Karlsruhe Institute of Technology

Wave phenomena
analysis and numerics

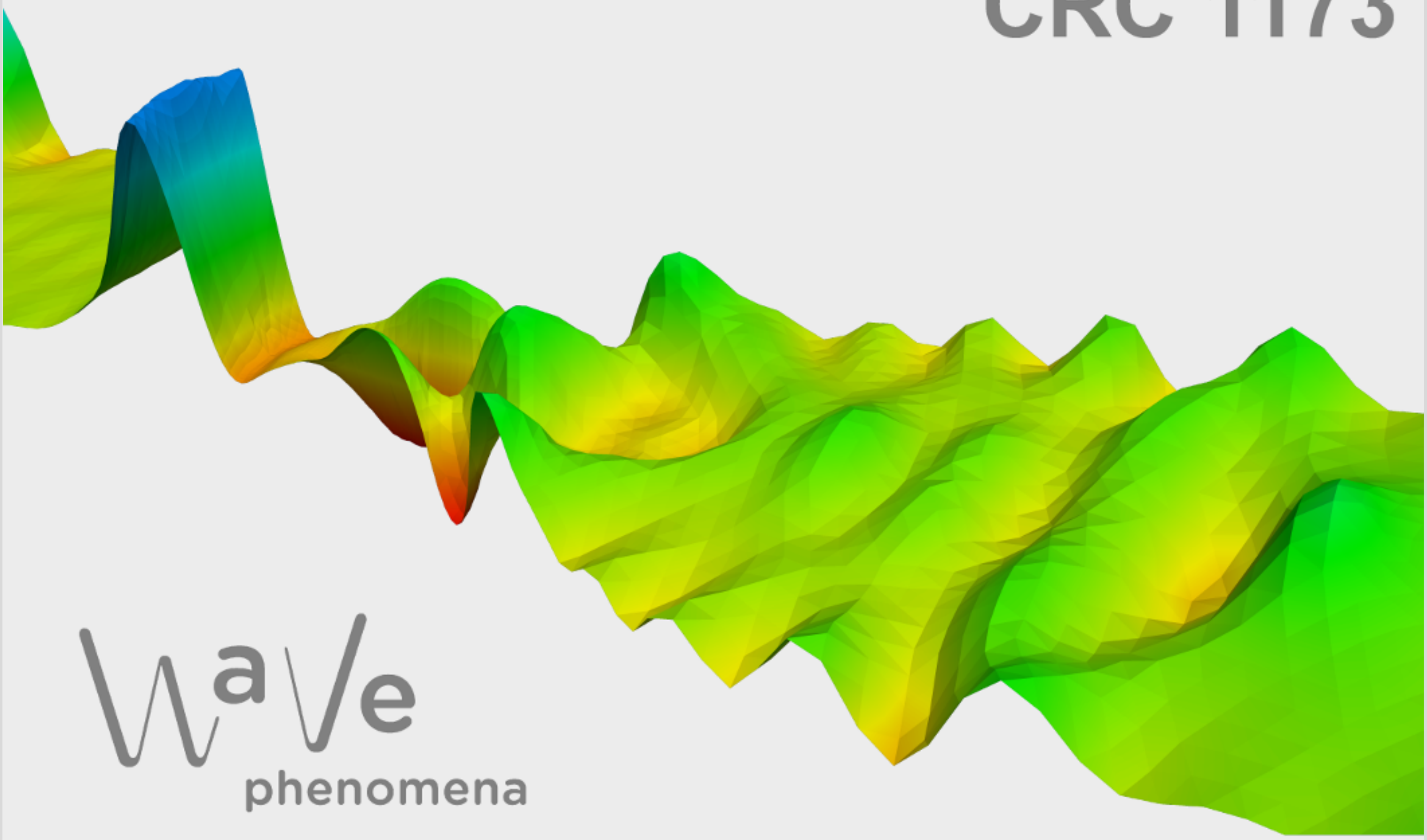
On the mathematical foundation of full waveform inversion in viscoelastic vertically transverse isotropic media

Andreas Rieder

CRC Preprint 2024/26 (revised), April 2025

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Wave phenomena

Participating universities

Funded by

ISSN 2365-662X

ON THE MATHEMATICAL FOUNDATION OF FULL WAVEFORM INVERSION IN VISCOELASTIC VERTICALLY TRANSVERSE ISOTROPIC MEDIA

ANDREAS RIEDER

Abstract. We present a mathematical framework for viscoelastic full waveform inversion (FWI) in vertically transverse isotropic media. FWI can be formulated as the nonlinear inverse problem of identifying parameters in the underlying attenuating anisotropic wave equation given partial wave field measurements (seismograms). From a mathematical point of view, one has to solve an operator equation for the full waveform forward operator, which is the corresponding parameter-to-state map. We give a rigorous definition of this operator, show its Fréchet differentiability, and explicitly characterize the adjoint operator of its Fréchet derivative. Thus, we provide the main ingredients to implement Newton-type/gradient-based regularization schemes for FWI. Our approach can be directly applied to other concepts of anisotropy.

1. Introduction

In seismic imaging we explore the earth's subsurface. The goal is to determine material parameters, such as the directional variations of velocities of compression and shear waves as well as their attenuations, from partial measurements of wave fields, which have been excited by artificial or natural sources. Here, full waveform inversion (FWI) refers to the corresponding fully nonlinear inverse and ill-posed problem involving the complete parameter-to-state map of the underlying wave propagation model without any further simplifications.

Accurate mathematical models describing the physics of wave propagation are essential to the success of FWI. Since most real rock formations exhibit a directional dependence of wave velocities and attenuation, the resulting anisotropic effects must be taken into account, as has been demonstrated by, e.g., [16, 18, 20]. The most realistic model to date is the viscoleastic wave equation with an anisotropic material law. Several such material laws have been described in the literature modeling various real-world media, see [5] for an overview. As an example, we will limit our analysis to the widely used *Vertical Transverse Isotropy* (VTI) [22], which involves rotational symmetry about the vertical axis and accounts for anisotropy caused by thin horizontal layers.

Department of Mathematics, Karlsruhe Institute of Technology, D-76128 Karlsruhe, Germany

E-mail address: andreas.rieder@kit.edu.

Date: April 22, 2025.

2020 *Mathematics Subject Classification.* 35F10, 35R30, 86A22.

Key words and phrases. full waveform seismic inversion, anisotropic viscoelastic wave equation, adjoint state method, nonlinear inverse and ill-posed problem.

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 258734477 - SFB 1173.

Since VTI media are characterized by three dimensionless parameters, the Thomsen parameters, FWI in the viscoelastic regime under VTI anisotropy entails the reconstruction of 11 parameter functions, which are in general spatially dependent: bulk density, the vertical velocities of the P- and S-wave, Thomsen parameters, and 5 scaling factors, which specify the attenuation anisotropy. In this work, we rigorously define the corresponding nonlinear operator $\Phi\colon \mathsf{D}(\Phi) \subset L^\infty(D)^{11} \to \mathcal{W}$ as the parameter-to-state map of the governing wave equation, where $\mathsf{D}(\Phi)$ is the admissible parameter set and $\mathcal{W}$ is a suitable Hilbert space containing the wave fields. Evaluating $\Phi(\boldsymbol{m})$ for a given parameter vector $\boldsymbol{m}$ means solving the wave equation, which we formulate as a first-order hyperbolic system and show to fit the abstract form studied in [15]. Thus, we can explicitly provide the Fréchet derivative $\Phi'(\boldsymbol{m}) \in \mathcal{L}\big(L^\infty(D)^{11}, \mathcal{W}\big)$[1] and its adjoint operator $\Phi'(\boldsymbol{m})^* \in \mathcal{L}\big(\mathcal{W}, (L^\infty(D)^{11})'\big)$, which are the central building blocks for most FWI solvers relying on local linearizations, see, e.g., [4, 8, 17, 24].

The paper is organized as follows. In the next section we discuss the elastic wave equation for VTI media, where we recall the main results of [22] and introduce our notation. Furthermore, we give a sound mathematical formulation of the stiffness tensor, which is the basis for all subsequent considerations. Then we add viscosity to the elastic wave equation in Section 3, introducing damping tensors through a slight variation of the generalized standard linear solid rheology due to [25]. Here, we also show that the resulting viscoelastic anisotropic wave equation can be rewritten in an abstract hyperbolic evolution equation as studied in [15]. So we immediately obtain well-posedness and regularity (Theorem 3.1). Based on this well-posedness, in Section 4 we define the corresponding parameter-to-state map Φ and formulate the inverse problem of FWI, which is locally ill-posed everywhere. Finally, we prove Fréchet differentiability of Φ (Theorem 4.2) and express explicitly the adjoint operator of the Fréchet derivative (Theorem 4.6). In the last section, we conclude with a brief comment on other anisotropy concepts and the two-dimensional situation.

2. The elastic wave equation for VTI media

Let $D \subset \mathbb{R}^3$ be a Lipschitz domain. Denoting the velocity by $\boldsymbol{v}\colon [0,\infty) \times D \to \mathbb{R}^3$ and the stress by $\boldsymbol{\sigma}\colon [0,\infty) \times D \to \mathbb{R}^{3\times 3}_{\mathrm{sym}}$, the elastic wave equation for VTI media reads

$$\rho\,\partial_t \boldsymbol{v} = \operatorname{div} \boldsymbol{\sigma} + \boldsymbol{f} \quad \text{in } [0,\infty) \times D, \tag{1a}$$

$$\partial_t \boldsymbol{\sigma} = \boldsymbol{C}\,\boldsymbol{\varepsilon}(\boldsymbol{v}) \quad \text{in } [0,\infty) \times D, \tag{1b}$$

where $\boldsymbol{f}$ denotes the external volume force density and $\rho\colon D \to (0,\infty)$ is the mass density. Further,

$$\boldsymbol{\varepsilon}(\boldsymbol{v}) = \frac{1}{2}\big[(\nabla_x \boldsymbol{v})^\top + \nabla_x \boldsymbol{v}\big]$$

is the linearized strain rate.

The linear map $\boldsymbol{C}\colon \mathbb{R}^{3\times 3}_{\mathrm{sym}} \to \mathbb{R}^{3\times 3}_{\mathrm{sym}}$ implements Hooke's law and is given in the Voigt notation by

$$\boldsymbol{C} = \mathbf{T}(c_{3,3}, c_{5,5}, c_{1,1}, c_{6,6}, c_{1,3})^\top \tag{2}$$

[1]Throughout $\mathcal{L}(Y,Z)$ denotes the space of bounded linear operators between normed vector spaces Y and Z. Further, $\mathcal{L}(Y) := \mathcal{L}(Y,Y)$.

where $\mathbf{T} \in \mathcal{L}\big(\mathbb{R}^5, \mathcal{L}(\mathbb{R}^{3\times 3}_{\mathrm{sym}})\big)$ is the tensor-valued map

$$\mathbf{T}(t_{3,3}, t_{5,5}, t_{1,1}, t_{6,6}, t_{1,3})^\top := \begin{pmatrix} t_{1,1} & t_{1,1} - 2t_{6,6} & t_{1,3} & 0 & 0 & 0 \\ t_{1,1} - 2t_{6,6} & t_{1,1} & t_{1,3} & 0 & 0 & 0 \\ t_{1,3} & t_{1,3} & t_{3,3} & 0 & 0 & 0 \\ 0 & 0 & 0 & t_{5,5} & 0 & 0 \\ 0 & 0 & 0 & 0 & t_{5,5} & 0 \\ 0 & 0 & 0 & 0 & 0 & t_{6,6} \end{pmatrix}.$$

The 5 independent real-valued entries of $\boldsymbol{C}$ are determined to describe layered media that are isotropic in the x_1-x_2-plane and anisotropic in all other planes containing the x_3-direction. These media are called *vertically transverse isotropic* (VTI). In this work, we express the entries of $\boldsymbol{C}$ by the dimensionless Thomsen parameters ε, δ, and γ, as well as v_{p} and v_{s}, which are the P- and S-wave velocities along the x_3-direction, respectively, see, e.g., [5, Chap. 1.2.1]. To model inhomogeneous material, these 5 quantities, like the mass density ρ, are real-valued functions on D, with both velocities attaining only positive values.

The content of the next two paragraphs is basically taken from Thomsen [22]. In VTI media the velocities of plane waves are directional dependent and, for weak anisotropy, are approximated in terms of the Thomsen parameters by

$$\begin{aligned} v_{\mathrm{p}}(\vartheta) &= v_{\mathrm{p}}\big(1 + \delta \sin^2(\vartheta)\cos^2(\vartheta) + \varepsilon \sin^4(\vartheta)\big), \\ v_{\mathrm{SV}}(\vartheta) &= v_{\mathrm{s}}\Big(1 + \frac{v_{\mathrm{p}}^2}{v_{\mathrm{s}}^2}\,(\varepsilon - \delta)\sin^2(\vartheta)\cos^2(\vartheta)\Big), \\ v_{\mathrm{SH}}(\vartheta) &= v_{\mathrm{s}}\big(1 + \gamma \sin^2(\vartheta)\big), \end{aligned}$$

where ϑ is the angle between the x_3-axis and the direction of wave propagation. The occurrence of shear-wave splitting is the most reliable evidence for the presence of anisotropy.

Figure 1 shows the directional variation of P- and S-waves in Mesaverde shale and Biotite crystal. The P-wave velocity increases monotonically towards the horizontal direction. The S-wave velocity depends on the polarity: horizontally polarized S-waves (SH) are fastest in the horizontal direction, whereas vertically polarized S-waves (SV) have the maximum velocity at about $\vartheta = \pi/4$. Observe that for the Biotite crystal there are directions of propagation where both S-waves travel faster than the P-wave.

We introduce the P- and S-wave moduli M[2] and μ, which are connected to the P- and S-wave velocities by

$$M = \rho\, v_{\mathrm{p}}^2 \quad \text{and} \quad \mu = \rho\, v_{\mathrm{s}}^2. \tag{3}$$

The relations of the Thomsen parameters and the entries of $\boldsymbol{C}$ are

$$\varepsilon = \frac{c_{1,1} - c_{3,3}}{2c_{3,3}}, \quad \delta = \frac{(c_{1,3} + c_{5,5})^2 - (c_{3,3} - c_{5,5})^2}{2c_{3,3}(c_{3,3} - c_{5,5})}, \quad \gamma = \frac{c_{6,6} - c_{5,5}}{2c_{5,5}}.$$

Additionally,

$$c_{3,3} = M, \quad c_{5,5} = \mu, \tag{4a}$$

[2]We have that $M = K + 4\mu/3$, where K is the bulk modulus.

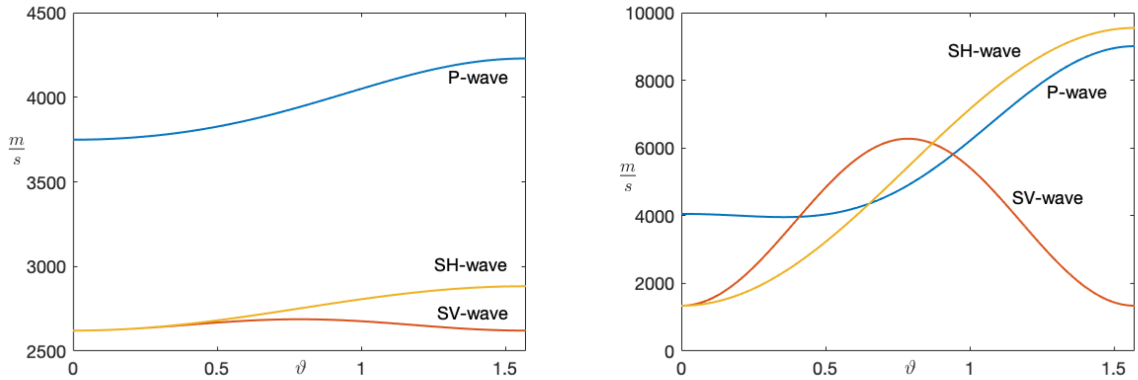


FIGURE 1. Approximations to P- and S-wave phase velocities for two homogeneous media. The angle $\vartheta \in [0, \pi/2]$ denotes direction of wave propagation relative to the vertical axes.
Left: Mesaverde shale at a depth of 3.9km with $v_\mathrm{p} = 3749$m/s, $v_\mathrm{s} = 2621$m/s, $\varepsilon = 0.128$, $\gamma = 0.1$, and $\delta = 0.078$.
Right: Biotite crystal with $v_\mathrm{p} = 4054$m/s, $v_\mathrm{s} = 1341$m/s, $\varepsilon = 0.1222$, $\gamma = 6.12$, and $\delta = -0.388$.
The numerical values for both materials are taken from [22].

yielding

$$\begin{aligned} c_{1,1} &= (2\varepsilon + 1)M, \quad c_{6,6} = (2\gamma + 1)\mu, \\ c_{1,3} &= -\mu \pm \sqrt{(M-\mu)\big((2\delta+1)M - \mu\big)}, \end{aligned} \tag{4b}$$

provided the radicand defining $c_{1,3}$ is non-negative. This is actually a restriction on δ: since $M > \mu$ (see explanation below following (5)) we must have that $2\delta + 1 \geq \mu/M$.

To resolve the ambiguity of which sign to choose in front of the square root in $c_{1,3}$, we consider the case $\delta = 0$ and choose the plus sign: we get $c_{1,3} = M - 2\mu = \lambda$, which is the (first) Lamé coefficient, and this happens to be the physically meaningful case.

The parameters are further subject to the following restrictions which guarantee the positive definiteness of $\boldsymbol{C}$ as a 6×6 matrix (all principal minors are positive):

$$M > 0, \quad \mu > 0, \quad 2\varepsilon + 1 > 0, \quad 2\gamma + 1 > 0, \quad 2\delta + 1 \geq \frac{\mu}{M}, \quad \frac{M}{\mu} > \frac{2\gamma+1}{2\varepsilon+1}, \tag{5a}$$

$$\big((2\varepsilon+1)M - (2\gamma+1)\mu\big)M > \Big(\sqrt{(M-\mu)\big((2\delta+1)M-\mu\big)} - \mu\Big)^2. \tag{5b}$$

Observe that $(2\varepsilon+1)M - (2\gamma+1)\mu > 0$ by the lower bound on M/μ. Setting $\varepsilon = \delta = \gamma = 0$ in (5b) yields $3M > 4\mu$ (or $3v_\mathrm{p}^2 > 4v_\mathrm{s}^2$, in particular $v_\mathrm{p} > v_\mathrm{s}$) which indicates that compression waves travel significantly faster than shear waves in isotropic media.

Remark 2.1. *In this remark we comment further on the physical aspects of the conditions in (5). As already mentioned above, they originate from the positive definiteness of the Hooke tensor $\boldsymbol{C}$. This property of $\boldsymbol{C}$ is a physical requirement so that the strain-energy volume density $\frac{1}{2}\sum_{i,j}(\boldsymbol{C}\boldsymbol{\varepsilon}(\boldsymbol{u}))_{i,j}\varepsilon(\boldsymbol{u})_{i,j}$ attains its minimum at zero strain state ($\boldsymbol{u} = 0$) and always increases when the medium is deformed. As far as the author knows,* *the resulting condition* $v_\mathrm{p} > \sqrt{4/3}\, v_\mathrm{s}$ *is satisfied for real material.*[3] *Therefore, we can conclude by continuity that* (5) *holds for natural materials and small absolute values of the Thomsen parameters. Most of the measured parameters listed in* [22] *have rather small absolute magnitudes (in relation to* M/μ*), so it seems that they satisfy* (5) *(samples taken have confirmed this conjecture, for example both sets of values reported in the caption of Figure* 1*). A full physical interpretation of the parameters is beyond the scope of the present work. In a nutshell, they have the following meaning:* ε *describes the level of P-wave anisotropy,* γ *the level of SH-wave anisotropy, and* δ *the level of P-wave moveout anisotropy; for all details we refer to the Section "Weak Anisotropy" of* [22, Page 1957 ff] *and to* [23, Section 1.2.2].

Finally, the inequalities in (5) *are invariant under the same positive scaling of* M *and* m*. This was to be expected since the validity of* (5) *should be independent of the physical units in which the P- and S-wave moduli are expressed. Consequently, we may replace* M *and* μ *by* v_p^2 *and* v_s^2*, respectively.*

Stiffness tensor as a function of the parameters. From here on, we treat the stiffness matrix (2) as a mapping taking 5 real values as arguments:

$$\boldsymbol{C}\colon \mathsf{D}(\boldsymbol{C}) \subset \mathbb{R}^5 \to \mathcal{L}(\mathbb{R}^{3\times 3}_{\mathrm{sym}}), \quad \boldsymbol{q} = (q_1, q_2, q_3, q_4, q_5)^\top \mapsto \boldsymbol{C}(\boldsymbol{q}),$$

where the entries of $\boldsymbol{C}(\boldsymbol{q})$ are set in agreement with (3) and (4) (with the plus sign in front of the root in $c_{1,3}$):

$$c_{1,1}(\boldsymbol{q}) = (2q_3+1)\,q_1, \quad c_{3,3}(\boldsymbol{q}) = q_1, \quad c_{5,5}(\boldsymbol{q}) = q_2, \quad c_{6,6}(\boldsymbol{q}) = (2q_4+1)\,q_2, \tag{6a}$$

$$c_{1,3}(\boldsymbol{q}) = q_2 + \sqrt{(q_1-q_2)\big((2q_5+1)q_1 - q_2\big)}\,\Big). \tag{6b}$$

Thus, $\boldsymbol{C}(\boldsymbol{q}) = \mathbf{T}\big(c_{3,3}(\boldsymbol{q}), c_{5,5}(\boldsymbol{q}), c_{1,1}(\boldsymbol{q}), c_{6,6}(\boldsymbol{q}), c_{1,3}(\boldsymbol{q})\big)^\top$.

We will define the domain of definition $\mathsf{D}(\boldsymbol{C})$ of $\boldsymbol{C}$ so that it includes material parameters $(M, \mu, \varepsilon, \gamma, \delta)$ which satisfy (5). Moreover, $\mathsf{D}(\boldsymbol{C})$ will be compact with an open interior. To this end, choose $\underline{q}_i, \overline{q}_i \in \mathbb{R}$ with $\underline{q}_i < \overline{q}_i$ for $i = 1, \dots, 5$, where $0 < \underline{q}_i$, $i = 1, 2$, and $-1/2 < \underline{q}_i$, $0 < \overline{q}_i$, $i = 3, 4, 5$. Now, with $r > 0$ define

$$\mathsf{D}(\boldsymbol{C}) := \Bigg\{ \boldsymbol{q} \in \bigtimes_{i=1}^{5} \big[\underline{q}_i, \overline{q}_i\big] \ :\ \frac{q_1}{q_2} \geq r + \max\Big\{\frac{2q_4+1}{2q_3+1}, \frac{1}{2q_5+1}\Big\},$$
$$\big((2q_3+1)q_1 - (2q_4+1)q_2\big)q_1 \geq r + \Big(\sqrt{(q_1-q_2)\big((2q_5+1)q_1-q_2\big)} - q_2\Big)^2 \Bigg\}.$$

Obviously, $\mathsf{D}(\boldsymbol{C})$ is compact and we can assume a non-empty interior: for instance, in case of $3\underline{q}_1 > 4\overline{q}_2$ we fix an $r > 0$ such that $r \leq \max\{(3\underline{q}_1 - 4\overline{q}_2)\underline{q}_2, \underline{q}_1/\overline{q}_2 - 1\}$. Then, $\mathfrak{Q} = (\underline{q}_1, \overline{q}_1) \times (\underline{q}_2, \overline{q}_2) \times \{0\} \times \{0\} \times \{0\} \subset \mathsf{D}(\boldsymbol{C})$. Moreover, by continuity, for each element in $\mathfrak{Q}$ there exists a neighborhood which is contained in $\mathsf{D}(\boldsymbol{C})$. Note that the limiting values $\underline{q}_i, \overline{q}_i$ can be adjusted so that $\mathsf{D}(\boldsymbol{C})$ includes the relevant material parameters for a certain physical setting, see, e.g., Table 1 in [22] for measured anisotropy values in sedimentary rocks.

[3] If Poisson's ratio of the homogeneous material is in $[0, 0.5)$ then $v_\mathrm{p} \geq \sqrt{2}\, v_\mathrm{s} > \sqrt{4/3}\, v_\mathrm{s}$, see, e.g., [11, (9.30)]. This is the case for many minerals and rock types, see, e.g.,[10, Table 3 and Figure 4].

One immediate and important consequence of the properties of $\mathsf{D}(\boldsymbol{C})$ is the uniform positive definiteness of $\boldsymbol{C}(\boldsymbol{q})$ for $\boldsymbol{q} \in \mathsf{D}(\boldsymbol{C})$ (as a 6×6 matrix): there are constants $0 < \mathrm{c} \leq \mathrm{C}$ such that

$$\mathrm{c}|w|^2 \leq w^\top \boldsymbol{C}(\boldsymbol{q}) w \leq \mathrm{C}|w|^2 \quad \text{for any } w \in \mathbb{R}^6 \text{ uniformly in } \boldsymbol{q} \in \mathsf{D}(\boldsymbol{C}). \tag{7}$$

Observe that c tends to zero as $r \searrow 0$.

Further, $\boldsymbol{C}$ is Fréchet differentiable at any $\boldsymbol{q} \in \operatorname{int}\big(\mathsf{D}(\boldsymbol{C})\big)$. In fact, for $\boldsymbol{h} \in \mathbb{R}^5$, $\boldsymbol{C}'(\boldsymbol{q})\boldsymbol{h} \in \mathcal{L}(\mathbb{R}^{3\times 3}_{\mathrm{sym}})$ is represented by

$$\begin{aligned}\boldsymbol{C}'(\boldsymbol{q})\boldsymbol{h} &= \begin{pmatrix} c'_{1,1}(\boldsymbol{q})\boldsymbol{h} & (c'_{1,1} - 2c'_{6,6})(\boldsymbol{q})\boldsymbol{h} & c'_{1,3}(\boldsymbol{q})\boldsymbol{h} & 0 & 0 & 0 \\ (c'_{1,1} - 2c'_{6,6})(\boldsymbol{q})\boldsymbol{h} & c'_{1,1}(\boldsymbol{q})\boldsymbol{h} & c'_{1,3}(\boldsymbol{q})\boldsymbol{h} & 0 & 0 & 0 \\ c'_{1,3}(\boldsymbol{q})\boldsymbol{h} & c'_{1,3}(\boldsymbol{q})\boldsymbol{h} & h_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & h_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & c'_{6,6}(\boldsymbol{q})\boldsymbol{h} \end{pmatrix} \\ &= \mathbf{T}\big(h_1, h_2, c'_{1,1}(\boldsymbol{q})\boldsymbol{h}, c'_{6,6}(\boldsymbol{q})\boldsymbol{h}, c'_{1,3}(\boldsymbol{q})\boldsymbol{h}\big)^\top\end{aligned} \tag{8a}$$

with

$$c'_{1,1}(\boldsymbol{q})\boldsymbol{h} = (2q_3 + 1)h_1 + 2q_1 h_3, \qquad c'_{6,6}(\boldsymbol{q})\boldsymbol{h} = (2q_4 + 1)h_2 + 2q_2 h_4, \tag{8b}$$

$$c'_{1,3}(\boldsymbol{q})\boldsymbol{h} = \frac{\big(q_1(2q_2+1) - q_2(q_5+1)\big)h_1 + \big(q_2 - q_1(q_5+1)\big)h_2 + q_1(q_1 - q_2)h_5}{\sqrt{(q_1 - q_2)\big((2q_5+1)q_1 - q_2\big)}} - h_2. \tag{8c}$$

For $\boldsymbol{q} \in \mathsf{D}(\boldsymbol{C})$ the inverse matrix

$$\widetilde{\boldsymbol{C}}(\boldsymbol{q}) := \boldsymbol{C}(\boldsymbol{q})^{-1} \tag{9}$$

exists and is Fréchet differentiable in the interior of $\mathsf{D}(\boldsymbol{C})$ according to

$$\widetilde{\boldsymbol{C}}'(\boldsymbol{q})\boldsymbol{h} = -\widetilde{\boldsymbol{C}}(\boldsymbol{q})\boldsymbol{C}'(\boldsymbol{q})\boldsymbol{h}\widetilde{\boldsymbol{C}}(\boldsymbol{q}) \tag{10}$$

for which we have applied the chain rule and the derivative of matrix inversion, see, e.g., [6, Example 16.14].

3. Adding viscosity to VTI media

In contrast to elastic materials, viscoelastic materials possess a memory effect, whereby the stress state at a given instant is subject to the cumulative deformation history of the material [7].

Instead of (1b) we allow a retarded material law

$$\partial_t \boldsymbol{\sigma}(t,x) = \mathsf{C}(0)\boldsymbol{\varepsilon}(\boldsymbol{v}(t,x)) + \int_0^t \partial_t \mathsf{C}(t-s)\boldsymbol{\varepsilon}\big(\boldsymbol{v}(s,x)\big)\,\mathrm{d}s, \quad (t,x) \in [0,\infty) \times D,$$

where $\mathsf{C}\colon [0,\infty) \to \mathcal{L}(\mathbb{R}^{3\times 3}_{\mathrm{sym}})$ is the time-dependent Hooke tensor.

In the *generalized standard linear solid* rheology, see, e.g., [21, 1, 19, 9], one defines

$$\mathsf{C}(t) := \boldsymbol{C}(\boldsymbol{p}) + \sum_{l=1}^{L} \exp\Big(-\frac{t}{\mathrm{t}_{\boldsymbol{\sigma},l}}\Big)\,\boldsymbol{C}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})$$

with relaxation times $\mathrm{t}_{\boldsymbol{\sigma},l} > 0$, $l = 1, \dots, L \in \mathbb{N}$, where $\boldsymbol{p} = (M, \mu, \varepsilon, \gamma, \delta)^\top$ contains the material parameters and $\boldsymbol{\tau} := (\tau_\mathrm{p}, \tau_\mathrm{s}, \tau_\mathrm{e}, \tau_\mathrm{g}, \tau_\mathrm{d})^\top$ collects the positive scaling factors of the respective material parameters. The entries of the tensor $\boldsymbol{C}_\mathrm{u}(\boldsymbol{p}, \boldsymbol{\tau}) \in \mathcal{L}(\mathbb{R}^{3\times 3}_\mathrm{sym})$ are the unrelaxed moduli, that is,

$$\boldsymbol{C}_\mathrm{u}(\boldsymbol{q}, \boldsymbol{\tau}) := \mathbf{T}\big(\boldsymbol{\tau} \odot \boldsymbol{c}(\boldsymbol{q})\big), \quad \boldsymbol{c}(\boldsymbol{q}) := (c_{3,3}, c_{5,5}, c_{1,1}, c_{6,6}, c_{1,3})^\top, \quad \boldsymbol{q} \in \mathsf{D}(\boldsymbol{C}), \tag{11}$$

where the $c_{i,j}$'s are functions of $\boldsymbol{q}$ via (6) and the binary operator symbol $\odot$ denotes component-wise multiplication.

Following the presentation in [4] and [6, Chap. 1.5], which is based on [25], we introduce $L \in \mathbb{N}$ damping tensors $\boldsymbol{\sigma}_l \colon [0, \infty) \times D \to \mathbb{R}^{3\times 3}_\mathrm{sym}$,

$$\boldsymbol{\sigma}_l(t) := \boldsymbol{\sigma}_{l,0} + \int_0^t \exp\Big(\frac{s-t}{\mathrm{t}_{\boldsymbol{\sigma},l}}\Big) \boldsymbol{C}_\mathrm{u}(\boldsymbol{p}, \boldsymbol{\tau}) \boldsymbol{\varepsilon}\big(\boldsymbol{v}(s)\big)\, \mathrm{d}s, \qquad l = 1, \dots, L,$$

and the corresponding stress decomposition $\boldsymbol{\sigma} = \boldsymbol{\sigma}_0 + \sum_{l=1}^L \boldsymbol{\sigma}_l$. This decomposition yields the first order system for viscoelastic waves in VTI media

$$\varrho\, \partial_t \boldsymbol{v} = \operatorname{div}\Big(\sum_{l=0}^{L} \boldsymbol{\sigma}_l\Big) + \boldsymbol{f} \qquad \text{in } [0, \infty) \times D, \tag{12a}$$

$$\partial_t \boldsymbol{\sigma}_0 = \boldsymbol{C}\big(\boldsymbol{p}\big)\, \boldsymbol{\varepsilon}(\boldsymbol{v}) \qquad \text{in } [0, \infty) \times D, \tag{12b}$$

$$\mathrm{t}_{\boldsymbol{\sigma},l}\, \partial_t \boldsymbol{\sigma}_l = \mathrm{t}_{\boldsymbol{\sigma},l}\, \boldsymbol{C}_\mathrm{u}(\boldsymbol{p}, \boldsymbol{\tau})\, \boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{\sigma}_l, \qquad l = 1, \dots, L, \qquad \text{in } [0, \infty) \times D, \tag{12c}$$

with initial conditions

$$\boldsymbol{v}(0) = \boldsymbol{v}_0 \quad \text{and} \quad \boldsymbol{\sigma}_l(0) = \boldsymbol{\sigma}_{l,0}, \quad l = 0, \dots, L. \tag{12d}$$

By continuity, we can find bounds $0 < \underline{\tau}_i \leq 1 < \overline{\tau}_i$, $i = 1, \dots, 5$, such that

(13a) $\boldsymbol{C}_\mathrm{u}(\boldsymbol{p}, \boldsymbol{\tau})$ is positive definite and satisfies (7) with adjusted constants

$$\text{for any } \boldsymbol{\tau}(\cdot) \in \bigtimes_{i=1}^{5} \big[\underline{\tau}_i, \overline{\tau}_i\big] \text{ and any } \boldsymbol{p}(\cdot) \in \mathsf{D}(\boldsymbol{C}) \text{ a.e. in } D.$$

Further, we require bounds for the bulk density:

$$0 < \rho_\mathrm{min} \leq \rho(\cdot) \leq \rho_\mathrm{max} < \infty \quad \text{a.e. in } D. \tag{13b}$$

Under these assumptions, the existence and uniqueness theory for abstract evolution equations developed in [14, 15] can be applied to (12) when expressed as

$$Bu'(t) + (A + BQ)u(t) = f(t), \quad t \in [0, \infty), \quad u(0) = u_0, \tag{14}$$

where

$$u(t) = (\boldsymbol{v}(t, \cdot), \boldsymbol{\sigma}_0(t, \cdot), \dots, \boldsymbol{\sigma}_L(t, \cdot))^\top, \quad f(t) = (\boldsymbol{f}(t, \cdot), \mathbf{0}, \dots, \mathbf{0})^\top,$$
$$\text{and} \quad u_0 = (\boldsymbol{v}_0, \boldsymbol{\sigma}_{0,0}, \dots, \boldsymbol{\sigma}_{L,0})^\top.$$

The operators A, B, and Q are defined on the Hilbert space

$$X := L^2(D, \mathbb{R}^3) \times L^2(D, \mathbb{R}^{3\times 3}_\mathrm{sym})^{1+L},$$

which carries the inner product

$$\big\langle (\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L)^\top, (\boldsymbol{w}, \boldsymbol{\psi}_0, \dots, \boldsymbol{\psi}_l)^\top \big\rangle_X := \int_D \Big(\boldsymbol{v} \cdot \mathbf{w} + \sum_{l=0}^{L} \boldsymbol{\sigma}_l : \boldsymbol{\psi}_l\Big)\, \mathrm{d}x$$

with the colon denoting the Frobenius inner product on $\mathbb{R}^{3\times 3}$.

For $w = (\boldsymbol{w}, \boldsymbol{\psi}_0, \dots, \boldsymbol{\psi}_L)^\top \in X$ we define, recalling (9),

$$
(15) \qquad Bw := \begin{pmatrix} \varrho\, \boldsymbol{w} \\ \widetilde{\boldsymbol{C}}(\boldsymbol{p})\boldsymbol{\psi}_0 \\ \widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p}, \boldsymbol{\tau})\boldsymbol{\psi}_1 \\ \vdots \\ \widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p}, \boldsymbol{\tau})\boldsymbol{\psi}_L \end{pmatrix} \quad \text{and} \quad Qw := \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \frac{1}{\mathrm{t}_{\boldsymbol{\sigma},1}}\, \boldsymbol{\psi}_1 \\ \vdots \\ \frac{1}{\mathrm{t}_{\boldsymbol{\sigma},L}}\, \boldsymbol{\psi}_L \end{pmatrix}.
$$

In view of (7), the collection of selfadjoint operators

$$
(16) \qquad \{B = B(\rho, \boldsymbol{p}, \boldsymbol{\tau}) : \rho, \boldsymbol{p}, \boldsymbol{\tau} \text{ satisfy (13)}\} \subset \mathcal{L}(X)
$$

is uniformly positive definite and uniformly bounded.

Finally, the differential operator

$$
(17a) \qquad Aw := - \begin{pmatrix} \operatorname{div}\left(\sum_{l=0}^{L} \boldsymbol{\psi}_l\right) \\ \boldsymbol{\varepsilon}(\boldsymbol{w}) \\ \vdots \\ \boldsymbol{\varepsilon}(\boldsymbol{w}) \end{pmatrix}
$$

is defined on[4]

$$
(17b) \qquad \mathsf{D}(A) := H_0^1(D, \mathbb{R}^3) \times \left\{\boldsymbol{\sigma} \in L^2\big(D, \mathbb{R}^{3\times 3}_{\mathrm{sym}}\big) : \operatorname{div} \boldsymbol{\sigma}_{*,j} \in L^2(D),\ j = 1,2,3\right\}^{1+L}.
$$

Within this setting, $A \colon \mathsf{D}(A) \subset X \to X$ is a maximal monotone operator [15, Lem. 4.1].

Now, all assumptions of the abstract framework from [15, Sec. 3] are met and we can conclude the following well-posedness results for the viscoelastic wave equation in VTI media.

Theorem 3.1. *Assume that* $\boldsymbol{p}(\cdot) \in \mathsf{D}(\boldsymbol{C})$ *a.e. in* D. *Let* (7) *and* (13) *hold.*

If $(\boldsymbol{v}_0, \boldsymbol{\sigma}_{0,0}, \dots, \boldsymbol{\sigma}_{L,0})^\top \in \mathsf{D}(A)$ *and* $\boldsymbol{f} \in W^{1,1}\big([0,\infty), L^2(D,\mathbb{R}^3)\big)$, *then* (12) *admits a unique classical solution* $(\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L)^\top \in \mathcal{C}([0,\infty), \mathsf{D}(A)) \cap \mathcal{C}^1([0,\infty), X)$.

Under less regularity, $(\boldsymbol{v}_0, \boldsymbol{\sigma}_{0,0}, \dots, \boldsymbol{\sigma}_{L,0})^\top \in X$ *and* $\boldsymbol{f} \in L^1_{\mathrm{loc}}([0,\infty), L^2(D,\mathbb{R}^3))$, (12) *admits a unique mild/weak solution* $(\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L)^\top \in \mathcal{C}([0,\infty), X)$ *(see, e.g., [12] for the definition of mild/weak solutions).*

Figure 2 illustrates the effects of damped wave propagation under different anisotropies in two spatial dimensions; see Section 5.2 for the two-dimensional situation.

The frequency dependence of wave propagation in real media about a center frequency ω_0 is modeled via a constant quality factor Q, which is the ratio of the full energy to the dissipated energy, see [1]. In this way the relaxations times $\mathrm{t}_{\boldsymbol{\sigma},l} = \mathrm{t}_{\boldsymbol{\sigma},l}(\omega_0) > 0$ are determined by a least squares method [2, 3]. The frequency-dependent P- and S-wave moduli are now given by

$$
(18) \qquad M = \frac{\rho\, v_{\mathrm{p}}^2}{1 + \tau_{\mathrm{p}}\alpha} \quad \text{and} \quad \mu = \frac{\rho\, v_{\mathrm{s}}^2}{1 + \tau_{\mathrm{s}}\alpha} \quad \text{with } \alpha := \sum_{l=1}^{L} \frac{\omega_0^2\, \mathrm{t}_{\boldsymbol{\sigma},l}^2}{1 + \omega_0^2\, \mathrm{t}_{\boldsymbol{\sigma},l}^2},
$$

replacing (3).

[4] In our definition of $\mathsf{D}(A)$, we impose zero Dirichlet boundary conditions for the velocity on all of $\mathrm{bd}(D)$. However, we could also split the boundary into a part where we have a Dirichlet condition for the velocity and a part where a Neumann condition applies to the stress, see [15, Sec. 4].

[5] FDTD: finite difference time domain

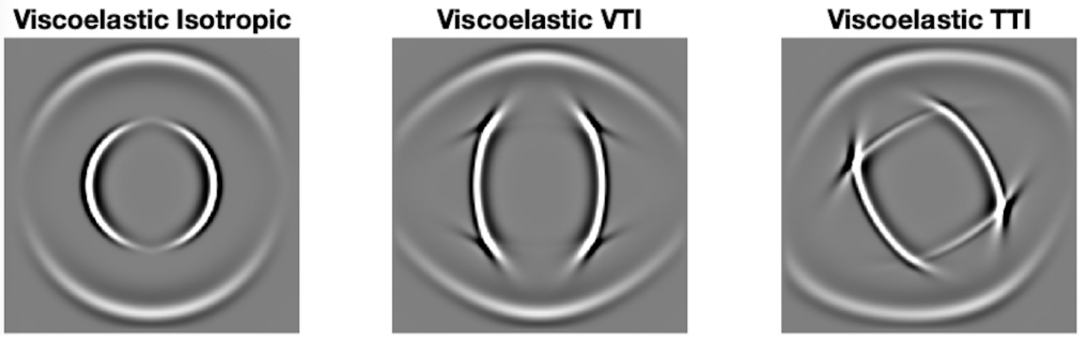


FIGURE 2. Numerical simulations of the vertical component of the particle velocity excited by a vertical point force in the middle. The snapshots show P- and SV-waves and the effects of attenuation for VTI and tilted transverse isotropic anisotropy (TTI), which is a rotated version of VTI, see Section 5.1. The homogeneous anisotropic VTI model is that of Greenhorn shale [13]. In the TTI model the symmetry axis has been rotated counter-clockwise by the angle $\theta_{\rm v} = \pi/6$. Attenuation is assumed to be isotropic with $\tau_{\rm p} = \tau_{\rm s} = 0.25$ and $L = 1$. The simulations were kindly provided by Thomas Bohlen (Geophysical Institute, KIT) and computed with the FDTD[5] code SOFI2D, which is available from https://gitlab.kit.edu/kit/gpi/ag/software/sofi2d.

4. The inverse problem

Full waveform inversion in VTI media under the viscoelastic regime entails the reconstruction of the 11 parameter functions ρ, $\boldsymbol{\pi} = (v_{\rm p}, v_{\rm s}, \varepsilon, \gamma, \delta)^\top$, and $\boldsymbol{\tau} = (\tau_{\rm p}, \tau_{\rm s}, \tau_{\rm e}, \tau_{\rm g}, \tau_{\rm d})^\top$ from (partial) wave field measurements.

4.1. **The setting.** To formulate this seismic inverse problem we need to define the full waveform forward operator Φ. In a first step we introduce the auxiliary operator $\widetilde{\Phi}$ with domain

$$\mathsf{D}(\widetilde{\Phi}) := \big\{(\rho, \boldsymbol{p}, \boldsymbol{\tau})^\top \in L^\infty(D)^{11} : (\rho, \boldsymbol{p}, \boldsymbol{\tau})^\top \text{ satisfy } (13)\big\}^6$$

according to

$$\widetilde{\Phi} \colon \mathsf{D}(\widetilde{\Phi}) \subset L^\infty(D)^{11} \to L^2\big([0,T], X\big), \quad (\rho, \boldsymbol{p}, \boldsymbol{\tau})^\top \mapsto (\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L)^\top|_{[0,T]},$$

where $T > 0$ is the observation period and $(\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L)^\top$ is the unique classical solution of (12) for a given $\boldsymbol{f} \in W^{1,1}\big([0,\infty), L^2(D, \mathbb{R}^3)\big)$ and initial values $(\boldsymbol{v}_0, \boldsymbol{\sigma}_{0,0}, \dots, \boldsymbol{\sigma}_{L,0})^\top \in \mathsf{D}(A)$. This mapping is well defined, see Theorem 3.1.

The final ingredient for Φ is the parameter transformation

$$P \colon \mathsf{D}(P) \subset L^\infty(D)^{11} \to L^\infty(D)^{11},$$

$$(p_0, p_1, p_2, \dots, p_{10})^\top \mapsto \Big(p_0, \frac{p_0\, p_1^2}{1 + \alpha p_6}, \frac{p_0\, p_2^2}{1 + \alpha p_7}, p_3, \dots, p_{10}\Big)^\top,$$

where $\mathsf{D}(P) = L^\infty(D)^6 \times \mathcal{T}$ with

$$\mathcal{T} := \Big\{\boldsymbol{w} \in L^\infty(D)^5 : \boldsymbol{w}(\cdot) \in \bigtimes_{i=1}^{5} \big[\underline{\tau}_i, \overline{\tau}_i\big] \text{ a.e. in } D\Big\}.$$

[6]For simplicity, here and later we use $(\rho, \boldsymbol{p}, \boldsymbol{\tau})^\top$ instead of the correct expression $(\rho, \boldsymbol{p}^\top, \boldsymbol{\tau}^\top)^\top$. Further, we will write $\widetilde{\Phi}(\rho, \boldsymbol{p}, \boldsymbol{\tau})$ for $\widetilde{\Phi}((\rho, \boldsymbol{p}, \boldsymbol{\tau})^\top)$ (likewise for other mappings).

Note that P implements the change of parameters due to (18):

$$P(\rho, v_{\rm p}, v_{\rm s}, \varepsilon, \gamma, \delta, \tau_{\rm p}, \tau_{\rm s}, \tau_{\rm e}, \tau_{\rm g}, \tau_{\rm d}) = (\rho, M, \mu, \varepsilon, \gamma, \delta, \tau_{\rm p}, \tau_{\rm s}, \tau_{\rm e}, \tau_{\rm g}, \tau_{\rm d})^\top,$$

in short, $P(\rho, \boldsymbol{\pi}, \boldsymbol{\tau}) = (\rho, \boldsymbol{p}, \boldsymbol{\tau})^\top$.

Remark 4.1. *A change in model parameters can be accounted for by adjusting the transformation P.*

Now, we define the full waveform forward operator $\Phi := \widetilde{\Phi} \circ P$ on $\mathsf{D}(\Phi) := P^{-1}(\mathsf{D}(\widetilde{\Phi}))$, the preimage of $\mathsf{D}(\widetilde{\Phi})$ under P, by

$$\Phi \colon \mathsf{D}(\Phi) \subset L^\infty(D)^{11} \to L^2\big([0,T], X\big), \quad (\rho, \boldsymbol{\pi}, \boldsymbol{\tau})^\top \mapsto \widetilde{\Phi}\big(P(\rho, \boldsymbol{\pi}, \boldsymbol{\tau})\big). \tag{19}$$

Since P is continuous, we can ensure, by appropriate choices of $\rho_{\min}$, $\rho_{\max}$, and the limiting values appearing in the definitions of $\mathsf{D}(\boldsymbol{C})$ and $\mathcal{T}$, that $\mathsf{D}(\Phi)$ has a non-empty interior, which contains all 11 physically relevant parameters $(\rho, \boldsymbol{\pi}, \boldsymbol{\tau})^\top$. A more explicit expression for $\mathsf{D}(\Phi)$ is not needed in what follows. It is hard to find anyway and would not gain us any deeper insight.

For the complete formulation of the inverse problem we model the measurement process by the linear seismogram operator $S \colon L^2([0,T], X) \to \mathbb{R}^N$, which samples certain components of $\big(\sum_{l=0}^L \boldsymbol{\sigma}_l, \mathbf{v}\big)$ at finitely many times in $[0,T]$ and receiver positions in D (or at its boundary). Here, N is the number of sample points (time points $\times$ number of receivers). Recall that $\sum_{l=0}^L \boldsymbol{\sigma}_l$ adds up to the stress $\boldsymbol{\sigma}$, some components of which can actually be observed, unlike the individual $\boldsymbol{\sigma}_l$'s which have no physical meaning. The images of S are called seismograms. Now, *full waveform inversion* (FWI) in VTI media under the viscoelastic regime consists of solving the nonlinear equation

$$S\Phi(\rho, \boldsymbol{\pi}, \boldsymbol{\tau}) = s \tag{20}$$

for a given seismogram $s \in \mathbb{R}^N$.

The inverse problem (20) is locally ill-posed at any parameter point $\boldsymbol{m}^+ \in \operatorname{int}\big(\mathsf{D}(\Phi)\big)$ in the following sense: in any neighborhood of $\boldsymbol{m}^+$ there exists a sequence $\{\boldsymbol{m}_k\}$ with

$$\lim_{k\to\infty} \|\Phi(\boldsymbol{m}_k) - \Phi(\boldsymbol{m}^+)\|_{L^2([0,T],X)} = 0 \quad \text{but} \quad \boldsymbol{m}_k \not\to \boldsymbol{m}^+ \text{ in } L^\infty(D)^{11}. \tag{21}$$

The validity of (21) can be shown by reproducing the proof of Theorem 4.3 in [15].

4.2. Differentiability and adjoint.

The adequate solution of the ill-posed seismic inverse problem (20) requires regularization, typically by iterative schemes based on local linearization. Required ingredients are Fréchet differentiability and the adjoint operator of the Fréchet derivative. We provide both for the full waveform forward operator Φ, where we will rely on the results from [15] for the abstract evolution equation (14).

4.2.1. *The derivate.* We start by giving the Fréchet derivative of P. For $(\rho, \boldsymbol{\pi}, \boldsymbol{\tau})^\top \in \mathsf{D}(P)$ and $(\widehat{\rho}, \widehat{\boldsymbol{\pi}}, \widehat{\boldsymbol{\tau}})^\top \in L^\infty(D)^{11}$ we have

$$P'(\rho, \boldsymbol{\pi}, \boldsymbol{\tau}) \begin{bmatrix} \widehat{\rho} \\ \widehat{\boldsymbol{\pi}} \\ \widehat{\boldsymbol{\tau}} \end{bmatrix} = \Bigg(\widehat{\rho},\ \frac{v_{\rm p}^2}{1+\alpha\tau_{\rm p}}\,\widehat{\rho} + \frac{2\rho v_{\rm p}}{1+\alpha\tau_{\rm p}}\,\widehat{v}_{\rm p} - \frac{\rho v_{\rm p}^2 \alpha}{(1+\alpha\tau_{\rm p})^2}\,\widehat{\tau}_{\rm p},$$

$$\frac{v_{\rm s}^2}{1+\alpha\tau_{\rm s}}\,\widehat{\rho} + \frac{2\rho v_{\rm s}}{1+\alpha\tau_{\rm s}}\,\widehat{v}_{\rm s} - \frac{\rho v_{\rm s}^2 \alpha}{(1+\alpha\tau_{\rm s})^2}\,\widehat{\tau}_{\rm s},\ \widehat{\varepsilon},\ \widehat{\gamma},\ \widehat{\delta},\ \widehat{\tau}_{\rm p},\ \widehat{\tau}_{\rm s},\ \widehat{\tau}_{\rm e},\ \widehat{\tau}_{\rm g},\ \widehat{\tau}_{\rm d} \Bigg)^\top \in L^\infty(D)^{11}.$$

Further, we differentiate $B\colon \mathsf{D}(\widetilde{\Phi}) \subset L^\infty(D)^{11} \to \mathcal{L}(X)$, as defined in (15) and (16), to get

$$B'(\rho,\boldsymbol{p},\boldsymbol{\tau})\begin{bmatrix}\widehat{\rho}\\ \widehat{\boldsymbol{p}}\\ \widehat{\boldsymbol{\tau}}\end{bmatrix}\begin{pmatrix}\boldsymbol{w}\\ \boldsymbol{\psi}_0\\ \vdots\\ \boldsymbol{\psi}_L\end{pmatrix} = \begin{pmatrix}\widehat{\rho}\,\boldsymbol{w}\\ \widetilde{\boldsymbol{C}}'(\boldsymbol{p})\widehat{\boldsymbol{p}}\,\boldsymbol{\psi}_0\\ \widetilde{\boldsymbol{C}}'_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\begin{bmatrix}\widehat{\boldsymbol{p}}\\ \widehat{\boldsymbol{\tau}}\end{bmatrix}\boldsymbol{\psi}_1\\ \vdots\\ \widetilde{\boldsymbol{C}}'_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\begin{bmatrix}\widehat{\boldsymbol{p}}\\ \widehat{\boldsymbol{\tau}}\end{bmatrix}\boldsymbol{\psi}_L\end{pmatrix} \in X$$

for $(\boldsymbol{w},\boldsymbol{\psi}_0,\dots,\boldsymbol{\psi}_L)^\top \in X$ where, by analogy to (10),

$$\widetilde{\boldsymbol{C}}'_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\begin{bmatrix}\widehat{\boldsymbol{p}}\\ \widehat{\boldsymbol{\tau}}\end{bmatrix} = -\widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{C}'_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\begin{bmatrix}\widehat{\boldsymbol{p}}\\ \widehat{\boldsymbol{\tau}}\end{bmatrix}\widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau}). \tag{22}$$

In view of (11) we get by the chain rule and the linearity of $\mathbf{T}$ that

$$\boldsymbol{C}'_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\begin{bmatrix}\widehat{\boldsymbol{p}}\\ \widehat{\boldsymbol{\tau}}\end{bmatrix} = \mathbf{T}\big(\widehat{\boldsymbol{\tau}}\odot\boldsymbol{c}(\boldsymbol{p}) + \boldsymbol{\tau}\odot\boldsymbol{c}'(\boldsymbol{p})\widehat{\boldsymbol{p}}\big) = \boldsymbol{C}_{\mathrm{u}}(\boldsymbol{p},\widehat{\boldsymbol{\tau}}) + \mathbf{T}\big(\boldsymbol{\tau}\odot\boldsymbol{c}'(\boldsymbol{p})\widehat{\boldsymbol{p}}\big) \tag{23a}$$

with, see (8),

$$\boldsymbol{c}'(\boldsymbol{p})\widehat{\boldsymbol{p}} = \big(\widehat{p}_1, \widehat{p}_2, c'_{1,1}(\boldsymbol{p})\widehat{\boldsymbol{p}}, c'_{6,6}(\boldsymbol{p})\widehat{\boldsymbol{p}}, c'_{1,3}(\boldsymbol{p})\widehat{\boldsymbol{p}}\big)^\top. \tag{23b}$$

Finally, we compose B and P to obtain

$$V := B\circ P\colon \mathsf{D}(\Phi)\subset L^\infty(D)^{11} \to \mathcal{L}(X), \quad V(\rho,\boldsymbol{\pi},\boldsymbol{\tau}) = B(P(\rho,\boldsymbol{\pi},\boldsymbol{\tau})). \tag{24}$$

By the chain rule,

$$\begin{aligned} V'(\rho,\boldsymbol{\pi},\boldsymbol{\tau})\begin{bmatrix}\widehat{\rho}\\ \widehat{\boldsymbol{\pi}}\\ \widehat{\boldsymbol{\tau}}\end{bmatrix}\begin{pmatrix}\boldsymbol{w}\\ \boldsymbol{\psi}_0\\ \vdots\\ \boldsymbol{\psi}_L\end{pmatrix} &= B'(P(\rho,\boldsymbol{\pi},\boldsymbol{\tau}))P'(\rho,\boldsymbol{\pi},\boldsymbol{\tau})\begin{bmatrix}\widehat{\rho}\\ \widehat{\boldsymbol{\pi}}\\ \widehat{\boldsymbol{\tau}}\end{bmatrix}\begin{pmatrix}\boldsymbol{w}\\ \boldsymbol{\psi}_0\\ \vdots\\ \boldsymbol{\psi}_L\end{pmatrix}\\ &= \begin{pmatrix}\widehat{\rho}\,\boldsymbol{w}\\ \widetilde{\boldsymbol{C}}'(\boldsymbol{p})\widehat{\boldsymbol{p}}_\pi\,\boldsymbol{\psi}_0\\ \widetilde{\boldsymbol{C}}'_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\begin{bmatrix}\widehat{\boldsymbol{p}}_\pi\\ \widehat{\boldsymbol{\tau}}\end{bmatrix}\boldsymbol{\psi}_1\\ \vdots\\ \widetilde{\boldsymbol{C}}'_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\begin{bmatrix}\widehat{\boldsymbol{p}}_\pi\\ \widehat{\boldsymbol{\tau}}\end{bmatrix}\boldsymbol{\psi}_L\end{pmatrix}\end{aligned} \tag{25}$$

with the following abbreviations

$$\boldsymbol{p} = (M,\mu,\varepsilon,\gamma,\delta)^\top, \quad \widehat{\boldsymbol{p}}_\pi = (p_{\pi,1},\, p_{\pi,2},\, \widehat{\varepsilon},\, \widehat{\gamma},\, \widehat{\delta})^\top, \tag{26a}$$

and

$$p_{\pi,1} = \frac{M}{\rho}\,\widehat{\rho} + \frac{2M}{v_{\mathrm{p}}}\,\widehat{v}_{\mathrm{p}} - \frac{\alpha M}{1+\alpha\tau_{\mathrm{p}}}\,\widehat{\tau}_{\mathrm{p}}, \qquad p_{\pi,2} = \frac{\mu}{\rho}\,\widehat{\rho} + \frac{2\mu}{v_{\mathrm{s}}}\,\widehat{v}_{\mathrm{s}} - \frac{\alpha\mu}{1+\alpha\tau_{\mathrm{s}}}\,\widehat{\tau}_{\mathrm{s}}. \tag{26b}$$

After this preparatory work, we can state the derivative of the full waveform forward operator Φ as defined in (19).

Theorem 4.2. *Under the assumptions of this section, the full waveform forward operator* Φ *is Fréchet differentiable at any interior point* $(\rho, \boldsymbol{\pi}, \boldsymbol{\tau})^\top$ *of* $\mathsf{D}(\Phi)$*:*

For $(\widehat{\rho}, \widehat{\boldsymbol{\pi}}, \widehat{\boldsymbol{\tau}})^\top \in L^\infty(D)^{11}$ *we have* $\Phi'(\rho, \boldsymbol{\pi}, \boldsymbol{\tau}) \begin{bmatrix} \widehat{\rho} \\ \widehat{\boldsymbol{\pi}} \\ \widehat{\boldsymbol{\tau}} \end{bmatrix} = \overline{u}$ *where* $\overline{u} = (\overline{\boldsymbol{v}}, \overline{\boldsymbol{\sigma}}_0, \dots, \overline{\boldsymbol{\sigma}}_L)^\top \in \mathcal{C}([0,T], X)$ *with* $\overline{u}(0) = 0$ *is the mild solution of*

$$\rho\, \partial_t \overline{\boldsymbol{v}} = \operatorname{div}\Big(\sum_{l=0}^{L} \overline{\boldsymbol{\sigma}}_l\Big) - \widehat{\rho}\, \partial_t \boldsymbol{v}, \tag{27a}$$

$$\partial_t \overline{\boldsymbol{\sigma}}_0 = \boldsymbol{C}(\boldsymbol{p})\varepsilon(\overline{\boldsymbol{v}}) + \boldsymbol{C}'(\boldsymbol{p})\widehat{\boldsymbol{p}}_\pi\, \varepsilon(\boldsymbol{v}), \tag{27b}$$

$$\begin{aligned} \partial_t \overline{\boldsymbol{\sigma}}_l &= \boldsymbol{C}_{\mathrm{u}}(\boldsymbol{p}, \boldsymbol{\tau})\varepsilon(\overline{\boldsymbol{v}}) \\ &\quad - \frac{1}{\tau_{\boldsymbol{\sigma},l}} \overline{\boldsymbol{\sigma}}_l + \boldsymbol{C}'_{\mathrm{u}}(\boldsymbol{p}, \boldsymbol{\tau}) \begin{bmatrix} \widehat{\boldsymbol{p}}_\pi \\ \widehat{\boldsymbol{\tau}} \end{bmatrix} \varepsilon(\boldsymbol{v}), \quad l = 1, \dots, L, \end{aligned} \tag{27c}$$

where $\boldsymbol{v} \in H_0^1(D, \mathbb{R}^3)$ *is the first component of the classical solution of* (12)*,* $\boldsymbol{p}$*,* $\widehat{\boldsymbol{p}}_\pi$*,* $\boldsymbol{C}'$*, and* $\boldsymbol{C}'_{\mathrm{u}}$ *are defined in* (26)*,* (8)*, and* (23)*, respectively.*

For the proof of the theorem above we quote Theorem 3.2 from [15] which applies to the parameter-to-solution map F with respect to (14):

$$F \colon \mathsf{D}(F) \subset \mathcal{L}^*(X) \to \mathcal{C}([0,T], X), \quad B \mapsto u,$$

where

$$\mathsf{D}(F) = \big\{ B \in \mathcal{L}^*(X) : \beta_- \|x\|_X^2 \le \langle Bx, x\rangle_X \le \beta_+ \|x\|_X^2 \big\}$$

for given $0 < \beta_- < \beta_+ < \infty$ and $\mathcal{L}^*(X) = \{ J \in \mathcal{L}(X) : J^* = J\}$.

Theorem 4.3. *Let* $T > 0$*,* $f \in W^{1,1}\big([0,T], X\big)$*, and* $u_0 \in \mathsf{D}(A)$*. Then,* F *is Fréchet differentiable at* $B \in \operatorname{int}\big(\mathsf{D}(F)\big)$ *with* $F'(B)H = \overline{u}$*,* $H \in \mathcal{L}^*(X)$*, where* $\overline{u} \in \mathcal{C}\big([0,T], X\big)$ *is the mild solution of*

$$B\overline{u}'(t) + A\overline{u}(t) + BQ\overline{u}(t) = -H\big(u'(t) + Qu(t)\big), \ t \in [0,T], \quad \overline{u}(0) = 0,$$

with $u = F(B)$ *being the classical solution of* (14)*.*

Proof of Theorem 4.2. By an appropriate choice of β_- and β_+, V from (24) maps $\mathsf{D}(\Phi)$ into $\mathsf{D}(F)$. Then, $\Phi'(\rho, \boldsymbol{\pi}, \boldsymbol{\tau}) \begin{bmatrix} \widehat{\rho} \\ \widehat{\boldsymbol{\pi}} \\ \widehat{\boldsymbol{\tau}} \end{bmatrix} = F'(V(\rho, \boldsymbol{\pi}, \boldsymbol{\tau}))V'(\rho, \boldsymbol{\pi}, \boldsymbol{\tau}) \begin{bmatrix} \widehat{\rho} \\ \widehat{\boldsymbol{\pi}} \\ \widehat{\boldsymbol{\tau}} \end{bmatrix}$ and an application of

Theorem 4.3 with $H = V'(\rho, \boldsymbol{p}, \boldsymbol{\tau}) \begin{bmatrix} \widehat{\rho} \\ \widehat{\boldsymbol{p}} \\ \widehat{\boldsymbol{\tau}} \end{bmatrix}$ yields

$$\begin{pmatrix} \rho\, \partial_t \overline{\boldsymbol{v}} \\ \widetilde{\boldsymbol{C}}(\boldsymbol{p})\partial_t \overline{\boldsymbol{\sigma}}_0 \\ \widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\partial_t \overline{\boldsymbol{\sigma}}_1 \\ \vdots \\ \widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\partial_t \overline{\boldsymbol{\sigma}}_L \end{pmatrix} = \begin{pmatrix} \operatorname{div}\big(\sum_{l=0}^{L} \overline{\boldsymbol{\sigma}}_l\big) \\ \boldsymbol{\varepsilon}(\overline{\boldsymbol{v}}) \\ \vdots \\ \boldsymbol{\varepsilon}(\overline{\boldsymbol{v}}) \end{pmatrix} - \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \frac{1}{\tau_{\boldsymbol{\sigma},1}} \widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\overline{\boldsymbol{\sigma}}_1 \\ \vdots \\ \frac{1}{\tau_{\boldsymbol{\sigma},L}} \widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\overline{\boldsymbol{\sigma}}_L \end{pmatrix}$$
$$- V'(\rho,\boldsymbol{\pi},\boldsymbol{\tau}) \begin{bmatrix} \widehat{\rho} \\ \widehat{\boldsymbol{\pi}} \\ \widehat{\boldsymbol{\tau}} \end{bmatrix} \left[\begin{pmatrix} \partial_t \boldsymbol{v} \\ \partial_t \boldsymbol{\sigma}_0 \\ \partial_t \boldsymbol{\sigma}_1 \\ \vdots \\ \partial_t \boldsymbol{\sigma}_L \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \frac{1}{\tau_{\boldsymbol{\sigma},1}} \boldsymbol{\sigma}_1 \\ \vdots \\ \frac{1}{\tau_{\boldsymbol{\sigma},L}} \boldsymbol{\sigma}_L \end{pmatrix} \right]$$

which can be casted into (27) by (25), (22) and (12b), (12c). □

4.2.2. *The adjoint.* We will derive a rather explicit expression for the adjoint operator of $\Phi'(\rho, \boldsymbol{\pi}, \boldsymbol{\tau})$, which requires some preparation. Again, we build upon an abstract result of [15, Thm. 3.3], which we quote below for completeness.

Theorem 4.4. *Under the notation and assumptions of Theorem 4.3 we have*

$$\big[F'(B)^* g\big] H = \int_0^T \big\langle H(u'(t) + Qu(t)), w(t) \big\rangle_X \mathrm{d}t, \quad g \in L^2([0,T], X),\ H \in \mathcal{L}^*(X), \tag{28}$$

where $w \in \mathcal{C}([0,T], X)$ is the mild solution of the adjoint evolution equation

$$Bw'(t) - A^* w(t) - Q^* B w(t) = g(t),\ t \in [0,T], \quad w(T) = 0. \tag{29}$$

To apply the abstract formulation to our concrete setting, it will be convenient to express the Fréchet derivative (8) of $\boldsymbol{C}$ as the sum

$$\boldsymbol{C}'(\boldsymbol{q})\boldsymbol{h} = \sum_{j=1}^{5} h_j \boldsymbol{C}_j'(\boldsymbol{q}) \quad \text{with} \;\; \boldsymbol{C}_j'(\boldsymbol{q}) = \boldsymbol{C}'(\boldsymbol{q})\boldsymbol{e}_j, = \mathbf{T}\big(\boldsymbol{c}'(\boldsymbol{p})\boldsymbol{e}_j\big), \tag{30}$$

where $\boldsymbol{e}_j \in \mathbb{R}^5$ is the j-th canonical unit vector and the meaning of $\boldsymbol{c}'(\boldsymbol{p})\boldsymbol{e}_j$ is given in (23b). For instance,

$$\boldsymbol{C}_1'(\boldsymbol{q}) = \mathbf{T}(1, 0, 2q_3 + 1, 0, t_{1,3})^\top \quad \text{with} \;\; t_{1,3} = \frac{q_1(2q_2+1) - q_2(q_5+1)}{\sqrt{(q_1 - q_2)\big((2q_5+1)q_1 - q_2\big)}}. \tag{31}$$

Similarly we decompose

$$\boldsymbol{C}(\boldsymbol{q}) \overset{(11)}{=} \mathbf{T}\big(c_{3,3}(\boldsymbol{q}), c_{5,5}(\boldsymbol{q}), c_{1,1}(\boldsymbol{q}), c_{6,6}(\boldsymbol{q}), c_{1,3}(\boldsymbol{q})\big)^\top = \sum_{j=1}^{5} \boldsymbol{C}_j(\boldsymbol{q}) \tag{32}$$

where $\boldsymbol{C}_1(\boldsymbol{q}) = \mathbf{T}\big(c_{3,3}(\boldsymbol{q})\boldsymbol{e}_1\big)$, $\boldsymbol{C}_2(\boldsymbol{q}) = \mathbf{T}\big(c_{5,5}(\boldsymbol{q})\boldsymbol{e}_2\big)$, and so on.

For a compact notation, we finally introduce the definitions

$$\boldsymbol{D}_i(\boldsymbol{q},\boldsymbol{t}) := \boldsymbol{C}_i'(\boldsymbol{q})\widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{q},\boldsymbol{t}) \quad \text{and} \quad \boldsymbol{E}_i(\boldsymbol{q},\boldsymbol{t}) := \boldsymbol{C}_i(\boldsymbol{q})\widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{q},\boldsymbol{t})$$
$$\text{for } \boldsymbol{q} \in \mathsf{D}(\boldsymbol{C}),\ \boldsymbol{t} \in \mathcal{T}, \quad i = 1,\dots,5.$$

Note that $\widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{q},\boldsymbol{1}) = \widetilde{\boldsymbol{C}}(\boldsymbol{q})$ for $\boldsymbol{1} = (1,1,1,1,1)$.

Remark 4.5. *In principle, the matrix inverse $\widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{q},\boldsymbol{t}) = \boldsymbol{C}_{\mathrm{u}}(\boldsymbol{q},\boldsymbol{t})^{-1}$ can be expressed explicitly in terms of the entries of $\boldsymbol{C}_{\mathrm{u}}(\boldsymbol{q},\boldsymbol{t})$ since only the two 3×3 diagonal blocks of $\boldsymbol{C}_{\mathrm{u}}(\boldsymbol{q},\boldsymbol{t})$ have to be inverted.*

We are now ready to state and prove the proposed adjoint operator of Φ'.

Theorem 4.6. *Let the assumptions of Theorem 4.2 hold. Then, the adjoint*

$$\Phi'(\rho,\boldsymbol{\pi},\boldsymbol{\tau})^* \in \mathcal{L}\big(L^2([0,T],X),(L^\infty(D)^{11})'\big)$$

at $(\rho,\boldsymbol{\pi},\boldsymbol{\tau})^\top \in \mathsf{D}(\Phi)$ is given by

$$\Phi'(\rho,\boldsymbol{\pi},\boldsymbol{\tau})^*\boldsymbol{g} = -\begin{pmatrix} \int_0^T \big(\frac{1}{\rho}\,\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}^\rho - \partial_t\boldsymbol{v}\cdot\boldsymbol{w}\big)\mathrm{d}t \\ \frac{2M}{v_{\mathrm{p}}}\int_0^T \boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_1^{\boldsymbol{\pi}}(\tau_{\mathrm{p}})\,\mathrm{d}t \\ \frac{2\mu}{v_{\mathrm{s}}}\int_0^T \boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_2^{\boldsymbol{\pi}}(\tau_{\mathrm{s}})\,\mathrm{d}t \\ \int_0^T \boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_3^{\boldsymbol{\pi}}(\tau_{\mathrm{e}})\,\mathrm{d}t \\ \int_0^T \boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_4^{\boldsymbol{\pi}}(\tau_{\mathrm{g}})\,\mathrm{d}t \\ \int_0^T \boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_5^{\boldsymbol{\pi}}(\tau_{\mathrm{d}})\,\mathrm{d}t \\ \int_0^T \boldsymbol{\varepsilon}(\boldsymbol{v}) : \big(\boldsymbol{\Sigma}_1^{\boldsymbol{\tau}} - \frac{\alpha M}{1+\alpha\tau_{\mathrm{p}}}\boldsymbol{\Sigma}_1^{\boldsymbol{\pi}}(\tau_{\mathrm{p}})\big)\,\mathrm{d}t \\ \int_0^T \boldsymbol{\varepsilon}(\boldsymbol{v}) : \big(\boldsymbol{\Sigma}_2^{\boldsymbol{\tau}} - \frac{\alpha \mu}{1+\alpha\tau_{\mathrm{s}}}\boldsymbol{\Sigma}_2^{\boldsymbol{\pi}}(\tau_{\mathrm{s}})\big)\,\mathrm{d}t \\ \int_0^T \boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_3^{\boldsymbol{\tau}}\,\mathrm{d}t \\ \int_0^T \boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_4^{\boldsymbol{\tau}}\,\mathrm{d}t \\ \int_0^T \boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_5^{\boldsymbol{\tau}}\,\mathrm{d}t \end{pmatrix} \in L^1(D)^{11} \tag{33}$$

for $\boldsymbol{g} = (\boldsymbol{g}_{-1},\boldsymbol{g}_0,\dots,\boldsymbol{g}_L)^\top \in L^2\big([0,T],L^2(D,\mathbb{R}^3)\times L^2(D,\mathbb{R}^{3\times 3}_{\mathrm{sym}})^{1+L}\big)$, where $\boldsymbol{v}$ is the first component of the solution of (12),

$$\boldsymbol{\Sigma}^\rho = \big(M\,\boldsymbol{D}_1(\boldsymbol{p},\boldsymbol{1}) + \mu\,\boldsymbol{D}_2(\boldsymbol{p},\boldsymbol{1})\big)\boldsymbol{\varphi}_0 + \big(M\tau_{\mathrm{p}}\,\boldsymbol{D}_1(\boldsymbol{p},\boldsymbol{\tau}) + \mu\tau_{\mathrm{s}}\,\boldsymbol{D}_2(\boldsymbol{p},\boldsymbol{\tau})\big)\sum_{l=1}^L \boldsymbol{\varphi}_l,$$

$$\boldsymbol{\Sigma}_i^{\boldsymbol{\pi}}(\sigma) = \boldsymbol{D}_i(\boldsymbol{p},\boldsymbol{1})\boldsymbol{\varphi}_0 + \sigma\,\boldsymbol{D}_i(\boldsymbol{p},\boldsymbol{\tau})\sum_{l=1}^L \boldsymbol{\varphi}_l,\ \sigma\in\mathbb{R}, \quad \boldsymbol{\Sigma}_i^{\boldsymbol{\tau}} = \boldsymbol{E}_i(\boldsymbol{p},\boldsymbol{\tau})\sum_{l=1}^L \boldsymbol{\varphi}_l, \quad i = 1,\dots,5,$$

and $w = (\boldsymbol{w}, \boldsymbol{\varphi}_0, \ldots, \boldsymbol{\varphi}_L)^\top \in \mathcal{C}([0,T], X)$ *uniquely solves the adjoint state equation*

$$\partial_t \boldsymbol{w} = \frac{1}{\rho} \operatorname{div}\Big(\sum_{l=0}^{L} \boldsymbol{\varphi}_l\Big) + \frac{1}{\rho} \boldsymbol{g}_{-1}, \tag{34a}$$

$$\partial_t \boldsymbol{\varphi}_0 = \boldsymbol{C}(\boldsymbol{p})\big(\boldsymbol{\varepsilon}(\boldsymbol{w}) + \boldsymbol{g}_0\big), \tag{34b}$$

$$\partial_t \boldsymbol{\varphi}_l = \boldsymbol{C}_{\mathrm{u}}(\boldsymbol{p}, \boldsymbol{\tau})\big(\boldsymbol{\varepsilon}(\boldsymbol{w}) + \boldsymbol{g}_l\big) + \frac{1}{\mathrm{t}_{\boldsymbol{\sigma},l}} \boldsymbol{\varphi}_l, \qquad l = 1, \ldots, L, \tag{34c}$$

with end condition $w(T) = \boldsymbol{0}$.

As a product of two $L^2(D)$ functions, each component of the right-hand side of (33) is actually a function in $L^1(D)$ which is a subspace of $L^\infty(D)'$.

Proof of Theorem 4.6. Since the operators from (15) and (17) satisfy $A^* = -A$, $Q^* = Q$, and $QB = BQ$, we observe that (34) is (29) when formulated for viscoelastic VTI media. Moreover, by (28),

$$\begin{aligned}
&\left\langle \Phi'(\rho, \boldsymbol{\pi}, \boldsymbol{\tau})^* \boldsymbol{g}, \begin{pmatrix} \widehat{\rho} \\ \widehat{\boldsymbol{\pi}} \\ \widehat{\boldsymbol{\tau}} \end{pmatrix} \right\rangle_{(L^\infty(D)^{11})' \times L^\infty(D)^{11}} \\
&\qquad = \left\langle F'\big(V(\rho, \boldsymbol{\pi}, \boldsymbol{\tau})\big)^* \boldsymbol{g}, V'(\rho, \boldsymbol{\pi}, \boldsymbol{\tau}) \begin{pmatrix} \widehat{\rho} \\ \widehat{\boldsymbol{\pi}} \\ \widehat{\boldsymbol{\tau}} \end{pmatrix} \right\rangle_{\mathcal{L}(X)' \times \mathcal{L}(X)} \\
&\qquad = \int_0^T \left\langle V'(\rho, \boldsymbol{\pi}, \boldsymbol{\tau}) \begin{pmatrix} \widehat{\rho} \\ \widehat{\boldsymbol{\pi}} \\ \widehat{\boldsymbol{\tau}} \end{pmatrix} \big(u'(t) + Qu(t)\big), w(t) \right\rangle_X \mathrm{d}t,
\end{aligned}$$

where $u = (\boldsymbol{v}, \boldsymbol{\sigma}_0, \ldots, \boldsymbol{\sigma}_L)^\top$ is the classical solution of (12).

With (25), the integrand can be evaluated to yield

$$\begin{aligned}
\left\langle V'(\rho, \boldsymbol{\pi}, \boldsymbol{\tau}) \begin{pmatrix} \widehat{\rho} \\ \widehat{\boldsymbol{\pi}} \\ \widehat{\boldsymbol{\tau}} \end{pmatrix} (u' + Qu), w \right\rangle_X &= \int_D \Big[\widehat{\rho}\, \partial_t \boldsymbol{v} \cdot \boldsymbol{w} + \widetilde{\boldsymbol{C}}'(\boldsymbol{p}) \widehat{\boldsymbol{p}}_\pi \partial_t \boldsymbol{\sigma}_0 : \boldsymbol{\varphi}_0 \\
&\quad + \sum_{l=1}^{L} \widetilde{\boldsymbol{C}}_{\mathrm{u}}'(\boldsymbol{p}, \boldsymbol{\tau}) \begin{bmatrix} \widehat{\boldsymbol{p}}_\pi \\ \widehat{\boldsymbol{\tau}} \end{bmatrix} \Big(\partial_t \boldsymbol{\sigma}_l + \frac{\boldsymbol{\sigma}_l}{\mathrm{t}_{\boldsymbol{\sigma},l}} \Big) : \boldsymbol{\varphi}_l \Big] \mathrm{d}x.
\end{aligned}$$

We continue to evaluate the terms of the integrand. By (10) and the selfadjointness of the involved tensors with respect to the Frobenius inner product, we get

$$\widetilde{\boldsymbol{C}}'(\boldsymbol{p}) \widehat{\boldsymbol{p}}_\pi \partial_t \boldsymbol{\sigma}_0 : \boldsymbol{\varphi}_0 = -\boldsymbol{C}'(\boldsymbol{p}) \widehat{\boldsymbol{p}}_\pi \widetilde{\boldsymbol{C}}(\boldsymbol{p}) \partial_t \boldsymbol{\sigma}_0 : \widetilde{\boldsymbol{C}}(\boldsymbol{p}) \boldsymbol{\varphi}_0 \overset{(12b)}{=} -\boldsymbol{C}'(\boldsymbol{p}) \widehat{\boldsymbol{p}}_\pi \boldsymbol{\varepsilon}(\boldsymbol{v}) : \widetilde{\boldsymbol{C}}(\boldsymbol{p}) \boldsymbol{\varphi}_0.$$

We apply (30) with (26) to the term on the right. Then,

$$\begin{aligned}
\widetilde{\boldsymbol{C}}'(\boldsymbol{p}) \widehat{\boldsymbol{p}}_\pi \partial_t \boldsymbol{\sigma}_0 : \boldsymbol{\varphi}_0 &= -\Big[\Big(\frac{M}{\rho}\, \widehat{\rho} + \frac{2M}{v_{\mathrm{p}}}\, \widehat{v}_{\mathrm{p}} - \frac{\alpha M}{1 + \alpha \tau_{\mathrm{p}}}\, \widehat{\tau}_{\mathrm{p}} \Big) \boldsymbol{C}_1'(\boldsymbol{p}) \\
&\quad + \Big(\frac{\mu}{\rho}\, \widehat{\rho} + \frac{2\mu}{v_{\mathrm{s}}}\, \widehat{v}_{\mathrm{s}} - \frac{\alpha \mu}{1 + \alpha \tau_{\mathrm{s}}}\, \widehat{\tau}_{\mathrm{s}} \Big) \boldsymbol{C}_2'(\boldsymbol{p}) + \widehat{\varepsilon}\, \boldsymbol{C}_3'(\boldsymbol{p}) + \widehat{\gamma}\, \boldsymbol{C}_4'(\boldsymbol{p}) + \widehat{\delta}\, \boldsymbol{C}_5'(\boldsymbol{p}) \Big] \boldsymbol{\varepsilon}(\boldsymbol{v}) : \widetilde{\boldsymbol{C}}(\boldsymbol{p}) \boldsymbol{\varphi}_0,
\end{aligned}$$

which we regroup into

$$\begin{aligned}\widetilde{\boldsymbol{C}}'(\boldsymbol{p})\widehat{\boldsymbol{p}}_\pi\partial_t\boldsymbol{\sigma}_0 : \boldsymbol{\varphi}_0 = -\Big[&\widehat{\rho}\Big(\frac{M}{\rho}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_1(\boldsymbol{p},\boldsymbol{1})\boldsymbol{\varphi}_0 + \frac{\mu}{\rho}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_2(\boldsymbol{p},\boldsymbol{1})\boldsymbol{\varphi}_0\Big)\\
&+ \widehat{v}_{\mathrm{p}}\Big(\frac{2M}{v_{\mathrm{p}}}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_1(\boldsymbol{p},\boldsymbol{1})\boldsymbol{\varphi}_0\Big) + \widehat{v}_{\mathrm{s}}\Big(\frac{2\mu}{v_{\mathrm{s}}}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_2(\boldsymbol{p},\boldsymbol{1})\boldsymbol{\varphi}_0\Big)\\
&+ \widehat{\varepsilon}\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_3(\boldsymbol{p},\boldsymbol{1})\boldsymbol{\varphi}_0\big) + \widehat{\gamma}\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_4(\boldsymbol{p},\boldsymbol{1})\boldsymbol{\varphi}_0\big) + \widehat{\delta}\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_5(\boldsymbol{p},\boldsymbol{1})\boldsymbol{\varphi}_0\big)\\
&- \widehat{\tau}_{\mathrm{p}}\Big(\frac{\alpha M}{1+\alpha\tau_{\mathrm{p}}}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_1(\boldsymbol{p},\boldsymbol{1})\boldsymbol{\varphi}_0\Big) - \widehat{\tau}_{\mathrm{s}}\Big(\frac{\alpha\mu}{1+\alpha\tau_{\mathrm{s}}}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_2(\boldsymbol{p},\boldsymbol{1})\boldsymbol{\varphi}_0\Big)\Big].\end{aligned}$$

Now we consider

$$\begin{aligned}\widetilde{\boldsymbol{C}}_{\mathrm{u}}'(\boldsymbol{p},\boldsymbol{\tau})\begin{bmatrix}\widehat{\boldsymbol{p}}_\pi\\ \widehat{\boldsymbol{\tau}}\end{bmatrix}\Big(\partial_t\boldsymbol{\sigma}_l + \frac{\boldsymbol{\sigma}_l}{\mathrm{t}_{\boldsymbol{\sigma},l}}\Big) : \boldsymbol{\varphi}_l &= -\boldsymbol{C}_{\mathrm{u}}'(\boldsymbol{p},\boldsymbol{\tau})\begin{bmatrix}\widehat{\boldsymbol{p}}_\pi\\ \widehat{\boldsymbol{\tau}}\end{bmatrix}\widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\Big(\partial_t\boldsymbol{\sigma}_l + \frac{\boldsymbol{\sigma}_l}{\mathrm{t}_{\boldsymbol{\sigma},l}}\Big) : \widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\\
&\overset{(23),(12c)}{=} -\boldsymbol{C}_{\mathrm{u}}(\boldsymbol{p},\widehat{\boldsymbol{\tau}})\boldsymbol{\varepsilon}(\boldsymbol{v}) : \widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l - \mathbf{T}\big(\boldsymbol{\tau}\odot\boldsymbol{c}'(\boldsymbol{p})\widehat{\boldsymbol{p}}_\pi\big)\boldsymbol{\varepsilon}(\boldsymbol{v}) : \widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l.\end{aligned}$$

We can handle $\mathbf{T}\big(\boldsymbol{\tau}\odot\boldsymbol{c}'(\boldsymbol{p})\widehat{\boldsymbol{p}}_\pi\big)\boldsymbol{\varepsilon}(\boldsymbol{v}) : \widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l$ by analogy with $\boldsymbol{C}'(\boldsymbol{p})\widehat{\boldsymbol{p}}_\pi\boldsymbol{\varepsilon}(\boldsymbol{v}) : \widetilde{\boldsymbol{C}}(\boldsymbol{p})\boldsymbol{\varphi}_0$ from above. Thus,

$$\begin{aligned}\mathbf{T}\big(\boldsymbol{\tau}\odot\boldsymbol{c}'(\boldsymbol{p})\widehat{\boldsymbol{p}}_\pi\big)\boldsymbol{\varepsilon}(\boldsymbol{v}) : \widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l = \widehat{\rho}\Big(&\frac{M\tau_{\mathrm{p}}}{\rho}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_1(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l + \frac{\mu\tau_{\mathrm{s}}}{\rho}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_2(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\Big)\\
&+ \widehat{v}_{\mathrm{p}}\Big(\frac{2M\tau_p}{v_{\mathrm{p}}}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_1(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\Big) + \widehat{v}_{\mathrm{s}}\Big(\frac{2\mu\tau_{\mathrm{s}}}{v_{\mathrm{s}}}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_2(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\Big)\\
&+ \widehat{\varepsilon}\big(\tau_{\mathrm{e}}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_3(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\big) + \widehat{\gamma}\big(\tau_{\mathrm{g}}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_4(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\big) + \widehat{\delta}\big(\tau_{\mathrm{d}}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_5(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\big)\\
&- \widehat{\tau}_{\mathrm{p}}\Big(\frac{\alpha M\tau_{\mathrm{p}}}{1+\alpha\tau_{\mathrm{p}}}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_1(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\Big) - \widehat{\tau}_{\mathrm{s}}\Big(\frac{\alpha\mu\tau_{\mathrm{s}}}{1+\alpha\tau_{\mathrm{s}}}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{D}_2(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\Big).\end{aligned}$$

Using (11) and (32) we get

$$\begin{aligned}\boldsymbol{C}_{\mathrm{u}}(\boldsymbol{p},\widehat{\boldsymbol{\tau}})\boldsymbol{\varepsilon}(\boldsymbol{v}) : \widetilde{\boldsymbol{C}}_{\mathrm{u}}(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l = \widehat{\tau}_{\mathrm{p}}&\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{E}_1(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\big) + \widehat{\tau}_{\mathrm{s}}\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{E}_2(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\big)\\
&+ \widehat{\tau}_{\mathrm{e}}\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{E}_3(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\big) + \widehat{\tau}_{\mathrm{g}}\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{E}_4(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\big) + \widehat{\tau}_{\mathrm{d}}\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{E}_5(\boldsymbol{p},\boldsymbol{\tau})\boldsymbol{\varphi}_l\big).\end{aligned}$$

By these evaluations we find that

$$\begin{aligned}&\left\langle V'(\rho,\boldsymbol{\pi},\boldsymbol{\tau})\begin{pmatrix}\widehat{\rho}\\ \widehat{\boldsymbol{\pi}}\\ \widehat{\boldsymbol{\tau}}\end{pmatrix}(u'+Qu), w\right\rangle_X\\
&\quad= \int_D \Big[\widehat{\rho}\Big(\partial_t\boldsymbol{v}\cdot\boldsymbol{w} - \frac{1}{\rho}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}^\rho\Big) - \widehat{v}_{\mathrm{p}}\Big(\frac{2M}{v_{\mathrm{p}}}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_1^{\boldsymbol{\pi}}(\tau_{\mathrm{p}})\Big) - \widehat{v}_{\mathrm{s}}\Big(\frac{2\mu}{v_{\mathrm{s}}}\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_2^{\boldsymbol{\pi}}(\tau_{\mathrm{s}})\Big)\\
&\qquad - \widehat{\varepsilon}\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_3^{\boldsymbol{\pi}}(\tau_{\mathrm{e}})\big) - \widehat{\gamma}\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_4^{\boldsymbol{\pi}}(\tau_{\mathrm{g}})\big) - \widehat{\delta}\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_5^{\boldsymbol{\pi}}(\tau_{\mathrm{d}})\big)\\
&\qquad + \widehat{\tau}_{\mathrm{p}}\Big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \Big(\boldsymbol{\Sigma}_1^{\boldsymbol{\tau}} - \frac{\alpha M}{1+\alpha\tau_{\mathrm{p}}}\boldsymbol{\Sigma}_1^{\boldsymbol{\pi}}(\tau_{\mathrm{p}})\Big)\Big) + \widehat{\tau}_{\mathrm{s}}\Big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \Big(\boldsymbol{\Sigma}_2^{\boldsymbol{\tau}} - \frac{\alpha\mu}{1+\alpha\tau_{\mathrm{s}}}\boldsymbol{\Sigma}_2^{\boldsymbol{\pi}}(\tau_{\mathrm{s}})\Big)\Big)\\
&\qquad - \widehat{\tau}_{\mathrm{e}}\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_3^{\boldsymbol{\tau}}\big) - \widehat{\tau}_{\mathrm{g}}\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_4^{\boldsymbol{\tau}}\big) - \widehat{\tau}_{\mathrm{d}}\big(\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\Sigma}_5^{\boldsymbol{\tau}}\big)\Big]\,\mathrm{d}x.\end{aligned}$$

Integrating both sides in time from 0 to T and changing the order of integration leads to the stated expression for $\Phi'(\rho,\boldsymbol{\pi},\boldsymbol{\tau})^*\boldsymbol{g}$. □

Remark 4.7. *We point out that the tensors $\boldsymbol{C}_i(\boldsymbol{p})$ and $\boldsymbol{C}_i'(\boldsymbol{p})$ being part of $\boldsymbol{E}_i(\boldsymbol{p},\boldsymbol{\tau})$ and $\boldsymbol{D}_i(\boldsymbol{p},\boldsymbol{\tau})$, respectively, can be further decomposed into linear combinations of the 5 basis tensors $\mathbf{T}(\boldsymbol{e}_i)$, $i = 1,\dots,5$. Then, only the scalars in front of the basis tensors depend on the material parameters. For instance,*

$$\boldsymbol{C}_4(\boldsymbol{p}) \overset{(32)}{=} c_{6,6}(\boldsymbol{p})\mathbf{T}(\boldsymbol{e}_4) = (2\gamma+1)\mu\,\mathbf{T}(\boldsymbol{e}_4)$$

and

$$\boldsymbol{C}_1'(\boldsymbol{p}) \overset{(31)}{=} \mathbf{T}(\boldsymbol{e}_1) + (2\varepsilon+1)\mathbf{T}(\boldsymbol{e}_3) + \frac{M(2\mu+1)-\mu(\delta+1)}{\sqrt{(M-\mu)\big((2\delta+1)M-\mu\big)}}\,\mathbf{T}(\boldsymbol{e}_5).$$

For a compact and clearer presentation of the previous theorem, we have dispensed with this granular decomposition.

5. Concluding remarks

5.1. **Other anisotropic media.** The techniques to obtain the Fréchet derivative and its adjoint for the viscoelastic FWI forward operator, which we have demonstrated in the previous sections, transfer to the corresponding operators for other anisotropic media, such as tilted transverse isotropic (TTI) and monoclinic media, see, e.g., [5]. In a TTI medium, for example, the vertical rotational symmetry axis of the VTI medium is tilted by the angles θ_v and θ_h in the vertical x_1-x_3-plane and the horizontal x_1-x_2-plane, respectively. In such a way, TTI media model the tilt of strata due to tectonic movement. The corresponding stiffness tensor can be obtained from the VTI tensor by a local rotation $O(\theta_\mathrm{v},\theta_\mathrm{h})$ (Bond transformation, see, e.g., [5]), that is, the tilt angles are spatially dependent functions and

$$\boldsymbol{C}_\mathrm{u}^{\mathrm{TTI}}(\boldsymbol{p},\boldsymbol{\tau},\theta_\mathrm{v},\theta_\mathrm{h}) = O(\theta_\mathrm{v},\theta_\mathrm{h})\,\boldsymbol{C}_\mathrm{u}(\boldsymbol{p},\boldsymbol{\tau})\,O(\theta_\mathrm{v},\theta_\mathrm{h})^\top.$$

Hence, three-dimensional FWI in attenuating TTI media inverts for 13 parameter functions. See [18] for a two-dimensional example, however, without attenuation.

5.2. **The two-dimensional case.** Our results include the two-dimensional situation as well. We only need to set the partial derivatives of the stress components with respect to x_2 in (12) to zero. Then the wave equation decomposes into two independent systems describing the P/SV case and the SH case. For instance, the stiffness tensor for the P/SV case is the 3×3 matrix

$$\boldsymbol{C} = \begin{pmatrix} c_{1,1} & c_{1,3} & 0 \\ c_{1,3} & c_{3,3} & 0 \\ 0 & 0 & c_{5,5} \end{pmatrix},$$

whose entries are still given by (4). Horizontally polarized S-waves occur only in the SH case, where the stiffness tensor is the 2×2 diagonal matrix with entries $c_{6,6}$ and $c_{5,5}$.

Acknowledgment

The author thanks Thomas Bohlen from the Geophysical Institute at KIT for sharing his expertise.

References

[1] J. O. Blanch, J. O. A. Robertsson, and W. W. Symes, *Modeling of a constant Q: Methodology and algorithm for an efficient and optimally inexpensive viscoelastic technique*, Geophysics, 60 (1995), pp. 176–184, https://doi.org/10.1190/1.1443744.

[2] T. Bohlen, *Viskoelastische FD-Modellierung seismischer Wellen zur Interpretation gemessener Seismogramme*, PhD thesis, Christian-Albrechts-Universität zu Kiel, 1998, https://bit.ly/2LMOSWr.

[3] T. Bohlen, *Parallel 3-D viscoelastic finite difference seismic modelling*, Comput. Geosci., 28 (2002), pp. 887–899, https://doi.org/10.1016/S0098-3004(02)00006-7.

[4] T. Bohlen, M. R. Fernandez, J. Ernesti, C. Rheinbay, A. Rieder, and C. Wieners, *Visco-acoustic full waveform inversion: From a DG forward solver to a Newton-CG inverse solver*, Comput. Math. Appl., 100 (2021), pp. 126–140, https://doi.org/10.1016/j.camwa.2021.09.001.

[5] J. Carcione, *Wave Fields in Real Media: Wave Propagation in Anisotropic, Anelastic, Porous and Electromagnetic Media*, Elsevier Science, 3rd ed., 2015, https://doi.org/10.1016/C2021-0-00938-X.

[6] W. Dörfler, M. Hochbruck, J. Köhler, A. Rieder, R. Schnaubelt, and C. Wieners, *Wave phenomena—mathematical analysis and numerical approximation*, vol. 49 of Oberwolfach Seminars, Birkhäuser/Springer, Cham, 2023, https://doi.org/10.1007/978-3-031-05793-9.

[7] G. Duvaut and J.-L. Lions, *Inequalities in mechanics and physics*, vol. 219 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin-New York, 1976. Translated from the French by C. W. John.

[8] I. Epanomeritakis, V. Akçelik, O. Ghattas, and J. Bielak, *A Newton-CG method for large-scale three-dimensional elastic full-waveform seismic inversion*, Inverse Problems, 24 (2008), pp. 034015, 26, https://doi.org/10.1088/0266-5611/24/3/034015.

[9] A. Fichtner, *Full seismic waveform modelling and inversion*, Advances in Geophysical and Environmental Mechanics and Mathematics, Springer-Verlag Berlin Heidelberg, 2011, https://doi.org/10.1007/978-3-642-15807-0.

[10] H. Gercek, *Poisson's ratio values for rocks*, International Journal of Rock Mechanics and Mining Sciences, 44 (2007), pp. 1–13, https://doi.org/10.1016/j.ijrmms.2006.04.011.

[11] P. L. Gould and Y. Feng, *Introduction to linear elasticity*, Springer, Cham, fourth ed., 2018, https://doi.org/10.1007/978-3-319-73885-7.

[12] K. Ito and F. Kappel, *Evolution equations and approximations*, vol. 61 of Series on Advances in Mathematics for Applied Sciences, World Scientific Publishing Co., Inc., River Edge, NJ, 2002, https://doi.org/10.1142/9789812777294.

[13] L. E. A. Jones and H. F. Wang, *Ultrasonic velocities in Cretaceous shales from the Williston basin*, Geophysics, 46 (1981), pp. 288–297, https://doi.org/10.1190/1.1441199.

[14] A. Kirsch and A. Rieder, *Inverse problems for abstract evolution equations with applications in electrodynamics and elasticity*, Inverse Problems, 32 (2016), pp. 085001, 24, https://doi.org/10.1088/0266-5611/32/8/085001.

[15] A. Kirsch and A. Rieder, *Inverse problems for abstract evolution equations II: Higher order differentiability for viscoelasticity*, SIAM J. Appl. Math., 79 (2019), pp. 2639–2662, https://doi.org/10.1137/19M1269403. Corrigendum under https://doi.org/10.48550/arXiv.2203.01309.

[16] V. Krampe, Y. Pan, and T. Bohlen, *Two-dimensional elastic full-waveform inversion of Love waves in shallow vertically transversely isotropic media: synthetic reconstruction tests*, Near Surface Geophysics, 17 (2019), pp. 449–461, https://doi.org/10.1002/nsg.12061.

[17] L. Métivier, R. Brossier, J. Virieux, and S. Operto, *Full waveform inversion and the truncated Newton method*, SIAM J. Sci. Comput., 35 (2013), pp. B401–B437, https://doi.org/10.1137/120877854.

[18] J.-W. Oh, Y. Shin, T. Alkhalifah, and D.-J. Min, *Multistage elastic full-waveform inversion for tilted transverse isotropic media*, Geophys. J. Int., 223 (2020), pp. 57–76, https://doi.org/10.1093/gji/ggaa295.

[19] J. O. Robertsson, J. O. Blanch, and W. W. Symes, *Viscoelastic finite-difference modeling*, Geophysics, 59 (1994), pp. 1444–1456, http://dx.doi.org/10.1190/1.1443701.

[20] S.-A. Sortan, *Effects of seismic anisotropy and attenuation on first-arrival waveforms recorded at the Asse II nuclear repository*, master's thesis, Karlsruhe Institute of Technology, Karlsruhe, 2022, https://doi.org/10.5445/IR/1000151635.

[21] A. Tarantola, *Theoretical background for the inversion of seismic waveforms including elasticity and attenuation*, Pure and Applied Geophysics, 128 (1988), pp. 365–399, https://doi.org/10.1007/BF01772605.

[22] L. Thomsen, *Weak elastic anisotropy*, Geophysics, 51 (1986), pp. 1954–1966, https://doi.org/10.1190/1.1442051.

[23] I. Tsvankin, *Signatures and Analysis of Reflection Data in Anisotropic Media*, vol. 29 of Handbook of Geophysical Exploration: Seismic Exploration, Pergamon, Amsterdam, 2001.

[24] P. Yang, R. Brossier, L. Métivier, J. Virieux, and W. Zhou, *A time-domain preconditioned truncated Newton approach to visco-acoustic multiparameter full waveform inversion*, SIAM J. Sci. Comput., 40 (2018), pp. B1101–B1130, https://doi.org/10.1137/17M1126126.

[25] U. Zeltmann, *The viscoelastic seismic model: existence, uniqueness and differentiability with respect to parameters*, PhD thesis, Karlsruhe Institute of Technology, 2018, http://dx.doi.org/10.5445/IR/1000093989.