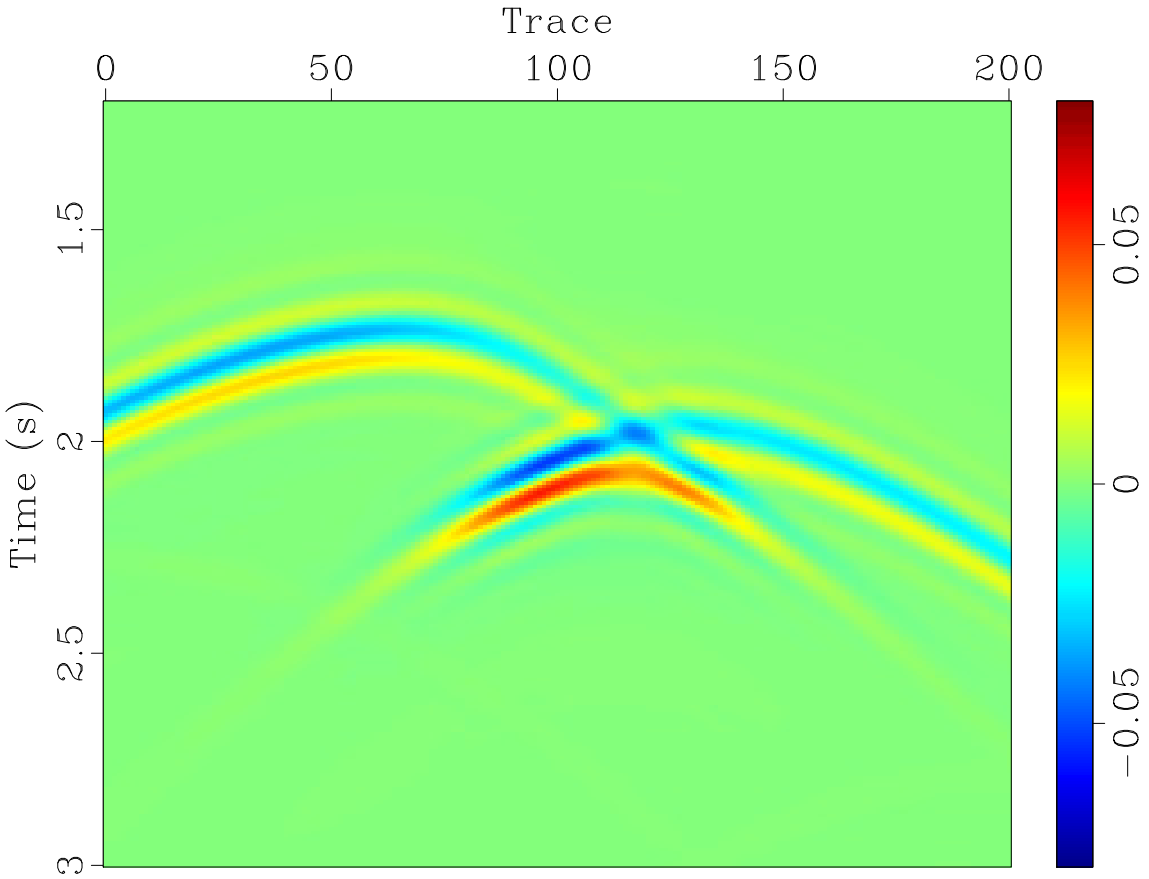

Figure 14: Vertical velocity gather, generated with a pressure source in the lens model, corresponding to pressure gather 5.

Economical computation of Λ in "thin" subdomain

Equation 57 suggests a thin-slab computation of the Λ action, which is both accurate and economical. This calculations place a receiver array at $z_s + \Delta z = 2900$ m depth, just 100 m above the source surface at $z_s = 3000$ m. For the discretization used to create the examples shown so far, that is just a 5 gridpoint difference in depth, as opposed to 100 gridpoints between the source and receiver depths for examples such as shown in Figure 4.

Asymptotically, $\Lambda^\pm_{z_s}$ depends only on the medium coefficients $\mathbf{c}$ in an arbirarily small region containing the source surface $z = z_s$. In this example, the homogeneous

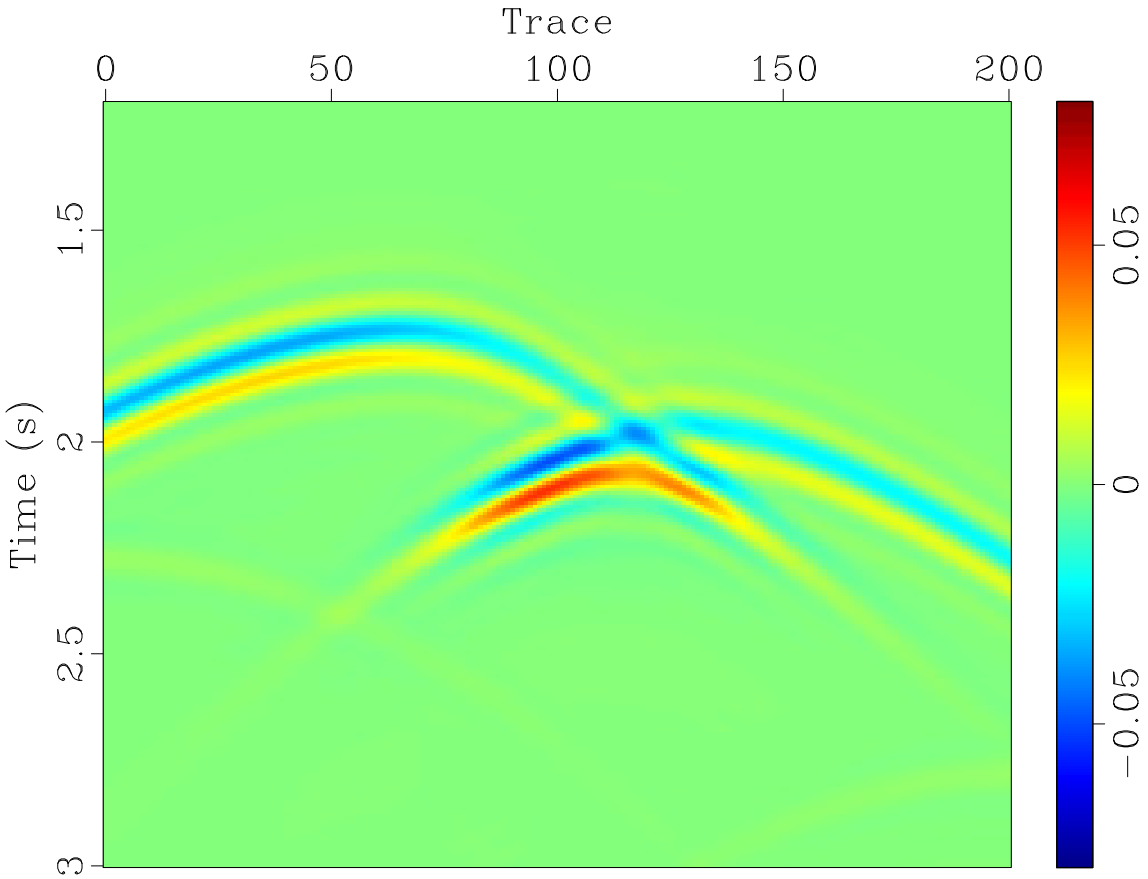

Figure 15: Vertical velocity gather, generated with a velocity source in the lens model, corresponding to pressure gather 6.

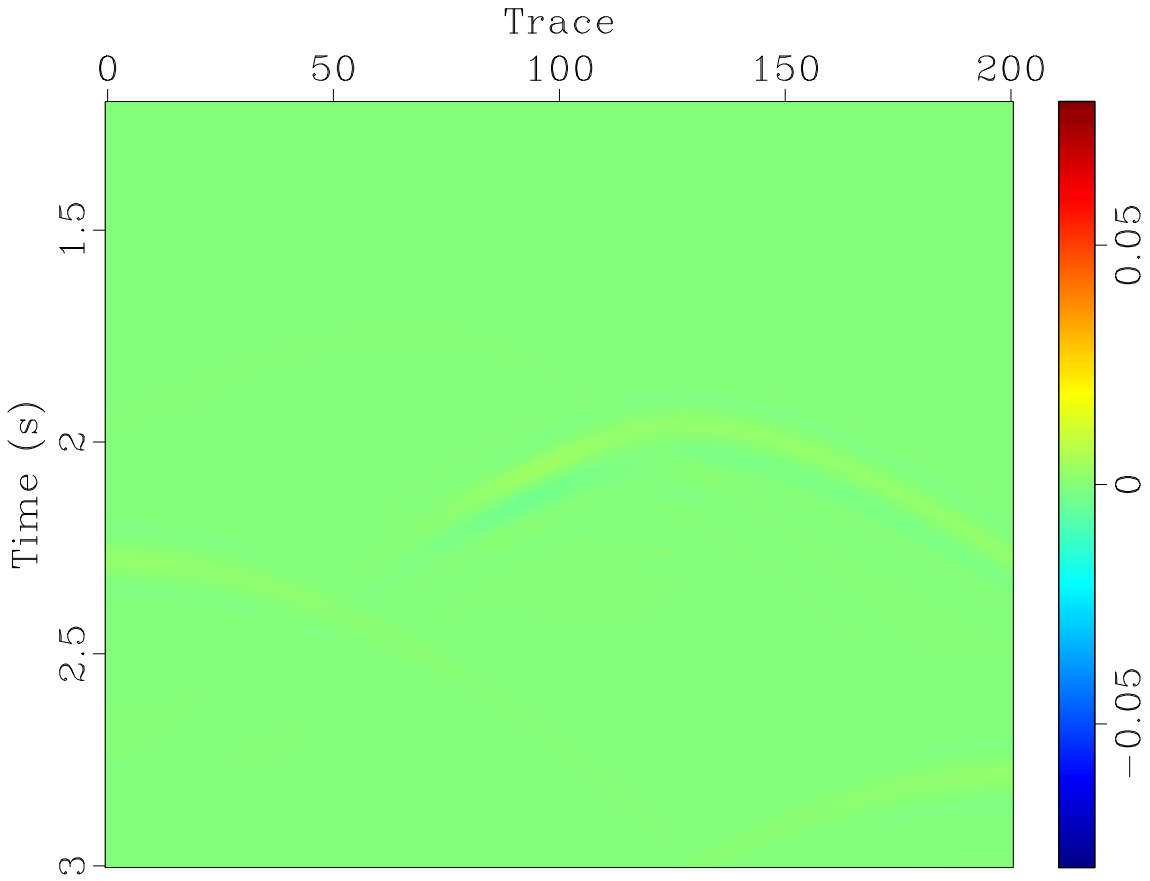

Figure 16: Difference between velocity gathers shown in Figures 14 and 15, plotted on the same color scale as these figures.

and lens models are identical in the depth range $2900 < z < 3000$ m, so the computed $\Lambda^+_{z_s}$ operators will be precisely the same for both models. Hence I show only results for the homogenous model.

The approximation to $\Lambda^+_{z_s}$ via equation 57 for this configuration is evaluated in Figures 17, 18, 19, and 20. The effect of aperture limitation is clearly diminished: the second figure in this series compares the full-aperture pressure source gather (Figure 9) with the image of the corresponding pressure gather (Figure 2) under the approximation to $\Lambda^+_{z_s}$, and the last figures show that the approximated source gathers accurately predict the point-source pressure gather at the receiver datum $z_r = 1000$ m.

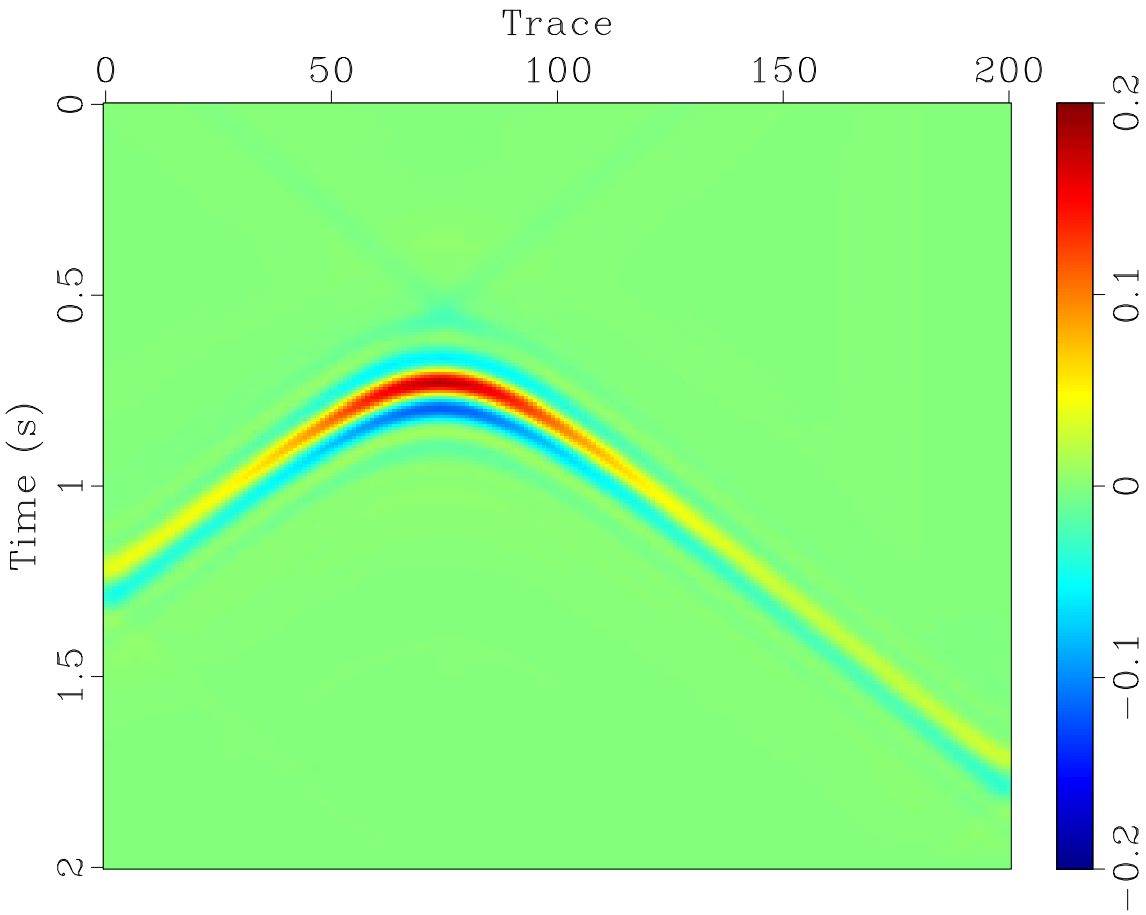

Figure 17: Pressure source gather = image under pressure-to-source operator $\Lambda^+_{z_s}$ of pressure gather shown in Figure 2, homogeneous model, using "near" receiver traces at $z = 2900$ m. Compare Figure 9: because the sources and receivers are close, little aperture is lost in this case.

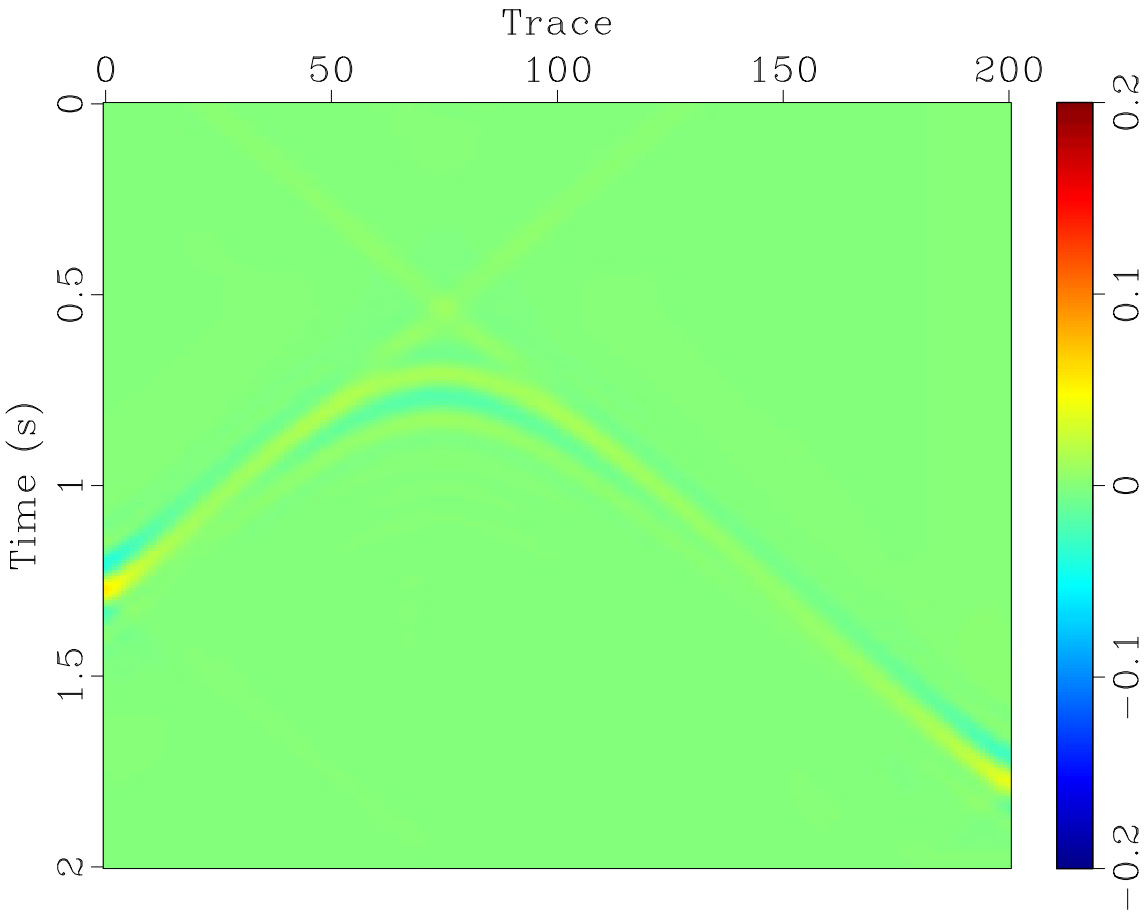

Figure 18: Difference between (a) image (Figure 17) of $\Lambda_{z_s}^+$ applied to pressure gather (Figure 2) using a near receiver array to implement formula 56, and (b) source gather (Figure 9) inferred from vertical velocity. Homogeneous model used in all propagations. Same color scale as in Figure 17.

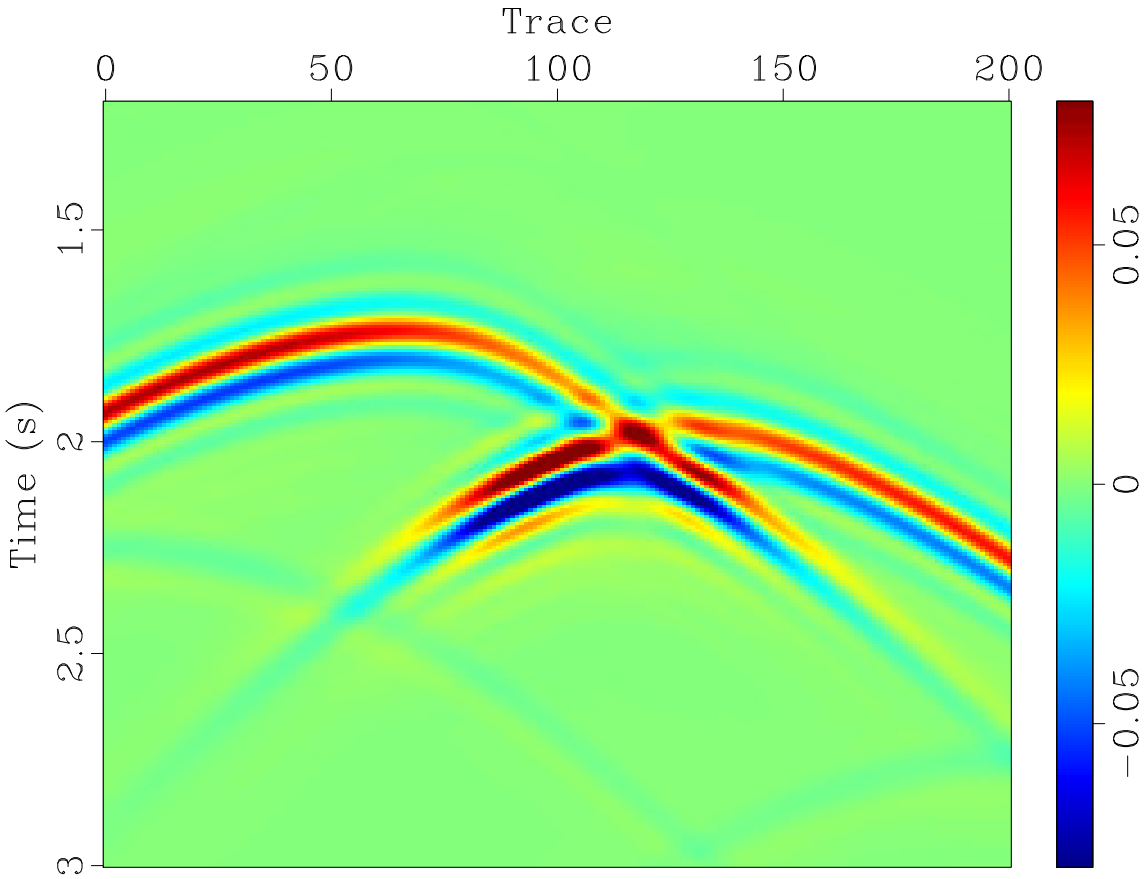

Figure 19: Pressure gather at receiver datum $z = z_r = 1000$ m simulated in lens model from source gather shown in Figure 17. Compare with point source pressure gather (Figure 4).

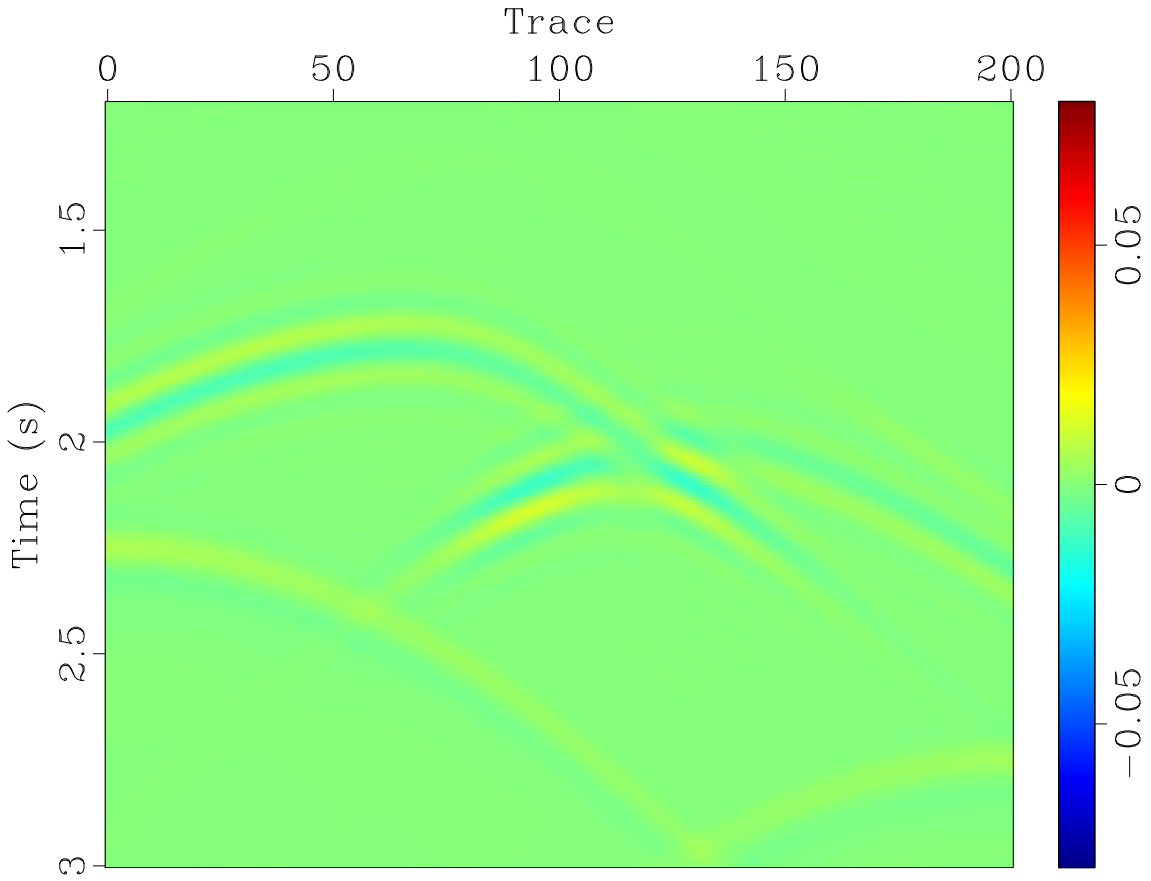

Figure 20: Plot of difference between data shown in Figures 4 and 19, plotted on same color scale as the latter two figures.

Symmetrizing Λ

Figure 21 shows the image of the pressure gather in Figure 2 under $(\tilde{\Lambda}^+_{z_s})^T$, using the "near" traces at $z = 2900$, that is, $\Delta z = 100$ m in expression 59, and propagation in the homogeneous model. Note the close resemblance to the image of the same pressure gather under $\tilde{\Lambda}^+_{z_s}$ displayed in Figure 17. The difference of these two images is displayed in 22, on the same color scale as the images themselves. Since the propagation takes place entirely in a region where all of the mechanical parameters are homogenous, I do not offer a similar comparison for the lens model.

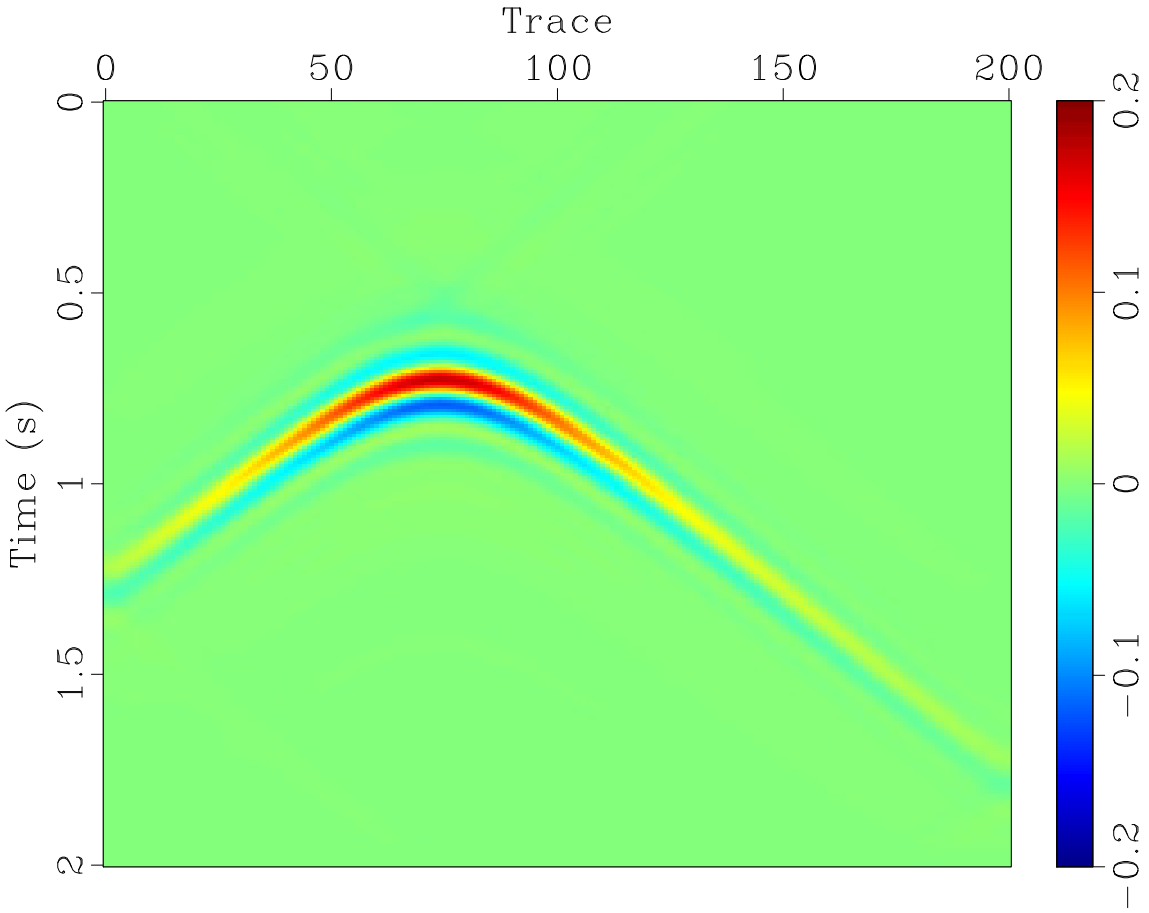

Figure 21: Pressure source gather = image under *transpose* of pressure-to-source operator $\Lambda^+_{z_s}$ of pressure gather shown in Figure 2, homogeneous model, using "near" receiver traces at $z = 2900$ m. Compare Figure 9 and 17: as noted in the text, $\Lambda^+_{z_s}$ is asymptotically symmetric, so the resemblance is not a surprise.

Asymptotic symmetry of Λ

Figure 23 shows the output of the symmetrized approximate source-to-pressure operator per equation 60, applied once again to the pressure data in Figure 2. Note the

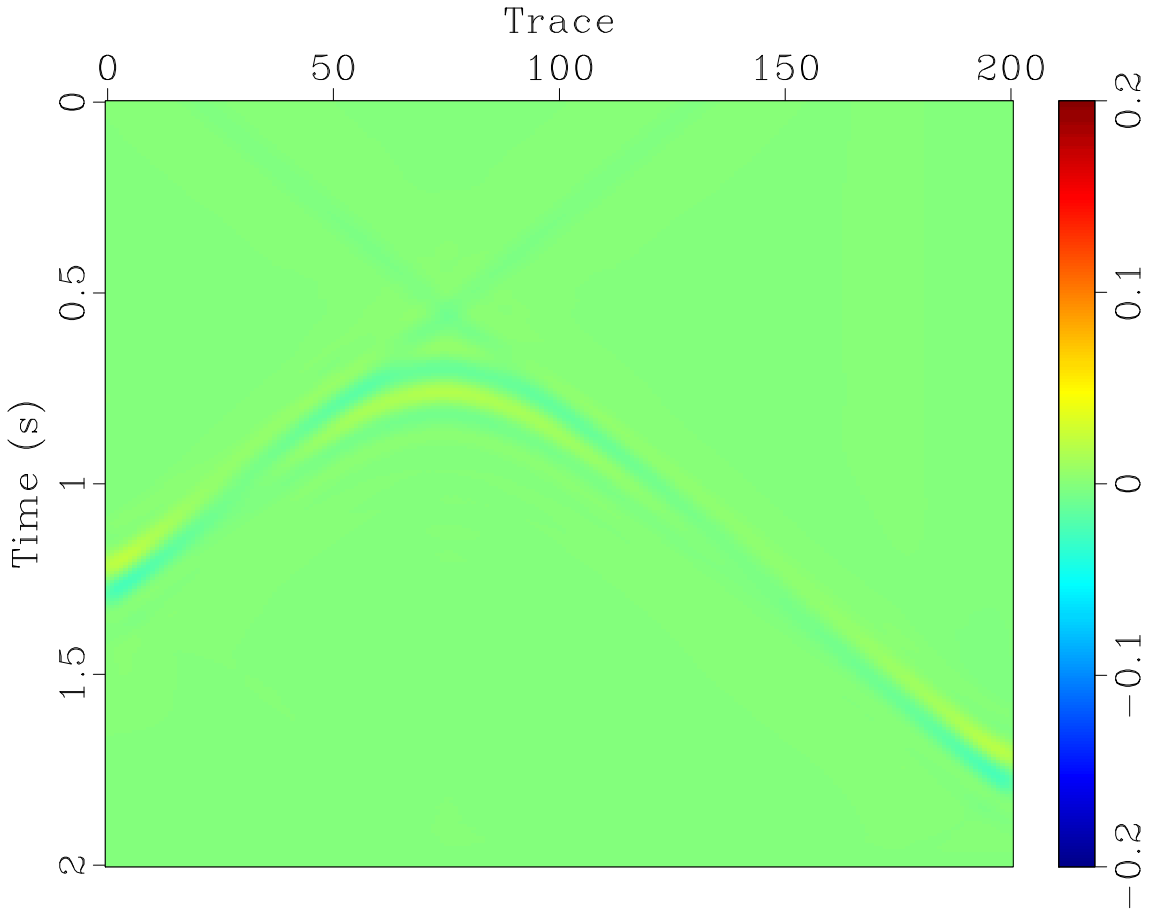

Figure 22: Difference between data in Figures 17 and 21, plotted on the same scale as these figures, showing that the asymptotic symmetry of $\Lambda^+_{z_s}$ is actually quantitative for the length, time and frequency scales of these examples.

resemblance to Figures 17 and 21. These are all asymptotic approximations of each other. Figure 24 shows the the difference between the pressure gather at $z = z_r$ produced from the pressure source output by the symmetrized Λ, and the point source simulation (Figure 5), plotted on the same scale as the latter, in both cases with all propagations in the lens model.

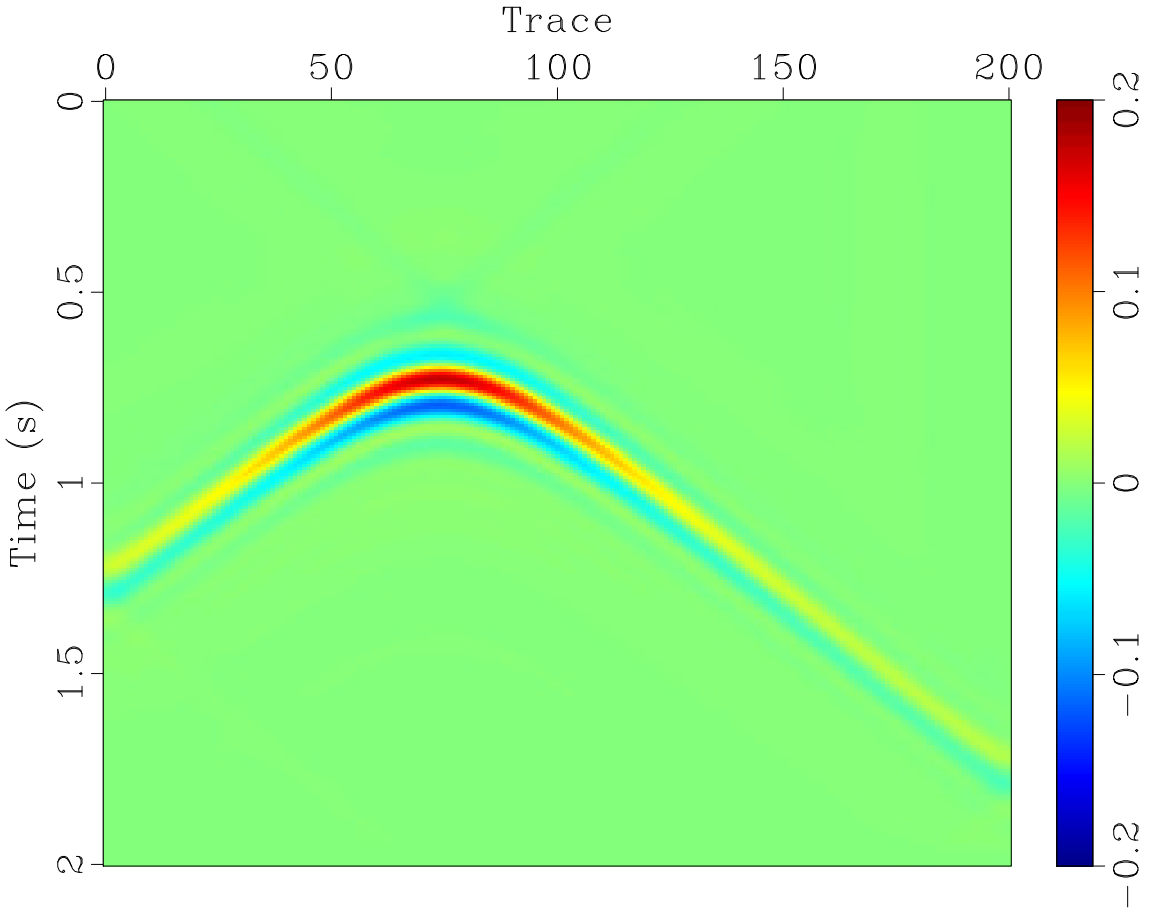

Figure 23: Pressure source gather = image under *symmetrized* pressure-to-source operator $\frac{1}{2}\left(\Lambda^+_{z_s} + (\Lambda^+_{z_s})^T\right)$ of pressure gather shown in Figure 2, homogeneous model, using "near" receiver traces at $z = 2900$ m. Compare Figure 17.

Unitary property of modeling operator

To illustrate this unitary property of $S^+_{z_s,z_r}$, I apply the operator

$$\frac{1}{2}((\Lambda^+_{z_s})^T + \Lambda^+_{z_s})(S^+_{z_s,z_r})^T \frac{1}{2}((\Lambda^+_{z_r})^T + \Lambda^+_{z_r})$$

to the data $S^+_{z_s,z_r} h_s$ (Figure 12), in which h_s is the downgoing source created earlier (Figure 9)The operator above is computed via the technique explained in the preced-

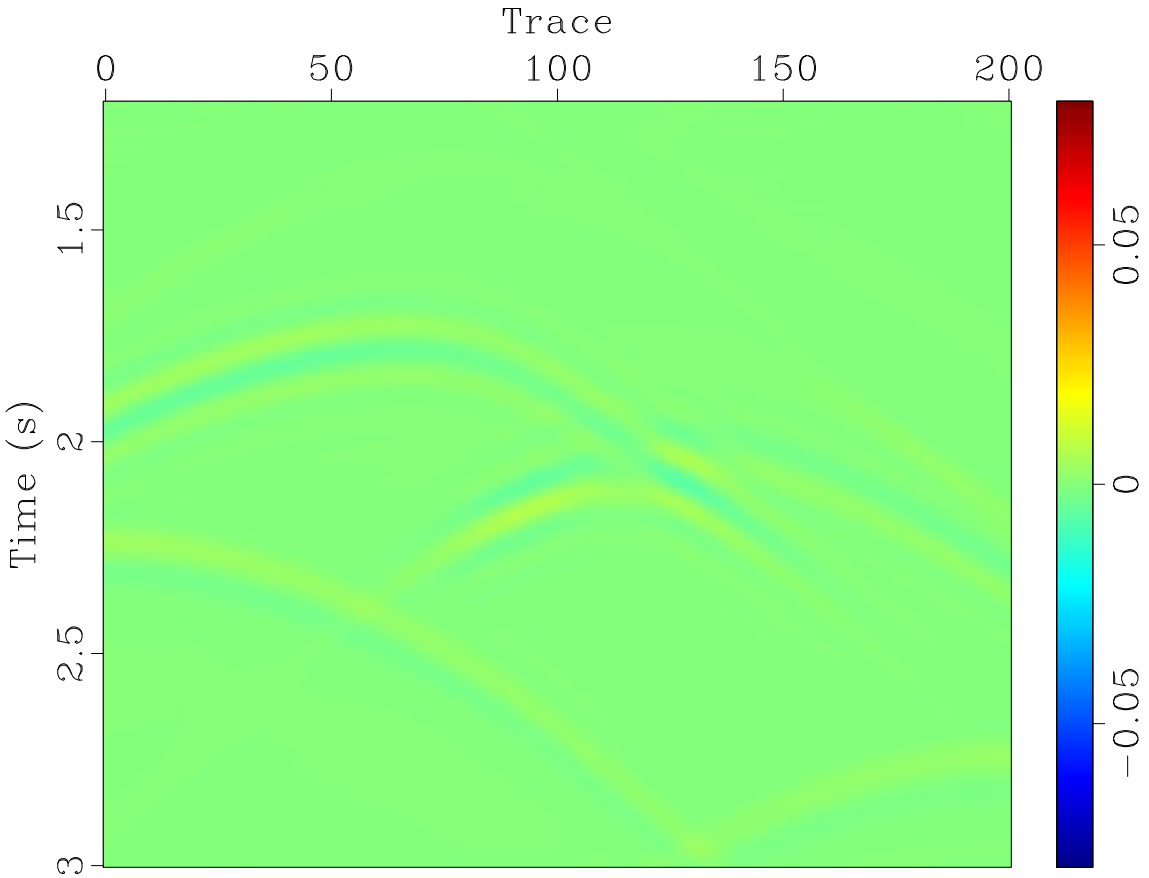

Figure 24: Difference between point source simulation (Figure 5) and pressure gather at $z = z_r = 1000$ m produced by simulation with the source shown in Figure 23, propagation in the lens model.

ing subsection, below, using auxiliary receiver arrays 100 m above the data source and receiver arrays.

The output is shown in Figure 25. The difference with the actual source is shown in Figure 26.

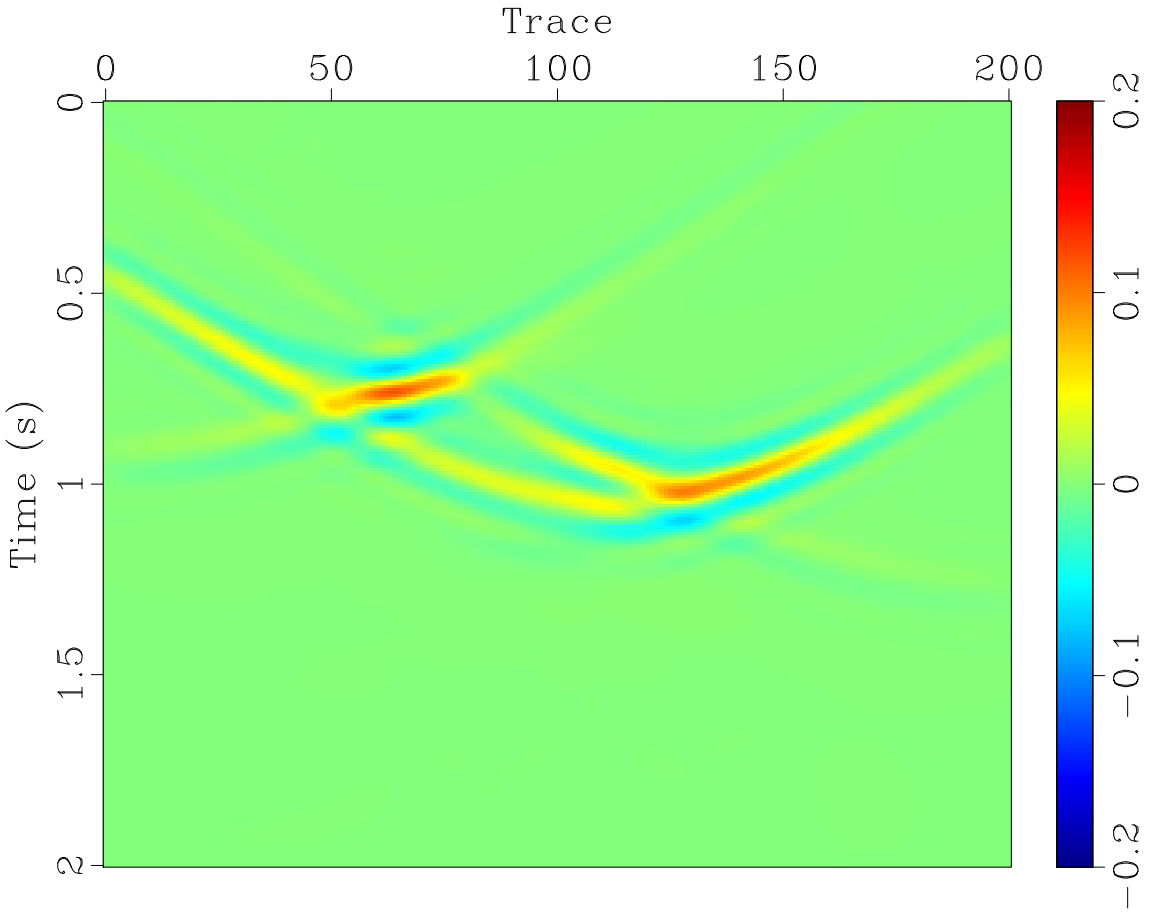

Figure 25: Inversion of data shown in Figure 4, simulated in lens model, using the approximate unitarity relation 42 and propagation in homogenous model.

Preconditioned CG iteration

This final subsection shows that result of Conjugate Gradient iteration, with and without preconditioning, applied to the source estimation problem 6, with zero and non-zero penalty weight α. The data d is the gather shown in 4, simulated using the lens model with source shown in Figure 9, or, alternatively, a point source with bandpass filter wavelet located at $x_d = 3500$ m, $z_d = 3500$ m. In the inversion, the material model is taken to be homogeneous, as has been the case in all of the previous examples.

Figure 27 shows the progress of the normal residual (Euclidean norm of the dif-

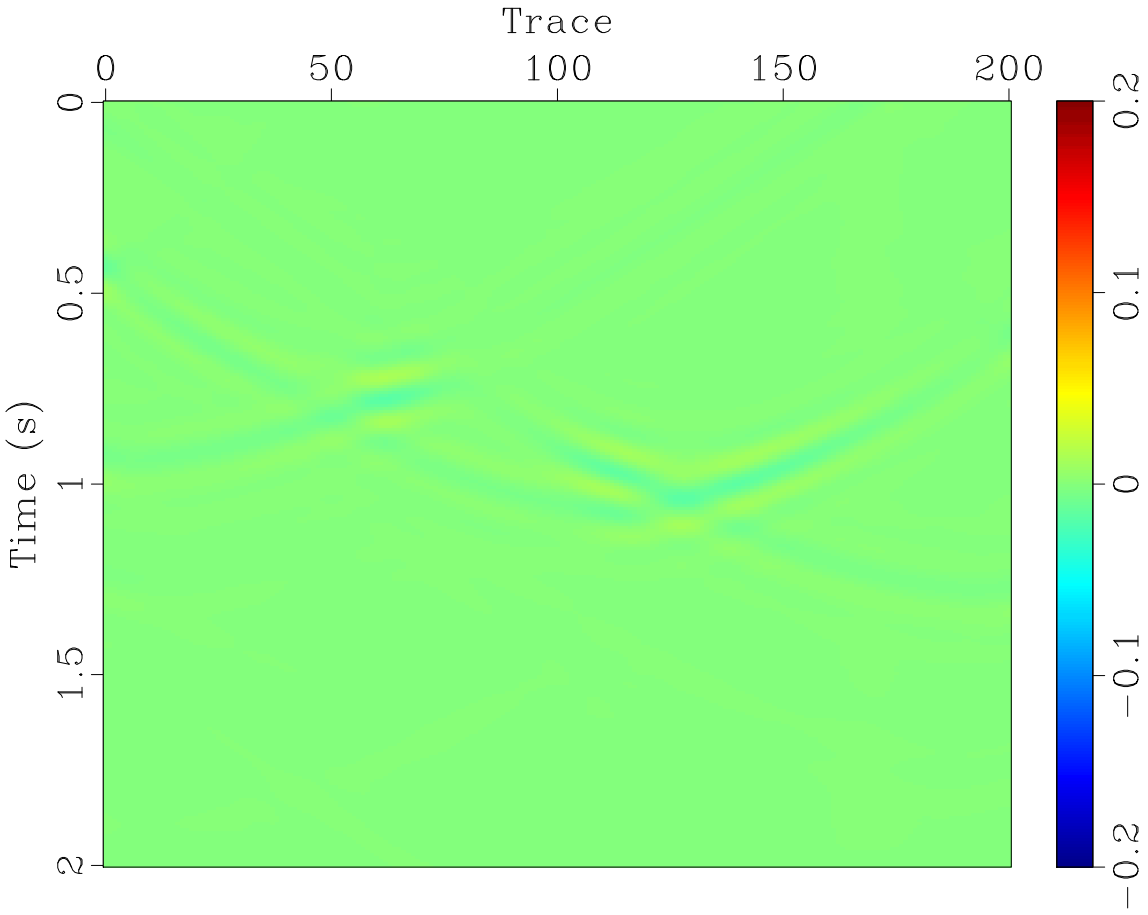

Figure 26: Difference between data displayed in Figures 11 and 25, plotted on the same color scale.

ference of the two sides of equation 50), for Conjugate Gradient and Preconditioned Conjugate Gradient (Algorithm 1) iterations, applied to solution of the optimization problem 46 with $\alpha = 0$. For CG, the norms are both the ordinary Euclidean norm, $W_m = W_d = I$. For PCG, W_m and W_d are given in display 43, with the symmetrized Λs computed as indicated in the preceding subsections. Convergence for the preconditioned algorithm is roughly 4 times as fast.

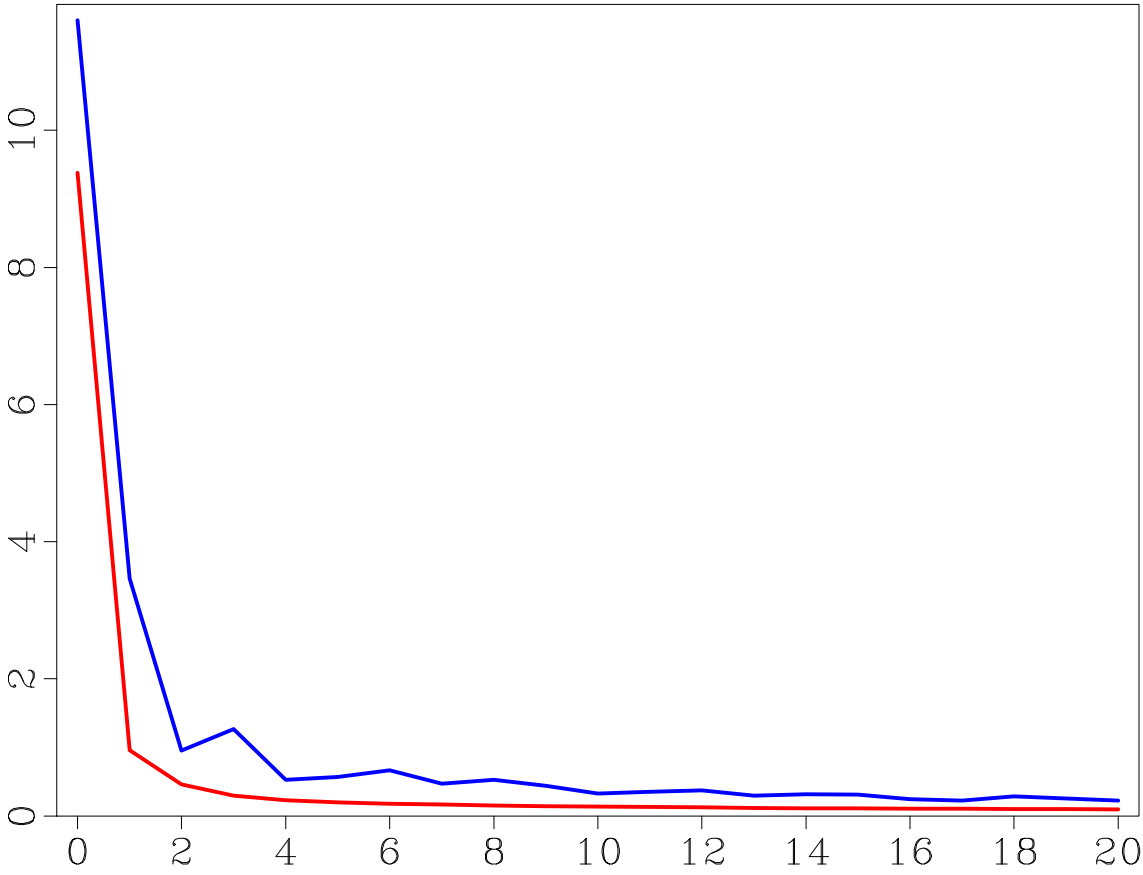

Figure 27: Comparison of normal residual (gradient) Euclidean norms: CG (blue), PCG (red), plotted vs. iteration. Data = lens model, point source (Figure 4), inversion in homogenous model. Penalty weight $\alpha = 0$.

Figure 28 shows the same comparison with non-zero penalty weight, $\alpha = 10^{-3}$. The PCG normal residual curve is almost identical with that in the $\alpha = 0$ case, wheras the CG convergence has slowed down noticeably, being about five times as slow as the preconditioned algorithm.

CONCLUSION

The linear modeling operator of surface source extended acoustic waveform inversion is approximately invertible, and this paper has shown how to approximately invert it.

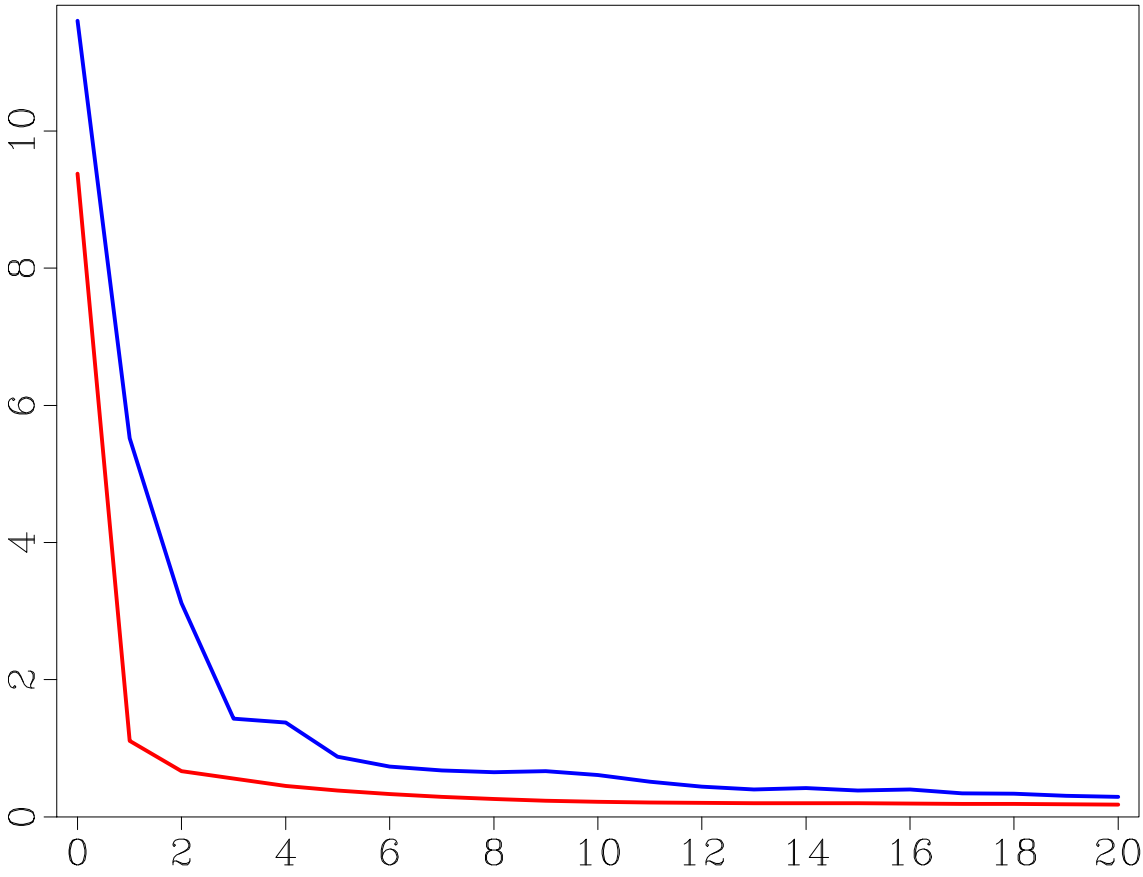

Figure 28: Comparison of normal residual (gradient) Euclidean norms: CG (blue), PCG (red), plotted vs. iteration. Data = lens model, point source (Figure 4), inversion in homogenous model. Penalty weight $\alpha = 10^{-3}$.

The construction is based on reverse time propagation of data, as inspired by the literature on photoacoustic tomography. However, since the input energy comes from a surface source, rather than a pressure boundary value, the pressure-to-source operator intervenes. It provides not just an approximate inverse, but a definition of weighted norms in domain and range spaces of the modeling operator, in terms of which that operator is approximately unitary. Accordingly, Krylov space iteration defined in terms of these weighted norms, or equivalently preconditioned Conjugate Gradient iteration, gives a rapidly convergent solution method for the linear subproblem.

The existence of an approximate unitary representation of the modeling operator is not merely a computational convenience, however. It reveals fundamental aspects of the operator's structure that enable an explanation for the mitigation of cycle-skipping, a feature of the *nonlinear* extended inverse problem. This fact echoes earlier observations concerned a reflected wave inverse problem, involving a modeling operator with a similar approximate inverse (ten Kroode, 2014; Symes, 2014). Also, the approximate inverse leads to a stable computation of the gradient of the nonlinear objective function 4, resolving a difficulty first noted also for reflected wave inversion (Kern and Symes, 1994).

The transmission inverse problem figuring most prominently in contemporary applied seismology is surely the Full Waveform Inversion (FWI) of diving wave data. This is essentially the same problem as the one discussed in this paper, and can be formulated and treated the same way, at least for acoustic wave physics. All of the topics treated here are open for elastic wave physics - the analogue of the pressure-to-source map would is the map from surface velocity field to corresponding constitutive defect, analogous to the elastic Dirichlet-to-Neumann map investigated by Rachele (2000).

The underlying tool in the ideas developed here is geometric optics (or ray theory), without which the very concept of downgoing waves would be meaningless. The physics of actual earth materials includes material heterogeneity on all scales, which appears to leave little room for the assumption of scale separation underlying geometric optics. Moreover, earth materials are anelastic, with elastic wave energy being converted to and from thermal excitation, pore fluid motion, and so on. A truly satisfactory understanding of inverse wave problems will eventually need to accommodate heterogenity and anelasticity beyond the current capabilities of the ray-based theory.

APPENDIX A

ADJOINT COMPUTATION

The adjoint of $\mathcal{S}^+_{z_s,z_r}$ can be computed by a variant of the adjoint state method, in this case a by-product of the conservation of energy. This calculation leads to equation 15, from which the other statements about adjoints made in the second section of the paper follow.

Suppose that $p^-, \mathbf{v}^-$ solve 9 with $(h_s, f_s\mathbf{e_z})\delta(\mathbf{z}-\mathbf{z_s})$ replaced by $(h_r, f_r\mathbf{e_z})\delta(\mathbf{z}-\mathbf{z_r})$. Then

$$0 = \left(\int dx\,dy\,dz\,\frac{p^+p^-}{\kappa} + \rho\mathbf{v}^+\cdot\mathbf{v}^-\right)|_{t\to\infty} - \left(\int dx\,dy\,dz\,\frac{p^+p^-}{\kappa} + \rho\mathbf{v}^+\cdot\mathbf{v}^-\right)|_{t\to-\infty}$$

$$= \int_{-\infty}^{\infty} dt\,\frac{d}{dt}\left(\int dx\,dy\,dz\,\frac{p^+p^-}{\kappa} + \rho\mathbf{v}^+\cdot\mathbf{v}^-\right)$$

$$= \int_{-\infty}^{\infty} dt\,\left(\int dx\,dy\,dz\,\frac{1}{\kappa}\frac{\partial p^+}{\partial t}p^- + p^+\frac{1}{\kappa}\frac{\partial p^-}{\partial t}\right.$$

$$\left. + \rho\frac{\partial\mathbf{v}^+}{\partial t}\cdot\mathbf{v}^- + \rho\mathbf{v}^+\cdot\frac{\partial\mathbf{v}^-}{\partial t}\right)$$

$$= \int_{-\infty}^{\infty} dt\,\left(\int dx\,dy\,dz\,\left(-\nabla\cdot\mathbf{v}^+ + h_s\delta(z-z_s)\right)p^- + p^+\left(-\nabla\cdot\mathbf{v}^- + h_r\delta(z-z_r)\right)\right.$$

$$\left. + (-\nabla p^+ + f_s\mathbf{e_z})\cdot\mathbf{v}^- + \mathbf{v}^+\cdot(-\nabla\mathbf{p}^- + \mathbf{f_r e_z})\right)$$

$$= \int_{-\infty}^{\infty} dt\,\left(\int dx\,dy\,dz\,\left(-\nabla\cdot\mathbf{v}^+ + h_s\delta(z-z_s)\right)p^- + p^+\left(-\nabla\cdot\mathbf{v}^- + h_r\delta(z-z_r)\right)\right.$$

$$\left. + p^+(\nabla\cdot\mathbf{v}^-) + (\nabla\cdot\mathbf{v}^+)p^- + f_s\delta(z-z_s)v_z^- + v_z^+f_r\delta(z-z_r)\right)$$

after integration by parts in the last two terms. Most of what is left cancels, leaving

$$0 = \int_{-\infty}^{\infty} dt\,dx\,dy\,(h_sP_sp^- + f_zP_sv_z^-) + (h_rP_rp^+ + f_rP_rv_z^+)$$

$$= \langle(h_s, f_s), \mathcal{S}^-(h_r, f_r)\rangle + \langle(h_r, f_r), \mathcal{S}^+_{z_s,z_r}(h_s, f_s)\rangle$$

whence 15 follows immediately.

DECLARATIONS

Funding

The author did not receive support from any organization for the submitted work.

Competing Interests

The author certifies that he has no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Data, Material, and Code Availability

The computational examples reported in this work were written in the Madacascar reproducible research framework (`http:www.reproducibility.org`). Code and data source is available from the author on request.

REFERENCES

Bao, G., and W. Symes, 1991, A trace theorem for solutions of linear partial differential equations: Mathematical Methods in the Applied Sciences, **14**, 553–562.

Courant, R., and D. Hilbert, 1962, Methods of mathematical physics, volume ii: Wiley-Interscience.

Dafni, R., and W. Symes, 2018, Accelerated acoustic least-squares inversion: 88th Annual International Meeting, Expanded Abstracts, Society of Exploration Geophysicists, 4291–4295.

Golub, G. H., and C. F. van Loan, 2012, Matrix computations, 4th ed.: Johns Hopkins University Press.

Griewank, A., 2000, Evaluating derivatives: Principles and techniques of algorithmic differentiation: Society for Industrial and Applied Mathematics (Frontiers in Applied Mathematics 19).

Hascoët, L., and V. Pascual, 2013, The Tapenade Automatic Differentiation tool: Principles, Model, and Specification: ACM Transactions On Mathematical Software, **39**.

Hou, J., and W. Symes, 2016a, Accelerating extended least-squares migration with weighted conjugate gradient iteration: Geophysics, **81**, no. 4, S165–S179.

———, 2016b, Accelerating least squares migration with weighted conjugate gradient iteration: 78th Annual International Conference and Exhibition, Expanded Abstract, European Association for Geoscientists and Engineers, P104.

Hu, W., A. Abubakar, and T. Habashy, 2007, Application of the nearly perfectly matched layer in acoustic wave modeling: Geophysics, **72**, SM169–SM176.

Huang, G., R. Nammour, W. Symes, and M. Dolliazal, 2019, Waveform inversion by source extension: 89th Annual International Meeting, Expanded Abstracts, Society of Exploration Geophysicists, 4761–4765.

Kern, M., and W. Symes, 1994, Inversion of reflection seismograms by differential semblance analysis: Algorithm structure and synthetic examples: Geophysical Prospecting, **99**, 565–614.

Lasiecka, I., 1986, Sharp regularity results for mixed hyperbolic problems of second order: Springer Verlag, volume **1223** *of* Springer Lecture notes in Mathematics.

Lasiecka, I., J.-L. Lions, and R. Triggiani, 1986, Non-homogeneous boundary value problems for second order hyperbolic operators: Journal de Mathématiques Pures et Appliquées, **65**, 149–192.

Lasiecka, I., and R. Trigianni, 1989, Trace regularity of the solutions of the wave equation with homogeneous boundary conditions and compactly supported data: Journal of Mathematical Analysis and Applications, **141**, 49–71.

Lax, P. D., 2006, Hyperbolic partial differential equations (Courant Lecture Notes): American Mathematical Society.

Payne, L., 1975, Improperly posed problems in partial differential equations: Lecture Note 22, CBMS, Society for Industrial and Applied Mathematics, Philadelphia.

Rachele, L., 2000, Boundary determination for an inverse problem in elastodynamics: Communications in Partial Differential Equations, **25**, 1951–1996.

Stefanov, P., and G. Uhlmann, 2005, Stable determination of generic simple metrics from the hyperbolic Dirichlet-to-Neumann map: International Mathematics Research Notices, **17**, 1047–1061.

———, 2009, Thermoacoustic tomography with variable sound speed: Inverse Problems, **25**, 075011.

Symes, W., 2007, Reverse time migration with optimal checkpointing: Geophysics, **72**, SM213–222.

———, 2014, Seismic inverse problems: recent developments in theory and practice: Inverse Problems - from Theory to Application, Proceedings, Institute of Physics, 2–5.

Symes, W., H. Chen, and S. Minkoff, 2020, Full waveform inversion by source extension: why it works: 90th Annual International Meeting, Expanded Abstracts, Society of Exploration Geophysicists, 765–769.

Symes, W., and L. E. Payne, 1983, Trace theorem for solutions of the wave equation and the remote determination of acoustic sources: Mathematical Methods in the Applied Sciences, **5**, 131–152.

Symes, W., D. Sun, and M. Enriquez, 2011, From modelling to inversion: designing a well-adapted simulator: Geophysical Prospecting, **59**, 814–833. (DOI:10.1111/j.1365-2478.2011.00977.x).

Tang, B., S. Xu, and Y. Zhang, 2013, 3D angle gathers with plane-wave reverse time migration: Geophysics, **78**, no. 2, S117–S123.

ten Kroode, F., 2012, A wave-equation-based Kirchhoff operator: Inverse Problems, 115013:1–28.

———, 2014, A Lie group associated to seismic velocity estimation: Inverse Problems - from Theory to Application, Proceedings, Institute of Physics, 142–146.

Virieux, J., 1984, SH-wave propagation in heterogeneous media: Velocity stress finite-difference method: Geophysics, **49**, 1933–1957.

Virieux, J., and S. Operto, 2009, An overview of full waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC127–WCC152.

Xu, S., D. Wang, F. Chen, G. Lambaré, and Y. Zhang, 2012, Inversion on reflected seismic wave: SEG Technical Program Expanded Abstracts, 1–7.

Xu, S., Y. Zhang, and B. Tang, 2011, 3D angle gathers from reverse time migration: Geophysics, **76**, no. 2, S77–S92.

Zhang, Y., A. Ratcliffe, G. Roberts, and L. Duan, 2014, Amplitude-preserving reverse time migration: from reflectivity to velocity and impedance inversion: Geophysics, **79**, S271–S283.

Zhang, Y., and J. Sun, 2009, Practical issues of reverse time migration: True amplitude gathers, noise removal and harmonic-source encoding: Beijing International Geophysical Conference and Exposition, Expanded Abstracts, Society of Explo-

ration Geophysicists, 204–209.