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. For an acoustic formulation for sources supported on a surface, with a soft contraint enforcing concentration at a point, a variant of the time reversal method from photoacoustic tomography provides an approximate solution. This approximate inverse can be used to precondition Krylov space iteration for rapid convergence to the solution of the core subproblem in this setting. Numerical examples illustrate 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 data is fit and wave physics are honored, that is,

$$P\mathbf{u} \approx d \text{ and } L[\mathbf{c}]\mathbf{u} = \mathbf{f}. \tag{2}$$

The source function $\mathbf{f}$ may be given, or to be determined along with $\mathbf{c}$. The task 2 may be cast as a nonlinear least squares problem:

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

Practical variants of the least squares problem 3 typically augment the objective with additive penalties or other constraints (Virieux and Operto, 2009; Fichtner, 2010; Schuster, 2017).

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 at uninformative model estimates due to "cycle-skipping". This phenomenon is often interpreted as the capture of optimizing sequences at local minima of the nonlinear objective function in 3 far from the global minimum, or at least not possessing the characteristics of a useful solution. See for example Gauthier et al. (1986); Virieux and Operto (2009); Pladys et al. (2021).

Source extension is one approach to avoiding this"cycle-skipping" obstacle. It consists in imposing the wave equation as a soft constraint, 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 linear operator A penalizes deviation from known (or assumed) characteristics of the source function - its null space consists of feasible (or "physical") source models.

Well-studied examples of extended source approaches to FWI include Wavefield Reconstruction Inversion (WRI) (van Leeuwen and Herrmann, 2013, 2016; Li et al., 2018; Aghmiry et al., 2020; Louboutin et al., 2020) and Adaptive Waveform Inversion (AWI) (Warner and Guasch, 2016; Guasch et al., 2019, 2020; Yong et al., 2021; Warner et al., 2021). Huang et al. (2019) present an overview of the recent literature on source extension methods.

The present paper concerns *surface source extension*: physical sources are presumed to be concentrated at points in space, whereas their extended counterparts are permitted to spread energy over a *source surface*. Similarly, receivers are located on a *receiver surface*. Assuming that the presumed point source is at location $\mathbf{x}_s$, a simple (though certainluy not the only) 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 sub-*

problem of problem 4:

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

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 modeling operator $PL[\mathbf{c}]^{-1}$ and the penalty operator A defined in 5 are 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 ("precondition") Krylov space methods for the solution of the surface source subproblem 6. 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. Numerical examples in this setting suggest the effectiveness of this acceleration.

I will use two 2D numerical models throughout to illustrate the theory. In both, horizontal lines serve as source and receiver surfaces. The first is a "horizontal crosswell" or slab configuration with an acoustic lens positioned between a deep source and a shallower line of receivers, resulting in markedly triplicated arrivals. The goal in this first example is to construct a surface source that explains the data in a homogeneous medium (that is, inversion in a wrong velocity, emulating the early iterations of FWI based on extended modeling). The second is a layered model in which velocity increases with depth, resulting the formation of diving waves. This configuration simulates an ocean-bottom node and a line of near-surface sources: the roles of source and receiver are switched for computational convenience. In this second example, the diving wave arrivals are isolated (by an appropriate mute operator) and a source constructed that explains them alone.

Since the constructions described here involve a significant number of components and ideas, I begin the paper with an overview, mapping out the key steps and their relation to other work. The next section defines the modeling operator $PL[\mathbf{c}]^{-1}$ and important specializations (pressure vs. normal velocity sources and data). It also introduces the two 2D examples mentioned above. The following section constructs an approximate inverse of the modeling operator by time reversal (as suggested by work in photoacoustic tomography), and illustrates its efficacy via the two examples. This construction requires extraction of velocity data, or equivalently a surface source, from pressure data. The subsequent sections describe this pressure-to-source operator, 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. I use the 2D examples to illustrate each of these developments. 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.

OVERVIEW

The preconditioner construction is partly inspired by the time reversal method in photoacoustic tomography (Stefanov and Uhlmann, 2009; Hristova, 2009). This problem clearly has strong similarities to, but also differences from, the problem studied here.

A simplified mathematical translation of this medical imaging task is to infer the initial excess pressure distribution over a fluid-containing region at time $t = 0$ from measurements of the pressure on a surface enclosing the fluid over a time interval $0 \leq t \leq t_{\max}$. The time reversal method presumes that the pressure field has returned to equilibrium (zero excess pressure), or close to it, in the fluid-containing region at the final time $t = t_{\max}$ Then the field can be (at least approximately) viewed as the solution of the backwards-in-time initial boundary value problem with zero final conditions at $t = t_{\max}$, and boundary values given by the measurements. Evolving the field backwards in time to $t = 0$ thus solves the problem. Except in special circumstances, the pressure field never actually vanishes at finite time, so the solution is approximate.

The seismic surface source extension proplem 6 differs in several obvious ways from the photoacoustic setting. The measurement or receiver surface does not surround the region of wave propagation. It is not the initial pressure time-slice at $t = 0$ that is to be determined, but a time-extended source $\mathbf{f}$ confined to a surface. The penalty operator A has no analogue in the basic photoacoustic problem description. Finally, once these obstacles are overcome, using the approximate solution operator to accelerate Krylov iteration for solution of the optimization problem 6 requires that the operator be identified as the adjoint of the modeling operator with respect to suitable inner products in its domain and range. The next few paragraphs sketch resolutions of each of these issues.

First, reverse-time propagation can be localized via ray theory, within high-frequency asymptotic approximation. This step requires some assumptions about the ray fields: all rays responsible for significant energy in the receiver data must arrive from the source surface, and must cross the source and receiver surfaces transversally, that is, making non-zero angles. The surfaces must be oriented, with a choice of ("outward") unit normal vector. The energetic rays must emanate from the source surface inwards ("incoming" rays), and arrive at the receiver surface outwards ("outgoing" rays). Then both surfaces can be treated locally as boundaries of propagation domains. With these assumptions, reverse-time propagation of the receiver data closely

approximates the acoustic fields near the source surface. Next, I show that for a causal field with a surface pressure source (with a continuous pressure field across the surface), the source amplitude is proportional to the jump in the (particle) velocity field. Moreover, to leading order in frequency, the velocity field switches sign at the surface, so the jump is just twice the sampled value on the surface. Thus constructing a source from the velocity trace on the receiver surface, reverse-time propagation of this source, and reading off the velocity field on the source surface yields a source that reproduces the pressure data, within an asymptotically small error.

This is the essence of the approximate solution of the problem 6 for penalty weight $\alpha = 0$. To see how Krylov iteration might be accelerated, I note that reverse-time propagation of the receiver surface source to pressure on the source surface is the transpose (adjoint) of the forward-time source-to-pressure propagation - this is a very simple version of the adjoint state construction (Plessix, 2006). So the approximate inverse may be described as: conversion of pressure to source at the receiver surface, followed by application of the modeling operator adjoint, followed by the conversion of pressure to source at the source surface. This sequence precisely describes the adjoint of the modeling operator with respect to weighted norms in the spaces of source and receiver data, with the weight operators being the pressure-to-source operators on the two surfaces. That this construction results in an approximate inverse means that the modeling operator is approximately unitary with respect to these weighted norms, which in turn implies that a properly constructed preconditioned conjugate gradient algorithm should converge rapidly.

Finally, the $\alpha = 0$ case is not sufficient: the penalty operator A is an essential component of the nonlinear inverse propblem 4. As it turns out, A commutes with the other operators involved - approximately, but that is enough. Therefore the effect of A can be compensated with an easily-computed factor, whence the preconditioned acceleration extended to the case $\alpha > 0$.

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). In this terminology, "Dirichlet" refers to pressure restricted to a surface, and "Neumann" to the normal derivative of pressure on that surface. Since the latter is proportional to normal particle acceleration thanks the equations of motion, and normal particle velocity is proportion to source, the relation to the pressure-to-source map as described in this paper is clear. A number of authors have described preconditioners for Least Squares Migration with subsurface offset extension in which the Dirichlet-to-Neumann or pressure-to-source map plays a central role (ten Kroode, 2012; Hou and Symes, 2015; Chauris and Cocher, 2017). It 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 role of the pressure-to-source map in these constructions is quite similar to the role described here. Reverse Time Migration, in its basic form, broadcasts pressure traces into the propagation region, backwards in time. Where this backward

propagated energy intersects the forward propagated Green's function, an imaging condition outputs an image amplitude map, which is not however an approximate inversion, as is widely understood, except in a crude kinematic sense. The various "true amplitude" or approximate least squares migrations, apply the pressure-to-source operator (either explicitly, or implicitly via switching boundary conditions) *before* reverse-time broadcasting. Since the data are presumed to be traces of the forward scattered (Born) pressure field, the backward propagated wavefield must approximate the forward scattered field, time-reversed. This reconstructed scattered field combines at scattering points with the Green's function to yield an approximate inversion for the materal parameter perturbations. Moreover, this approximate Born inversion is also adjoint to the Born scattering operator in weighted norms, one of the weight operators being the pressure-to-source map, hence an effective preconditioner for Krylov space methods for linear least squares inversion, that is, Least Squares Migration (Hou and Symes, 2016). Thus the present construction bears a considerable resemblance to these "true-amplitude" reverse-time migration methods.

The discussion in this paper is formal and incomplete, in the sense that some important mathematical underpinnings are only sketched. The transport of high-frequency energy along rays is a crucial ingredient in the constructions explained here. Of course, such transport only occurs in material parameter models that vary slowiy in space on the wavelength scale, and even then only approximately. Rather than provide explicit smoothness constraints and asymptotic error estimates, I will use these concepts in an intuitive way, using "approximate" and "asymptotic" and the symbol "$\approx$" to suggest asymptotically negligible errors not explicitly assessed, and refer to such errors as "negligible".

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 receiver surface is not well-defined for arbitrary finite-energy acoustic fields. Thus the modeling operator envisioned above is not well-defined, strictly speaking. 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). A number of authors have described precise forms of the ray conditions described above, and shown how these conditions lead to the desired behaviour of the modeling operator, that is, mapping of finite energy sources to finite energy data, or equivalent properties (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 incoming/outgoing ray constrants..

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 if it is restricted to either causal or anti-causal vector wavefields.

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, $L = L[\mathbf{c}]$.

Most of what follows is valid for any space dimension $n > 0$. I will describe the theory for $n = 3$, write $\mathbf{x} = (x, y, z)^T$ for the spatial coordinate vector, and similarly for the particle velocity $\mathbf{v} = (v_x, v_y, v_z)^T$. For computational convenience, the examples are two-dimensional.

The surface source extension replaces point sources on or near a surface Σ_s in $\mathbf{R}^3$ with source functions confined to the surface . The surface should be supplied with a choice of unit normal vector field $\mathbf{n}_s$. Since the surface separates space into two non-overlapping parts, at least locally, designate the part into which the normal vector points as the "outside", so that $\mathbf{n}_s$ is the outward normal.

For acoustic modeling, surface sources are combinations of constitutive law defects and loads normal to the surface, localized on Σ_s. That is, right-hand sides in the system $L\mathbf{u} = \mathbf{f}$ take the form $\mathbf{f}(\mathbf{x}, t) = (h_s\delta_{\Sigma_s}, f_s\mathbf{n}_s\delta_{\Sigma_s})^T$ for scalar defect h_s and normal load f_s. With the choice L given in 8, the causal system $L\mathbf{u} = \mathbf{f}$ takes the form

$$\begin{aligned} \frac{1}{\kappa}\frac{\partial p}{\partial t} &= -\nabla \cdot \mathbf{v} + h_s\delta_{\Sigma_s}, \\ \rho\frac{\partial \mathbf{v}}{\partial t} &= -\nabla p + f_s\mathbf{n}_s\delta_{\Sigma_s}, \\ p &= 0 \text{ for } t \ll 0, \\ \mathbf{v} &= 0 \text{ for } t \ll 0. \end{aligned} \tag{9}$$

in which δ_{Σ_s} is the delta function on Σ_s, and similarly for Σ_r.

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

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

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 a receiver surface Σ_r with (outward) unit normal field $\mathbf{n}_r$. P_s and P_r are the sampling operators on Σ_s and Σ_r respectively. In practice, sampling necessarily occurs at a discrete array of points (trace locations). In this theoretical discussion, I will neglect the finite sample rate, and regard the data, for example $P_r p$, as continuously sampled. The output samples are necessarily muted, that is, non-zero only over a space-time domain of finite extent. This mute, and any tapering applied to the data traces, are regarded as part of the sampling operators P_s, P_r.

The vector modeling operator $\mathcal{S}$ relates the source amplitudes h_s, f_s on Σ_s to the sampled pressure and normal velocity p, $\mathbf{v} \cdot \mathbf{n}_r$ for the solution $(p, \mathbf{v})$ of the systems 9 by

$$\mathcal{S}(h_s, f_s)^T = (P_r p, P_r(\mathbf{v} \cdot \mathbf{n}_r))^T, \tag{10}$$

Denote by $\Pi_i, i = 0, 1$ the projection on the first,respectively second, component of a vector in $\mathbf{R}^2$:

$$\Pi_0(a, b)^T = a,\ \Pi_1(a, b)^T = b. \tag{11}$$

The forward modeling operator from pressure source to pressure trace is

$$S = \Pi_0 \mathcal{S} \Pi_0^T \tag{12}$$

With these conventions, we can write the version of the source subproblem 6 studied in this paper as: given a pressure gather d on Σ_r,

$$\text{find } h_s \text{ to minimize } \|Sh_s - d\|^2 + \alpha^2 \|Ah_s\|^2. \tag{13}$$

The two examples mentioned in the introduction illustrate the setting just described. The first example (Figure 1) embeds an acoustic (low-velocity) lens between a source surface $\Sigma_s = \{z = z_s = 3000 \text{ m}\}$ (orange horizontal line) m and receiver surface $\Sigma_r = \{z = z_r = 1000 \text{ m}\}$ (yellow horizontal line). The data to be used in this example results from a point source at $(3500, 3500)$ m (orange star). I have overlain the rays connecting this source point with the data portion of the receiver surface: evidently all of the rays involved cross the source and receiver surfaces transversally. The outward normal vectors are chosen as $\mathbf{n}_s = \mathbf{e}_z = (0, 1)^T$ on the source surface and $\mathbf{n}_r = -\mathbf{e}_z$ on the receiver surface. One sees from the Figure that the ray field carrying high frequency energy is incoming at the source surface, outgoing at the receiver surface.

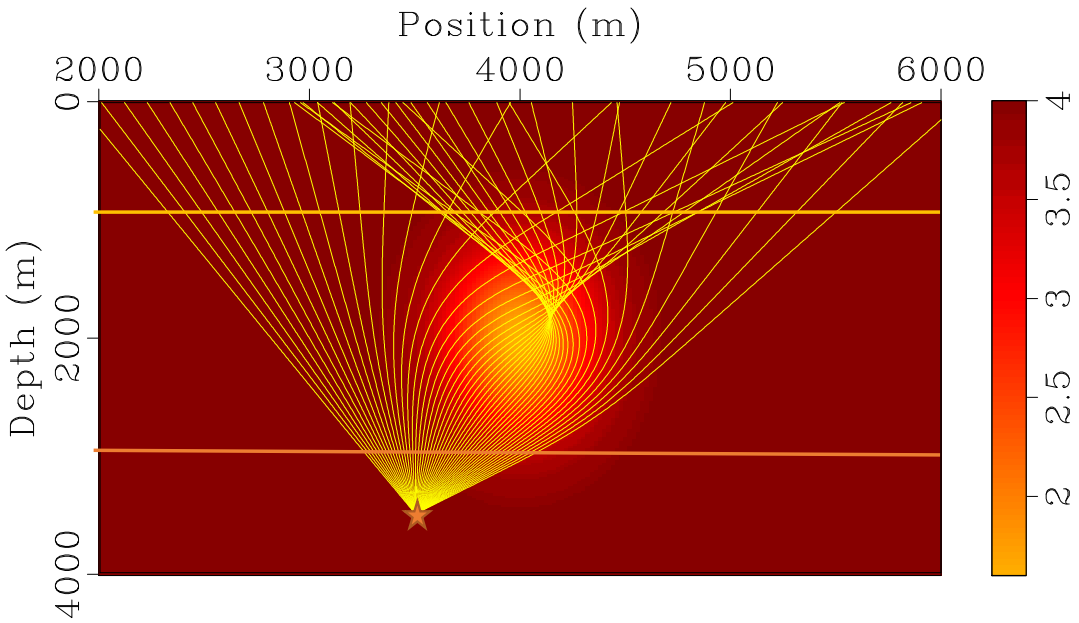
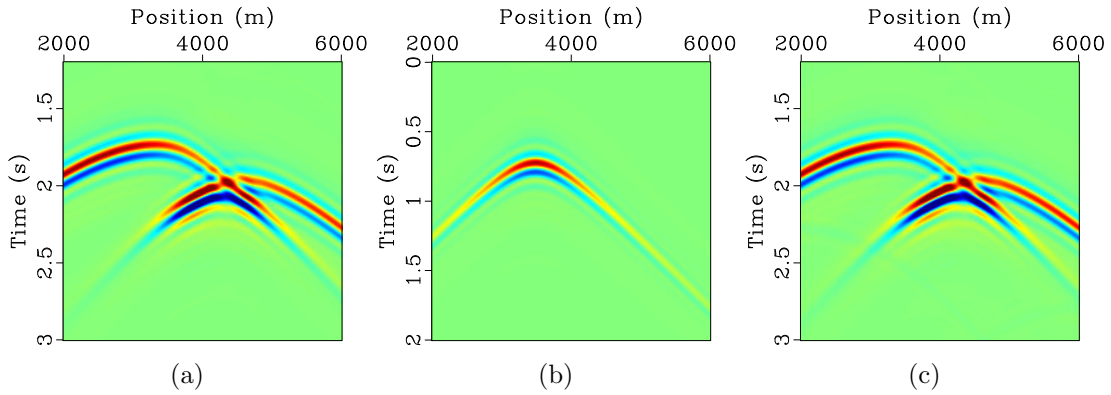


Figure 1: Bulk modulus with acoustic lens. Color scale unit is GPa. Orange horizontal line is source surface at depth $z_s = 3000$ m, yellow horizontal line is receiver surface at depth $z_r = 1000$ m. Point source location for data generation at $(3500, 3500)$ m indicated with star. Overlain with rays from point source to receiver surface. Outward normal points down at source surface, up at receiver surface, so rays are incoming at source surface, outgoing at receiver surface.



(a) (b) (c)

Figure 2: Data generated using configuration in Figure 1. (a) traces from point source at $(3500, 3500)$ with $[1, 2, 7.5, 12.5]$ Hz zero phase trapezoidal bandpass filter wavelet, delayed by 0.5 s. (b) Equivalent extended source on the source surface at depth $z_s = 3000$ m. (c) Gather generated by equivalent source shown in (b). Color scale is same for (a) and (c).

Figure 2a shows the gather $P_r p_{\text{pt}}$, where p_{pt} is causal field generated by the point source at $(3500, 3500)$ m with a trapezoidal bandpass filter wavelet having significant energy between 1 and 12.5 Hz. The mute embedded in P_r limits trace positions to 2000 m $\leq x_r \leq$ 6000 m, and times to 1.2 s $\leq t \leq$ 3 s.

Figure 2b displays an extended (pressure) source h_s, in the form of traces in the spatial range 2000 m $\leq x_r \leq$ 6000 m and time range 0 s $\leq t \leq$ 2 s. The modeling operator output Sh_s, obtained from the causal solution $(p, \mathbf{v})$ of 9 with this choice of h_s and $f_s = 0$, by application of the sampling operator P_r to p, is shown in Figure 2c.

Remark: The close resemblance between Figures 2a and 2c is not accidental - the extended source h_s shown in Figure 11a is constructed so that Sh_s (Figure 2c) closely approximates the point source gather $P_r p_{\text{pt}}$ (Figure 2a). This construction will be explained in the next section.

The second example is intended as a cartoon of long-offset node acquisition. It features a depth-dependent increasing velocity. Figure 3 shows bulk modulus field (once again, the density is constant and $= 1$ g/cm^3). The source and receiver surfaces are horizontal lines, as in the first example, at depths $z_s = 500$ m and $z_r = 100$ m respectively. A point source used to generate a data traces is positioned at $z_s = 500, x_s = 10000$ m. The source wavelet is the same bandpass filter as in the previous example.

A diving wave arrival is clearly visible in Figure 4a, which displays data traces over the positive offset range $10000 \leq x_r \leq 20000$ m. This part of the wavefield is connected to some refracted rays from the source, those shown in blue in Figure 3. Other refracted rays, shown in red in that Figure, either do not arrive at the receiver surface within the time range of the data, or are associated with another branch of

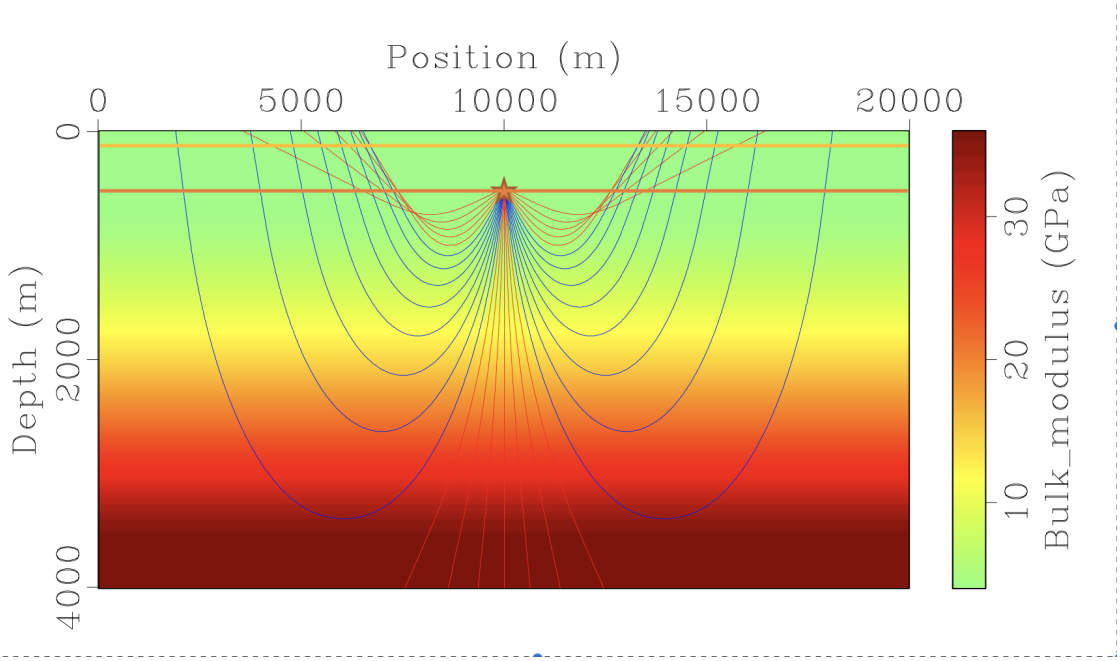
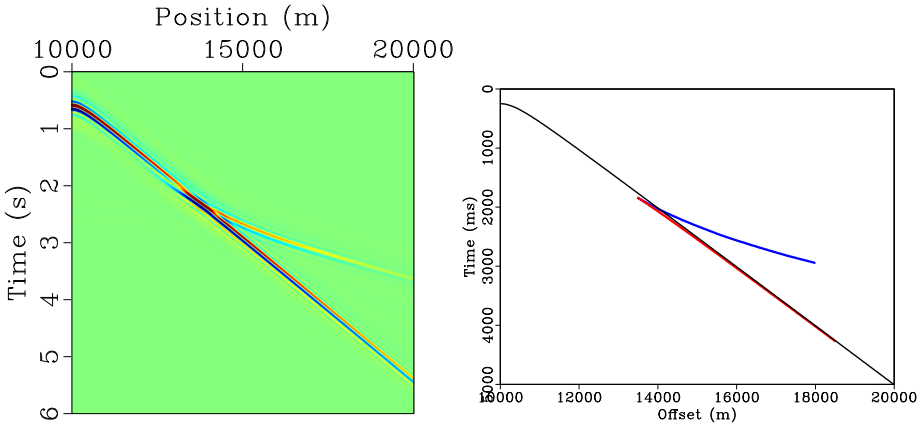


Figure 3: Bulk modulus generating diving waves. Color scale unit is GPa. Orange horizontal line is source surface at depth $z_s = 500$ m, yellow horizontal line is receiver surface at depth $z_r = 100$ m. Point source location for data generation at $(500, 10000)$ m indicated with star. Overlain with rays from point source to receiver surface. Outward normal points up at both source and receiver surfaces, so depicted rays are incoming at source surface, outgoing at receiver surface. Significance of red vs. blue coloring explained in text and caption of Figure 4b.



(a) (b)

Figure 4: Data generated using configuration in Figure 3. (a) Traces from point source at $(500, 10000)$ m with $[1, 2, 7.5, 12.5]$ Hz zero phase trapezoidal bandpass filter wavelet, delayed by 0.4 s. Note that direct arrival waveform appears to change where it intersects the diving wave arrival visible on the right. (b) Arrival times at surface $z = 0$ of rays leaving source point ($(500, 10000)$ m) at $t = 0$. Blue and red branches corresponds to blue and red rays in Figure 3, respectively. The diving wave carried by the red rays arrives at the red time branch, which is nearly coincident with the direct arrival branch (the black line). This part of the refracted wave field combines with the direct wave to produce the apparent wavelet change shown in subfigure (a). The blue arrival time branch corresponds to the diving wave energy carried by the blue rays in Figure 3, displayed in Figure 5b.

refracted arrivals. These two collections of refracted rays result in two arrival time branches, shown in blue and red in Figure 4b. The refracted arrival times shown in red, corresponding to some of the rays shown in red in Figure 3, corresponds to a refracted arrival that almost overlaps the direct wave, as is clearly visible in Figure 4a. The arrival combines with the direct wave to result in an apparent waveform change in the latter. The refracted arrival times shown in blue in Figure 4b correspond to the blue rays in Figure 3, and to the faster event extending to the right of the direct wave.

Note that at the source, all of the rays shown in Figure 3 are incoming if the source surface orientation is chosen with inwards = down, in contrast to the previous example. The "blue" rays, as well as those of the "red" rays that are refracted back towards $z = 0$, are outgoing at the receiver surface if it is oriented so that inwards = down also.

The "blue" arrival can be isolated via a mute, depicted in Figure 5a. The mute includes time truncation at 6 s, and offsets between -10000 and 10000 m. The isolated diving wave (for positive offsets) is shown in Figure 5b. This muted field can be inverted via the techniques exploained below, that is, a source field on $z = 500$ m constructed that regenerates approximately the data shown in 5b. In principle the

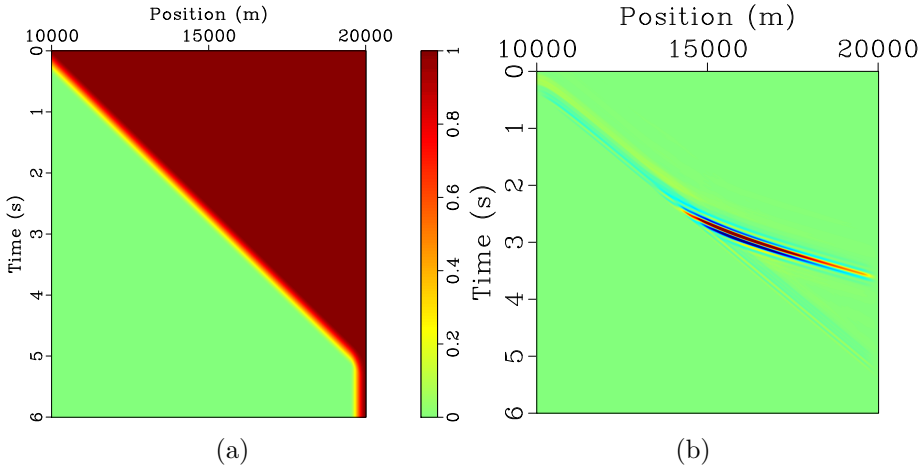


Figure 5: (a) Mute function, incorporated in the sampling operator P_r, designed to isolate the diving wave branch visible in Figure 4a. (b) Diving wave gather, mute in (a) applied to data in Figure 4a. Color scale is 5 tmies that of Figure 4a.

"red" arrival, or the "red" and "blue" arrivals together, could also be inverted by the same technique, however the "red" arrival can only be isolated by subtraction of the direct wave, which is not itself incoming (or outgoing). In this paper, the simpler option is pursued, of inverting the muted diving wave arrival in Figure 5b, in order to illustrated the localization mechanism at the heart of the inversion.

TIME REVERSAL

There is no measurement surface surrounding the region of propagation in the surface source inversion problem. Therefore the time reversal construction, as found in the literature on photoacoustic tomography, must be localized in space-time, so that an artificial surrounding region may be constructed. It must also be modified to produce a source, rather than a slice of the pressure field. The assumption of incoming rays at the source surface and outgoing rays at the receiver surface makes both modifications possible.

The construction takes place in three stages: first, an acoustic field with high-frequency energy propagating along an incoming ray family is identified as the causal solution of an inhomogeneous boundary value problem *locally and on the inside*. The boundary value is simply the restriction of the pressure field to the surface. Second, the same causal solution is identified as the causal solution of an acoustic problem with source, singular on the surface in the fashion of equation 9. The source amplitude is proportional to the velocity field sampled on the surface. Third, this identification is used twice, first to propagate the receiver data backwards in time as the solution of an anti-causal problem with source on the receiver surface; the velocity field of which is sampled on the source surface; then the causal problem is solved with the source

built from the velocity samples. This solution must asymptotically approximate the field generating the data, along the ray fields carrying all of the high-frequency energy. Therefore the reconstructed source is an approximate solution of the problem 6 for $\alpha = 0$.

Field localization

The first task is to view an acoustic field locally as a solution of a boundary value problem with prescribed pressure on the boundary of a domain, as is the case in the photoacoustic tomography configuration.

In this subsection $(p, \mathbf{v})$ is a solution of the homogeneous acoustic system

$$\begin{aligned} \frac{1}{\kappa}\frac{\partial p}{\partial t} &= -\nabla \cdot \mathbf{v}, \\ \rho\frac{\partial \mathbf{v}}{\partial t} &= -\nabla p, \end{aligned} \tag{14}$$

The source surface Σ is given a choice of unit normal field $\mathbf{n}$. Rays carrying high-frequency energy in $(p, \mathbf{v})$ pass through Σ, and intersect Σ only at times > 0. Also, the velocity vectors of all such rays have a negative $\mathbf{n}$ component where they cross Σ. Construct a time interval $[0, \tau]$ and a region in space Ω^+ containing Σ so that Σ forms part of the boundary $\partial\Omega^+$ of Ω^+, $\mathbf{n}$ is the outward unit normal to Ω^+ on Σ, and the energetic rays cross $\partial\Omega^+$ *only* at points of Σ in the time interval $[0, \tau]$ - that is, in time less than τ the rays do not reach other parts of $\partial\Omega^+$. Note that the rays are *incoming* to Ω^+ in the terminology introduced earlier. See Figure 6.

Consequent to this ray geometry, conclude that the field $(p, \mathbf{v})$ is asymptotically negligible both throughout Ω^+ at $t = 0$ and at all points of $\partial\Omega^+$ except for Σ, for $0 \leq t \leq \tau$. So $(p, \mathbf{v})$ differs negligibly from the solution of the boundary value problem

$$\begin{aligned} \frac{1}{\kappa}\frac{\partial \tilde{p}}{\partial t} &= -\nabla \cdot \tilde{\mathbf{v}}, \\ \rho\frac{\partial \tilde{\mathbf{v}}}{\partial t} &= -\nabla \tilde{p}, \\ \tilde{p}(\mathbf{x}, t) &= p(\mathbf{x}, t) \text{ for } \mathbf{x} \in \Sigma, 0 \leq t \leq \tau, \\ \tilde{p}(\mathbf{x}, t) &= 0 \text{ for } \mathbf{x} \in \partial\Omega^+ \setminus \Sigma, 0 \leq t \leq \tau. \end{aligned} \tag{15}$$

Source representation

The fields involved in the surface source extension are causal and source-generated, rather than solutions of boundary value problems such as the system 15. To transition to a source representation, begin by extending the domain Ω^+, lying on the incoming side of the surface Σ, to a full neighborhood Ω of Σ by adding a domain Ω^- on the outside (see Figure 6). Once again, construct Ω^- so that none of the rays carrying

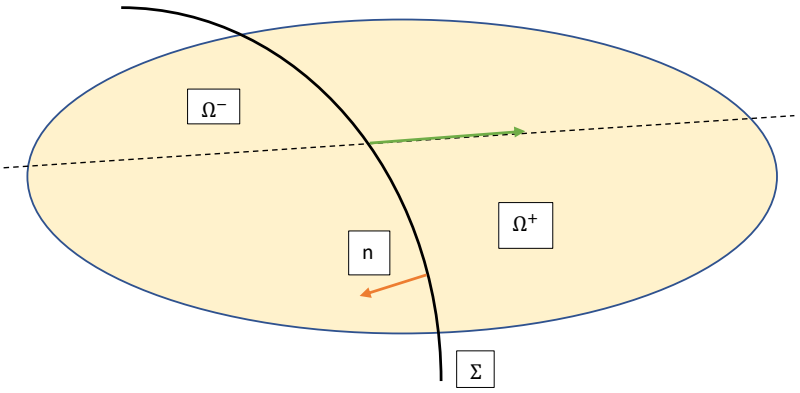


Figure 6: Localization near crossing points of rays. Σ is source surface. Dashed line is part of a ray carrying high frequency energy. Ray velocity vector has negative component in direction of unit normal $\mathbf{n}$ (outward, in chosen orientation). Ω^+ contains all ray segments for times $0 \leq t \leq \tau$ (green arrow). Rays intersect boundary $\partial\Omega^+$ only in Σ for this time range. Ω^- lies on outward side of Σ, which is part of its boundary. The domain Ω is the union of Ω^+ and Ω^-.

high-frequency energy cross the boundary of Ω^- except at Σ. Extend the field $(\tilde{p}, \tilde{\mathbf{v}})$ to the full neigborhood Ω by solving the system 15 with Ω^+ replaced by Ω^-. Thus the extended field is pieced together out of causal fields solving the boundary value problem on either side of the boudary, with *the same* boundary condition, that is, the trace of p on Σ. Thus the extended field is continuous in p. it is not however continuous in $\mathbf{v}$, as will be seen shortly. More precisely, the normal component $\mathbf{v} \cdot \mathbf{n}$ has a discontinuity, while the tangential components are continuous. The jump in $\mathbf{v} \cdot \mathbf{n}$ contributes a delta function to the gradient: as a field in the whole neigborhood Ω, $(\tilde{p}, \tilde{\mathbf{v}})$ solves

$$\begin{aligned}
\frac{1}{\kappa}\frac{\partial \tilde{p}}{\partial t} &= [\mathbf{v} \cdot \mathbf{n}]_\Sigma \delta_\Sigma - \nabla \cdot \tilde{\mathbf{v}}, \\
\rho \frac{\partial \tilde{\mathbf{v}}}{\partial t} &= -\nabla \tilde{p}, \\
\tilde{p}(\mathbf{x}, t) &= 0 \text{ for } \mathbf{x} \in \partial\Omega, 0 \leq t \leq \tau, \\
\tilde{p} &= 0 \text{ for } \mathbf{x} \in \Omega^+, t, \\
\tilde{\mathbf{v}} &= \mathbf{0} \text{ for } \mathbf{x} \in \Omega^+, t = 0.
\end{aligned} \tag{16}$$

The square bracket denotes the jump in the quantity enclosed across Σ, from the outside to the inside. Since $\mathbf{n}$ points from the inside to the outside, this convention contributes another minus sign.

This representation is even better than it looks: since the ray families in $\Omega^\pm$ cross the boundary only at Σ within the prescribed time interval, the solution changes only

by an asymptotically negligible amount in Ω for $0 \leq t \leq \tau$ if the boundary condition is removed and the system is solved for all of space-time.

There is a downside, namely the apparent necessity of constructing the solution in Ω^- so that the jump of $\mathbf{v}\cdot\mathbf{n}$ may be computed. That is actually not necessary, however. This is best seen by analyzing a generic geometric optics component associated to the phase function ψ in Ω^+ and its family of rays. The acoustic field is assembled by a sum (in the continuum, an integral) over such solutions. A particular summand looks like

$$a(\mathbf{x},t)e^{i\omega(t-\psi(\mathbf{x}))}. \tag{17}$$

for $\mathbf{x} \in \Omega^+$. The phase ψ solves the eikonal equation, the amplitude a the related transport equation. The restriction of p to Σ determines ψ (and a) restricted to Σ, essentially by Fourier synthesis. Since the eikonal equation is even in $\nabla\psi$, and ψ restricted to Σ determines the tangential derivatives of ψ, the normal component of $\nabla\psi$ is determined up to sign by the boundary values of p. The spatial component of a ray $X(t)$ associated to ψ solves the ordinary differential equation

$$\frac{dX}{dt} = \nabla\psi(X). \tag{18}$$

The normal component of this ray must be negative (in Ω^+) - this is the incoming condition, and guarantees that the ray does not intersect $t = 0$ in Ω^+ near Σ, justifying the initial condition in 15. Similarly, the normal component must be positive in Ω^- near Σ. But both must have the same magnitude, since ψ solves the eikonal equation - so they are negatives of each other. For causal solutions such as $(\tilde{p},\tilde{\mathbf{v}})$, the Newton's law equation in 15 is equivalent to

$$\tilde{\mathbf{v}}(\mathbf{x},t) = -\frac{1}{\rho}\int_{-\infty}^{t} ds\nabla\tilde{p}(\mathbf{x},s), \tag{19}$$

except at Σ. it follows that the normal components of $\tilde{\mathbf{v}}$ have limits at point of Σ that are opposite in sign: for $\mathbf{x} \in \Sigma$,

$$\lim_{\zeta\to 0}(\tilde{\mathbf{v}}(\mathbf{x}+\zeta\mathbf{n}(\mathbf{x}),t)\cdot\mathbf{n}(\mathbf{x})) = -\lim_{\zeta\to 0}(\tilde{\mathbf{v}}(\mathbf{x}-\zeta\mathbf{n}(\mathbf{x}),t)\cdot\mathbf{n}(\mathbf{x}))$$

whence

$$[\tilde{\mathbf{v}}\cdot\mathbf{n}] = 2\lim_{\zeta\to 0}(\tilde{\mathbf{v}}(\mathbf{x}-\zeta\mathbf{n}(\mathbf{x}),t)\cdot\mathbf{n}(\mathbf{x})) \approx 2\mathbf{v}(\mathbf{x})\cdot\mathbf{n}(\mathbf{x}) \tag{20}$$

That is, at least over the short time interval $0 \leq t \leq \tau$, $(p,\mathbf{v})$ is asymptotically

approximated in Ω^+ by $(\tilde{p}, \tilde{\mathbf{v}})$, which solves

$$\begin{aligned} \frac{1}{\kappa}\frac{\partial \tilde{p}}{\partial t} &= h\delta_\Sigma - \nabla \cdot \tilde{\mathbf{v}}, \\ \rho\frac{\partial \tilde{\mathbf{v}}}{\partial t} &= -\nabla \tilde{p}, \\ \tilde{p} &= 0 \text{ for } \mathbf{x} \in \Omega, t = 0, \\ \tilde{\mathbf{v}} &= \mathbf{0} \text{ for } \mathbf{x} \in \Omega, t = 0. \end{aligned} \tag{21}$$

with $h = 2\mathbf{v} \cdot \mathbf{n}$ on Σ.

This result appears limited by the short time assumption (on τ) and the corresponding spatial localization to Ω^+. However neither of these apparent limitations is real. The short time assumption may be relaxed by windowing the fields to short time windows then adding up the results. The approximating fields $(\tilde{p}, \tilde{\mathbf{v}})$ are themselves approximated by ray theory solutions. The assumptions on rays crossing source and receiver surfaces are global, and extend the approximation 21 globally. Thus the last two conditions may be replaced by the basic causality condition

$$\begin{aligned} \frac{1}{\kappa}\frac{\partial \tilde{p}}{\partial t} &= h\delta_\Sigma - \nabla \cdot \tilde{\mathbf{v}}, \\ \rho\frac{\partial \tilde{\mathbf{v}}}{\partial t} &= -\nabla \tilde{p}, \\ \tilde{p} &= 0 \text{ for } t \ll 0, \\ \tilde{\mathbf{v}} &= \mathbf{0} \text{ for } t \ll 0. \end{aligned} \tag{22}$$

The effect of this change on the receiver surface traces of $\tilde{p}$ and $\tilde{\mathbf{v}}$ is negligible.

An example has already appeared in Figure 2b. In this example, the field $(p, \mathbf{v})$ is the point source solution with the point source depicted in Figure 1, located at $(3500, 3500)$ m. The role of Σ is played by the source surface $\Sigma_s = \{z = z_s = 3000 \text{ m}\}$, and near this surface the field solves the source-free system 14. The "inside" in this example is the region $z > z_s$, and the outward unit normal is therefore $\mathbf{n} = (0, 0, -1)^T$. Thus $(p, \mathbf{v})$ should be asymptotically close to the solution $(\tilde{p}, \tilde{\mathbf{v}})$ of the system 21 in the region $z < z_s$, at least at locations connected by the ray field shown in Figure 1 to the source surface, provided that the source amplitude h is chosen $= -2v_z$ on $\Sigma_s = \{z = z_s\}$. Figure 2b displays exactly this h, and 2c the traces of $\tilde{p}$ on the receiver surface $z = z_r = 1000$ m, which satisfies the ray condition. The extremely close match between Figures 2a (traces of p on Σ_r) and 2c is consistent with the conclusion reached here.

Approximate Inversion

The time-reversal construction can now be translated into the context of seismic surface source inversion. Assume that $(p, \mathbf{v})$ is a solution of the free-space acoustic

system 14, with high-frequency energy carried by a ray family that satisfies the criteria described earlier: incoming at the source surface Σ_s with normal field $\mathbf{n}_s$, outging at the receiver surface Σ_r with normal field $\mathbf{n}_r$, and every ray crossing Σ_r meets Σ_s. The goal of this subsection is to describe a procedure that constructs a source field h_s on Σ_s so that $Sh_s \approx P_r p$, thus approximately inverting the modeling operator S.

The sampled fields $(P_r p, P_r \mathbf{v})$ on Σ_r is negligible for large t. Due the outgoing assumption, $(p, \mathbf{v})$ are negligible for large t near Σ_r. The integrated Newton's law 19 acquires a minus sign if the integration from $-\infty$ to t is replaced by integration from t to ∞. Apply the construction of the last subsection in reverse time, to conclude that $(p, \mathbf{v})$ is approximated by $(\hat{p}, \hat{\mathbf{v}})$, where the latter solves the time-reversed problem

$$\begin{aligned} \frac{1}{\kappa}\frac{\partial \hat{p}}{\partial t} &= h_r \delta_{\Sigma_r} - \nabla \cdot \hat{\mathbf{v}}, \\ \rho \frac{\partial \hat{\mathbf{v}}}{\partial t} &= -\nabla \hat{p}, \\ \hat{p} &= 0 \text{ for } t \gg 0, \\ \hat{\mathbf{v}} &= \mathbf{0} \text{ for } t \gg 0. \end{aligned} \tag{23}$$

with $h_r = -2P_r(\mathbf{v} \cdot \mathbf{n}_r)$.

By construction, high-frequency energy in $(\hat{p}, \hat{\mathbf{v}})$ is carried by the same ray family as for $(p, \mathbf{v})$. Every ray passes over Σ_s and is incoming there. Since $(p, \mathbf{v})$ vanishes for $t \ll 0$ near Σ_s, $(\hat{p}, \hat{\mathbf{v}})$ is negligible there, So the construction detailed in the last two sections is applicable again, and the reverse-time field $(\hat{p}, \hat{\mathbf{v}})$ is approximated by the forward-time field $(\tilde{p}, \tilde{\mathbf{v}})$ that solves

$$\begin{aligned} \frac{1}{\kappa}\frac{\partial \tilde{p}}{\partial t} &= h_s \delta_{\Sigma_s} - \nabla \cdot \tilde{\mathbf{v}}, \\ \rho \frac{\partial \tilde{\mathbf{v}}}{\partial t} &= -\nabla \tilde{p}, \\ \tilde{p} &= 0 \text{ for } t \ll 0, \\ \tilde{\mathbf{v}} &= \mathbf{0} \text{ for } t \ll 0. \end{aligned} \tag{24}$$

with $h_s = 2P_s(\hat{\mathbf{v}} \cdot \mathbf{n}_r) \approx 2P_s(\mathbf{v} \cdot \mathbf{n}_r)$.

Summary

The surface source version of the time reversal construction involves three steps: given pressure data $P_r p$ on the receiver surface Σ_r for a solution $(p, \mathbf{v})$ of the system 14 satisfying the ray conditions,

1. create a source h_r on Σ_r by multiplying the corresponding normal velocity trace by -2: $h_r = -2P_r(\mathbf{v} \cdot \mathbf{n}_r)$;
2. propagate h_r backwards in time by solving the system 23 for the acoustic field

$(\hat{p}, \hat{\mathbf{v}})$;

3. create a source h_s on Σ_s by multiplying tne normal velocity trace of $\hat{\mathbf{v}} \cdot \mathbf{n}_s$ by 2: $h_s = 2P_s(\hat{\mathbf{v}} \cdot \mathbf{n}_s)$.

Then $P_r p \approx S h_s$.

The surface source representation part of the construction shows that under the assumed ray conditions, a solution of 14 is asymptotically the same near the receiver surface Σ_r as the solution of the system 9 with a suitable source h_s (and zero velocity source, $f_s = 0$). Therefore this construction also inverts the modeling operator S in the sense that $Sh_s \approx S\tilde{h_s}$ if $P_r p \approx S\tilde{h_s}$. That is, this construction produces an asymptotic right inverse of S: it does not necessily (approximately) recover the source that generates the data, but recovers some source that generates the data.

Example: Lens

The first example uses the acoustic lens model (Figure 1), with Σ_s = horizontal line at depth 3000 m, and Σ_r = horizontal line at depth 1000 m. The pressure data to be inverted is the sampling of the pressure field on Σ_r, generated by the point source shown in the figure, at $(3500, 3500)$ m (Figure 2a). However, I will invert this data in a homogeneous medium with $\kappa \equiv 4$ GPa. This choice illustrates two facts about surface source inversion and approximate inversion based on time reversal. First, the success of the inversion demonstrates the insensitivity of the time reversal method to ray multipathing (triplication), evident in the data (Figure 2a). Second, it is entirely possible to invert data in a material model other than the one in which it was produced (in the case of synthetic data, of course). This capability is critically important in the application of approximate source inversion in nonlinear extended 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 (and visible in Figure 1), the acoustic field in this example satisfies the ray conditions on which the time reversal construction is based: The rays are incoming at Σ_s, outgoing at Σ_r and all rays carrying significant energy at points of Σ_r pass over Σ_s. The outward unit normal at Σ_s is $\mathbf{n}_s = \mathbf{e}_z = (0, 1)$; at Σ_r, it's $\mathbf{n}_r = -\mathbf{e}_z = (0, -1)$. So the source on the receiver surface Σ_r, input to reverse time propagation (system 23, generating the fields $(\hat{p}, \hat{\mathbf{v}})$) is $h_r = 2P_r v_z$ (the field $P_r v_z$ is depicted in Figure 7a). After reverse time propagation , the approximate inversion $\hat{h}_s = 2P_s \hat{v}_z$ is shown in Figure 7b. It differs considerably from the source field used to generate the data (Figure 11a), however when used as the source in forward modeling (system 24) *with the same material parameters used in the inversion*, it generates a pressure gather on Σ_r (Figure 8a) very closely approximating the input gather (Figure 2c). The difference is plotted on the same color scale in Figure 8b.

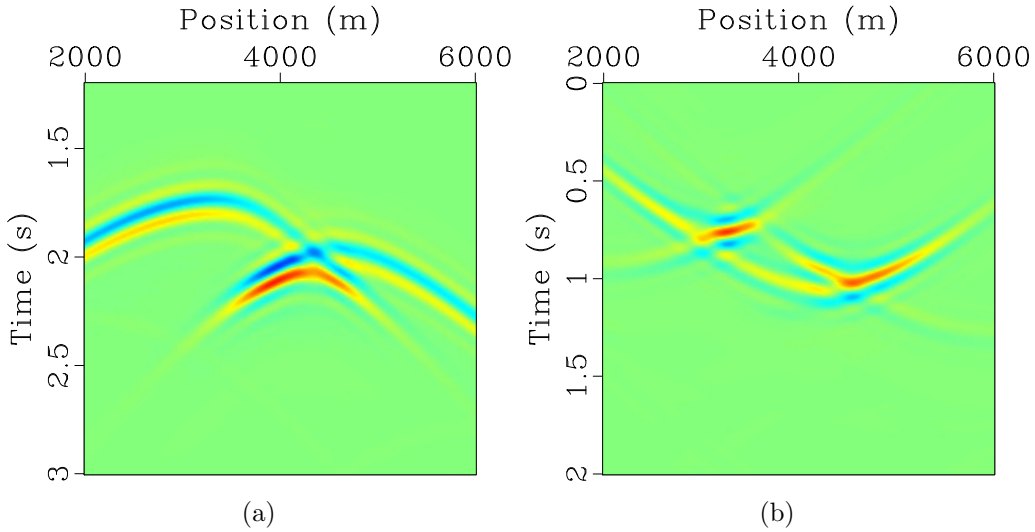


Figure 7: (a) Normal velocity data corresponding to pressure data in Figure 2c. (b) Source function resulting from time-reversal inversion of pressure data in Figure 2c. Propagation in homogeneous bulk modulus model (different from the model used to generate the data!). Plotted with same color scale as Figure 11a.

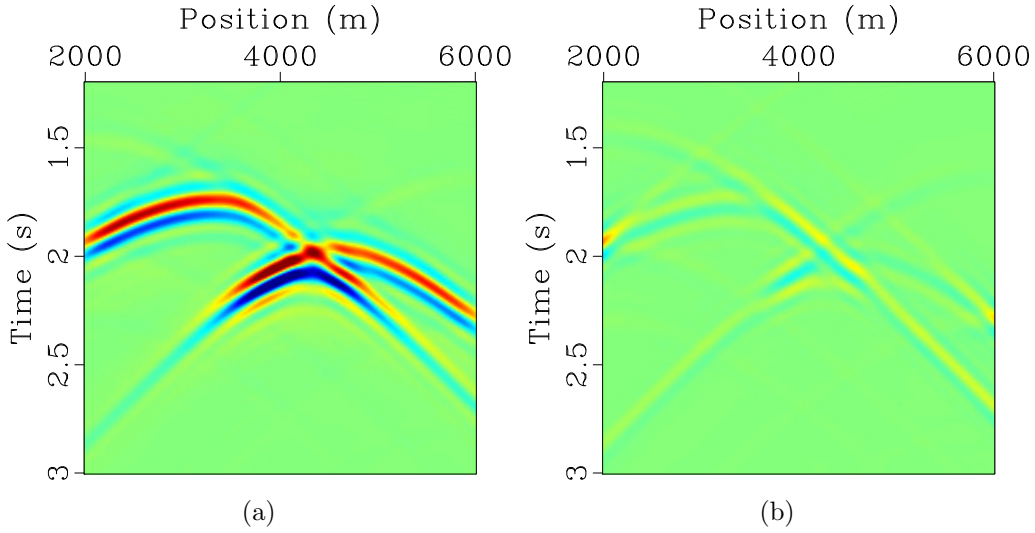


Figure 8: (a) Resimulated data: pressure gather generated from inverted source (Figure 7b), using same homogeneous bulk modulus as used in inversion. (b) Difference between data depicted in Figure 8a and in Figure 2c. Both (a) and (b) plotted with same color scale as Figure 2c.

Example: Diving Wave

The second example is based on the diving wave model (Figure 3). In this example, unlike the previous one, the source and receiver surfaces do not bound the propagating region, even in the vertical direction. Thus this exercise differs even more than the first from time reversal in photoacoustic tomography. Inspection of the blue part of the ray field in Figure 3 reveals that the rays are incoming at the source surface, and outgoing at the receiver surface, if the outward normal is chosen as $-\mathbf{e}_z$. The mute shown in Figure 5a eliminates the rest of the ray field. What remains satisfies the condition that every ray passing the receiver surface originates in the source surface.

In this example, the inversion is performed with the same model as used to generate the data. Note that sufficiently small changes in the model would result in similar ray configurations, so time reversal inversion could proceed as is shown here. However sufficiently large changes in the bulk modulus model could result in fundamental changes in the ray field: for instance, change to a homogenous model would completely eliminate the diving rays and the corresponding data component, thus making an inversion of the type presented here impossible.

The inversion proceeds otherwise as before. The source $h_r = -2P_r v_z$ on Σ_r is built from the normal component of velocity, shown in Figure 9a. Reverse time propagation to create the fields $(\tilde{p}, \tilde{\mathbf{v}})$, extraction of the normal component of velocity (note the sign difference with the previous example!) and scaling by the factor 2 produces the inverted source shown in Figure 9b. Forward propagation yields the pressure gather shown in Figure 10a, which closely matches the original diving wave pressure data (Figure 5b). The difference is displayed on the same color scale in Figure 10b.

PRESSURE-TO-SOURCE

The result of the last section has a couple of apparent defects. First, the approximate inversion of pressure data requires knowledge of velocity data, which is not prescribed in the problem statement 13. Second, the significance of the reverse-time field $(\hat{p}, \hat{\mathbf{v}})$ is obscure. In this section these two issues are addressed.

Pressure-to-source operator

The first step is to formalize the relation between pressure and source (or equally well, pressure and normal velocity) on as surface Σ as the application of an operator.

Observe that in principle it's not necessary to supply the normal velocity at the source (or receiver) surface, since it's determined locally by the pressure on the surface, or causal (or anti-causal) fields. This follows from the observation that the system 15 has a unique solution, a by-product of the conservation of energy (see for example Lax (2006), section 4.3). That is, in the language of that discussion, knowledge of p

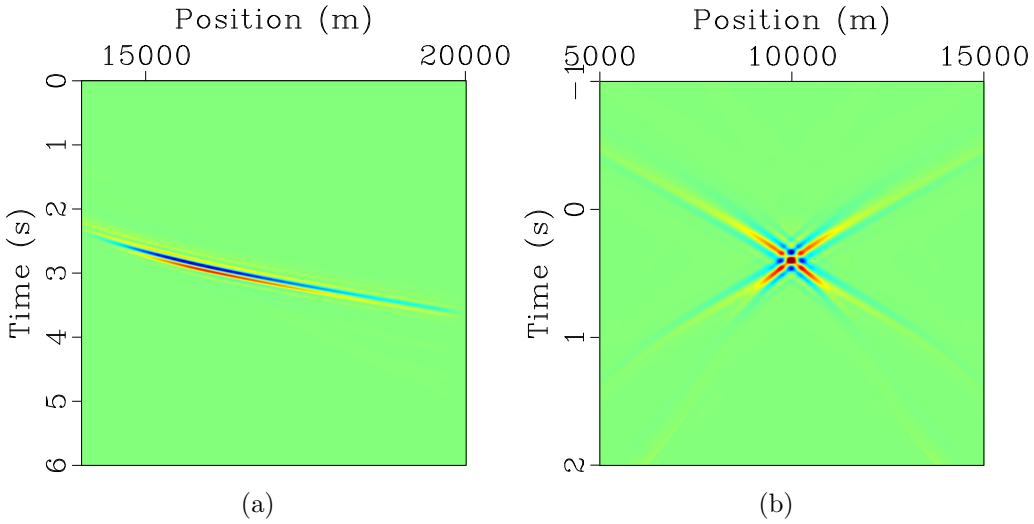


Figure 9: (a) Normal velocity data corresponding to pressure data in Figure 5b. (b) Source function resulting from time-reversal inversion of pressure data in Figure 5b. Propagation in same model (Figure 3) as used to generate the data.

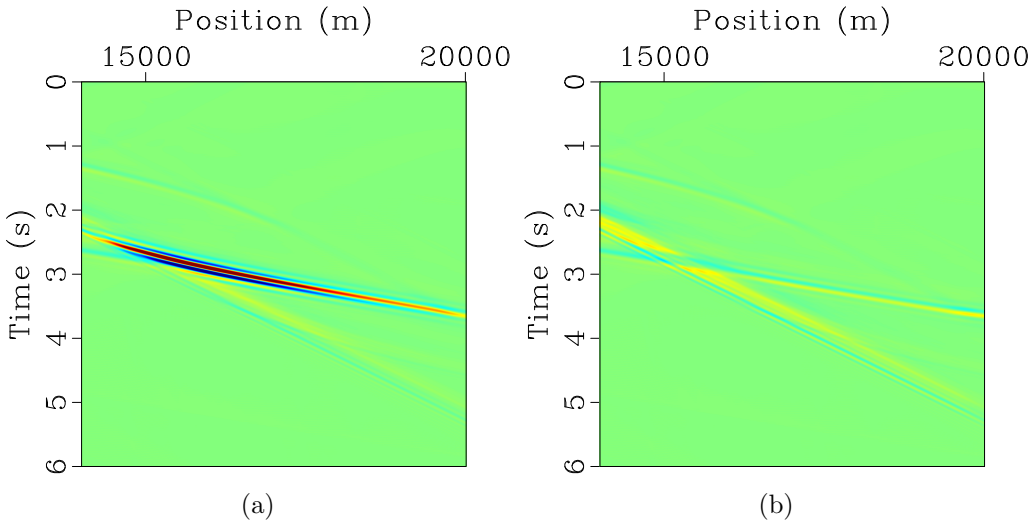


Figure 10: (a) Resimulated data: pressure gather generated from inverted source (Figure 9b), using same bulk modulus (Figure 3) as used in data generation and inversion. (b) Difference between data depicted in Figure 10a and in Figure 5b. Both (a) and (b) plotted with same color scale as Figure 5b. Note that low-amplitude "ghosts" of other arrival branches appear in the difference plot.

on Σ for $0 \le t \le \tau$ determines $(\tilde{p}, \tilde{\mathbf{v}})$ in the entire domain Ω^+, $0 \le t \le \tau$. Therefore knowledge of p on Σ for $0 \le t \le \tau$ determines $\mathbf{v} \cdot \mathbf{n}$, up to a negligible error. As in the last section, it follows that knowledge of p on Σ, any t, determines a source h on Σ for which the pressure field $\tilde{p}$ (solution of the system 22) differs negligibly from p.

This relation defines a causal *pressure-to-source operator* $\Lambda : p \mapsto h$, yielding the source on a surface Σ that reproduces the (causal) pressure field from its trace on Σ, at least locally on the inward side as defined by the selection of normal field on Σ, under the ray conditions defined earlier. Since $h \approx 2\mathbf{v} \cdot \mathbf{n}$ (limiting value of $\mathbf{v}$ from the inward direction), Λ is effectively the pressure-to-normal-velocity operator.

With this convention, two of the three steps in the approximate inversion procedure are operator applications. The operators involved are the causal pressure-to-source operator for the source surface Σ_s, denoted Λ_s, and an analogous anti-causal operator for the receiver surface Σ_r. The characterization of the the source field h_r in equation 23 shows that this anti-causal pressure-to-source operator differs from its causal analogue only by sign. Denote the causal operator by Λ_r.

Time Reversal as Adjoint State

The middle step is solution of the reverse-time system 23 for the field $(\hat{p}, \hat{\mathbf{v}})$, with input pressure source h_r on Σ_r and output $P_s\hat{p}$ on Σ_s. To understand the significance of this mapping, it is just as easy, and will be useful later, to include a velocity source f_r, that is, modify the second equation in 23 by adding $f_r\mathbf{n}_r\delta_{\Sigma_r}$, so that it becomes a reverse-time version of the basic system 9. Recall the definition of energy in the acoustic field $(p, \mathbf{v})$:

$$E[p, \mathbf{v}](t) = \frac{1}{2}\int dx\, dy\, dz\, \left(\frac{p(\mathbf{x}, t)^2}{\kappa(\mathbf{x})} + \rho(\mathbf{x})\mathbf{v}(\mathbf{x}, t) \cdot \mathbf{v}(\mathbf{x}, t)\right) \tag{25}$$

This is a quadratic form. Polarization makes a bilinear form, into which one can plug a pair of fields. Use the solutions $(p, \mathbf{v})$ of the forward-time system 9 for one of the pair, and the solution $(\hat{p}, \hat{\mathbf{v}})$ of system 23 (augmented with the velocity surface source) for the other. Since $(p, \mathbf{v})$ vanishes for large negative time, $(\hat{p}, \hat{\mathbf{v}})$ for large positive time,

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

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

Carry out the differentiation under the integral sign, replace time derivatives by space derivatives and source terms using the equations of motion in the systems 9 and 23, and integrate by parts to move all of the space derivatives onto p and $\mathbf{v}$. Most of the

remaining expressions cancel, leaving

$$= \int_{-\infty}^{\infty} dt \left[\int d\sigma_s(\mathbf{x}) \left(h_s P_s \hat{p} + f_s P_s(\hat{\mathbf{v}} \cdot \mathbf{n}_s) \right) + \int d\sigma_r(\mathbf{x}) \left(h_r P_r p + f_r P_r(\mathbf{v} \cdot \mathbf{n}_r) \right) \right]$$

Here σ_s and σ_r are the area elements on Σ_s and Σ_r. These integrals are inner products of time traces on Σ_s and Σ_r. Denote these inner products by $\langle \cdot, \cdot \rangle_s$ and $\langle \cdot, \cdot \rangle_r$ respectively. Then the above is written in terms of the vector modeling operator $\mathcal{S}$ (definition 10) as

$$= \langle (h_s, f_s)^T, (P_s \hat{p}, P_s(\hat{\mathbf{v}} \cdot \mathbf{n}_s))^T \rangle_s + \langle (h_r, f_r)^T, \mathcal{S}(h_s, f_s)^T \rangle_r,$$

whence identify the adjoint of $\mathcal{S}$ as

$$\mathcal{S}^T(h_r, f_r) = -(P_s \hat{p}, P_s(\hat{\mathbf{v}} \cdot \mathbf{n}_s))^T. \tag{26}$$

Specializing to the pressure modeling operator S (equation 12),

$$S^T = (\Pi_0 \mathcal{S} \Pi_0^T)^T = \Pi_0 \mathcal{S}^T \Pi_0^T \tag{27}$$

so to calculate the adjoint of S applied to h_r, solve system 23 (with $f_r = 0$) and extract the pressure at Σ_s:

$$S^T(h_r) = -P_s \hat{p}.$$

Combined with the definitions of the pressure-to-source operators Λ_s and Λ_r, these results enable a translation of the approximate inverse construction in the last section into a chain of operator applications: to invert a pressure gather d on Σ_r for a source h_s on Σ_s that approximately reproduces it ($d \approx S h_s$),

1. create a source $h_r = -2\Lambda_r d$ on Σ_r;
2. apply the adjoint S^T (solve the reverse time system 23) to h_r;
3. apply Λ_s to $-S^T d$.

Putting this all together, an approximate right inverse operator $S^\dagger$ to S may be expressed as

$$S^\dagger = \Lambda_s S^T \Lambda_r, \tag{28}$$

that is,

$$S S^\dagger \approx I. \tag{29}$$

Positivity of Λ

Equation 28 is remarkable, in that the form of $S^\dagger$ is that of an adjoint with respect to weighted norms, defined by the operators Λ_s^{-1} in the domain of S and Λ_r in the

range. This only makes sense if the pressure-to-source operators are positive definite and symmetric, the necessary conditions to define weighted norms.

In fact the pressure-to-source operator is positive semi-definite, and that fact has fundamental physical significance. Suppose that $(p, \mathbf{v})$ solves the basic system 9 with $f_s = 0$. From the definition of energy 25 and causality of the fields,

$$E[p,\mathbf{v}](t_1) = \frac{1}{2}\int_{-\infty}^{t_1} dt \int dx\, dy\, dz\, \frac{d}{dt}\left(\frac{p(\mathbf{x},t)^2}{\kappa(\mathbf{x})} + \rho(\mathbf{x})\mathbf{v}(\mathbf{x},t)\cdot\mathbf{v}(\mathbf{x},t)\right)$$

$$= \int_{-\infty}^{t_1} dt \int dx\, dy\, dz\, \left(\frac{p(\mathbf{x},t)\frac{\partial p}{\partial t}(\mathbf{x},t)^2}{\kappa(\mathbf{x})} + \rho(\mathbf{x})\mathbf{v}(\mathbf{x},t)\cdot\partial\mathbf{v}\partial t(\mathbf{x},t)\right)$$

Using 9, replace the time derivatives with spatial derivatives and the delta function on ths source surface, integrate by parts to shift the space derivatives to p. Everything cancels except for

$$= \int_{-\infty}^{t_1} dt \int dx\, dy\, dz\, p(\mathbf{x},t)h_s(\mathbf{x},t)\delta_{\Sigma_s}(\mathbf{x}) = \int_{-\infty}^{t_1} dt \int_{\Sigma_s} d\sigma_s\, p h_s.$$

Take the limit as $t_1 \to \infty$ to get

$$\lim_{t_1\to\infty} E[p,\mathbf{v}](t_1) = \langle h_s, P_s p\rangle_s = \langle P_s p, \Lambda_s P_s p\rangle_s.$$

That is, the value of the quadratic form defined by Λ_s, evaluated at the pressure trace on Σ_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 Λ_s is positive semi-definite.

Approximate symmetry of Λ

Except in special cases, $\Lambda_{s,r}$ is not symmetric. However, it is *approximately symmetric* in the high-frequency sense. This fact follows from a high-frequency asymptotic analysis, a sketch of which follows for the special case of a planar source surface $\Lambda_s = \{z = z_s\}$, as in the two examples. Since the same reasoning applies to Λ_r, I will omit the subscript for the remainder of this subsection.

The pressure trace $p(x, y, t)$ can be expressed by Fourier synthesis as

$$p(x,y,t) = \frac{1}{(2\pi)^3}\int d\omega\, d\xi\, d\eta\, \omega^2\, d\eta\, \hat{p}(\omega,\xi,\eta)e^{i\omega(t-x\xi-y\eta)} \tag{30}$$

in which the spatial frequencies have been expresed as slownesses ξ, η scaled by temporal frequency ω. Choose $a_0(x, y, t)$ so that $a_0(x, y, t) = 1$ if $p(x, y, t) \neq 0$. Then

$$p(x,y,t) = a_0(x,y,t)p(x,y.t)$$

$$= \frac{1}{(2\pi)^3}\int d\omega\, d\xi\, d\eta\, \omega^2\, d\eta\, \hat{p}(\omega,\xi,\eta)[a_0(x,y,t)e^{i\omega(t-x\xi-y\eta)}] \tag{31}$$

This identity expresses $p(x,y,t)$ as a sythesis of oscillatory wave packets of the form

$$a_0(x,y,t)e^{i\omega(t-x\xi-y\eta)} \tag{32}$$

in which the spatial frequencies have been expressed as slownesses ξ,η scaled by temporal frequency ω. The function a_0 is a frequency-independent envelope, equal to 1 where h is non-zero.

The wave packet $a_0(x,y,t)e^{i\omega(t-x\xi-y\eta)}$ can be viewed as the boundary value on Σ of a high-frequency asymptotic solution of the form 17, provided that the slowness vector (ξ,η) satisfies the propagating condition $c^2(\mathbf{x})(\xi^2+\eta^2) < 1$ at all $\mathbf{x}$ for which $a_0(\mathbf{x},t)$ is non-zero for any t. [In general, this is a strong condition, and might not even have any solutions. However it always has solutions if $a_0(x,y,t)$ is nonzero only over a sufficiently small region, which is the case is $p(x,y,t)$ is non-zero over a sufficiently small region. An arbitrary $p(x,y,t)$ can always be broken up into a sum, the summands of which are non-zero only over small regions, so in fact the requirement mentioned above is not a significant constraint: just carry out the following construction for summands, and add up the results.] In that case, two solutions ψ of the eikonal equation $|\nabla\psi| = c$ exist, at least locally near Σ, for which $\psi(x,y,0) = x\xi + y\eta$. Choose the solution that increases in the direction of the outward unit normal $\mathbf{n}$, and solve the corresponding transport equation for an amplitude a satisfying $a(x,y,0,t) = a_0(x,y,t)$. Then the geometric optics solution $ae^{i\omega(t-\psi)}$ is asymptotic to the pressure field of an incoming solution. From the evolution equation for $\mathbf{v}$, integrated in time as in equation 19, an expression follows for the corresponding normal velocity:

$$-\mathbf{n}\cdot\frac{1}{\rho}\int_{-\infty}^{t} ds\nabla(ae^{i\omega(s-\psi)}) = \frac{1}{\rho}\int_{-\infty}^{t} ds\mathbf{n}\cdot(i\omega a\nabla\psi - \nabla a)e^{i\omega(s-\psi)})$$

Now use the identity

$$e^{i\omega(s-\psi)} = \frac{1}{i\omega}\frac{d}{ds}e^{i\omega(s-\psi)}$$

and integrate by parts in s to obtain the expression for the normal velocity field

$$= \frac{a}{\rho}(\mathbf{n}\cdot\nabla\psi)e^{i\omega(t-\psi)} + O(1/\omega).$$

The terms not written explicitly begin with one of order -1 in ω. The remainder appears to be also of order -1, but application of the integration-by-parts trick described above yields another term of order -2 plus another remainder of apparent order -2. Continuing this process develops the $O(1/\omega)$ remainder in an asymptotic series of homogeneous terms in $1/\omega$ of arbitrary length.

An expression for the normal velocity on Σ follows by restriction. Note that

$\mathbf{n} \cdot \nabla\psi = (c^{-2}(x,y,z) - (\xi^2 + \eta^2))^{1/2}$ because of the choice of sign and the eikonal equation, and the root is real because of the propagating assumption.

The upshot of this calculation is that to leading order in frequency, the normal velocity corresponding to a pressure wave packet differs by a position- and slowness-dependent factor. Accounting for the factor of 2 difference between the source amplitude and velocity trace, this calculation may be summarized:

$$\Lambda(a_0 e^{i\omega(t-x\xi-y\eta)}) = \lambda(a_0 e^{i\omega(t-x\xi-y\eta)}), \tag{33}$$

where $\lambda(x,y,t,\xi,\eta,\omega)$ is defined by the asymptotic series in ω developed above. Its dominant term, $O(1)$ (order 0) in ω, is

$$\begin{aligned}\lambda^0(x,y,t,\xi,\eta,\omega) &= 2\frac{(c^{-2}(x,y,z) - \xi^2 + \eta^2)^{1/2}}{\rho} + O(1/\omega) \\ &= 2(\kappa\rho)^{-1/2}\left(1 - \frac{\kappa(k_x^2 + k_y^2)}{\rho\omega^2}\right)^{1/2} + O(1/\omega).\end{aligned} \tag{34}$$

In the final expression, the slownesses ξ, η have been replaced by the Fourier variables $k_x = x\xi, k_y = y\eta$. Note that as a function of k_x, k_y, ω, λ^0 is homogeneous of order 0. It is easy to see that the Nth term of the asymptotic series defining λ is of homogeneous order $-N$.

Operators that affect oscillatory wave packets by multiplication with functions having the properties of λ have come to be called "pseudodifferential" (Nirenberg, 1972; Taylor, 1981; Saint Raymond, 1991). The multiplier (λ in this case) is called the *symbol* of the operator, and the dominant term in an asymptotic development of the multiplier (for instance λ^0) is called the *principal symbol*. Amongst other special properties of these operators, their adjoints are operators of the same type. In particular, the principal symbol of the adjoint is the Hermitian (matrix) adjoint of the principal symbol. For scalar operators like Λ (mapping scalar-valued functions to scalar-valued functions), the symbols are scalar (1×1), and the Hermitian adjoint is simply the complex conjugate (Taylor (1981), section II.4). Therefore scalar operators with real principal symbol, like Λ, have the same principal symbols as their adjoints. So Λ differs from Λ^T by a pseudodifferential operator whose symbol is an asymptotic sum of terms of negative order in frequency. Therefore the difference is negligible in the limit of high frequencies. This observation establishes the conclusion of this subsection:

$$\Lambda^T \approx \Lambda, \tag{35}$$

Before leaving this topic, recall that the computation outlined above was predicated on the propagating assumption, namely that the slowness parameters ξ, η of the wave packet defined in 32 satisfy $c(x,y,z_s)^2(\xi^2 + \eta^2) < 1$ whereever the amplitude $a_0(x,y,t)$ is non-zero. Therefore the conclusion 35 holds only for application to such wave packets. However, the restriction of propagating waves (such as those that are output by the modeling operator S or its adjoint S^T) are well-approximated

by sums over propagating wave packets: this is part of the assumption that the rays carrying high-frequency energy cross the source and receiver surfaces at non-zero angles. To make the definition of Λ complete, it must be extended to non-propagating wave packets, for instance by making its effect negligible via a dip filter. I will ignore this refinement, since the application to non-propagating components appears to play little role in approximate inversion and its applications as developed here.

COMPUTING Λ

The missing ingredient in the preceding discussion is any indication of how the action of the pressure-to-source operators Λ is to be computed. As will be seen in the next section, an (approximate) inverse of this operator is also required in the formulation of preconditioned Krylov iteration.

Direct computation of the pressure-to-source operator Λ_s, for instance, by solving 9 and reading off $P_s v_z$, turns out to be numerically ill-behaved. An alternate approach begins with the observation that a pressure gather on Σ_r produced by a pressure source h_s on Σ_s can also be well-approximated by the field from a velocity source. To see this, recall the first major step in the time reversal construction, that a solution of the wave equation near Σ_s can be represented as the solution of a boundary value problem with p given on Σ_s (system 15). The derivation combined the unique determination of the field by its pressure boundary values on an enclosing domain with the consequences of the standing assumptions about rays carrying high frequency energy. Precisely the same logic could have been applied to the normal velocity boundary values, as these also uniquely determine the field in an enclosing domain. Proceeding further, a source representation based on continuous extension of normal velocity, with a jump in the pressure follows by logic precisely analogous to that leading to the system 21: instead, obtain a field $(\tilde{p}, \tilde{\mathbf{v}})$ differing negligibly from the field of the same name in the earlier discussion, solving

$$\begin{aligned} \frac{1}{\kappa}\frac{\partial \tilde{p}}{\partial t} &= -\nabla \cdot \tilde{\mathbf{v}}, \\ \rho\frac{\partial \tilde{\mathbf{v}}}{\partial t} &= f_s \mathbf{n}\delta_\Sigma - \nabla \tilde{p}, \\ \tilde{p} &= 0 \text{ for } t \ll 0, \\ \tilde{\mathbf{v}} &= \mathbf{0} \text{ for } t \ll 0. \end{aligned} \tag{36}$$

(that is, system 9 with $h = 0$) with $f_s = 2P_s p$.

This fact can usefully be rewritten using the vector modeling operator $\mathcal{S}$ (definition 10), and the projection operators Π_i (definition 11), which insert scalar sources into the pairs that occur in the acoustic system 9: explicitly, $\Pi_0(h, f)^T = h, \Pi_1(h, f)^T =$

$f, \Pi_0^T h = (h, 0), \Pi_1^T f = (0, f)$. Then

$$\begin{aligned} P_r\tilde{p} &= \Pi_0 \mathcal{S} \Pi_1^T f_s, \\ P_r(\mathbf{n}_r \cdot \tilde{\mathbf{v}}) &= \Pi_1 \mathcal{S} \Pi_1^T f_s. \end{aligned} \tag{37}$$

Since the field $(\tilde{p}, \tilde{\mathbf{v}})$ constructed here differs negligibly from that constructed earlier with the same name, which in turn differs negligibly from p, it follows that $P_r\tilde{p} \approx P_r p = Sh_s$. By definition of S (12), this is the same as $= \Pi_0 \mathcal{S} \Pi_0^T h_s$. Since the velocity fields also differ negligibly, obtain

$$\begin{aligned} P_r\tilde{p} &\approx \Pi_0 \mathcal{S} \Pi_0^T h_s, \\ P_r(\mathbf{n}_r \cdot \tilde{\mathbf{v}}) &\approx \Pi_1 \mathcal{S} \Pi_0^T h_s. \end{aligned} \tag{38}$$

From the definition of Λ_s, $h_s = \Lambda_s P_s p = \frac{1}{2}\Lambda_s f_s$. Combining 37, 38, and this last identity, obtain

$$\mathcal{S}\left(-\frac{1}{2}\Lambda_{z_s} f_s, f_s\right)^T \approx 0. \tag{39}$$

That is, the same data gather on Σ_r can be obtained either with a pressure source h_s or with a velocity source f_s, and there is a definite relation betwen the two. To illustrate this relation, recall that for the first of the two examples introduced above, the velocity source f_s is twice the pressure trace shown in Figure 11a, and the pressure source (Figure 2b) is twice the normal velocity trace shown in Figure 11b. Figure 12b shows $\Pi_0 \mathcal{S} \Pi_1^T f_s$, Figure 12a shows $\Pi_0 \mathcal{S} \Pi_0^T h_s$, and Figure 12c shows the difference on the same color scale.

One other important consequence of this duality is an alternate view of the time reversal construction of the approximate inverse. The entire construction can be carried out with traces of the normal velocity in place of trace of the pressure. The steps are precisely analogous, and justified in exactly the same way. Therefore I will simply summarize the result: to invert a pressure gather d on Σ_r for a source f_s on Σ_s that approximately reproduces it ($d \approx \Pi_0 \mathcal{S} \Pi_1^T f_s$):

1. create a source $f_r = -2d$ on Σ_r;
2. apply the adjoint $(\Pi_1 \mathcal{S} \Pi_1^T)^T = \Pi_1 \mathcal{S}^T \Pi_1^T$ (solve the reverse time system

$$\begin{aligned} \frac{1}{\kappa}\frac{\partial \hat{p}}{\partial t} &= -\nabla \cdot \hat{\mathbf{v}}, \\ \rho \frac{\partial \hat{\mathbf{v}}}{\partial t} &= f_r \mathbf{n} r \delta_{\Sigma_r} - \nabla \hat{p}, \\ \hat{p} &= 0 \text{ for } t \gg 0, \\ \hat{\mathbf{v}} &= \mathbf{0} \text{ for } t \gg 0. \end{aligned} \tag{40}$$

3. then extract $h_s = 2\mathbf{n_s} \cdot P_s \hat{\mathbf{v}}. = 2\Pi_1 \mathcal{S}^T \Pi_1^T f_r = -4\Pi_1 \mathcal{S}^T \Pi_1^T d$.

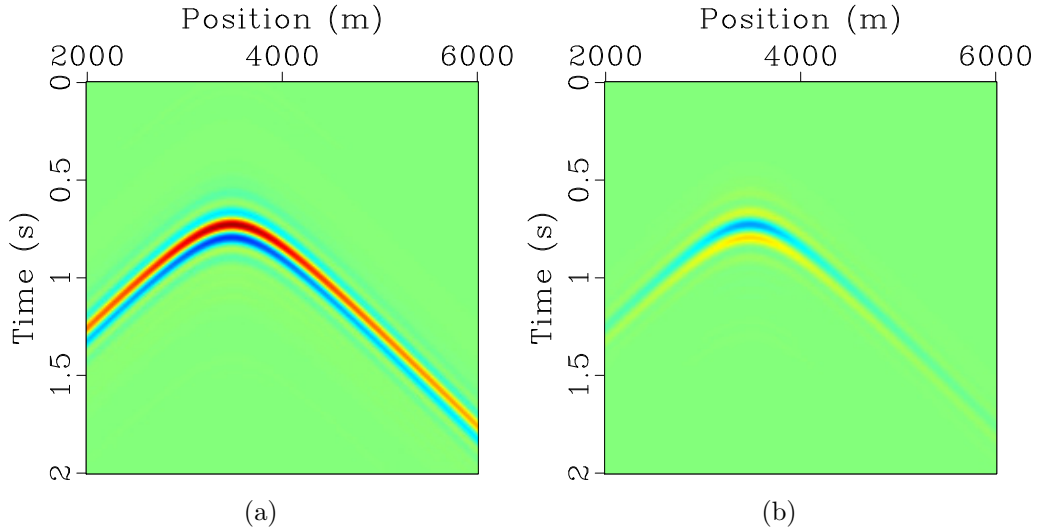


Figure 11: (a) Pressure gather for Lens example on source surface $\Sigma_s = \{z = 1000 \text{ m}\}$, generated by point source at $(3500, 3500)$ m. (b) Normal velocity gather for Lens example on source surface $\Sigma_s = \{z = 1000 \text{ m}\}$, generated by point source at $(3500, 3500)$ m.

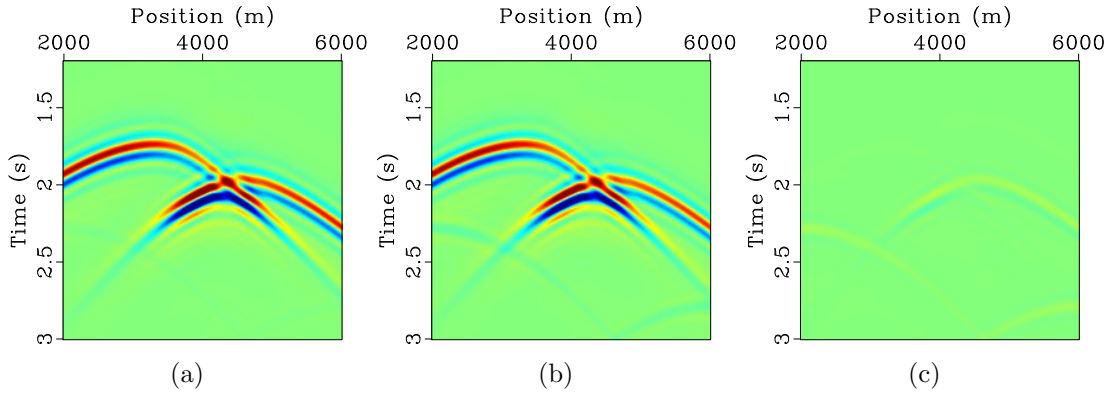


Figure 12: (a) Pressure gather on receiver surface $\Sigma_r = \{z = 1000 \text{ m}\}$ due to extended pressure source = 2 × normal velocity gather shown in Figure 11b. This is the same gather as shown in Figure 2c, and the extended pressure source has been shown in Figure 2b. (b) Pressure gather on receiver surface $\Sigma_r = \{z = 1000 \text{ m}\}$ due to extended velocity source = 2 × pressure gather shown in Figure 11b. (c) Difference (b)-(a), plotted on same color scale.

The upshot of this construction is the alternate computation for the approximate inverse:

$$S^\dagger \approx -4\Pi_1 \mathcal{S}^T \Pi_1^T. \tag{41}$$

The first row of 39, slightly rearranged, reads

$$\Pi_0 \mathcal{S} \Pi_1^T f_s \approx \frac{1}{2} S \Lambda_s f_s. \tag{42}$$

Apply the approximate inverse construction in the form 41 to solve 42 for Λ_s:

$$\Lambda_s \approx -8\Pi_1 \mathcal{S}^T \Pi_1^T \Pi_0 \mathcal{S} \Pi_1^T. \tag{43}$$

This identity is the major result of this section: it shows how to compute that action of Λ_s by propagating the input pressure trace, identified as a source for the velocity evolution, forward in time reading off the pressure trace on Σ_r, identifying it once more as a point load (source for velocity), propagating it backwards in time, and finally reading off the velocity trace, scaled to be a pressure evolution source on Σ_s.

The importance of this result lies in the failure of the obvious method for computing the action of Λ_s, namely to employ the pressure trace as a source in the velocity equation (f_s, in the notation used above) at Σ_s, and read off the velocity field also at Σ_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 43 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.

It is implicit in the discussion above of the pressure-to-source operator depends only on the model coefficients near the source surface Σ_s. However an even more useful observation is that the calculations in the approximation 43 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. Take advantage of this observation by replacing - for the purpose of evaluating the expression 43 only - the receiver surface by an auxiliary receiver surface Σ_a near Σ_s but a small positive distance from it. Define the auxiliary operator $\mathcal{S}_a$ by sampling solution of the system 9 on Σ_a rather than Σ_r. Thus obtain an approximation for Λ_s

$$\Lambda_s \approx -8\Pi_1 \mathcal{S}_a^T \Pi_1^T \Pi_0 \mathcal{S}_a \Pi_1^T. \tag{44}$$

Using an auxiliary receiver datum closer to the source surface to compute that pressure-to-source maps (only!) 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 43. If the domain is small enough, the ray conditions imply that no ray re-enters the region between the source suface and the auxiliary surface, therefore an absorbing boundary condition can be applied outside of

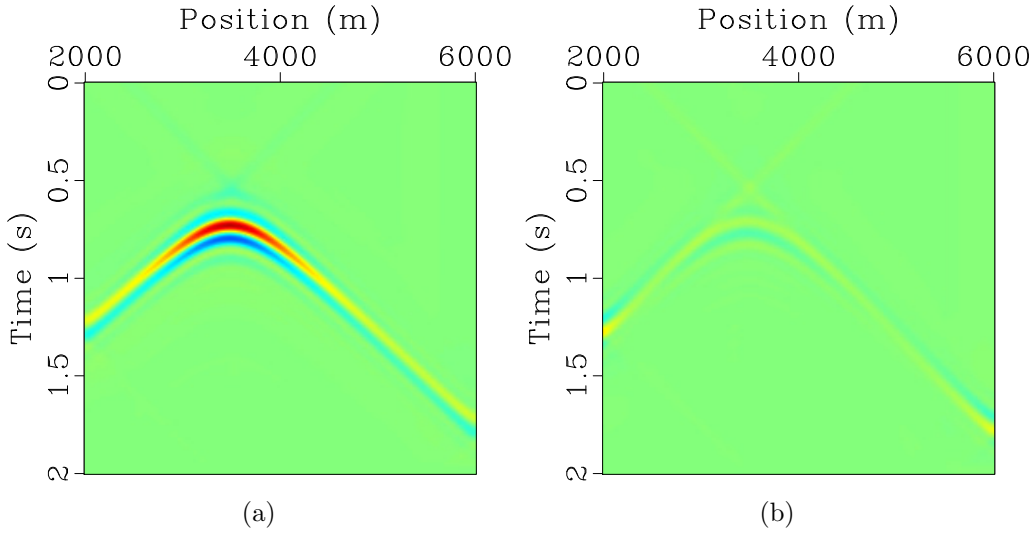


Figure 13: (a) Pressure source gather = image under pressure-to-source operator Λ_s of pressure gather shown in Figure 11a, computed using auxiliary receiver surface"near" at $z = 2900$ m and equation 44. Compare Figure 2b: because the sources and receivers are close, little aperture is lost in this case. (b) Difference between data plotted in (a) and source gather inferred from vertical velocity at source surface ($z = 3000$ m) (Figure 2b). Same color scale as in Figure (a).

this region, considerably shrinking the computational domain and substantial improving computational efficiency. The results shown below used such a truncation strategy.

- 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 43 can be much larger, producing an estimated source gather much less affected by aperture limitation.

Figures 13a and 13b show an example of this construction.

In the iterative methods to be discussed in the next section, it will be important to be able to compute the transpose of the pressure-to-source operator - not approximately, but to machine precision. The prescription 44 only defines an approximation. However simply substituting the right-hand side of equation 44 - in effect, substituting the approximation for the operator that it approximates - gives an operator variant of Λ_s with an accessible machine-precision adjoint:

$$\Lambda_s^T \approx -8\Pi_1 \mathcal{S}_a^T \Pi_0^T \Pi_1 \mathcal{S}_a \Pi_1^T. \tag{45}$$

From now on, when Λ_s or Λ_s^T appear, the symbols are understood to mean the approximations 44 and 45.

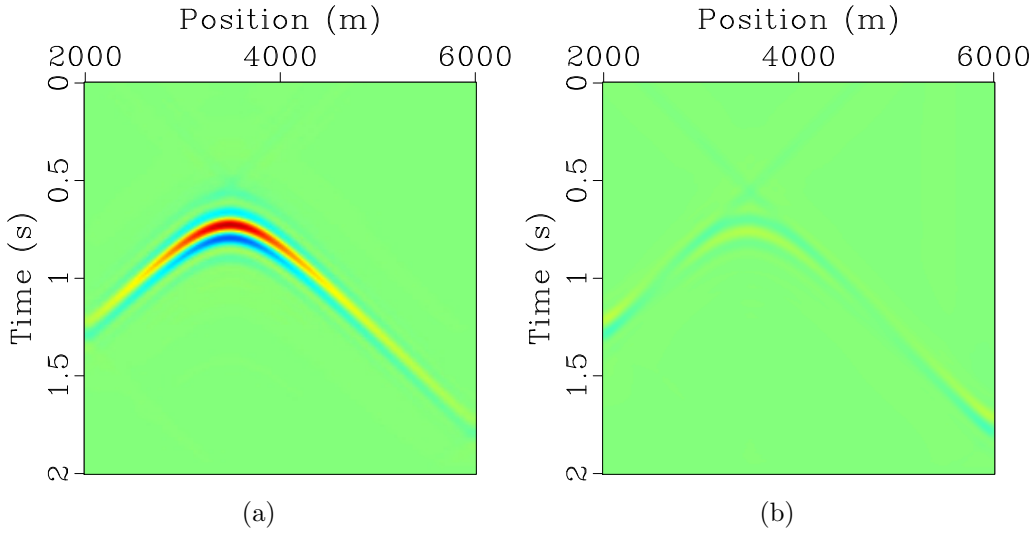


Figure 14: (a) Pressure source gather = image under transposed pressure-to-source operator Λ_s^T of pressure gather shown in Figure 11a, computed using auxiliary receiver surface"near" at $z = 2900$ m and equation 45. Compare Figure 2b: because the sources and receivers are close, little aperture is lost in this case. (b) Difference between (a) and source gather inferred from vertical velocity at source surface ($z = 3000$ m) (Figure 2b). Same color scale as (a).

Figure 14a shows the image of the pressure gather in Figure 11a under Λ_s^T approximated as in equation 45, using the auxiliary receiver surface at $z = 2900$. Note the close resemblance to the image of the same pressure gather under Λ_s displayed in Figure 13a. The difference of these two images is displayed in 14b, on the same color scale as the images themselves. Note that propagation takes place entirely in a region where the mechanical parameters are homogenous.

ACCELERATED ITERATIVE INVERSION

The preceding sections provide an approximate inverse for the modeling operator S and a means to compute it effectively. These constitute the basic ingredients of an effective preconditioning strategy. The remaining tasks are: the construction of weighted norms making S approximately unitary (insofar as possible), definition of a formal preconditioned conjugate gradient algorithm; and modification of the symmetrized approximate inverse to account for the presence of the penalty term. Examples illustrate the effectiveness of the resulting preconditioning strategy.

Symmetrization and Weighted Norms

Since Λ_s approximates its transpose (equation 35), it also approximates its symmetrization:

$$\frac{1}{2}(\Lambda_s^T + \Lambda_s) \approx \Lambda_s. \tag{46}$$

and similarly for Λ_r.

Define

$$\begin{aligned} W_m^{-1} &= \frac{1}{2}(\Lambda_s^T + \Lambda_s), \\ W_d &= \frac{1}{2}(\Lambda_r^T + \Lambda_r) \end{aligned} \tag{47}$$

(the subscripts signify "model space" and "data space" of S). These operators define (semi)-norms in the spaces of finite energy pressure fields on Σ_s and Σ_r:

$$\|p_s\|^2 = \langle p_s, W_m p_s\rangle, \|p_r\|^2 = \langle p_r, W_d p_r\rangle,$$

The identification of the symmetrized Λ_s as the inverse of another operator W_m is formal at this point, since the former operator is likely to have null (or nearly-null) vectors due to aperture-related amplitude loss, and certainly due to asymptotic vanishing at non-propagating waves. 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.

The adjoint of S with respect to the weighted norms is given approximately by

$$S^\dagger \approx W_m^{-1} S^T W_d, \tag{48}$$

As we have shown (equations 28, 29), $S^\dagger$ is an approximate right inverse to S: repeating equation 29 for convenience,

$$SS^\dagger \approx I. \tag{49}$$

This is not quite enough to establish that S is approximately unitary with respect to the weighted norms defined by W_m and W_d, since that would require that $S^\dagger$ also be an approximate left inverse. The ray conditions assumed throughout include the presumption that all rays carrying high frequency energy arriving at the receiver surface (within the support of the sampling operator) pass over the source surface. Consequently a source supported on the source surface exists that produces this data. However it was not assumed that all rays passing over the source surface arrive at the receiver surface. Therefore high frequency energy can be emitted from the source surface without arriving at the receiver surface. We have seen examples of this phenomenon already. The extended source shown in Figure 9b generates the diving wave data shown in Figure 5b to close approximation, even though the latter was generated by a point source. The difference between the extended source in Figure

9b and the point source depicted in Figure 3 is an approximate null vector for the modeling operator S as defined for this example.

In fact, that is true for the difference between $S^\dagger Sh$ and h for any extended source h:

$$S(S^\dagger Sh - h) = (SS^\dagger - I)Sh \approx 0.$$

Suppose that n is an approximate null vector for S, and h is any other extended source. Then

$$\langle S^\dagger Sh, n\rangle_m = \langle Sh, Sn\rangle_d \approx 0.$$

That is, $S^\dagger S$ is the (approximate) orthogonal projector onto the orthocomplement of the null space of S. So iterations such as CG based on $S^\dagger S$ construct approximation sequences lying close to this orthocomplement, as if the null space did not exist. This observation allows us to treat S as if it were unitary in discussing the behaviour of Krylov space algorithms based on the weighted spaces, even though it actually is not.

Similar relations have been derived for other wave-theoretic modeling operators, and have been used to accelerate iterative solutions of inverse scatering problems: Dafni and Symes (2018) review some of this literature.

Concerning computation, just as for the pressure-to-source operators, the weight operators are to be understood as replaced by their approximations via equations 44 and 45:

$$W_m^{-1} = \frac{1}{2}(\Lambda_s + \Lambda_s^T)$$
$$-4\Pi_1 \mathcal{S}_a^T(\Pi_0^T\Pi_1 + \Pi_1^T\Pi_0)\mathcal{S}_a\Pi_0. \tag{50}$$

with a similar approximate realization of W_d.

Figure 15a shows the output of the symmetrized approximate pressure-to-source operator per equation 50, applied once again to the pressure data in Figure 11a. Note the resemblance to Figures 13a and 14a. These are all asymptotic approximations of each other. Figure 15b 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 2c), plotted on the same scale as the latter, in both cases with all propagations in the lens model.

The identity 50 shows that only one forward and one adjoint simulation are necessary to compute the action of W_m^{-1} or W_d, and only over the truncated computational domain between the source or receiver surface and its auxiliary, nearby surface. 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 55, below: it is the weight matrix for both the domain and range norms for A. It is perfectly OK for one of

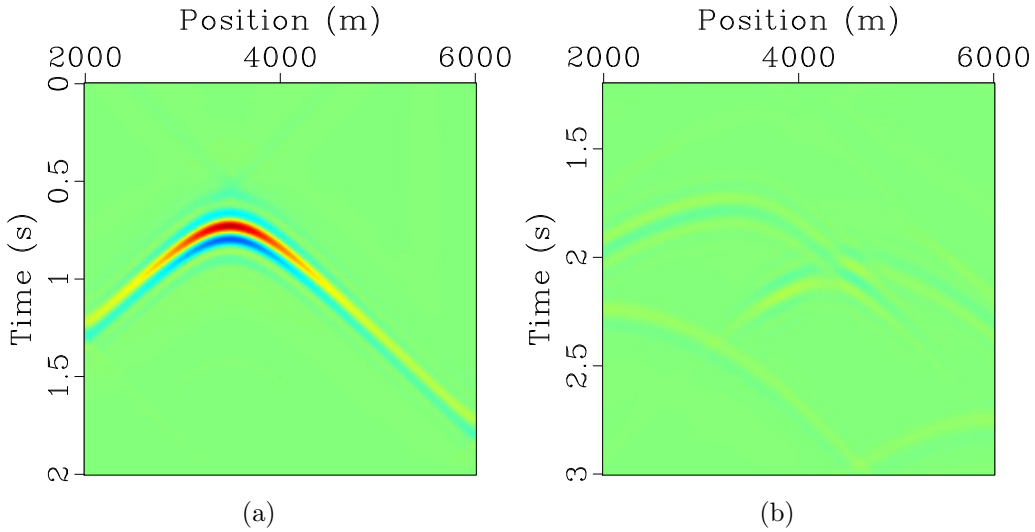


Figure 15: (a) Pressure source gather = image under symmetrized pressure-to-source operator W_m^{-1} applied to pressure gather shown in Figure 11a, computed using auxiliary receiver surface"near" at $z = 2900$ m and equation 50. Compare Figures 2b, 13a and 14a: these are all asymptotic approximations of each other. (b) Difference between pressure gather at the receiver surface Σ_r ($z = 1000$ m) modeled using symmetrized pressure-to-source gather in (a), and pressure gather shown in Figure 2c, propagated in lens model from source gather shown in Figure 2b (inferred from vertical velocity at source surface ($z = 3000$ m)). Same color scale as Figures 2a and 2c.

these to be replaced by an asymptotic approximation, so long as it is symmetric and computable (and at least semi-definite). Manipulations very similar to those used to obtain equation 50 lead to the approximation

$$W_m \approx -\frac{1}{16}\left(\Pi_0 \mathcal{S}_a^T (\Pi_1^T \Pi_0 + \Pi_0^T \Pi_1)\mathcal{S}_a \Pi_0^T\right) \tag{51}$$

Comparison with the definition 50 shows that W_m and 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 W_m^{-1} is inverse to W_m only in an approximate (asymptotic, aperture-limited) sense.

Preconditioned Conjugate Gradient Iteration

The most convenient arrangement the Conjugate Gradient (CG) algorithm taking advantage of the structure 48 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:

$$\text{minimize}_h \|Sh - d\|_d^2 + \alpha^2 \|Ah\|_m^2 \tag{52}$$

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 corresponding to d.

The minimizer of the objective defined in equation 52 solves the normal equation

$$(S^\dagger S + \alpha^2 A^\dagger A)h = S^\dagger d \tag{53}$$

where the weighted adjoint $S^\dagger$ has already been defined in equation 48, and $A^\dagger$ is the adjoint of A in the weighted model space norm defined by W_m, namely

$$A^\dagger = W_m^{-1} A^T W_m. \tag{54}$$

Rewrite the normal equation 53 as

$$W_m^{-1}(S^T W_d S + \alpha^2 A^T W_m A)h = W_m^{-1} S^T W_m d \tag{55}$$

Since W_m is self-adjoint and positive semidefinite, the common factor on both sides of 55 can be re-written as

$$Nh = (S^* S + \alpha^2 A^* A)h = S^* d \tag{56}$$

in which S^*, A^* are the adjoints with the original (Euclidean) inner product in the

domains but the weighted inner product in data space:

$$S^* = S^T W_d, \tag{57}$$
$$A^* = A^T W_m. \tag{58}$$

Note the S^*S and A^*A are symmetric in the Euclidean sense, so equation 56 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 56 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 $h_0 = 0$
$r_0 \leftarrow S^*d$
$p_0 \leftarrow M^{-1}r_0$
$g_0 \leftarrow p_0$
$q_0 \leftarrow Np_0$
$k \leftarrow 0$
repeat
  $\alpha_k \leftarrow \frac{\langle g_k, r_k\rangle}{\langle p_k, q_k\rangle}$
  $h_{k+1} \leftarrow h_k + \alpha_k p_k$
  $r_{k+1} \leftarrow r_k - \alpha_k q_k$
  $g_{k+1} \leftarrow M^{-1}r_{k+1}$
  $\beta_{k+1} \leftarrow \frac{\langle g_{k+1}, r_{k+1}\rangle}{\langle g_k, r_k\rangle}$
  $p_{k+1} \leftarrow g_{k+1} + \beta_{k+1}p_k$
  $q_{k+1} \leftarrow Np_{k+1}$
  $k \leftarrow k + 1$
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 $M^{-1/2}NM^{-1/2} \approx I$, which is in turn true if the eigenvalues of $M^{-1/2}NM^{-1/2}$ 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 if M is easy to invert.

Preconditioning with Penalty

Note that the normal operator appearing on the left-hand side of 53 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. We have already observed that the pressure-to-source operators Λ_s and Λ_r are pseudodifferential, and A is pseudodifferential as well. The rules for calculation of compositions and inverses for such operators show that the operators W_m^{-1} (properly interpreted) and W_m are

also pseudodifferential. Scalar pseudodifferential operators approximately commute, so $A^\dagger \approx A^T$. Therefore

$$S^\dagger S + \alpha^2 A^\dagger A \approx I + \alpha^2 A^T A \tag{59}$$

(restricted to the orthocomplement of the null space of S). 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.

From 59 and 55, 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

$$\begin{aligned} M &= (I + \alpha^2 A^T A)^{1/2} W_m (I + \alpha^2 A^T A)^{1/2}, \\ M^{-1} &= (I + \alpha^2 A^T A)^{-1/2} W_m^{-1} (I + \alpha^2 A^T A)^{-1/2}. \end{aligned} \tag{60}$$

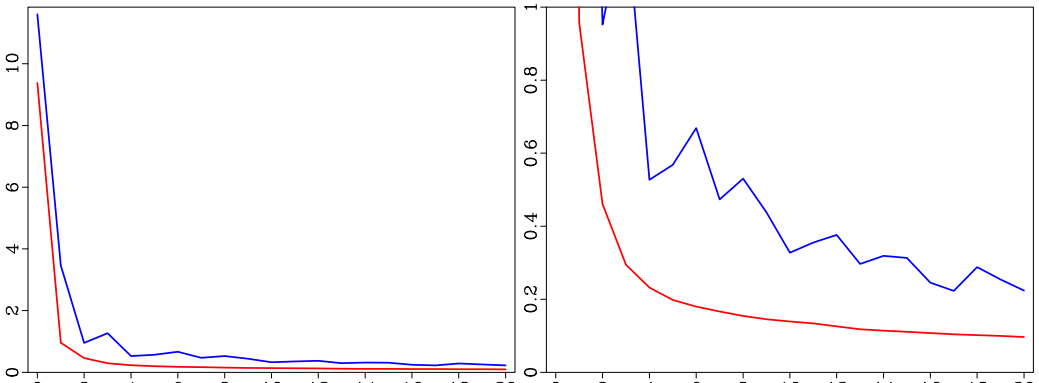
With this choice, 59 implies that $M^{-1}N \approx I$, also M is symmetric. As already mentioned, powers of $I + \alpha^2 A^T A$ are trivial to compute, given the choice of A made here.

Example

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 2a, simulated using the lens model (Figure 3) with source shown in Figure 11a, 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 16a shows the progress of the normal residual (Euclidean norm of the difference of the two sides of equation 55), for Conjugate Gradient and Preconditioned Conjugate Gradient (Algorithm 1) iterations, applied to solution of the optimization problem 52 with $\alpha = 0$. Because the normal residual is measured in the Euclidean norm for both CG and PCG, the results are comparable. For PCG, the preconditioner M is given by equation 60, with the symmetrized Λs computed as indicated in the preceding section with auxiliary source and receiver surfaces shifted by 100 m. Convergence for the preconditioned algorithm is roughly 4 times as fast: this speed comparison is much more obvious in a vertically stretched version of the figure (Figure 16b).

Figure 17a shows the same comparison with non-zero penalty weight, $\alpha = 10^{-3}$.



(a) (b)

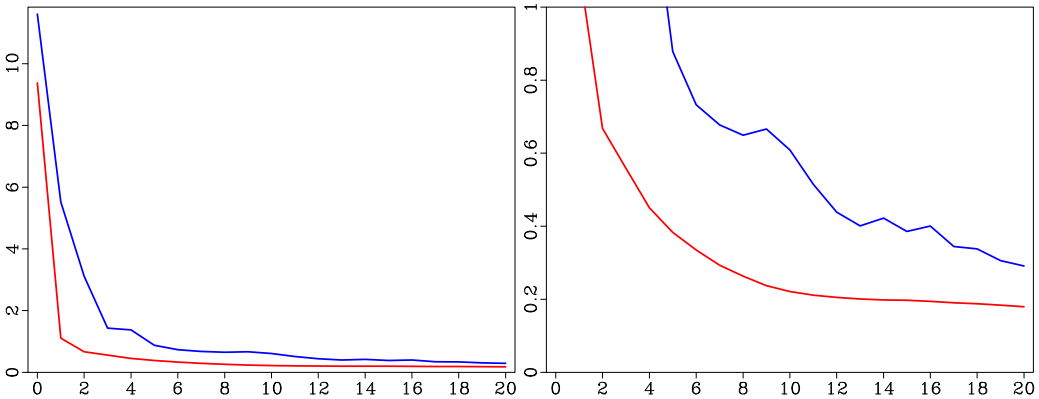
Figure 16: (a) Comparison of normal residual (gradient) Euclidean norms: CG (blue), PCG (red), plotted vs. iteration. Data = lens model, point source (Figure 2a), inversion in homogenous model. Penalty weight $\alpha = 0$. (b) Same comparison, with expanded vertical scale, to show more clearly that PCG ir roughly 4-5 times as fast as (in achieving a given normal residual) as CG.

Convergence for both CG and PCG is somewhat slower than in the $\alpha = 0$ case. However the relation is about the same, still 4-5 times as many iterations to achieve a given normal residual level with CG as with PCG.

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. 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



(a) (b)

Figure 17: (a) Comparison of normal residual (gradient) Euclidean norms: CG (blue), PCG (red), plotted vs. iteration. Data = lens model, point source (Figure 2a), inversion in homogenous model. Penalty weight $\alpha = 10^{-3}$. (b) Same comparison, with expanded vertical scale, to show more clearly that PCG ir roughly 4-5 times as fast as (in achieving a given normal residual) as CG.

objective function 4, resolving a difficulty first noted also for reflected wave inversion (Kern and Symes, 1994).

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 incoming/outgoing waves (on which these results rest) 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.

ACKNOWLEDGEMENTS

The author thanks the anonymous reviewers for their very helpful comments and suggestions.

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 Madagascar reproducible research framework (`http:www.reproducibility.org`). Code and data source is available from the author on request.

REFERENCES

Aghmiry, H., A. Gholami, and S. Operto, 2020, Accurate and efficient data-assimilated wavefield reconstruction in the time domain: Geophysics, **85**, no. 2, A7–A12.

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.

Chauris, H., and E. Cocher, 2017, From migration to inversion velocity analysis: Geophysics, **82**, no. 3, S207–S223.

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.

Fichtner, A., 2010, Full seismic waveform modelling and inversion: Springer Verlag.

Gauthier, O., A. Tarantola, and J. Virieux, 1986, Two-dimensional nonlinear inversion of seismic waveforms: Geophysics, **51**, 1387–1403.

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

Guasch, L., A. Calderón, M. Tang, P. Nachev, and M. Warner, 2020, Full-waveform inversion imaging of the human brain: NPJ Digital Medicine, **3**.

Guasch, L., M. Warner, and C. Ravaut, 2019, Adaptive waveform inversion: Practice: Geophysics, **84**, R447–R461.

Hou, J., and W. Symes, 2015, An approximate inverse to the extended Born modeling operator: Geophysics, **80**, no. 6, R331–R349.

———, 2016, Accelerating extended least-squares migration with weighted conjugate gradient iteration: Geophysics, **81**, no. 4, S165–S179.

Hristova, Y., 2009, Time reversal in thermoacoustic tomography - an error estimate: Inverse Problems, **25**, 055008.

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.

Li, Z., Y. Lin, and K. Zhang, 2018, Time-domain wavefield reconstruction inversion: Applied Geophysics, **14**, 523—-528.

Louboutin, M., G. Rizzuti, and F. Herrmann, 2020, Time-domain wavefield reconstruction inversion in a tti medium. (arXiv:2004.07355v1).

Nirenberg, L., 1972, Lectures on partial differential equations: Lecture Note 17, CBMS, AMS, Providence.

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

Pladys, A., R. Brossier, Y. Li, and L. Métivier, 2021, On cycle-skipping and misfit function modification for full-wave inversion: Comparison of five recent approaches: Geophysics, **86**, R563–R587.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503.

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

Saint Raymond, X., 1991, Elementary introduction to the theory of pseudodifferential operators: CRC Press. Studies in Advanced Mathematics.

Schuster, G., 2017, Seismic inversion: Society of Exploration Geophysicists. Investigations in Geophysics.

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., 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.

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

Taylor, M., 1981, Pseudodifferential Operators: Princeton University Press.

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.

van Leeuwen, T., and F. Herrmann, 2013, Mitigating local minima in full-waveform inversion by expanding the search space: Geophysical Journal International, **195**, 661–667.

———, 2016, A penalty method for pde-constrained optimization in inverse problems: Inverse Problems, **32**, 1–26.

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

Warner, M., and L. Guasch, 2016, Adaptive waveform inversion: theory: Geophysics, **81**, R429–R445.

Warner, M., T. Nangoo, A. Umpleby, N. Shah, D. Kahn, and M. Isernia, 2021, Adaptive reflection waveform inversion: Faster, tighter, deeper, smarter: 91st Annual International Meeting, Expanded Abstracts, Society of Exploration Geophysicists, 582–586.

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.

Yong, P., R. Brossier, L. Métivier, Y. Li, W. He, and J. Virieux, 2021, Improving adaptive waveform inversion by local matching filter: 82nd EAGE Annual Conference and Exhibition, Expanded Abstracts, European Association of Geoscientists and Engineers, 1–5.

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.