Instability of adjoint pseudo-acoustic anisotropic wave equations

Huy Le, Stewart A. Levin, and Biondo Biondi

ABSTRACT

We discover that the adjoint system of the second-order pseudo-acoustic anisotropic wave equations that we have implemented in previous work has unstable solutions and propose two possible alternative systems. The unstable solutions have magnitude that grows linearly with time and does not propagate spatially. This is not the kind of numerical instability that occurs when the propagation timestep does not satisfy the CFL condition. This instability is inherent to this particular type of wave equations. In fact, there is a family of unstable wave equations that all describe the same kinematics. The key to stability for this type of pseudo-acoustic anisotropic wave equations is not whether the system is self-adjoint but whether it reduces to the scalar wave equation in isotropic media.

INTRODUCTION

The pseudo-acoustic anisotropic wave equations were introduced by Alkhalifah (2000) to avoid having to process shear waves. This system of equations is derived by setting shear wave velocities along symmetry axes to zeros in the dispersion relation. The most computationally efficient form is the system of coupled second-order wave equations, which has been commonly used in reverse time migration (Fletcher et al., 2009; Duveneck and Bakker, 2011; Zhang et al., 2011). There are several properties to know about this kind of pseudo-acoustic anisotropic wave equations.

1. Even though shear wave velocity is set to zero, the solution space still encompasses residual shear waves. These residual shear waves manifest as dimond-shaped artifacts and can be easily reduced when the source injection location is in isotropic media, for example in marine environments (Alkhalifah, 2000). Complete suppression of this artifact requires projection onto a solution subspace (Le and Levin, 2014; Maharramov et al., 2015; Xu and Zhou, 2014).

2. There is actually a family of infinite number of equivalent systems of wave equations that all come from the same dispersion relation (Fowler et al., 2010). These systems similarly describe P-wave kinematics but behave differently in terms of amplitude.

3. All members of this family of pseudo-acoustic wave equations suffer from inherent weak instability (Bube et al., 2012). This instability is a result of setting shear velocity to zero combined with second-order time derivatives.

Bube et al. (2016) show that the instability of this type of second-order pseudo-acoustic wave equations can occur in practice when numerical computation is performed with low presicion. They also develop a stable and self-adjoint system by reintroducing finite shear wave velocity and rewriting second-order derivatives a cascaded first-orders. We recently find that this instability problem also occurs in the adjoint systems even with high precision numerics. The unstable solutions grow linearly with time and are stationary (i.e. does not propagate spatially), as analyzed by Bube et al. (2012). In this report, we present two alternative systems that do not suffer from this instability. Compared to those proposed by Bube et al. (2016), our solutions are computationally more efficient and do not require medium parameters on different staggered grids.

NON-SELF-ADJOINT SYSTEM

The coupled second-order systems of pseudo-acoustic anisotropic wave equations are commonly used in reverse time migration and waveform inversion due to its computational efficiency and capability to acurately describe P-wave kinematics. One particular form of such system is

$$\begin{cases} \partial_t^2 \sigma_x = c_{11}(\partial_x^2 + \partial_y^2)\sigma_x + c_{13}\partial_z^2 \sigma_z, \\ \partial_t^2 \sigma_z = c_{13}(\partial_x^2 + \partial_y^2)\sigma_x + c_{33}\partial_z^2 \sigma_z, \end{cases} \tag{1}$$

which can be written in matrix form as

$$\partial_t^2 \sigma = CD\sigma, \tag{2}$$

where σ is the stresses, C is the density-normalized stiffness matrix, and D is the derivative matrix

$$\sigma = \begin{bmatrix} \sigma_x \\ \sigma_z \end{bmatrix}, C = \begin{bmatrix} c_{11} & c_{13} \\ c_{13} & c_{33} \end{bmatrix}, D = \begin{bmatrix} \partial_x^2 + \partial_y^2 & \\ & \partial_z^2 \end{bmatrix}. \tag{3}$$

The elements of stiffness matrix C are related to vertical velocity, v, horizontal velocity, v_x, NMO velocity, v_n, and Thomsen parameters, ϵ and δ by

$$c_{11} = v^2(1 + 2\epsilon) = v_x^2, \tag{4}$$
$$c_{13} = v^2\sqrt{1 + 2\delta} = vv_n, \tag{5}$$
$$c_{33} = v^2. \tag{6}$$

Spatial derivatives on the right hand side of system 1 have form of $c\partial_i^2$, which is not self-adjoint (Appendix A). The adjoint system is

$$\partial_t^2 \lambda = DC\lambda, \tag{7}$$

or

$$\begin{cases} \partial_t^2 \lambda_x = (\partial_x^2 + \partial_y^2)(c_{11}\lambda_x + c_{13}\lambda_z), \\ \partial_t^2 \lambda_z = \partial_z^2(c_{13}\lambda_x + c_{33}\lambda_z), \end{cases} \tag{8}$$

where $\lambda = \begin{bmatrix} \lambda_x \\ \lambda_z \end{bmatrix}$ is the adjoint wave fields.

The gradients of the objective function χ with respect to c_{ij} are

$$\frac{\partial \chi}{\partial c_{11}} = \int_0^T \lambda_x(\partial_x^2 + \partial_y^2)\sigma_x dt, \tag{9}$$

$$\frac{\partial \chi}{\partial c_{13}} = \int_0^T \lambda_x \partial_z^2 \sigma_z + \lambda_z(\partial_x^2 + \partial_y^2)\sigma_x dt, \tag{10}$$

$$\frac{\partial \chi}{\partial c_{33}} = \int_0^T \lambda_z \partial_z^2 \sigma_z dt, \tag{11}$$

from which gradients with respect to (v, ϵ, δ) can be computed from chain rule

$$\frac{\partial \chi}{\partial v} = \frac{\partial \chi}{\partial c_{11}} 2v(1+2\epsilon) + \frac{\partial \chi}{\partial c_{13}} 2v\sqrt{1+2\delta} + \frac{\partial \chi}{\partial c_{33}} 2v, \tag{12}$$

$$\frac{\partial \chi}{\partial \epsilon} = \frac{\partial \chi}{\partial c_{11}} 2v^2, \tag{13}$$

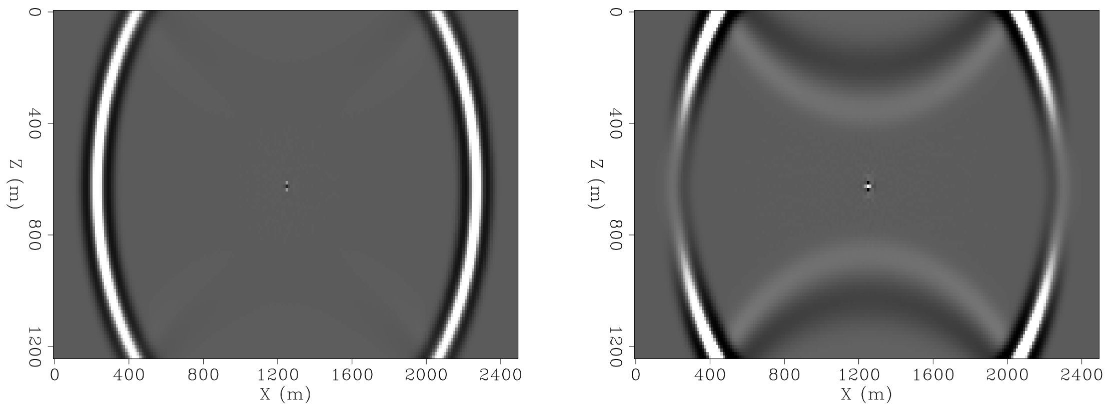
$$\frac{\partial \chi}{\partial \delta} = \frac{\partial \chi}{\partial c_{13}} \frac{v^2}{\sqrt{1+2\delta}}. \tag{14}$$

Unfortunately, the above adjoint equations have solutions that grow linearly with time. The unstable solutions seem to have zero wave speed and does not propagate spatially. This is the kind of weak instability that is inherent to this type of pseudo-acoustic wave equations (Bube et al., 2012). Figure 1 shows the XZ palnes of two adjoint wavefields in a homogeneous isotropic medium. Notice the stationary instability points at injection location. Figure 2 shows the time evolutions of these adjoint wavefields at every second. Interestingly, the two adjoint wavefields have opposite linearly growing solutions. From equations 8, it is easy to see that, for isotropic media ($c_{11} = c_{13} = c_{33}$), if (λ_x, λ_z) is a solution, so is $(\lambda_x + t, \lambda_z - t)$. In fact, in such medium, the adjoint equations do not reduce to the well-known scalar acoustic isotropic wave equation. All simulations are performed in three dimensions.

A SELF-ADJOINT SYSTEM

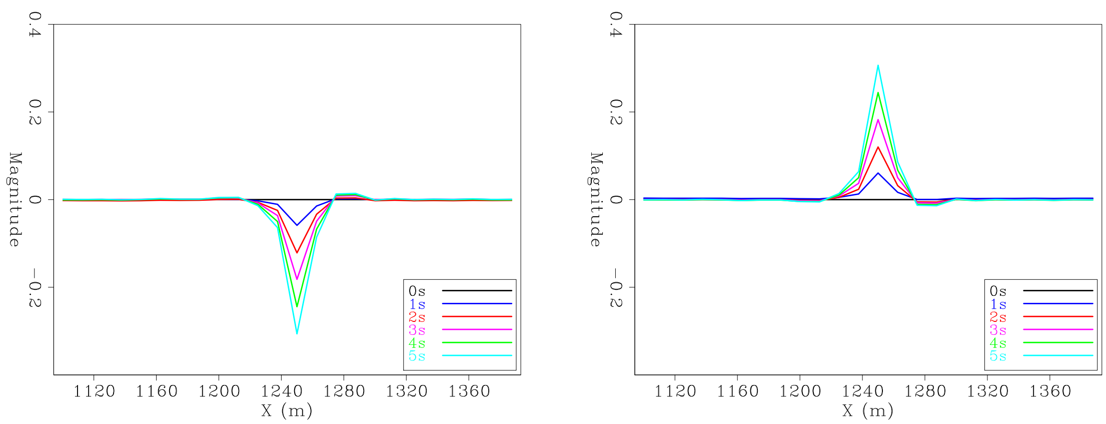
Bube et al. (2016) modify the forward equations making it self-adjoint in order to achieve stability. One possible way to make the system self-adjoint is to define

$$R = \sqrt{C} = \begin{bmatrix} r_{11} & r_{13} \\ r_{13} & r_{33} \end{bmatrix}, \tag{15}$$



(a) (b)

Figure 1: XZ planes of adjoint wavefields λ_x (a) and λ_z (b) at 0.8 seconds show stationary instability at the source locations. [**ER**]



(a) (b)

Figure 2: Time evolutions of adjoint wavefields λ_x (a) and λ_z (b) at every second show the linearly growing instability. [**ER**]

and rewrite the system as

$$\partial_t^2 \sigma = RDR\sigma, \tag{16}$$

or

$$\begin{cases} \partial_t^2 \sigma_x = r_{11}(\partial_x^2 + \partial_y^2)(r_{11}\sigma_x + r_{13}\sigma_z) + r_{13}\partial_z^2(r_{13}\sigma_x + r_{33}\sigma_z), \\ \partial_t^2 \sigma_z = r_{13}(\partial_x^2 + \partial_y^2)(r_{11}\sigma_x + r_{13}\sigma_z) + r_{33}\partial_z^2(r_{13}\sigma_x + r_{33}\sigma_z). \end{cases} \tag{17}$$

Because C is always symmetric and positive semi-definite when $\epsilon \geq \delta$, R exists and is also symmetric and positive semi-definite.

Equations 16 correctly captures the kinematics of the original equations 2. It can be shown that the two systems have the same dispersion relation

$$\det(C\hat{D} + \omega^2 I) = \det(R\hat{D}R + \omega^2 I), \tag{18}$$

where

$$\hat{D} = -\begin{bmatrix} k_x^2 + k_y^2 & \\ & k_z^2 \end{bmatrix}. \tag{19}$$

Moreover, if C is nonsingular, system 2 is equivalent to system 17 by a change of variable $\sigma \rightarrow R\sigma$. The adjoint system 8 is also equivalent to this newly derived system by a similar change of variable $\lambda \rightarrow R^{-1}\sigma$. Despite this equivalence, system 17 is stable. This is because both equations in this system reduce to the scalar acoustic wave equation in isotropic media. In fact, in isotropic media, the equivalence among these systems break because C is singular.

The gradients of the objective function with respect to r_{ij} are

$$\frac{\partial \chi}{\partial r_{11}} = \int_0^T \left[\lambda_x(\partial_x^2 + \partial_y^2)(r_{11}\sigma_x + r_{13}\sigma_z) + \sigma_x(\partial_x^2 + \partial_y^2)(r_{11}\lambda_x + r_{13}\lambda_z)\right] dt, \tag{20}$$

$$\begin{aligned} \frac{\partial \chi}{\partial r_{13}} = \int_0^T & \left[\lambda_x \partial_z^2(r_{13}\sigma_x + r_{33}\sigma_z) + \lambda_z(\partial_x^2 + \partial_y^2)(r_{11}\sigma_x + r_{13}\sigma_z) + \right. \\ & \left. \sigma_x\partial_z^2(r_{13}\lambda_x + r_{33}\lambda_z) + \sigma_z(\partial_x^2 + \partial_y^2)(r_{11}\lambda_x + r_{13}\lambda_z)\right] dt, \end{aligned} \tag{21}$$

$$\frac{\partial \chi}{\partial r_{33}} = \int_0^T \left[\lambda_z \partial_z^2(r_{13}\sigma_x + r_{33}\sigma_z) + \sigma_z\partial_z^2(r_{13}\lambda_x + r_{33}\lambda_z)\right] dt. \tag{22}$$

Elements of R are computed from trace and determinant of C

$$r_{11} = \frac{v}{t}(1 + 2\epsilon + s), \tag{23}$$

$$r_{13} = \frac{v}{t}\sqrt{1 + 2\delta}, \tag{24}$$

$$r_{33} = \frac{v}{t}(1 + s), \tag{25}$$

where

$$s = \sqrt{2(\epsilon - \delta)}, t = \sqrt{2(\epsilon + 1) + 2s}. \tag{26}$$

Since $s = \sqrt{2(\epsilon - \delta)}$ can be zero, derivatives of s with respect to either ϵ or δ can become unbounded. As a result, one has to define new variables

$$\alpha = 1 + s, \tag{27}$$

$$\beta = \sqrt{1 + 2\delta}. \tag{28}$$

so that

$$r_{11} = \frac{v}{t}(\beta^2 + \alpha^2 - \alpha), \tag{29}$$

$$r_{13} = \frac{v\beta}{t}, \tag{30}$$

$$r_{33} = \frac{v\alpha}{t}, \tag{31}$$

with

$$t = \sqrt{\alpha^2 + \beta^2}. \tag{32}$$

Now the gradients with respect to (v, α, β) can be computed from change rule

$$\frac{\partial \chi}{\partial v} = \frac{\partial \chi}{\partial r_{11}} \frac{\beta^2 + \alpha^2 - \alpha}{t} + \frac{\partial \chi}{\partial r_{13}} \frac{\beta}{t} + \frac{\partial \chi}{\partial r_{33}} \frac{\alpha}{t}, \tag{33}$$

$$\frac{\partial \chi}{\partial \alpha} = \frac{\partial \chi}{\partial r_{11}} \frac{v(\alpha^3 + \alpha\beta^2 - \beta^3)}{t^3} - \frac{\partial \chi}{\partial r_{13}} \frac{v\alpha\beta}{t^3} + \frac{\partial \chi}{\partial r_{33}} \frac{v\beta^2}{t^3}, \tag{34}$$

$$\frac{\partial \chi}{\partial \beta} = \frac{\partial \chi}{\partial r_{11}} \frac{v\beta(\alpha^2 + \alpha + \beta^2)}{t^3} + \frac{\partial \chi}{\partial r_{13}} \frac{v\alpha^2}{t^3} - \frac{\partial \chi}{\partial r_{33}} \frac{v\alpha\beta}{t^3}. \tag{35}$$

After the inversion, (ϵ, δ) can be obtained by

$$\delta = \frac{\beta^2 - 1}{2}, \tag{36}$$

$$\epsilon = \frac{(\alpha - 1)^2}{2} + \delta. \tag{37}$$

Using system 17 for waveform inversion is computationally inefficient due to expensive forward solutions and cumbersome gradient expressions involving a change of variables. Next section seeks an alternative solution.

ANOTHER SELF-ADJOINT SYSTEM

Another way to make system 2 self-adjoint is to split the differential operator D instead of the medium matrix C. However, taking the square root of the horizontal derivative, $\partial_x^2 + \partial_y^2$, which results in a pseudo-differential operator, requires either Fourier transform or spectral factorization and helix coordinates (Claerbout and Fomel, 2014). Instead, we expand system 1 to three equations

$$\begin{cases} \partial_t^2 \sigma_x = c_{11} \partial_x^2 \sigma_x + c_{11} \partial_y^2 \sigma_y + c_{13} \partial_z^2 \sigma_z, \\ \partial_t^2 \sigma_y = c_{11} \partial_x^2 \sigma_x + c_{11} \partial_y^2 \sigma_y + c_{13} \partial_z^2 \sigma_z, \\ \partial_t^2 \sigma_z = c_{13} \partial_x^2 \sigma_x + c_{13} \partial_y^2 \sigma_y + c_{33} \partial_z^2 \sigma_z, \end{cases} \tag{38}$$

which has the same matrix form as system 2, but with

$$\sigma = \begin{bmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \end{bmatrix}, C = \begin{bmatrix} c_{11} & c_{11} & c_{13} \\ c_{11} & c_{11} & c_{13} \\ c_{13} & c_{13} & c_{33} \end{bmatrix}, D = \begin{bmatrix} \partial_x^2 & & \\ & \partial_y^2 & \\ & & \partial_z^2 \end{bmatrix}. \tag{39}$$

The adjoint equations of system 38 are

$$\begin{cases} \partial_t^2 \lambda_x = \partial_x^2 (c_{11}\lambda_x + c_{11}\lambda_y + c_{13}\lambda_z), \\ \partial_t^2 \lambda_y = \partial_y^2 (c_{11}\lambda_x + c_{11}\lambda_y + c_{13}\lambda_z), \\ \partial_t^2 \lambda_z = \partial_z^2 (c_{13}\lambda_x + c_{13}\lambda_y + c_{33}\lambda_z). \end{cases} \tag{40}$$

This adjoint system suffers the same instability as system 8.

Now to make it self-adjoint, define

$$D_1 = \sqrt{D} = \begin{bmatrix} \partial_x & & \\ & \partial_y & \\ & & \partial_z \end{bmatrix}, \tag{41}$$

and rewrite the system 38 as

$$\partial_t^2 \sigma = D_1 C D_1 \sigma, \tag{42}$$

or

$$\begin{cases} \partial_t^2 \sigma_x = \partial_x (c_{11}\partial_x\sigma_x + c_{11}\partial_y\sigma_y + c_{13}\partial_z\sigma_z), \\ \partial_t^2 \sigma_y = \partial_y (c_{11}\partial_x\sigma_x + c_{11}\partial_y\sigma_y + c_{13}\partial_z\sigma_z), \\ \partial_t^2 \sigma_z = \partial_z (c_{13}\partial_x\sigma_x + c_{13}\partial_y\sigma_y + c_{33}\partial_z\sigma_z). \end{cases} \tag{43}$$

It can be verified that systems 2 and 42 have the same dispersion relation, i.e. they describe the same kinematics

$$\det(C\hat{D} + \omega^2 I) = \det(\hat{D}_1 C \hat{D}_1 + \omega^2 I), \tag{44}$$

where

$$\hat{D} = -\begin{bmatrix} k_x^2 & & \\ & k_y^2 & \\ & & k_z^2 \end{bmatrix}, \hat{D}_1 = \begin{bmatrix} ik_x & & \\ & ik_y & \\ & & ik_z \end{bmatrix}. \tag{45}$$

Moreover, if one applies D_1 on both sides of system 38 and make a change of variables $\sigma' = D_1\sigma$, they would arrive at system 43. This means that if $(\sigma_x, \sigma_x, \sigma_z)$ is a smooth solution of system 38, $(\partial_x\sigma_x, \partial_y\sigma_x, \partial_z\sigma_z)$ is a solution of system 43.

The gradients of the objective function with respect to c_{ij} now are

$$\frac{\partial \chi}{\partial c_{11}} = \int_0^T (\partial_x\lambda_x + \partial_y\lambda_y)(\partial_x\sigma_x + \partial_y\sigma_y) dt, \tag{46}$$

$$\frac{\partial \chi}{\partial c_{13}} = \int_0^T (\partial_x\lambda_x + \partial_y\lambda_y)\partial_z\sigma_z + (\partial_x\sigma_x + \partial_y\sigma_y)\partial_z\lambda_z dt, \tag{47}$$

$$\frac{\partial \chi}{\partial c_{33}} = \int_0^T \partial_z\lambda_z\partial_z\sigma_z dt, \tag{48}$$

from which gradients with respect to (v, ϵ, δ) can be computed from chain rule as before (equations 12, 13, and 14).

Solving system 43 requires staggered grid finite difference methods. Figure 3 shows grid locations of wavefield variables and medium parameters. Advantages of this system include efficient forward solution, simple expressions for gradients, and co-location of medium parameters. Numerial experiments, however, show that even though system 43 is self-adjoint, it is unstable. Indeed, if one integrates system 40, which is unstable, one can prove that $(\sigma_x, \sigma_y, \sigma_z) = (\int_x \lambda_x, \int_y \lambda_y, \int_z \lambda_z)$ is also a solution of system 38.

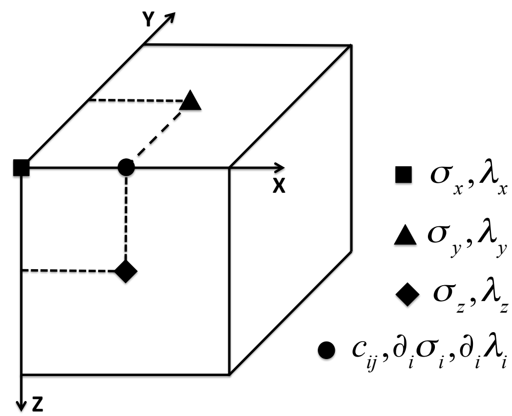


Figure 3: Staggered grids for system 38. [**NR**]

NON-SELF-ADJOINT BUT STABLE SYSTEM

Fowler et al. (2010) show that system 1 is a member of a family of infinitely many kinematically equivalent wave equations. This section attempts to find a stable system in this family. The general form of this family of equations is

$$\begin{cases} \partial_t^2 \sigma_x = \left[a_1(\partial_x^2 + \partial_y^2) + a_2 \partial_z^2\right] \sigma_x + \left[b_1(\partial_x^2 + \partial_y^2) + b_2 \partial_z^2\right] \sigma_z, \\ \partial_t^2 \sigma_z = \left[c_1(\partial_x^2 + \partial_y^2) + c_2 \partial_z^2\right] \sigma_x + \left[d_1(\partial_x^2 + \partial_y^2) + d_2 \partial_z^2\right] \sigma_z, \end{cases} \tag{49}$$

or

$$\partial_t^2 \sigma = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \sigma, \tag{50}$$

where $A = a_1(\partial_x^2 + \partial_y^2) + a_2 \partial_z^2$ and so on. The adjoint equations are

$$\partial_t^2 \lambda = \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix} \lambda, \tag{51}$$

or

$$\begin{cases} \partial_t^2 \lambda_x = (\partial_x^2 + \partial_y^2)(a_1 \lambda_x + c_1 \lambda_z) + \partial_z^2 (a_2 \lambda_x + c_2 \lambda_z), \\ \partial_t^2 \lambda_z = (\partial_x^2 + \partial_y^2)(b_1 \lambda_x + d_1 \lambda_z) + \partial_z^2 (b_2 \lambda_x + d_2 \lambda_z). \end{cases} \tag{52}$$

Derivatives with respect to these coefficients (a_i, b_i, c_i, d_i) are

$$\frac{\partial \chi}{\partial a_1} = \int_0^T \lambda_x(\partial_x^2 + \partial_y^2)\sigma_x dt, \tag{53}$$

$$\frac{\partial \chi}{\partial a_2} = \int_0^T \lambda_x \partial_z^2 \sigma_x dt, \tag{54}$$

$$\frac{\partial \chi}{\partial b_1} = \int_0^T \lambda_x(\partial_x^2 + \partial_y^2)\sigma_z dt, \tag{55}$$

$$\frac{\partial \chi}{\partial b_2} = \int_0^T \lambda_x \partial_z^2 \sigma_z dt, \tag{56}$$

$$\frac{\partial \chi}{\partial c_1} = \int_0^T \lambda_z(\partial_x^2 + \partial_y^2)\sigma_x dt, \tag{57}$$

$$\frac{\partial \chi}{\partial c_2} = \int_0^T \lambda_z \partial_z^2 \sigma_x dt, \tag{58}$$

$$\frac{\partial \chi}{\partial d_1} = \int_0^T \lambda_z(\partial_x^2 + \partial_y^2)\sigma_z dt, \tag{59}$$

$$\frac{\partial \chi}{\partial d_2} = \int_0^T \lambda_z \partial_z^2 \sigma_z dt. \tag{60}$$

For the this system to correctly describe the kinematics, its dispersion relation has to be equal to that of system 2, which results in five constraints on eight medium parameters (a_i, b_i, c_i, d_i)

$$a_1 + d_1 = c_{11}, \tag{61}$$

$$a_2 + d_2 = c_{33}, \tag{62}$$

$$a_1 d_1 - b_1 c_1 = 0, \tag{63}$$

$$a_2 d_2 - b_2 c_2 = 0, \tag{64}$$

$$a_1 d_2 + a_2 d_1 - b_1 c_2 - b_2 c_1 = c_{11} c_{33} - c_{13}^2. \tag{65}$$

These constraints form an under-determined system. There are infinitely many equivalent solutions. Appendix B systematically shows that members of this family of equations with have zero parameter coeffficients are all unstable. There are still multiple solutions with all non-zero coefficients to consider. To simplify algebra and latter computation, one choice is to have symmetric operators

$$b_1 = c_1 = \frac{1}{2} r v_x^2, \tag{66}$$

$$b_2 = c_2 = \frac{1}{2} r v^2. \tag{67}$$

Other coefficients can be expressed in terms of (r, v_x, v) as

$$(a_1, d_1) = \frac{1}{2} v_x^2 \left(1 \pm \sqrt{1 - r^2}\right), \tag{68}$$

$$(a_2, d_2) = \frac{1}{2} v^2 \left(1 \mp \sqrt{1 - r^2}\right). \tag{69}$$

Substitute the above expressions into the last constraint 65 and solve for

$$r = \frac{v_n}{v_x}. \tag{70}$$

To compute the gradients, define new variables

$$\alpha = \sqrt{2(\epsilon - \delta)}, \tag{71}$$
$$\beta = \sqrt{1 + 2\epsilon}, \tag{72}$$
$$\gamma = \sqrt{\beta^2 - \alpha^2}, \tag{73}$$

so that

$$1 + 2\delta = \gamma^2, v_x = v\beta, v_n = v\gamma, r = \frac{\gamma}{\beta}, \sqrt{1 - r^2} = \frac{\alpha}{\beta}, \tag{74}$$

and

$$(a_1, d_1) = \frac{1}{2}v^2\beta\left(\beta \pm \alpha\right), \tag{75}$$
$$(a_2, d_2) = \frac{1}{2}v^2\left(1 \mp \frac{\alpha}{\beta}\right), \tag{76}$$
$$b_1 = c_1 = \frac{1}{2}v^2\beta\gamma, \tag{77}$$
$$b_2 = c_2 = \frac{1}{2}v^2\frac{\gamma}{\beta}. \tag{78}$$

Now derivatives with respect to (v, α, β) can be computed from chain rule

$$\begin{aligned}\frac{\partial\chi}{\partial v} &= v\left\{\frac{\partial\chi}{\partial a_1}\beta(\beta+\alpha) + \frac{\partial\chi}{\partial d_1}\beta(\beta-\alpha) + \frac{\partial\chi}{\partial a_2}\left(1 - \frac{\alpha}{\beta}\right) + \frac{\partial\chi}{\partial d_2}\left(1 + \frac{\alpha}{\beta}\right)\right. \\ &\quad \left. +\gamma\left[\left(\frac{\partial\chi}{\partial b_1} + \frac{\partial\chi}{\partial c_1}\right)\beta + \left(\frac{\partial\chi}{\partial b_2} + \frac{\partial\chi}{\partial c_2}\right)\frac{1}{\beta}\right]\right\},\end{aligned} \tag{79}$$

$$\begin{aligned}\frac{\partial\chi}{\partial\alpha} &= \frac{1}{2}v^2\left\{\left(\frac{\partial\chi}{\partial a_1} - \frac{\partial\chi}{\partial d_1}\right)\beta - \left(\frac{\partial\chi}{\partial a_2} - \frac{\partial\chi}{\partial d_2}\right)\frac{1}{\beta}\right. \\ &\quad \left. -\frac{\alpha}{\gamma}\left[\left(\frac{\partial\chi}{\partial b_1} + \frac{\partial\chi}{\partial c_1}\right)\beta + \left(\frac{\partial\chi}{\partial b_2} + \frac{\partial\chi}{\partial c_2}\right)\frac{1}{\beta}\right]\right\},\end{aligned} \tag{80}$$

$$\begin{aligned}\frac{\partial\chi}{\partial\beta} &= \frac{1}{2}v^2\left\{\left(\frac{\partial\chi}{\partial a_1} - \frac{\partial\chi}{\partial d_1}\right)(2\beta+\alpha) + \left(\frac{\partial\chi}{\partial a_2} - \frac{\partial\chi}{\partial d_2}\right)\frac{\alpha}{\beta^2}\right. \\ &\quad \left. -\frac{1}{\gamma}\left[\left(\frac{\partial\chi}{\partial b_1} + \frac{\partial\chi}{\partial c_1}\right)(2\beta^2 - \alpha^2) + \left(\frac{\partial\chi}{\partial b_2} + \frac{\partial\chi}{\partial c_2}\right)\frac{\alpha^2}{\beta^2}\right]\right\},\end{aligned} \tag{81}$$

After the inversion, (ϵ, δ) can be retrieved from (α, β) by

$$\epsilon = \frac{\beta^2 - 1}{2}, \tag{82}$$
$$\delta = \epsilon - \frac{\alpha^2}{2}. \tag{83}$$

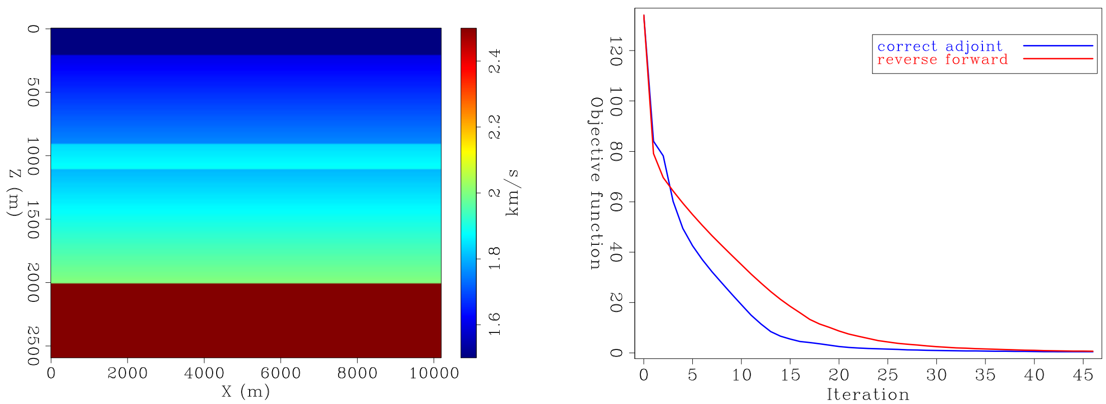
As with our stable self-adjoint system 17, using system 49 involves a change of variables and cumbersome expressions for the gradients, but the forward equations are computationally more efficient.

CORRECT ADJOINT AND TIME-REVERSED FORWARD

Both of the above proposed solutions (equations 17 and 49) are significantly more expensive than the original adjoint equations 8. Instead of using the correct adjoint equations, one might time-reverse the forward equations to compute the adjoint wavefields. In this section, we study how incorrect adjoints affect the convergence of FWI. In the two following 2D synthetic examples, we have uniformly distributed sources and receivers on the water surface and use a Ricker wavelet of 5 Hz central frequency to model our systhetic data. We use a LBFGS solver and invert simutaneously for velocity, v, and Thomsen parameters, ϵ and δ. We do not concern with parameterization and trade-off in these experiments.

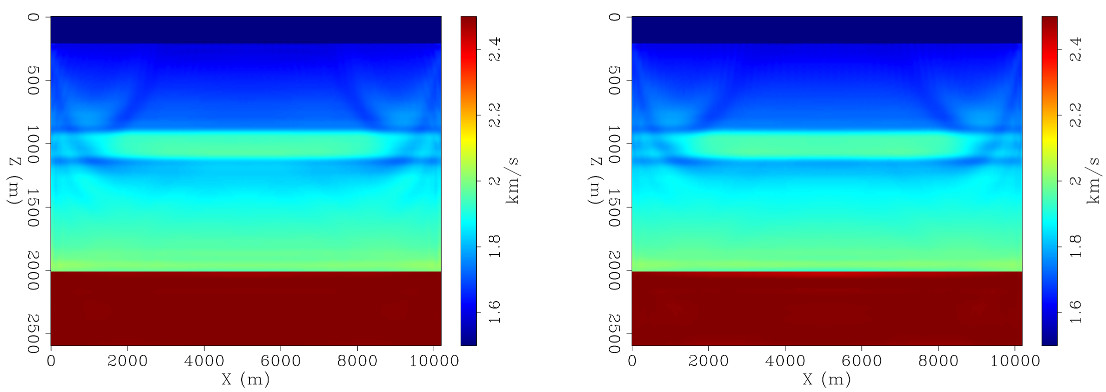
Figure 4a shows a simple velocity model (ϵ and δ models are similar). For this model, we try to recover the high velocity layer at 1000 meter depth. Figure 4b plots the objective functions of two inversions with correct adjoint and time-reversed forward. Figures 5a and 5b respectively show the inverted velocity models after 50 iterations using the correct adjoint equations and time-reversed forward equations. The instability of the correct adjoint equations does not affect our inversion because it is stationary at injection locations (receivers, for adjoint equations). These figures show that upon convergence, the incorrect adjoint performs just as well as the correct one. Figures 6a and 6b show the inverted models after only 12 iterations, at which the two objective functions are most different. As one expects, the correct adjoint gives a slightly better result. Although, the two inversions carry out the same number of objective function and gradient evaluations per iteration, a more truly cost comparison is between correct adjoint iteration 12^{th} (Figure 6a) and time-reversed forward iteration 18^{th} (Figure 6c) because solving the correct adjoint equations is about 1.5 times more expensive than reversing the forward equations. In this metric, using the time-reversed forward for adjoint wavefields actually leads to a better result.

In another experiment to understand the effect of incorrect adjoint equations, we use BP 2007 synthetic anisotropic models. Figures 7 show the true and initial velocity models. Figures 8a and 8b show the inverted models ater 100 iterations with the correct adjoint and time-reversed forward. We again observe that the two methods perform equally well if converged. This can also be seen in Figure 8c plotting the objective functions. For such complicated model as this one, the difference between correct and incorrect adjoints at early iterations are less significant. Figures 9a and 9b show the inverted models after 10 iterations and Figure 9c shows the inverted model with incorrect adjoint after 15 iterations. These three results are virtually indistinguishable.



(a) (b)

Figure 4: Simple velocity model (a) and objective functions from two inversions (b) using the correct adjoint and time-reversed forward. The initial model does not include the high velocity layer at 1000 m depth. [**CR**]



(a) (b)

Figure 5: Inverted velocity models after 50 iterations using the correct adjoint (a) and time-reversed forward (b). [**CR**]

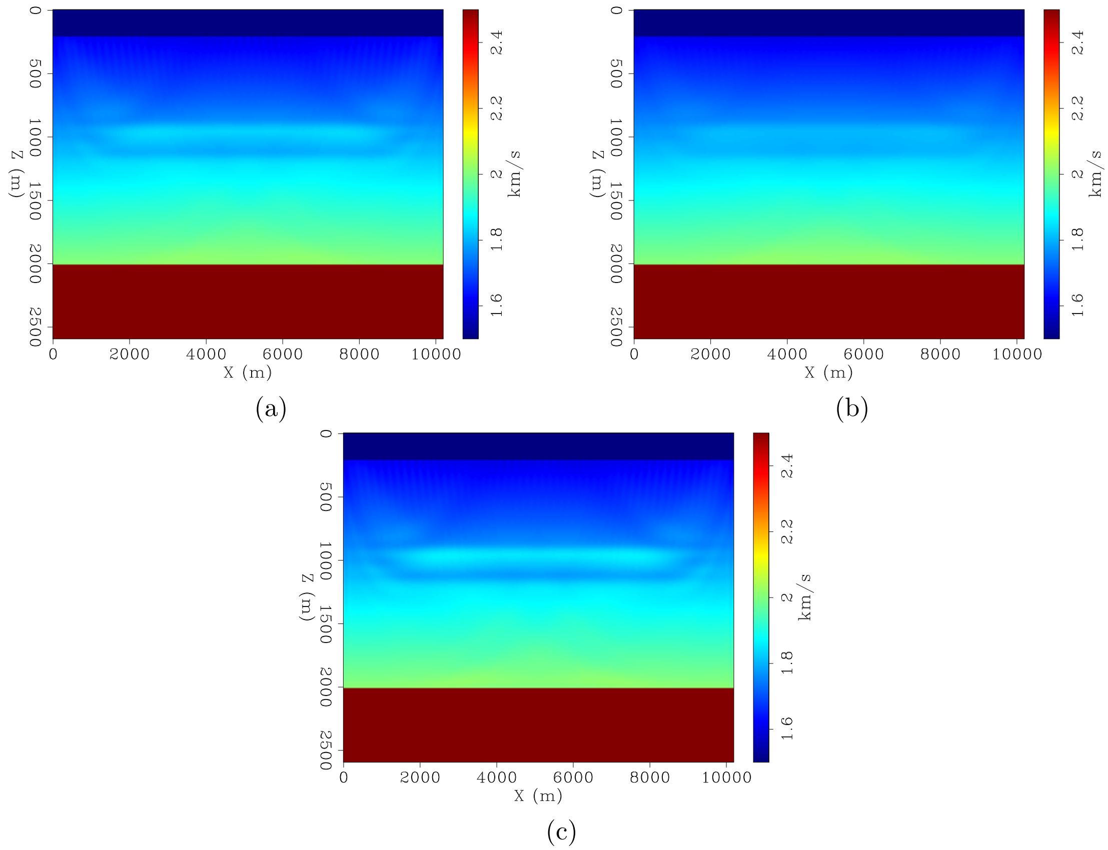


Figure 6: Inverted velocity models after 12 iterations using the correct adjoint (a) and time-reversed forward (b). (c) shows the inverted velocity model after 18 iterations using the time-reversed forward. True cost comparison is between (a) and (c). [**CR**]

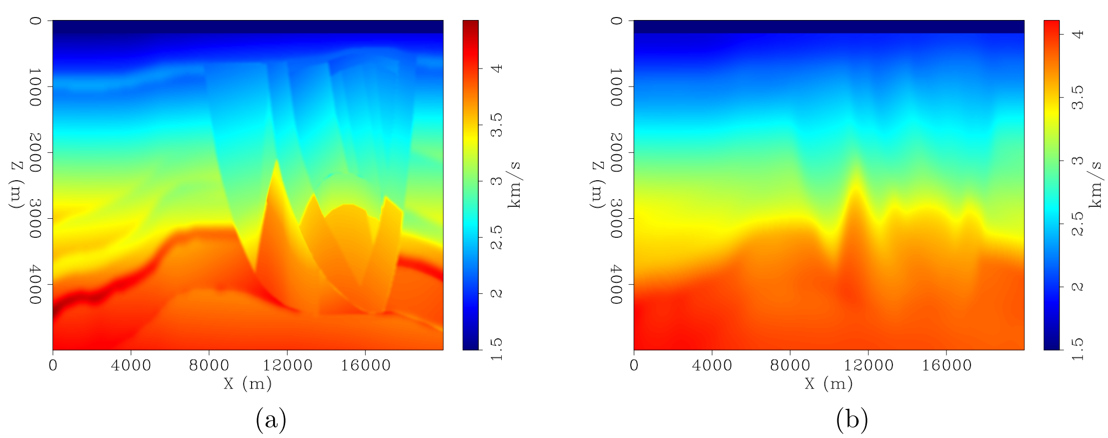


Figure 7: True (a) and initial (b) velocity models from BP2007 synthetic. [**CR**]

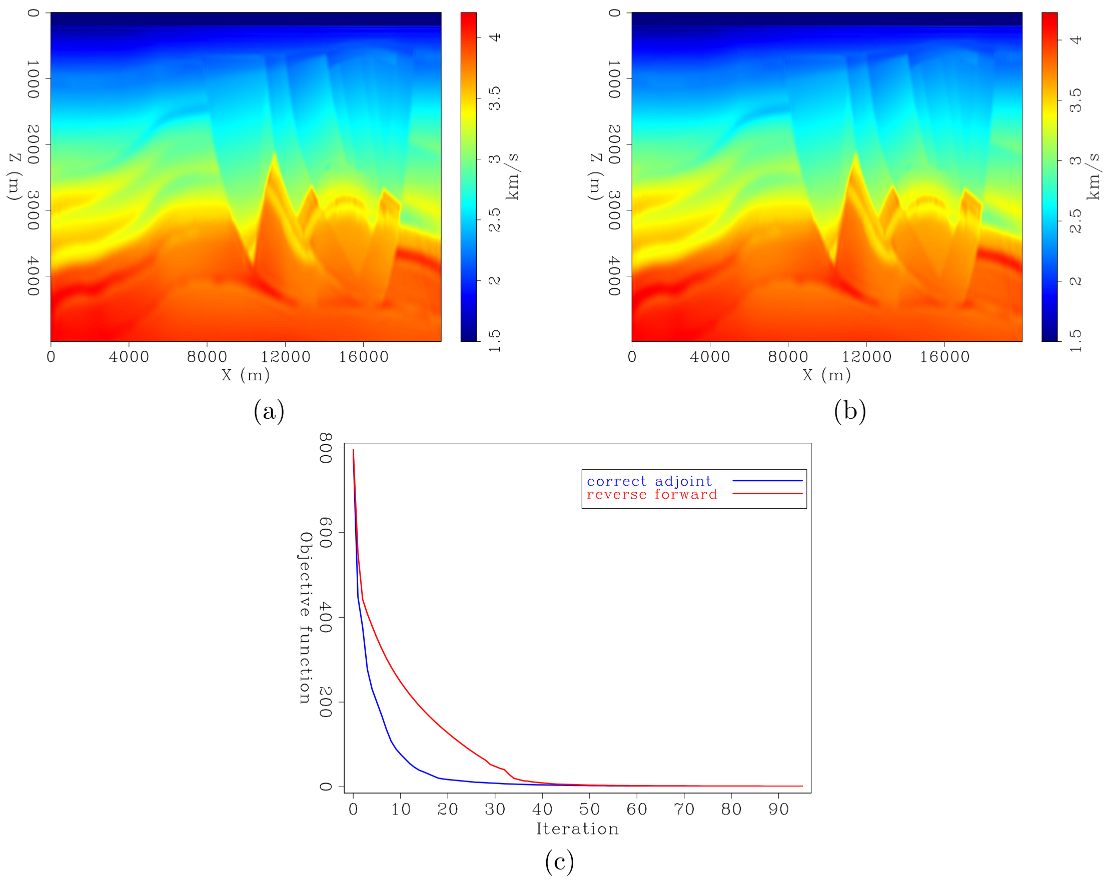


Figure 8: Inverted velocity models after 100 iterations using the correct adjoint (a) and time-reversed forward (b). Objective functions of the two inversions (c). [**CR**]

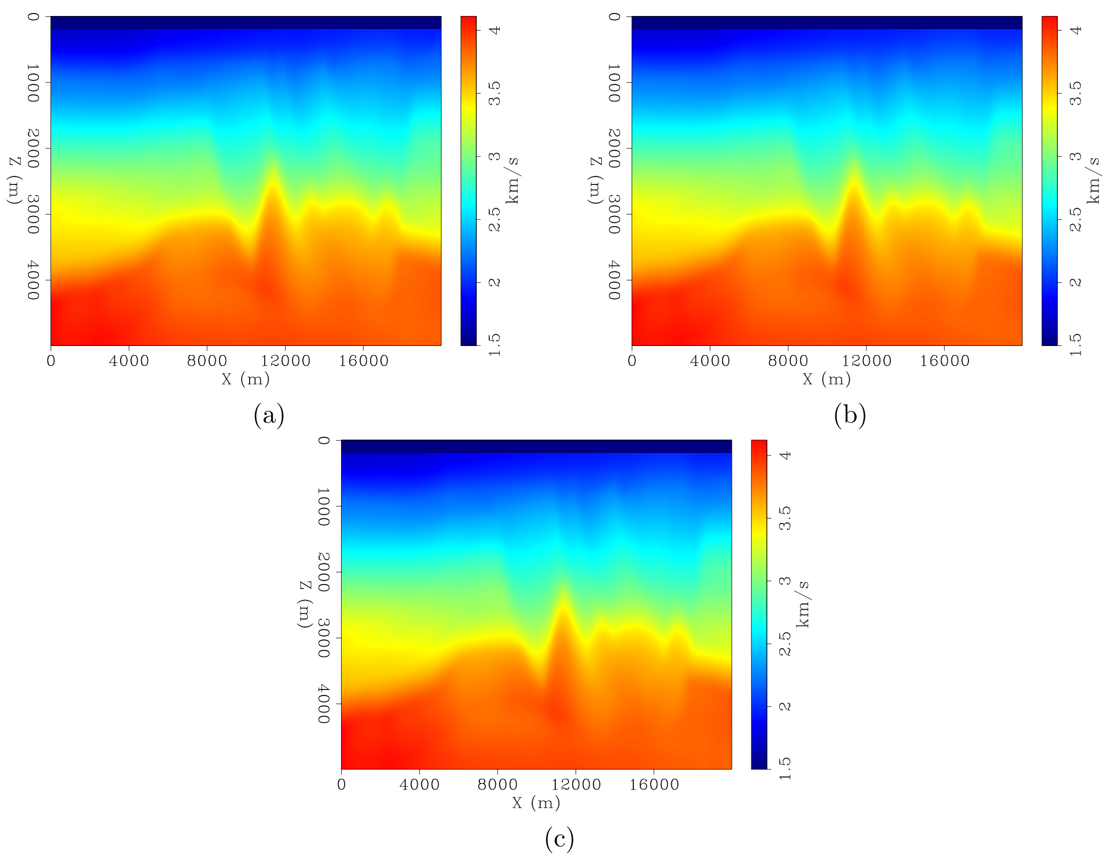


Figure 9: Inverted velocity models after 10 iterations using the correct adjoint (a) and time-reversed forward (b). (c) shows the inverted velocity model after 15 iterations using the time-reversed forward. True cost comparison is between (a) and (c). [**CR**]

CONCLUSIONS

The second-order pseudo-acoustic anisotropic wave equations can have unstable adjoint solutions with linearly growing magnitude and zero wave speed. This kind of weak instability is an inherent property of this particular type of wave equations. We propose two alternative systems of equations that are kinematically equivalent and free of this instability. The first system is self-adjoint and obtained by splitting the medium parameter matrix while the second system is a generalization of the original system with all non-zero coefficients. To use in waveform inversion, both of these systems involve a change of variables and are computationally more expensive than the original system. Instead, one can use the time-reversed forward wave equations to compute the gradients. Our numerical experiments indicate that in terms of convergence, the difference between the correct and incorrect adjoints is minimal.

APPENDIX A: ADJOINTS OF COMMON DERIVATIVES

Define functional inner product as

$$\langle u, v\rangle = \int_0^T \int_\Omega uv d\mathbf{x}dt. \tag{A-1}$$

Adjoint operator is defined by

$$\langle u, Lv\rangle = \langle L^* u, v\rangle. \tag{A-2}$$

For a particular form of $L = \partial_x^2$

$$\langle u, Lv\rangle = \langle u, \partial_x^2 v\rangle = \int_0^T \int_\Omega u\partial_x^2 v d\mathbf{x}dt \tag{A-3}$$

$$= \int_0^T \int_{\Omega_{yz}} u\partial_x v dydzdt|_{\Omega_x} - \int_0^T \int_{\Omega_{yz}} v\partial_x u dydzdt|_{\Omega_x} + \int_0^T \int_\Omega v\partial_x^2 u d\mathbf{x}dt \tag{A-4}$$

$$= \int_0^T \int_\Omega v\partial_x^2 u d\mathbf{x}dt = \langle \partial_x^2 u, v\rangle. \tag{A-5}$$

where integration by parts has been carried out twice in the x-direction and boundary conditions $u|_{\Omega_x} = v|_{\Omega_x} = 0$. So by definition, $L = \partial_x^2$ is self-adjoint $L^* = \partial_x^2$.

Similarly, for $L = c\partial_x^2$, where c can be velocity or any medium parameter,

$$\langle u, Lv\rangle = \langle u, c\partial_x^2 v\rangle = \langle cu, \partial_x^2 v\rangle = \langle \partial_x^2 cu, v\rangle. \tag{A-6}$$

As a result, this operator is not self-adjoint $L^* = \partial_x^2 c$.

One can easily find the adjoints of the following common differential operators

L	L^*	Self-adjoint
∂_i^2	∂_i^2	yes
$c\partial_i^2$	$\partial_i^2 c$	no
$a\partial_i^2 b$	$b\partial_i^2 a$	no
∂_i	$-\partial_i$	no
$a\partial_i b$	$-b\partial_i a$	no
$\partial_i c\partial_j$	$\partial_j c\partial_i$	no

APPENDIX B: MEMBER SYSTEMS WITH ZERO COEFFICIENTS

Because of constraint equations 63, $a_1 d_1 = b_1 c_1$, and 64, $a_2 d_2 = b_2 c_2$, if one of the coefficients, a_i, b_i, c_i, d_i, is zero, at least one other must also vanish.

Case 1: $a_1 = b_1 = 0$

System 49 becomes

$$\begin{cases} \partial_t^2 \sigma_x = a_2 \partial_z^2 \sigma_x + b_2 \partial_z^2 \sigma_z, \\ \partial_t^2 \sigma_z = \left[c_1(\partial_x^2 + \partial_y^2) + c_2 \partial_z^2\right] \sigma_x + \left[d_1(\partial_x^2 + \partial_y^2) + d_2 \partial_z^2\right] \sigma_z. \end{cases} \tag{B-1}$$

Case 1.1: $a_2 b_2 \neq 0$ and $c_2 d_2 \neq 0$

Because of equation 64, define $r = \frac{a_2}{c_2} = \frac{b_2}{d_2}$. Because $r \neq 0$, multiply the second equation in B-1 by r and subtract the first equation, resulting in an equivalent system

$$\begin{cases} \partial_t^2 \sigma_x = a_2 \partial_z^2 \sigma_x + b_2 \partial_z^2 \sigma_z, \\ \partial_t^2 (r\sigma_z - \sigma_x) = rc_1(\partial_x^2 + \partial_y^2)\sigma_x + rd_1(\partial_x^2 + \partial_y^2)\sigma_z, \end{cases} \tag{B-2}$$

which, after a change of variable $\sigma_z' = r\sigma_z - \sigma_x$ or $\sigma_z = \frac{\sigma_z' + \sigma_x}{r}$, becomes

$$\begin{cases} \partial_t^2 \sigma_x = (a_2 + \frac{b_2}{r})\partial_z^2 \sigma_x + \frac{b_2}{r}\partial_z^2 \sigma_z', \\ \partial_t^2 \sigma_z' = (rc_1 + d_1)(\partial_x^2 + \partial_y^2)\sigma_x + d_1(\partial_x^2 + \partial_y^2)\sigma_z'. \end{cases} \tag{B-3}$$

This system is a special case of the general system 49 with $a_1 = b_1 = c_2 = d_2 = 0$.

Case 1.2: $a_2 b_2 \neq 0$ and $c_2 d_2 = 0$

Constraint 64 dictates that $c_2 = d_2 = 0$, which leads us back to the above case of $a_1 = b_1 = c_2 = d_2 = 0$.

Case 1.3: $a_2 b_2 = 0$ *and* $c_2 d_2 \neq 0$

In this case, $a_2 = b_2 = 0$, resulting in no solution because constraint 65 is violated.

Case 1.4: $a_2 b_2 = c_2 d_2 = 0$

This results in four cases

- $a_2 = c_2 = 0$: this case has a solution.
- $a_2 = d_2 = 0$: violation of constraint 62.
- $b_2 = c_2 = 0$: the five constraints reduce to

$$d_1 = c_{11}, \tag{B-4}$$
$$a_2 + d_2 = c_{33}, \tag{B-5}$$
$$a_2 d_2 = 0, \tag{B-6}$$
$$a_2 d_1 = c_{11} c_{33} - c_{13}^2. \tag{B-7}$$

  The last constraint requires that $a_2 \neq 0$, so $d_2 = 0$, which, means $a_2 = c_{33}$. This does not satisfy the last constraints.

- $b_2 = d_2 = 0$: the five constraints reduce to

$$d_1 = c_{11}, \tag{B-8}$$
$$a_2 = c_{33}, \tag{B-9}$$
$$a_2 d_1 = c_{11} c_{33} - c_{13}^2, \tag{B-10}$$

  in which the last constraint is not satisfied.

In conclusion, for the case $a_1 = b_1 = 0$, there are two solutions $a_1 = b_1 = c_2 = d_2 = 0$ and $a_1 = b_1 = a_2 = c_2 = 0$. The first solution results in

$$a_2 = c_{33}, \tag{B-11}$$
$$d_1 = c_{11}, \tag{B-12}$$
$$b_2 c_1 = c_{13}^2. \tag{B-13}$$

If one chooses $b_2 = c_1 = c_{13}$ for example, one gets the system

$$\begin{cases} \partial_t^2 \sigma_x = c_{33} \partial_z^2 \sigma_x + c_{13} \partial_z^2 \sigma_z, \\ \partial_t^2 \sigma_z = c_{13} (\partial_x^2 + \partial_y^2) \sigma_x + c_{11} (\partial_x^2 + \partial_y^2) \sigma_z. \end{cases} \tag{B-14}$$

The second solution results in

$$d_1 = c_{11}, \tag{B-15}$$
$$d_2 = c_{33}, \tag{B-16}$$
$$b_2 c_1 = c_{13}^2 - c_{11} c_{33} = v^4 (\delta - \epsilon). \tag{B-17}$$

If one chooses $b_2 = v^2(\delta - \epsilon)$ and $c_1 = v^2$ for example, one gets the system

$$\begin{cases} \partial_t^2 \sigma_x = v^2(\delta-\epsilon)\partial_z^2\sigma_z, \\ \partial_t^2\sigma_z = v^2(\partial_x^2+\partial_y^2)\sigma_x + \left[v^2(1+2\epsilon)(\partial_x^2+\partial_y^2) + v^2\partial_z^2\right]\sigma_z. \end{cases} \tag{B-18}$$

Exchanging b_2 and c_1 results in

$$\begin{cases} \partial_t^2 \sigma_x = v^2\partial_z^2\sigma_z, \\ \partial_t^2\sigma_z = v^2(\delta-\epsilon)(\partial_x^2+\partial_y^2)\sigma_x + \left[v^2(1+2\epsilon)(\partial_x^2+\partial_y^2) + v^2\partial_z^2\right]\sigma_z. \end{cases} \tag{B-19}$$

Case 2: $a_1 = c_1 = 0$

This case is similar to the case $b_1 = d_1 = 0$, considered below, if one interchanges $\sigma_x \leftrightarrow \sigma_z$.

Case 3: $b_1 = d_1 = 0$

System 49 becomes

$$\begin{cases} \partial_t^2 \sigma_x = \left[a_1(\partial_x^2+\partial_y^2) + a_2\partial_z^2\right]\sigma_x + b_2\partial_z^2\sigma_z, \\ \partial_t^2\sigma_z = \left[c_1(\partial_x^2+\partial_y^2) + c_2\partial_z^2\right]\sigma_x + d_2\partial_z^2\sigma_z, \end{cases} \tag{B-20}$$

Similar analysis as in the first case applies.

Case 3.1: $a_2b_2 \neq 0$ and $c_2d_2 \neq 0$

Define $r = \frac{a_2}{c_2} = \frac{b_2}{d_2}$, multiply the second equation in B-20 by r, and subtract from the first equation

$$\begin{cases} \partial_t^2 \sigma_x = \left[a_1(\partial_x^2+\partial_y^2) + a_2\partial_z^2\right]\sigma_x + b_2\partial_z^2\sigma_z, \\ \partial_t^2(\sigma_x - r\sigma_z) = (a_1 - rc_1)(\partial_x^2+\partial_y^2)\sigma_x, \end{cases} \tag{B-21}$$

which, after a change of variable $\sigma_z' = \sigma_x - r\sigma_z$ or $\sigma_z = \frac{\sigma_x - \sigma_z'}{r}$, becomes

$$\begin{cases} \partial_t^2 \sigma_x = \left[a_1(\partial_x^2+\partial_y^2) + (a_2 + \frac{b_2}{r})\partial_z^2\right]\sigma_x - \frac{b_2}{r}\partial_z^2\sigma_z', \\ \partial_t^2\sigma_z' = (a_1 - rc_1)(\partial_x^2+\partial_y^2)\sigma_x. \end{cases} \tag{B-22}$$

This system is a special case of the general system 49 with $b_1 = d_1 = c_2 = d_2 = 0$.

Case 3.2: $a_2b_2 \neq 0$ and $c_2d_2 = 0$

In this case, $c_2 = d_2 = 0$, which leads us back to the above case of $b_1 = d_1 = c_2 = d_2 = 0$.

Case 3.3: $a_2 b_2 = 0$ *and* $c_2 d_2 \neq 0$

In this case, $a_2 = b_2 = 0$, the five constraints reduce to

$$a_1 = c_{11}, \tag{B-23}$$
$$d_2 = c_{33}, \tag{B-24}$$
$$a_1 d_2 = c_{11} c_{33} - c_{13}^2. \tag{B-25}$$

There is no solution because the last constraint is not satisfied.

Case 3.4: $a_2 b_2 = c_2 d_2 = 0$

Consider four cases

- $a_2 = c_2 = 0$: this case has a solution.
- $a_2 = d_2 = 0$: violation of constraint 62.
- $b_2 = c_2 = 0$: the five constraints reduce to

$$a_1 = c_{11}, \tag{B-26}$$
$$a_2 + d_2 = c_{33}, \tag{B-27}$$
$$a_2 d_2 = 0, \tag{B-28}$$
$$a_1 d_2 = c_{11} c_{33} - c_{13}^2. \tag{B-29}$$

  The last constraint requires that $d_2 \neq 0$, so $a_2 = 0$, which, means $d_2 = c_{33}$. This does not satisfy the last constraints.

- $b_2 = d_2 = 0$: violation of constraint 65.

In conclusion, for the case $b_1 = d_1 = 0$, there are two solutions $b_1 = d_1 = c_2 = d_2 = 0$ and $b_1 = d_1 = a_2 = c_2 = 0$. The first solution results in

$$a_1 = c_{11}, \tag{B-30}$$
$$a_2 = c_{33}, \tag{B-31}$$
$$b_2 c_1 = c_{13}^2 - c_{11} c_{33} = v^4(\delta - \epsilon). \tag{B-32}$$

If one chooses $b_2 = v^2$ and $c_1 = v^2(\delta - \epsilon)$ for example, one gets the system

$$\begin{cases} \partial_t^2 \sigma_x = \left[v^2(1+2\epsilon)(\partial_x^2 + \partial_y^2) + v^2 \partial_z^2 \right] \sigma_x + v^2 \partial_z^2 \sigma_z, \\ \partial_t^2 \sigma_z = v^2(\delta - \epsilon)(\partial_x^2 + \partial_y^2)\sigma_x, \end{cases} \tag{B-33}$$

Exchanging b_2 and c_1 gives

$$\begin{cases} \partial_t^2 \sigma_x = \left[v^2(1+2\epsilon)(\partial_x^2 + \partial_y^2) + v^2 \partial_z^2 \right] \sigma_x + v^2(\delta - \epsilon) \partial_z^2 \sigma_z, \\ \partial_t^2 \sigma_z = v^2(\partial_x^2 + \partial_y^2)\sigma_x, \end{cases} \tag{B-34}$$

The second solution results in

$$a_1 = c_{11}, \tag{B-35}$$
$$d_2 = c_{33}, \tag{B-36}$$
$$b_2 c_1 = c_{13}^2. \tag{B-37}$$

If one chooses $b_2 = c_1 = c_{13}$ for example, one gets back the system 1

$$\begin{cases} \partial_t^2 \sigma_x = c_{11}(\partial_x^2 + \partial_y^2)\sigma_x + c_{13}\partial_z^2 \sigma_z, \\ \partial_t^2 \sigma_z = c_{13}(\partial_x^2 + \partial_y^2)\sigma_x + c_{33}\partial_z^2 \sigma_z. \end{cases} \tag{B-38}$$

Case 4: $c_1 = d_1 = 0$

This case is similar to the case $a_1 = b_1 = 0$, already considered above, if one interchanges $\sigma_x \leftrightarrow \sigma_z$.

The above four cases consider what happens when two of four coefficients (a_1, b_1, c_1, d_1) is zero. Analysis of four other cases when two of (a_2, b_2, c_2, d_2) becomes zero is similar by just interchanging $(\partial_x^2 + \partial_y^2) \leftrightarrow \partial_z^2$. Unfortunately, none of the above analyzed solutions produces stable forward and adjoint systems. The only stable solution is one with all non-zero coefficients.

REFERENCES

Alkhalifah, T., 2000, An acoustic wave equation for anisotropic media: Geophysics, **65**, 1239–1250.

Bube, K. P., T. Nemeth, J. P. Stefani, R. Ergas, W. Liu, K. T. Nihei, and L. Zhang, 2012, On the instability in second-order systems for acoustic VTI and TTI media: Geophysics, **77**, No. 5, T171–T186.

Bube, K. P., J. Washbourne, R. Ergas, and T. Nemeth, 2016, Self-adjoint, energy-conserving second-order pseudo-acousti system for VTI and TTI media for reverse-time mingration and full-waveform inversion: SEG Annual International Meeting, Expanded Abstracts, 1110–1115, Society of Exploration Geophysicists.

Claerbout, J. and S. Fomel, 2014, Geophysical Image Estimation by Example: Stanford Exploration Project.

Duveneck, E. and P. M. Bakker, 2011, Stable P-wave modeling for reverse-time migration in tilted TI media: Geophysics, **76**, No. 2, WA3–WA11.

Fletcher, R. P., X. Du, and P. J. Fowler, 2009, Reverse time migration in tilted transversely isotropic (TTI) media: Geophysics, **74**, No. 6, WCA179–WCA187.

Fowler, P. J., X. Du, and R. P. Fletcher, 2010, Coupled equations for reverse time migration in transversely isotropic media: Geophysics, **75**, No. 1, S11–S22.

Le, H. and S. A. Levin, 2014, Removing shear artifacts in acoustic wave propagation in orthorhombic media: SEG Annual International Meeting, Expanded Abstracts, 486–490, Society of Exploration Geophysicists.

Maharramov, M., B. Biondi, and M. Meadows, 2015, Resolving the effects of production-induced overburden dilation using simultaneous tv-regularized time-lapse FWI: SEP-Report, **158**, 1–10.

Xu, S. and H. Zhou, 2014, Accurate simulations of pure quasi-P-waves in complex anisotropic media: Geophysics, **79**, no. 6, T341–T348.

Zhang, Y., H. Zhang, and G. Zhang, 2011, A stable TTI reverse time migration and its implementation: Geophysics, **76**, No. 3, WA3–WA11.